ICCWCS 2019

Third International Conference on Computing and Wireless Communication Systems

Faculty of Sciences, Ibn Tofaïl University -Kénitra- Morocco

April 24-25, 2019

EDITORS

Jamal Zbitou

Adil Echchelh

Mostafa Hefnawi

Ahmed Errkik

EAI

RESEARCH MEETS INNOVATION

Proceedings of the Third Edition of The International Conference on Computing and Wireless Communication Systems (ICCWCS'19)

April 24-25, 2019, Faculty of Sciences, Ibn Tofaïl University -Kénitra- Morocco

ICCWCS 2019

ICCWCS General Chair

Pr Jamal Zbitou, FST of Settat, University of Hassan 1st , Morocco

ICCWCS Co-Chair

Pr Mostafa Hefnawi, Royal Millitary College (RMCC), Canada

ICCWCS'19 Chair

Pr Adil Echchelh, Faculty of Sciences, Ibn Tofail University, Morocco

Conference Organization

Steering Committee

Jamal Zbitou	Fsts, Hassan 1st University, Morocco
Mostafa Hefnawi	Royal Military College (Rmcc), Canada
Mohamed Latrach	Eseo-Ietr Angers, France
Adil Echchelh	Faculty of Sciences, Ibn Tofail University, Morocco
Rachid Mandry	Fsts, Hassan 1st University, Morocco
Abdellah Lakhouili	Fsts, Hassan 1st University, Morocco
Mohammed Aboulfatah	Fsts, Hassan 1st University, Morocco
Ahmed Errkik	Fsts, Hassan 1st University, Morocco
Abdelali Tajmouati	Fsts, Hassan 1st University, Morocco
Larbi El Abdellaoui	Fsts, Hassan 1st University, Morocco
Abdelmounim Elhassane	Fsts, Hassan 1st University, Morocco
Houssine Bouayad	Fsts, Hassan 1st University, Morocco

Organizing Committee

General Chair
Jamal Zbitou — Fsts, Hassan 1st University, Morocco

General Co-Chair
Mostafa Hefnawi — Royal Millitary College (Rmcc), Canada

Local Chair
Adil Echchelh — Faculty of Sciences, Ibn Tofail University, Morocco

Web Chair
Jamal Zbitou — Fsts, Hassan 1st University, Morocco

Technical Program Committee

Adil Echchelh	Fs Of Kenitra, Ibn Tofail University, Morocco
Najiba El Amrani El Idrissi	Fst Of Fez, Morocco
Aboelmagd Noureldin	Royal Military College Of Canada
Ahmed Rhallabi	Pcm Imn Nantes, France
Naoufal Raissouni	Ensa Of Tetuan, Morocco
Bernard Dousset	Ups, Toulouse, France
Benboubker Mohamed Badr	Ensa Of Tetuan, Morocco
Donghyun Baek	Chung-Ang University, Korea
Abdelmajid Badri	Fst Of Mohammedia, Morocco
Santiago Sánchez Cortés	Institute of Structure of Matter, Spain
Said Kafhali	Fst of Settat, Morocco
Fumie Costen	School Of Electrical And Electronic Engineering, University Of Manchester, UK

Mohammed Benbrahim	Ensak, Ibn Tofail University, Morocco
Hossam Hassanein	Queen's University, Kingston, Ontario, Canada
Lhoussaine Masmoudi	Mohammed V University of Rabat, Morocco
Junwu Tao	Inpt Of Toulouse, France
Mohammed El Haj Tirari	Insea, Morocco
Mourad Nedil	Engineering School, University of Quebec, Abitibi-Témiscamingue (UQAT), Canada
Hassane Moustabchir	ENSA of FES, Morocco
Malika Boudraa	Usthb, Algeria
Omar Oussouaddi	University of Moulay Ismail, Morocco
Mostafa Hefnawi	Royal Military College, Canada
Hamid Bennis	Est Of Meknes, Moulay Ismail University, Morocco
Mohamed Latrach	Eseo-Ietr Angers, France
Adnane Latif	ENSA Of Marrakech, Morocco
Mohammed El Gibari	Ietr – University Of Nantes
Khalid Haddouch	Ensa El Hoceima, Morocco
Mohamed Ben Slima	Enet'com, Sfax, Tunisia
Mohammed Berrada	Ensa of Fes, Morocco
Qingfeng Zhang	South University Of Science And Technology, China
Abdelouahid Lyhyaoui	Ensa, Tangier, Uae, Morocco
Abdelmounim Elhassane	Fsts, Hassan 1st University, Morocco
Mounir Benaissa	University Of Sfax, Tunisia
Mohamed Chaouki Abounaima	Fst Of Fez, Morocco
Mohammed Jorio	Fst Of Fez, Morocco
Jerome Tissier	Eseo-Ietr Angers, France
Abdelmjid Saka	Ensa Of Fez, Morocco
Rashid Saleem	University Of Engineering And Technology
Abdeslam Ait Madi	Ensa Of Kenitra, Morocco
Abri Mehadji	Tlemcen University, Algeria
Houssine Bouayad	Fsts, Hassan 1st University, Morocco
Soumaya Cherkaoui	University Of Sherbrooke, (Québec) Canada
Abdelhamid Zouhair	Ensa Of Al Hoceima, Morocco
Mohamed Tarbouchi	Royal Military College Of Canada, Kingston, Ontario, Canada
Abdelkbir Lahrech	UH1 Settat, Morocco
Boubker Sbihi	School Of Information Sciences, Rabat, Morocco
Quazi Rahman	Western University London, Ontario, Canada
Bennani Dosse Saad	Ensa Of Fes, Morocco
Ahmed Sulyman	Embry-Riddle Aeronautical University, USA

El Mokhtar En-Naimi	Fst Of Tangier, List Laboratory, Uae, Morocco
Larbi Talbi	University Of Quebec, Outaouais, Canada
Hicham Rouijaa	Fst Of Settat, Morocco
Tayeb A. Denidni	Inrs University, Montréal (Quebec), Canada
Mohamed Chrayah	Ensa Of Tetuan, Morocco
Pierre Langlois	Polytechnique Montréal, (Québec) Canada
Mohammed Al Achab	Ensa, Tetuan, Uae, Morocco
Elkhider Sidi Ahmed	University Of Nantes, France
Mohamed Lazaar	Ensa, Tetuan, Uae, Morocco
Nabil El Akkad	Ensa Of Fes, Morocco
Juan Manuel Gorriz	University of Granada, Spain
Otman Chakkor	Ensa Of Tetuan, Morocco
Rafik Lasri	Polydisciplinary Faculty Of Larache, Morocco
Samira Khoulji	Ensa Of Tetuan, Morocco
Sergey Mosin	Kazan Federal University, Kazan, Russia
Setti Larbi	Polydisciplinary Faculty Of Larache, Morocco
Youssef Zaz	Fs Of Tetuan , Morocco
Tran Manh Cuong	Hanoi National University Of Education, Vietnam
Youness Tabii	Ensa Of Tetuan, Morocco
Ignacio Rojas	University Of Granada, Spain
Said Mazer	Ensa Of Fes, Morocco
Ali Gharsallah	University Of Tunis El Manar, Tunisia
Mustapha Elharoussi	Fst of Settat, Morocco
Jack Legrand	University Of Nantes, France
Abdelali Tajmouati	Fsts, Hassan 1st University, Morocco
Jose Francisco Monserrat Del Rio	Polytechnic University, Valencia, Spain
Hassan Ammor	Emi of Rabat, Morocco
Arhallias Abdellah	University Of Nantes, France
Alfidi Mohammed	Ensa Of Fes, Morocco
Ahmed Lakhssassi	University Of Quebec, Canada
Semma El Alami	Fsts, Hassan 1st University, Morocco
Catalin Iulian Pruncu	University Of Birmingham, United Kingdom (Uk)
Abdelmajid El Bakkali	FST of Errachidia, Morocco
Lotfi Osman	Sup'com Of Tunisia, University Of Carthage, Tunisia
Mohssin Aoutoul	STIC Lab, Chouaib Doukkali University (UCD), Morocco
Ankur Singh Bist	Kiet Ghaziabad, India
Larbi El Abdellaoui	Fsts, Hassan 1st University, Morocco

Luciano Lamberti	Politecnico di Bari, Italy
Vladescu Alina	National Institute of Optoelectronics, Romania
Karim El Moutaouakil	Ensa El Hoceima, Morocco
Karl Dearn	University of Birmingham, UK
Noura Aknin	Fs Of Tetuan, Morocco
Sorin Vlase	University of Transilvania, Romania
Benslimane Mohamed	EST Usmba, Morocco
Laurent Dubar	University of Valenciennes, France
Chalh Zakaria	Ensa Of Fes, Morocco
Federico Sciammarella	Northern Illinois University, USA
Hakim Naceur	University of Valenciennes, France
Hakim Allali	Fsts, Hassan 1st University, Morocco
Zitouni Azari	University of Lorraine, France
Pascal Dherbécourt	University of Rouen Normandie, France
Ahmed El Oualkadi	University Abdelmalek Essaadi, Morocco

Preface

The Third Edition of the International Conference on Computing and Wireless Communication Systems (ICCWCS 2019) has been held in the Faculty of Sciences, university Ibn Tofaïl Kenitra, Morocco. In collaboration with (Association Marocaine de l'Innovation et de la Recherche Scientifique , Morocco) on 24-25 April 2019. Its objective is to provide an ideal platform to discuss, to share experiences and to open opportunities for collaboration between the different existing working groups in the areas of engineering and information technology. It is also an excellent opportunity for young researchers to start in the work of publication and dissemination of their work and contributions of research.

The technical program of ICCWCS 2019 consisted of 93 full papers. The conference tracks were: Track 1 - Antennas & Propagation; Track 2 - Passive & Active Components, Circuits & Subsystems; Track 3 - Computer Science, and Information Technology; Track 4 - Multimedia and Signal Processing and Track 5 - Renewable Energy. Aside from the high quality technical paper presentations, the technical program also featured five keynote speeches. The keynote speeches were Arnberg Ronny, Project Manager at IVL Swedish Environmental Research Institute,in Sweden; Dr. Catalin Iulian Pruncu, University of Birmingham & IMI Truflo Marine Ltd in UK; Prof. Mostafa Hefnawi, Royal Military College (RMCC), in Canada; Prof. Mohamed Latrach, ESEO Angers, in France; and Prof. Ahmed Lakhssassi, Quebec University, Outaouais, in Canada.

Coordination with the steering chairs, Jamal Zbitou, Mostafa Hefnawi, Mohamed Latrach, Adil Echchelh, Rachid Mandry, Abdellah Lakhouili, Mohammed Aboulfatah, Ahmed Errkik. Abdelali Tajmouati, Larbi El Abdellaoui Abdelmounim Elhassane, , and Houssine Bouayad, was essential for the success of the conference. We sincerely appreciate their constant support and guidance. It was also a great pleasure to work with such an excellent organizing committee team for their hard work in organizing and supporting the conference. In particular, the Technical Program Committee.

We strongly believe that Third Edition of ICCWCS conference provides a good forum for all researcher, developers. We also expect that the future ICCWCS conference will be as successful, as indicated by the contributions presented in this volume.

Dr. Jamal Zbitou, Dr. Mostafa Hefnawi and Dr. Adil Echchelh

Contents

Antennas & Propagation

Multiband Planar Inverted F Antenna (PIFA) for ISM, WLAN and WiMAX Applications 1
Driss EL NABAOUI, Abdelali TAJMOUATI, Jamal Zbitou, Ahmed Errkik, Larbi Elabdellaoui, Hamid Bennis, Mohamed Latrach

Antenna with Defected Ground Structure for future mobile and Wireless communication 5G 11
Anouar Mondir, Larbi Setti

Analysis of a 38 GHz Band Antenna for 5G wireless communication using Wave Concept Iterative Procedure (WCIP) method 23
Doae EL HADRI, Asmaa ZUGARI, Alia ZAKRITI

A New Design of UHF RFID Tag Based on Split Ring Resonator For Moroccan Frequency Band 32
Latifa EL AHMAR, Ahmed ERRKIK, Jamal ZBITOU, Ilham BOUZIDA, Brahim LAKSSIR

New Design of a Multiband Microstrip Antenna 40
Elkilani Samir, El Abdellaoui Larbi, Tajmouati Abdlali, Errkik Ahmed, Latrach Mohamed

Dual-band UWB antenna for Surface Movement Radar 46
Salaheddine AOURIK, Ahmed ERRKIK, Larbi EL ABDELLAOUI, Abdelali TAJMOUATI, Mohamed LATRACH

New Design of a Microstrip-Fed Printed Slot UWB Antenna with Hilbert Fractal Defected Ground Structure 57
Khadidja Krim, Mehadji Abri, Hadjira Badaoui, Samira Fersiti, Fatima Mahieddine

Design of a Miniature Fractal Microstrip Antenna for RFID Tag 64
Mohamed Ihamji, El Hassan Abdelmounim, Hamid Bennis, Mohamed Latrach

Phased array antenna for millimeter-wave application 73
Oumaima Nayat-Ali, Otman El Mrabet, Mariam Aznabet, Fatima Khoutar

Double Layer Metamaterial Superstrate for Performance Enhancement of a Microstrip Patch Antenna 82
Fatima Zohra Khoutar, Mariam Aznabet, Otman EL Mrabet, Nayat-Ali Oumaima

Invention of an Original Tetra-Generations Patch Antenna for the New Generation of Mobile Telephony and the Study of the Thermal Effect of GSM on the Human Head and Hand 90
JALAL BAAYER, HASSAN AMMOR

A New Design of a Miniaturized UHF-RFID Passive Tag Antenna Based on L-Shape Radiators with Meandered Dipole 103
Yassine Gmih, Abdelmajid Farchi

A Novel Compact Antenna for Passive UHF RFID Tag Using π-Match and Meander Techniques 113
Younes EL HACHIMI, Abdelmajid FARCHI

A Metamaterial inspired elliptical slotted antenna for WLAN and uplink /downlink X-bands Rejection 122
Mourad Elhabchi, Mohamed Srifi, Raja Touahni

Passive & Active Components, Circuits and Subsystems

New Technologies of Power Transistors for Efficiency Increase of Power Converters: The Reliability Consideration 131
Pascal Dherbecourt, Ahmed El Oualkadi, Eric Joubert, Al Mehdi Bouchour, Wadia Jouha, Mohamed Masmoudi, Olivier Latry

Temperature Effects of GaN HEMTs on the Design of Power Converters 141
Al Mehdi BOUCHOUR, Pascal Dherbécourt, Olivier Latry, Ahmed El Oualkadi

Square Complementary Split Ring Resonator (CSRR) Low Pass Filter 151
Bendaoued Mohammed, Mandry Rachid, El Abdellaoui Larbi, Aytouna Fouad, Latrach mohamed

A Novel Periodic CPW UWB Low Pass Filter With Large Rejected Band 159
Fouad Aytouna, Chakkor Otman, Latrach Mohamed

A SPICE IGBT Model with easy parameters extraction 167
Yassine Hadini, Abdelghafour Galadi, Adil Echchelh

Design of a Microstrip Balanced Amplifier Using the Wilkinson Power Divider 176
amine rachakh, Larbi El Abdellaoui, Ahmed Errkik, Abdelali Tajmouati, Mohamed Latrach

A Novel Zero Bias Microstrip MESFET Power Limiter 185
Khalifa ECHCHAKHAOUI, Elhassane ABDELMOUNIM, Hamid BENNIS, Mohamed LATRACH

Broadband Solid State GaAs Power Amplifier for L and S Bands Applications. 196
Mohamed Ribate, Rachid Mandry, Larbi El Abdellaoui, Ahmed Errkik, Mohamed Latrach

Microwave Absorption Properties of Garnet Ferrite Filled Polymer Nanocomposites in X-Band 206
Souhail Alami Massmoudi, H. Ammor

Contribution to the numerical study of the thermal gradient during the fabrication of composite materials by the RTM process. 214
Soukayna ELYOUSSFI, Abdellatif EL OUJDI, SAAD aouatif, HATTABI Mohamed, ECHCHELH Adil

Computer Science & Information Technology

Optimisation of asset allocation and reinsurance strategy by the mean of a genetic algorithm : NSGA-II 221
Mehdi Echchelh

Extensive Analysis of Clustered Routing Protocols For Heteregeneous Sensor Networks 231
Yassine Yazid, Imad Ez-zazi, Marouan Salhaoui, Mounir Arioua, El Oualkadi Ahmed, Antonio González

Query expansion using Wikidata attributes' values 241
Sarah Dahir, Abderrahim El Qadi, Hamid Bennis

Numerical characterization of Symmetric Delamination in Orthotropic plate using Lamb waves 248
Salah Nissabouri, Mhammed El Allami, El hassan Boutyour

Data mining for road accident analysis in a big data context 255
Fatima Zahra El Mazouri, Mohammed Chaouki Abounaima, Said Najah, Khalid Zenkouar

Comparative Study of ZigBee and 6LoWPAN Protocols: Review 264
IMANE ALAOUI ISMAILI, Abdelilah Azyat, Naoufal Raissouni, Nizar Ben Achhab, Asaad Chahboun, Mohammed Lahraoua

Artificial Intelligence based Composition for E-Government Services 273
Amina adadi, Mohammed BERRADA, Nabil EL AKKAD

Innovative Models Proposals for the Implementation of Adaptive Hypermedia Systems 281
Mehdi TMIMI, Mohamed BENSLIMANE, Mohammed BERRADA, Kamar OUAZZANI

A New Approach to Integrate The RFID in the Internet of Things Using The MQTT Protocol and 6lo-RFID Framwork 293
Youssef Mouzouna, Mohamed Taouzari, Ahmed Mouhsen, Said Benalla, Jamal El Aoufi, Hanane Nasraoui

Application of game theory to the optimization of wireless systems Algorithmic solutions 302
Sara Riahi, Azzeddine Riahi

Comparative study of performing features applied in CNN architectures 313
SARA LAROUI, HICHAM OMARA, Mohamed LAZAAR, Oussama MAHBOUB

A comparative study of CHN-MNC, GA and PSO for solving constraints satisfaction problems 326
Adil Bouhouch, Chakir Loqman, Hamid Bennis, Abderrahim El Qadi

The Hybrid Machine Learning Model Based on Random Forest Optimized by PSO and ACO for Predicting Heart Disease 335
Youness KHOURDIFI, Mohamed BAHAJ

An Energy Balanced Algorithm for data communication in WSN 345
HAJAR LAGRAINI, HICHAM OULDZIRA, MOSTAFA CHHIBA, ABDELMOUMEN TABYAOUI, AHMED MOUHSEN

Investigating fading impact on AODV, OLSR and DSDV protocols in NS3 354
Hanene Brahmia, Cherif Tolba

Numerical approach of determining the Stress Intensity Factor (KI) for a curved crack problem using Extended Isogeometric Analysis 364
SAID ELFAKKOUSSI, Hassane MOUSTABCHIR, AHMED ELKHALFI

Reference Model of Open Distributed Processing Basic Modelling Concepts in Event-B 374
Abdessamad Jarrar, Nawfal FERFRA, Taoufiq Gadi, Youssef Balouki

GPU optimized parallel implementation of NaSch traffic model 383
Yassine El Hafid, Abdessamad El Rharras, Mohamed Wahbi, Rachid Saadane

Multimedia and Signal Processing

Non Negative Matrix Factorization for Blind Source Separation 395
Nabila Aoulass, Otman Chakkor

Reconstruction of three-dimensional scenes from a sequence of images of which only one is calibrated 406
Karima KARIM, Nabil El AKKAD, Khalid SATORI, Mohamed ES-Sabry, Mostafa MERRAS, Abderrahim SAAIDI

An approch for FPGA implementation of LST-SW algorithm 416
Assaad El Makhloufi, Samir El Adib, Naoufal Raissouni

Hybrid Beamforming for Relay-Aided mmWave Backhaul links 426
Mostafa Hefnawi, Jamal Zbitou, Mohamed Aboulfatah, Hassan Abdelmounim

Noise Reduction using DUET algorithm for dual-microphone mobile station 435
Nabila Aoulass, Otman Chakkor

Sharing a Secret Image Based in CMY model using RSA Method in Visual Cryptography 445
Asmaa HILMI, Soufiane MEZROUI, Ahmed EL OUALKADI

A Wavelet-Based ECG Delineation and Automated Diagnosis of Myocardial Infarction in PTB Database 454
RACHID HADDADI, Elhassane Abdelmounim, Mustapha El Hanine, Abdelaziz Belaguid

A Layout Investigation of Large Wind Farm in Akhfennir using Real Coded Genetic Algorithm 464
Mohammed Amine Hassoine, Fouad Lahlou, Adnane Addaim, Abdessalam Ait Madi

Renewable Energy

New Model and Sliding mode control of a three phase grid connected photovoltaic system without DC-DC Converter 474
BAHRI Hicham, Aboulfatah Mohamed, Guisser M'hamed, Elhasan Abdelmounim, El Malah Mohammed, Adekanle simon

Power Control of a Wind Energy Conversion System based on a Doubly Fed Induction Generator 484
Bouchaib RACHED, Mustapha ELHAROUSSI, Elhassane ABDELMOUNIM

Boost converter system modelling and incremental conductance algorithm for photovoltaic system via Matlab/Simulink 491
Abdessamad Benlafkih, Mohamed Chafik Elidrissi

Miniature 2.45 GHz Patch Antenna with Defected Ground Structure for Rectenna Applications 501
Abdellah TAYBI, Abdelali TAJMOUATI, Jamal ZBITOU, Ahmed ERRKIK, Larbi EL ABDELLAOUI, Mohamed LATRACH

Doubly Fed Induction Generator Wind Turbine Control Based on a Novel MRAS Observer 508
Oluwaseun Adekanle, M'hammed Guisser, Elhassane Abdelmounim, Mohamed Aboulfatah, Hicham Bahri, Adeola Balogun

A New Modeling of the Junction Metal Semiconductor for Rectenna System 519
Abdelhak Hamid allah, Ahmed Errkik, Jamal Zbitou, Larbi El Abdellaoui, Abdelali Tajamouati, Mohamed Latrach

Valorization of by-products of fish by reactive extrusion to produce biofuel 528
Dounia Makoure, Abdellah Arhaliass, Adil Echchelh, Jack Legrand

Estimation of the minimum mean square error channel MMSE And energy efficiency management 540
Sara Riahi, Azzeddine Riahi

Multiband Planar Inverted F Antenna (PIFA) for ISM, WLAN and WiMAX Applications

D. EL NABAOUI[1], A. TAJMOUATI[1], J. ZBITOU[1], H. BENNIS[2], A. ERRKIK[1], L. ELABDELAOUI[1], M. LATRACH[3]
[1] LMEET, FST of Settat, Hassan 1st Univeristy, Settat, Morocco
[2]TIM Research Team, EST of Meknes, Moulay Ismail University, Meknes, Morocco
[3]Microwave Group, ESEO, Angers, France

{ Driss.elnabaoui@gmail.com }

Abstract. This paper presents a new design of a multi-band inverted F antenna structure (PIFA) fed by a coaxial line. The design tool is the CST-MW software which uses the finite element method. To create new resonances frequencies it is necessary to insert slits of various forms on the radiation element. The final validated circuit is designed to operate in the Industrial Scientific and Medical (ISM 2.4 - 2.5GHz), Wireless Local Area Network (WLAN 5.15 – 5.35 GHZ) and Worldwide Interoperability for Microwave Acces (WiMAX 3.30 – 3.80). Obviously, it has been found that the radiation patterns of the antenna shown are similar to the bidirectional radiation pattern. For simulated The properties of the antenna, such as return losses and radiation patterns, we used CST-Mw Studio.

Keywords: *PIFA antenna, multibands antennas, slottedPIFAs, Simulation.*

1 Introduction

Currently, it requests wireless communication system technologies, which have the following attributes compact size, low profile, multi-band, high-bandwidth, etc…. In order to response the rapidly growing demand, an antenna should be responsible and operate in many frequency bands. Recently, multi-band antennas developments were improved thanks to the notion PIFA. Planar Inverted F Antenna (PIFA) is a linear Inverted F antenna (IFA) [1]. To increase the bandwidth and resonance frequency, the radiator element must be replaced by a plate. The reduced radiation behind the PIFA improves antenna performance in terms of gain and minimizes the wave produced due to power absorption [2-4]. We have a maximum gain in terms of polarization states like horizontal and vertical. The bandwidth can be adjusted if changing the height, length and width of the ground plane. To decrease the quality factor and to boost the bandwidth, many slits can insert in the ground plane. In medical applications, the main objective is to reduce the maximum size of the PIFA [5].

2 Planar Inverted Antenna F (PIFA)

Planar Inverted F antenna is developed from mono pole antenna. Inverted L is realized by folding down the mono pole in order to decrease the height of the antenna at the same time maintaining identical resonating length [6]. When feed is applied to the Inverted L, the antenna

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284080

appears as Inverted F. The thin top wire of Inverted F is replaced by planar element to get the Planar Inverted F antenna. This sequence is clearly observable in Fig 1. [7-9]

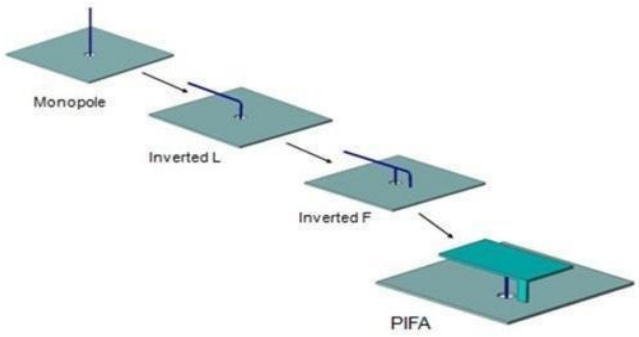


Fig. 1. PIFA from monopole

2.1 Pifa Design

PIFA consists of ground plane, radiating patch above the ground plane and shorting plane. A coaxial probe feed is given between the ground plane and patch element. Top radiating patch plane is folded at one edge of a patch and shorted to the ground plane to decrease the antenna length as shown in Fig 2 [10-12]. The size of the patch and resonating frequency can be determined by the following equations:

$$L_p + W_p - W = \frac{\lambda}{4} \tag{1}$$

$$f_r \times \lambda = \frac{C}{\sqrt{\varepsilon_r}} \tag{2}$$

$$f_r = \frac{C}{4\left(L_p + W_p - W\right)\sqrt{\varepsilon_r}} \tag{3}$$

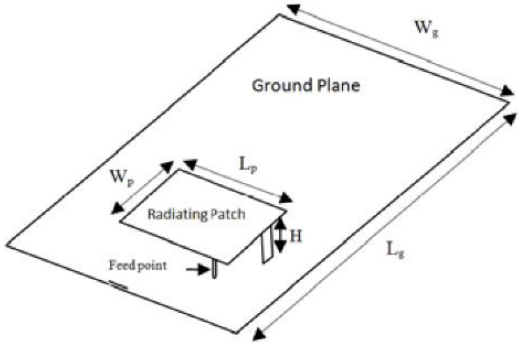


Fig. 2. Basic PIFA

Where Lp = length of the patch, Wp = Width of the patch, C = Velocity of light, εr = dielectric constant, ß = wavelength. PIFA has moderate (or) high gain in both horizontal and vertical polarization. Generally, most of the wireless systems use vertical polarization. Even if transmitter antenna polarization is not known, still the signal is received with good strength. When antenna orientation is not fixed, a signal with good gain (greater than 10 dB) is received and signal strength is calculated by summing up the horizontal and vertical components [13-15].

2 Antenna Description

The design of the proposed antenna is shown in Fig 3. It consists of patch plane, ground plane, shorting plate and feeding post connected to the ground plane. Between dielectric medium and patch plate, air is placed. Resonating frequency can be calculated if the initial patch and shorting pin sizes are known using equation (3).

The dimensions of PIFA are 50×46 mm^2 and are located 5 mm above the printed circuit board. The printed circuit board layer has a relative permittivity of 4.4 (FR4_epoxy) of size 93×60×0.5 mm^3. To provide an RF mass, the PCB is metallized on its back surface. Using optimization, PIFA is operating at 2.45 GHz, 3.6 GHz and 5.2 GHz resonance frequencies to cover ISM, WiMAX and WLAN bandwidths. The proposed antenna is powered by a coaxial probe. The U-shaped slot is introduced on the patch plane to obtain the multi-band.

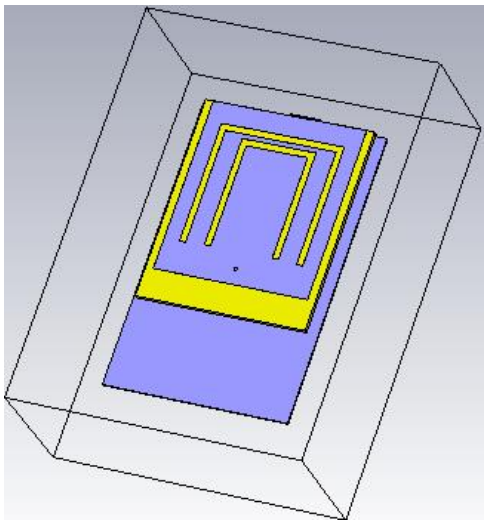


Fig. 3. PIFA design

Antenna geometry is shown in Table 1 and antenna description is shown in Table 2.

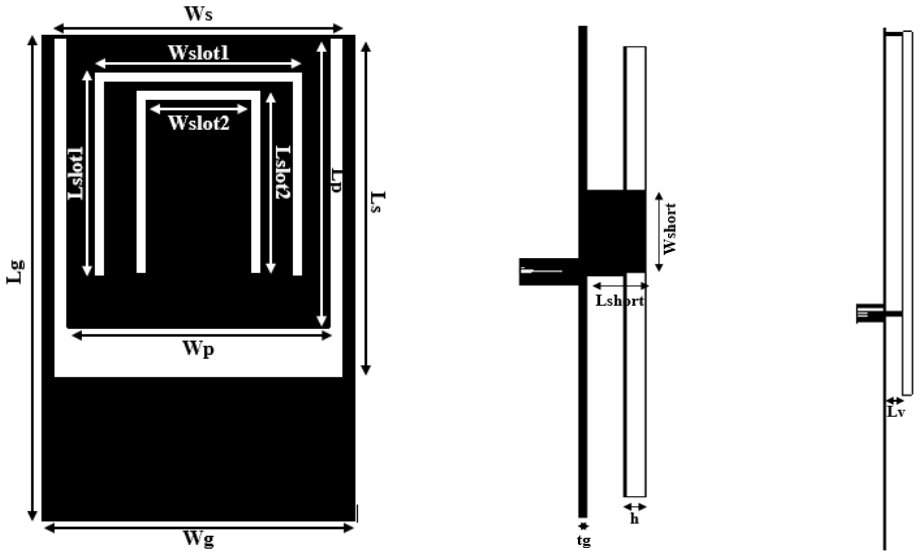
Table 1. Dimension of the proposed antenna

Parameter	Optimized Value (mm)
Ws	55
Ls	65
H	1.6
Wp	50
Lp	55.5
t	0.035
Wg	60
Lg	93
Lshort	5
tg	0.5
Wshort	10
Ri	0.5
Vcentre	18
Ro	1.637
Wslot1	40
Lslot1	39
Wslot2	24
Lslot1	35
Lv	3

Table 2. Antenna description

Shape	**Rectangular**
Frequency of operation	ISM WLAN WiMAX
Dielectric constant of the substrate	FR4 Epoxy (4.4)
Height of the dielectric substrate	1.6 mm
Feeding method	Coaxial
Gain	(2-7) dB

The dimensions of different slots are clearly mentioned in the front view of antenna as shown in Fig 4.



(a) (b) (c)

Fig. 4. Geometry of the proposed antenna: (a) front view, (b) top view, (c) side view

4 Simulation Results

The Planar Inverted F antenna was analyzed and optimized with the CST Microwave Studio. Since generally PIFA is a high frequency device driven model is used while designing antenna in CST Microwave Studio. The simulation results in all cases were swept over a frequency range between 2 and 7 GHz.

4.1 Return Loss

Reflection coefficient of an antenna is an important parameter in any antenna design analysis. It gives the measure of how much amount of power will be reflected back from the antenna. It should be kept as minimum as possible.

The simulated result is shown in Figure 5, as we can see we have reached the suitable frequency bands as ISM, WLAN and WiMAX.

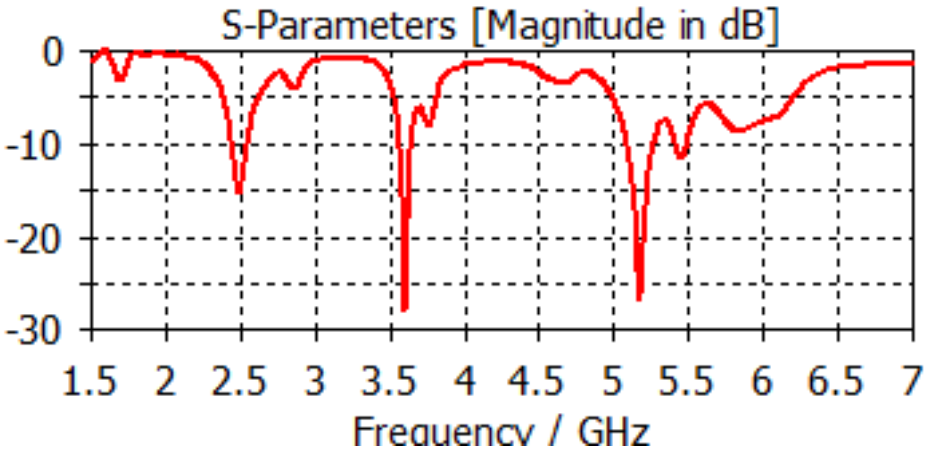


Fig. 5. Simulated Return Loss versus frequency

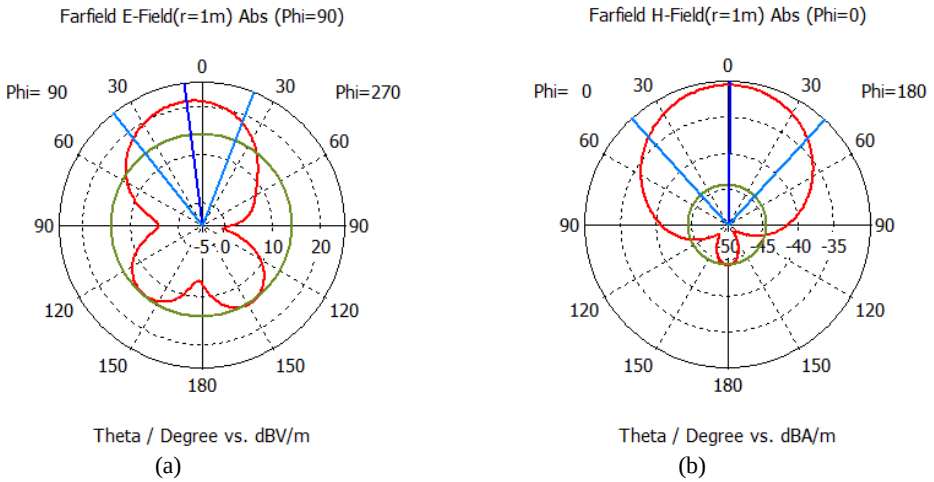
Table 3. Frequency bandwidth versus return-loss value

Frequency	Return Loss in dB
2.41 – 2.58 GHz	-15.26
3.51 – 3.69 GHz	-27.76
5.04 – 5.31 GHz	-26.62

4.2 Radiation Pattern

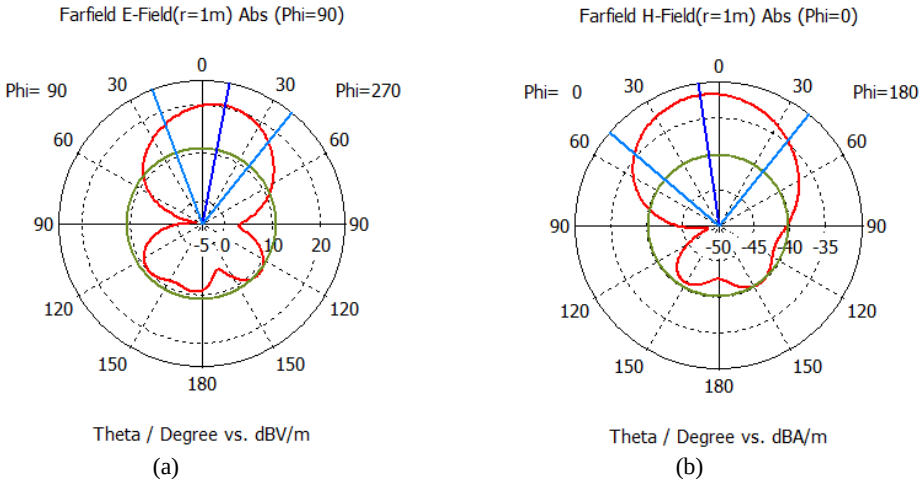
The radiation pattern of an antenna is used to describe the 3-dimensional radiation characteristics.

Figure 6, 7 and 8 presents the radiation pattern for different center frequencies, which give a stable radiation for different frequency bands.



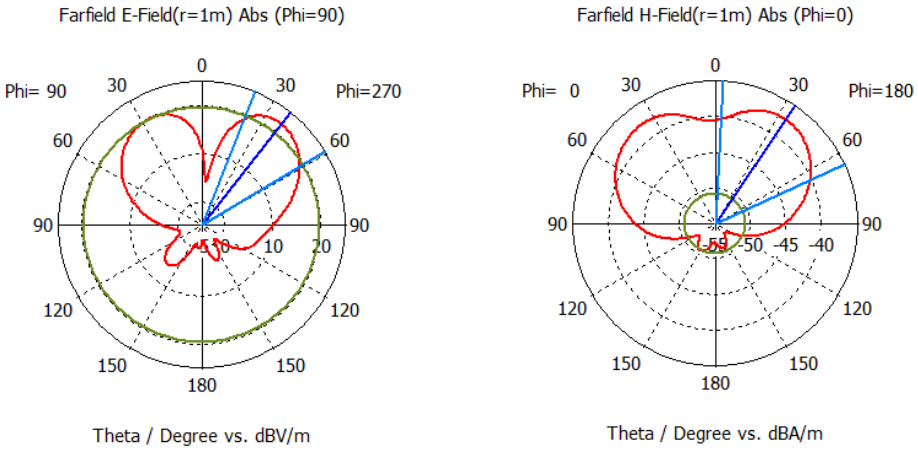
(a) (b)

Fig. 6. Radiation pattern of the proposed antenna at 2.45 GHz in the (a) E-plane and (b) H-plane



(a) (b)

Fig. 7. Radiation pattern of the proposed antenna at 3.60 GHz in the (a) E-plane and (b) H-plane



(a) (b)

Fig. 8. Radiation pattern of the proposed antenna at 5.20 GHz in the (a) E-plane and (b) H-plane

Figure 9, illustrates the variation of the gain versus frequency. After the simulation, we have obtained the gain 6.75dB at 2.45GHz, 5.58dB at 3.6GHz and 7.14dB at 5.2GHz.

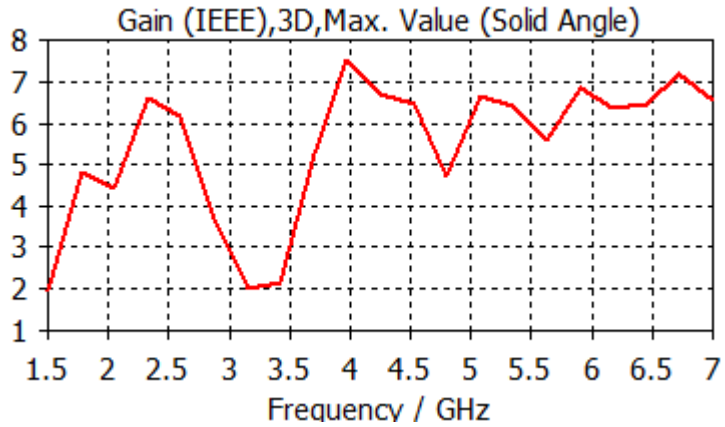
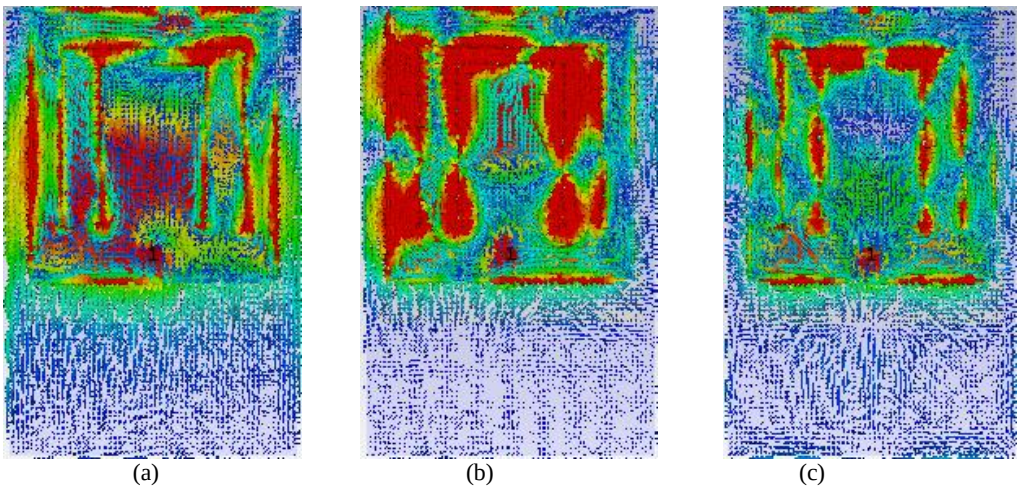


Fig. 9. Simulated gain of the proposed antenna

We present the simulated surface current distribution of the proposed antenna in the Fig. 10 at different frequencies.



(a) (b) (c)

Fig. 10. Simulated surface current distribution at (a) 2.45GHz, (b) 3.6GHz and (c) 5.2GHz

Conclusion

In this work, we have validated into simulation a Planar Inverted Antenna (PIFA) based on the use of a coaxial line. To validate this antenna structure, we have used CST-MW. The PIFA covers a bandwidth of 170MHz (2.41-2.58GHz), 180MHz (3.51-3.69GHz) and 270MHz (5.04-5.31GHz) and lower and higher bands with directivity of 7.22dB, 7.49dB and 8.24dB at lower and higher resonating frequencies 2.45GHz, 3.60GHz and 5.20GHz respectively. This antenna structure has a stable bidirectional radiation pattern and good input impedance matching with a significant bandwidth.

References

[1] David M. POZAR “Microwave Engineering” Fourth Edition.

[2] Saad Wasmi Luhaib, Kaydar M. Quboa and Bareq M. Abaoy "Design and Simulation Dual-Band PIFA Antenna for GSM Systems” 9 thin ternational multi conference on systems 2012.

[3] Chenhui Jiang “Microwave and Millimeter-Wave Integrated Circuit Systems in Packaging” January 19, 2010.

[4] Adnan Iftikhar, Muhammad Nadeem Raftiq, "A Dual band balanced planar inverted F antenna (PIFA) for mobile applications” IEEE Proc-Microwave, Antennas Propagation, Vol.149, No,2, pp. 85-91,2013.

[5] D. El Nabaoui, A. Tajmouati, J. Zbitou, A. Errkik, L. Elabdellaoui, A. Mediavilla “A New Design of a Low-Cost Multiband Fractal CPW-Fed Antenna”, International Journal of Microwave and Optical Technology (IJMOT), Vol. 13, No. 1, PP. 1-7, January 2018.

[6] Constantine A. Balanis “Antenna Theory Analysis and Design” Third Edition.

[7] P. Nepa, G. Manara, A.A. Serra, g. Nenna. “Multiband PIFA for WLAN mobile terminals”, IEEE Antennas Wireless Propagat Letters, 2005, vol.4, pp. 349 – 350.

[8] C. Mahatthanajatuphat, S. Saleekaw, P. Akkaraekthalin and M. Krairiksh “A Rhombic patch monopole Antenna with ModifiedMinkowski Fractal Geometry for UMTS, WLAN, and Mobile WIMAX Application”, Progress in Electromagnetics Research, PIER Vol. 89, pp. 57–74, 2009.

[9] G.Jeevagan navukarasu lenin , Ganesh babu.T, Rajkumar.R , Ramanathan.A “Design of an E Shaped Patch Antenna for GPS and IRNSS Application”, 2016 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp. 179-183, July 2016.

[10] Antennas for all applications" 3rd edition by John D Krauss
[11] P. Salonen, M. Keskilammi, and M. Kivikoski, "Single-Feed Dual-Band Planar Inverted-F Antenna with U-Shaped Slot", IEEE Trans. Antennas and Propagat., Vol. 48, No. 8, pp. 1262-1264, Aug. 2000.
[12] Yamina BELHADEF and Nourediene BOUKLI HACENE "Design of New Multiband Slotted PIFA Antennas", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011.
[13] Seyed Ehsan Hosseini, Amir Reza Attari and Aref Pourzadi "A Multiband PIFA with a Slot on the Ground Plane for Wireless Applications", International Journal of Information and Electronics Engineering, Vol. 3, No. 4, July 2013
[14] Naveen Kumar and Garima Saini « A Multiband Pifa with Slotted Ground Plane for Personal Communication Handheld Devices", International Journal of Engineering Research and Development, Volume 7, Issue 11, PP. 70-74, July 2013.

Antenna with Defected Ground Structure for future mobile and Wireless communication 5G

[1]Mondir Anouar, [2]Setti Larbi

{ [1]mondir.anouar06@gmail.com, [2]settilarbi@gmail.com }

FPL, Department of physic
Abdelmal Essaadi university Avenue Khenifra, Tétouan 93000,Morocco

Abstract. The scope of this study is to design and develop a miniature antenna for future mobile and wireless communication 5G.On this paper is designed a new type DGS antenna for future mobile. The size of this antenna is 5.33mm width and 6.67mm long, with dielectric constant ε_r=2. 2 and thickness H=0. 8mm. The new patch antenna has T&U slotted and notches.50 ohm single microstrip line feeding with radiation patch, which is resonant with the dual band frequencies. The design and simulation results are analyzed the return loss, VSWR, surface current distribution, Fairfield radiation pattern and Fairfield gain. This antenna obtained the return loss -20.16576dB and -28.7672dB which cover the27.58443GHz to 27.652 GHz and 29.284GHz to 29.318 GHz band. The Fairfield gain have a 7.35dB and 5.11dB.This antenna operates dual band frequency 27.626Gz and 29.303 GHz for 5G communication is presented..

Keywords: DGS antenna theory, filter DGS.T&U Slotted,notches,5G communication,mmwave

1 Introduction

Utilization of the millimeter wave (mmWave) band is a key technology for the development of the heterogeneous networks (HetNets) that will be used for 5th generation wireless cellular networks (5G) [1].

The beams generated in mmwave (band ka) are much more directive than in lower frequency bands. This why it's reserve for the communications satellite. Its frequency band between 26.5 GHz and 40 GHz, The big problem when used mmwave absorption and attenuation with environment , For example, the oxygen molecule (O2) normally absorb electromagnetic energy at 60Ghz to a much higher degree than in the regions 30-160Ghz . The rain attenuation and atmospheric absorption characteristics of mmwave propagation. The solution proposed for this type of frequency band it is of added by small base station with Distance less than 1Km. Moreover used (MIMO)multiple antennas at the transmitter and /or receive improve the global performance of wireless systems .

For working the huge of frequency band size of antenna change smaller and become a many defiance for fabrication**.** This paper, design a new antenna by depends on merging three ideas in [1] [2] [3] which works the dual band of f1=27. 632 GHz, f2=29.3032 GHz frequency use for future 5G wireless and mobile communication is presented**.** In [3] is created a novel topology of antenna patch with mesh-grille, and also is proposed a position for the antenna to be installed that avoids the problem of attenuation of the wave with the user. In [4] is developed a new concept of antenna patch with a slot with shape 'T' and 'L' to be used in band frequency 28Ghs. In [5] is proposed a circular patch with DGS that improves the performance of other antennas.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284110

In [6] is proposed an antenna with a 'U' slot bi-band in 28Ghz and 38Ghz, and added an array of an antenna to improve their performance.

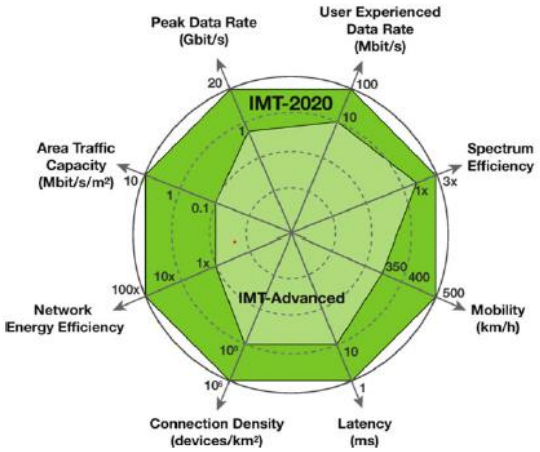


Fig. 1. Enhancement of key capabilities from IMT-Advanced to IMT-2020[13]

2 Geomrtry of new type

The geometry of the new antenna is presented in fig (1). The size of this antenna is 5mm width and 6.33mm long, with dielectric constant ε_r=2. 2 with thickness H=0. 8mm. The new patch antenna has T&U slotted and notches.50 ohm single microstrip line feeding with radiation patch, which is resonant with the dual band frequencies.

Before calculating the length of the patch it is necessary to calculate the effective dielectric constant, effective with and length extension of the resulting length by using the equation:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}[1 + 12\frac{h}{W}]^{-\frac{1}{2}} \qquad 1.1$$

$$L_{eff} = L + 2\Delta L \qquad 2.2$$

$$L_{eff} = \frac{c}{2f_0\sqrt{\varepsilon_{eff}}} \qquad 1.3$$

All the dimensions presented in table.1, table.2, table.3.

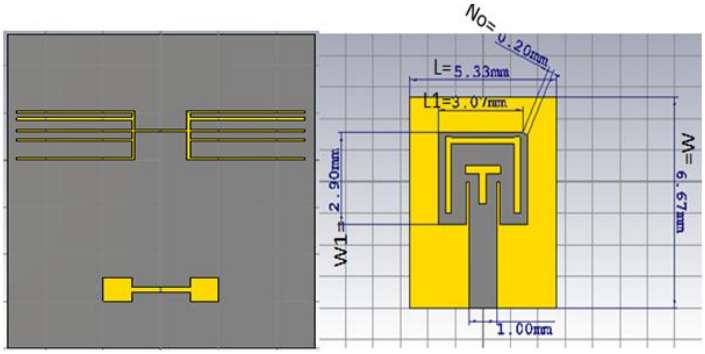
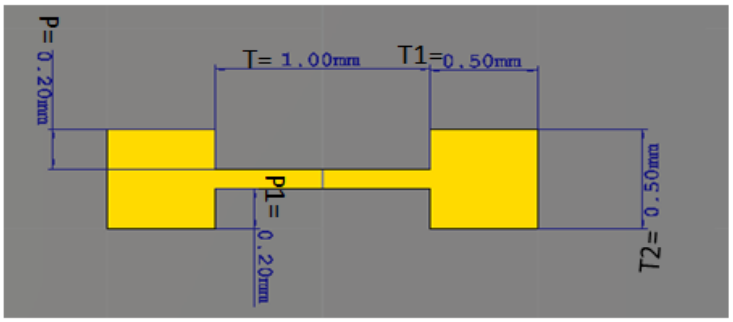


Fig. 2. Geometry of proposed antenna with DSG unit cell and DGS filter

Table 1: dimensions of patch antenna

Parameters	**Dimensions (mm)**
L	5.33
W	6.67
L1	3.07
W1	2.9
NO	0.2

Fig.3 and table.2 present the geometry and dimensions of the filter DGS .



. **Fig.3**. Geometry of proposed DGS filter

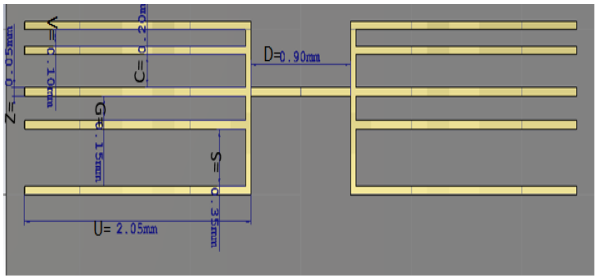


Fig.4. Geometry of proposed DGS unit cell

Table 2: dimensions of DGS filter.

Parameters	**Dimensions (mm)**
P	0.2
T	1
T1	0.5
T2	0.5
P1	0.2

And the sconde table.3. present the geometry of unit cell

Table 3: dimensions of DGS filter

Parameters	**Dimensions (mm)**
D	0.9
C	0.3
G	0.15
V	0.1
Z	0.05
U	2.05
S	0.35

3 DGS antenna theory

In general, circuit equivalent of this structure DGS to consist of a circuit LC parallel and series with the transmission line:
Fig.5 represent DGS equivalent .

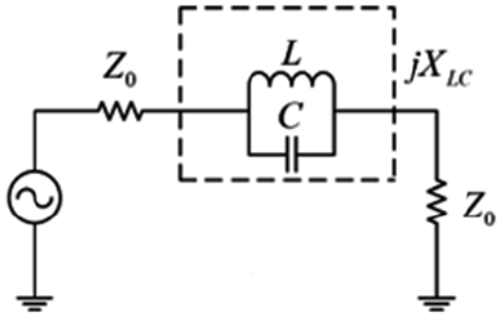


Fig.5. Equivalent circuit LC

The next calcul we take into consideration the lossless we insert resistance in parallel like illustrated in fig (2)

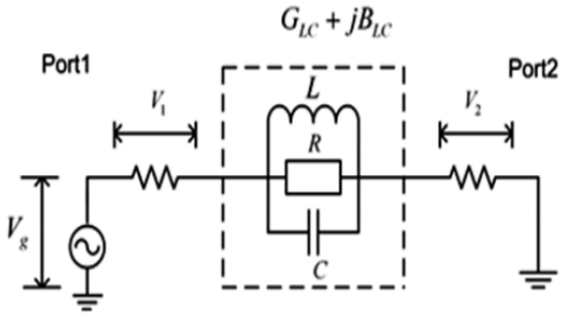


Fig.6. Equivalent circuit LRC

So the equivalent circuit is given by:

$$Y = \frac{1}{R} + j2\pi(Cf - \frac{1}{4\pi^2 Lf})$$

Because:

$$Z = \frac{1}{Y}$$

$$Z = \frac{1}{\frac{1}{R} + j2\pi(Cf - \frac{1}{4\pi^2 Lf})}$$

And we know:

$$S_{21} = \frac{1}{1 + \frac{1}{2} Z Z_0^{-1}}$$

At -3db we find:

$$| S_{21} | = \frac{2Z_0}{\sqrt{4Z_0^2 + (\frac{w_c}{w_0^2 - w_c^2})^2}} = \frac{1}{\sqrt{2}}$$

- wc is the angular cutoff frequency.

- w0 is the angular resonance frequency

At the resonance frequency we calculate the resistance R replacing Ze = R and we find:

$$S_{21}\ |_{w=w_0} = |\ \frac{2Z_0}{2Z_0 + Z_e}\ | = \frac{2Z_0}{2Z_0 + R}$$

This gives:

$$R = \frac{2Z_0(1 - S_{21}\ |_{w=w_0})}{S_{21}\ |_{w=w_0}}$$

The current distribution through the patch, by increasing the length electric current Which leads to increased capacitance and the inductor and consequently the antenna has characteristics slow wave.
The slow wave factor (SWF) shows the relationship between the number in free space, k0 and the propagation constant of the line of transmission.SWF is determined by:

$$SWF = \sqrt{\varepsilon_e}$$

With :

$$\varepsilon_e = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}[1 + 12\frac{h}{w}]^{-\frac{1}{2}}$$

In addition, the propagation constant is given by:

$$\beta = \sqrt{\varepsilon}_e.k_0$$

Which gives therefore:

$$SWF = \frac{\beta}{k_0}$$

4 Simulated results

The DGS filtre and unit cell MMW frequency antenna for 5G networks is presented in Fig.5.

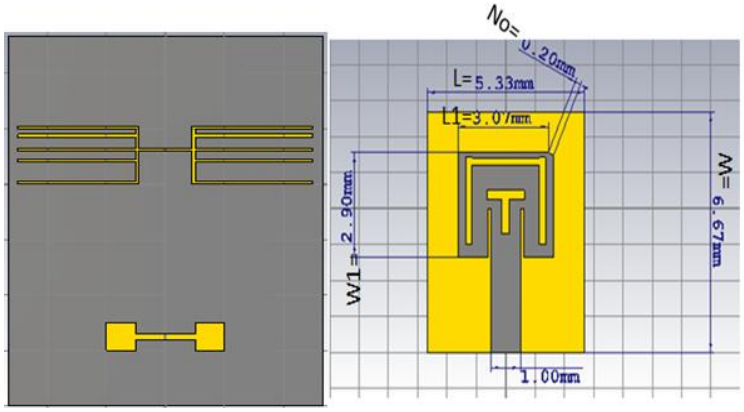


Fig.5. Geometry of proposed antenna

At first we begin with the introduction of a unit cell fig 6 and in one Second time we introduce the filter DGS fig.11.

4.1 Patch antenna with unit cell :

Fig.6 illustrates the patch scheme with the DGS cell unit engraved in the ground plane.

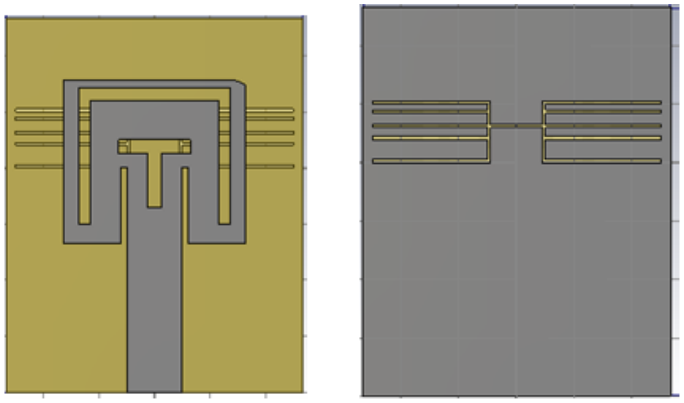


Fig.6. the DGS cell unit engraved in the ground plane.

The engraving of the slots and notches on the patch we note an improvement at the level of adaptation by keeping the resonant frequency Intact.

On the pretext that we want to further improve the adaptation of the antenna, we add breaks on the patch as shown in the fig6.
Fig.7. is present the simulated Sparameter of designed antenna with DGS unit cell and Fig.8 is present the simulated VSWR. Its result obtained at f1 for good matching impedance of dual band antenna. Notice that the resonance frequency of the antenna, with coefficient of reflection which is equal to $S11_{f1}$=-25.119848dB $S11_{f2}$=-16.4987dB and $VSWR_{f1}$=1.35 and $VSWR_{f2}$=1.2345. the gain of the proposed antenna is equal G_{f1}=7.16dB, G_{f1}=4.18dB Fig.9.

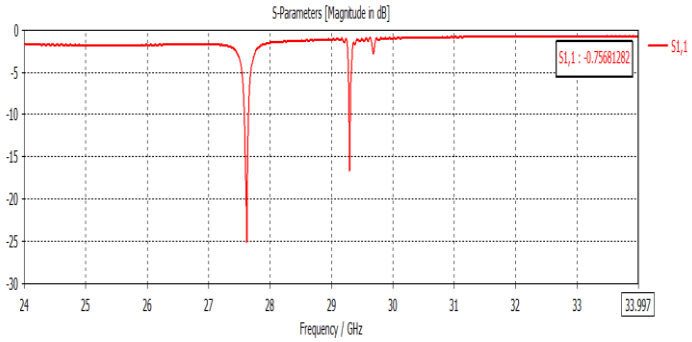


Fig.7: Simulated S-parameter of designed DGS unit cell

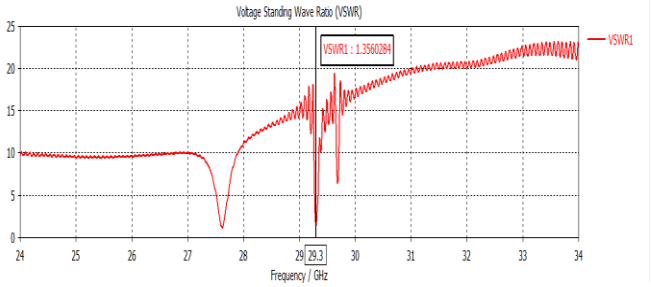


Fig.8: Simulated VSWR result of designed DGS unit cell

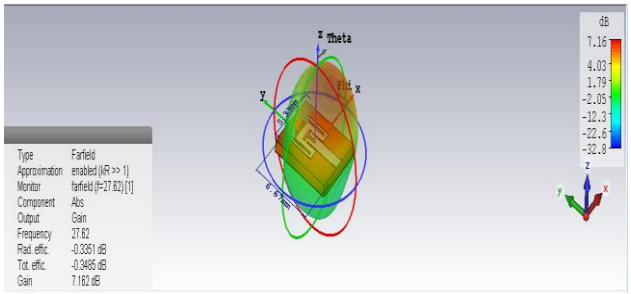


Fig. 11: Simulated 3D radiation pattern of DGS unit cell f1.

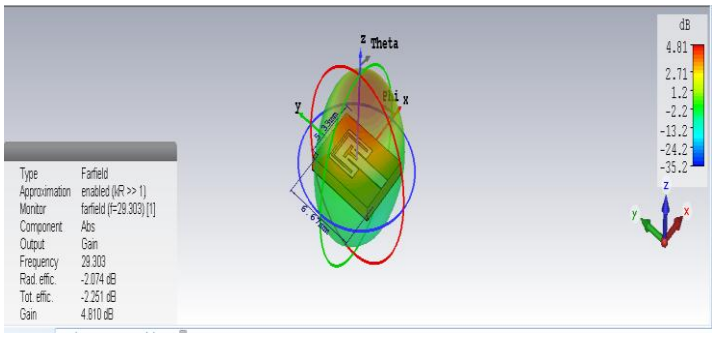


Fig. 12: Simulated 3D radiation pattern of DGS unit cell f2.

4.2 Patch antenna with DGS filter and unit cell :

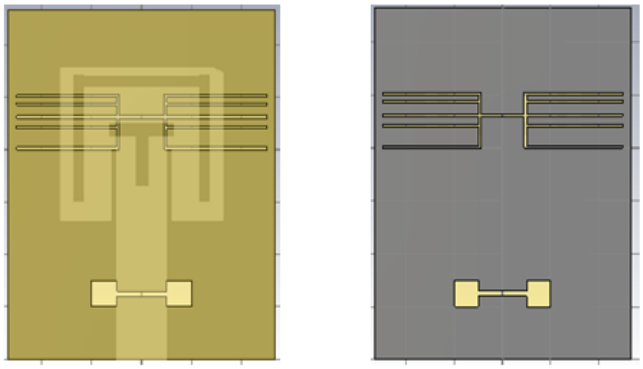


Fig.13: Geometry of proposed designed Unit cell and DGS filter

To understand the effect of introducing the DGS structure to the antenna patch we did a study same structure .but this time we engraved it on the ground plane of a transmission line the geometry of the filter is shown in fig.13.

Fig.(14,15,16,17) are present the simulated S parameter of designed microstrip patch antenna and VSWR and gain .This simulated result obtained the F1=27.636GHz and F2=29.323 GHz of resonant frequency have a return loss $S11_{f1}$=-20.119848dB $S11_{f2}$=-28.7672dB and $VSWR_{f1}$=1.2862and $VSWR_{f2}$=1.07.the gain of the proposed Antenna is equal $Gf1$=7.35dB,$Gf2$=5.11dB.

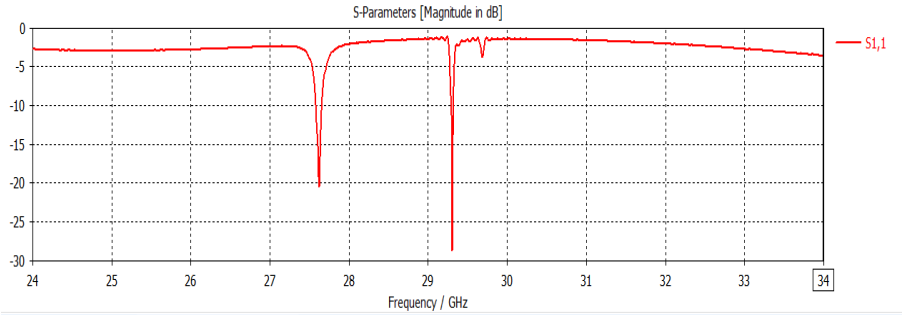


Fig. 14: Simulated S-parameter of designed DGS unit cell and DGS filter

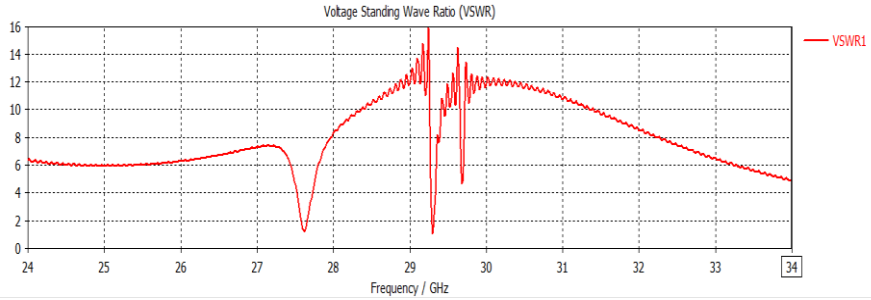


Fig.15: Simulated VSWR result of designed DGS unit cell and DGS filter

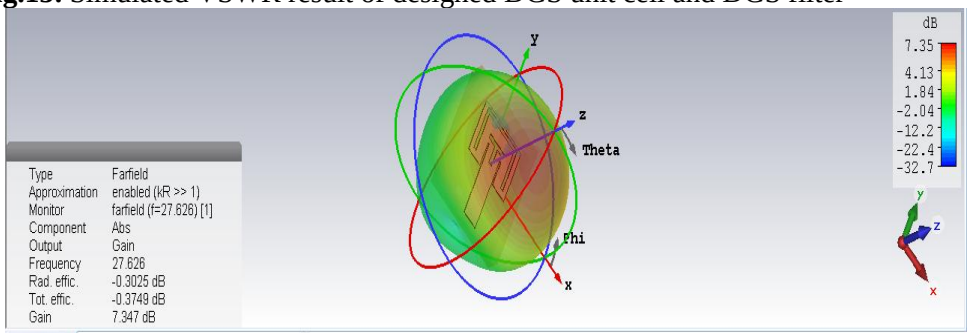


Fig. 11: Simulated 3D radiation pattern of DGS unit cell and DGS filter f1.

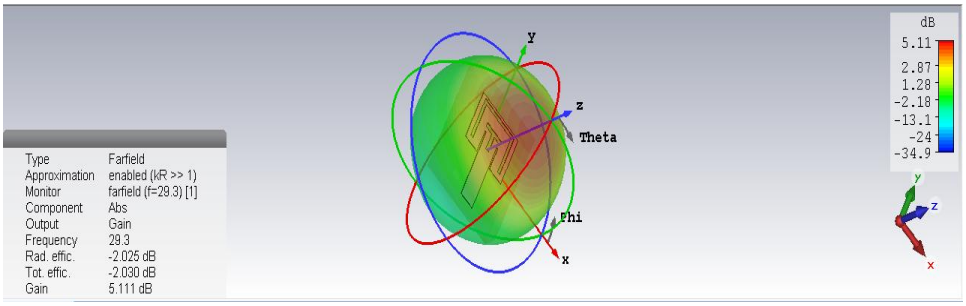


Fig. 11: Simulated 3D radiation pattern of DGS unit cell and DGS filter f2

Note that the gain of the antenna with DGS cells for the frequency f1 = 27.626 equal to 7.16dB and for a frequency f2 = 29.303 GHZ, Gain equal to 4.17dB.when adding DGS filter the gain increases from 7.16dB for f1 and 5.11dB for f2 in table (4) we have a comparison between antenna with unit cell and we adding the filter DGS .

Table 4: Comparison of designed antenna DGS unit cell with antenna DGS unit cell & DGS filter

	F(Ghz)	G(dB)	VSWR	S11(dB)
Unit cell DGS	F1=27.63 F2=29.323	G_{f1}=7.16 G_{f2}=4.18	$VSWR_{f1}$=1.35 $VSWR_{f2}$=1.2345	$S11_{f1}$=-25.11984 $S11_{f2}$=-16.4987
Add filter DGS	F1=27.63 F2=29.323	G_{f1}=7.35 G_{f2}=5.11	$VSWR_{f1}$=1.2862 $VSWR_{f2}$=1.07	$S11_{f1}$=-20.119848 $S11_{f2}$=-28.7672

Table 5: Comparison of designed antenna with other reference antenna

	[2]	[1]	[3]	Antenna proposed
Gain db	7.77	5.54, 6.4	15.6-10	5.35-7.11
S11(dB)	-31.24	-35.73,-47.53	-50,-40	-20.119848, -28.7672
Resonant Frequency(Ghz)	30	28,28	28-39.95	27.63, 29.323
size mm×mm	10×10	10×10	16×16	5.33×6.67

4 CONCLUSION

I n this paper,the DGS structure has been presented and we analysed different parameters by CST simulator.The detail of the design approach presented in the article may beprone to modification amid evolution of wireless devices in the future.

Nonethelss ,the authors predic the methodologies able to help to increases gaon of this antenna, we can also add array antenna for the same reason.
The antenna obtained is characterized by a better adaptation, more bandwidth and greater gain and directivity which make it suitable to respond to the majority of constraints imposed by the modern telecommunication systems.

ACKNOWLEDGMENT

The authors would like to thank my professor Larbi SETTI for valuable discussions and feedback.

REFERENCES

[1] saini and S. K. agarwal, "T and L slotted patch Antenna for future mobile and wireless communication". *in 2017 8th International Conference on Computing", Communication and Networking Technologies .in 2017*

[2] A. zaidi and A. . baghdad,A. badri "Design and optimization of an inset fed circular microstrip patch antenna using DGS structure for applications in millimeter wave band". International Conference onWireless Networks and Mobile Communications (WINCOM) .in 2016

[3] U. Rafique and H. . khalil,H. saif-Ur-Rehma "Dual-band microstrip patch antenna array for 5G mobile"communications.in 2017 Progress in Electromagnetics Research Symposium – Fall(PIERS - FALL) (2017).in

[4] F. Rainville and. P. Harackiewicz "Magnetic tuning of a microstrip patch antenna fabricated on a ferrite film". Articles 5(1992)..in 2017 .

[6] A. makoto miao . j. hirokawa . k. sakaguchi. tanighuchi. Makoto" Demonstration of mmWave Systems and Networks for the HetNet in 5G Mobile"Communication.in 2016

[7] Tadilo. Endeshaw Bogale Long .Bao" Massive MIMO and mmWave for 5G Wireless HetNet".in 2016

[8] M.khandelwal,B.kanaujia,S.kumar" Defected ground structure: fundamentals, analysis, and applications in modern wireless" s-International Journal of Antennas and Propagation , 2017 - hindawi.com

[8] Theodore S.Rappaport,shu sun,Rimma mayzus,Hang zhao,yaniv azar,Kevin wang,georef N,wong Jocelyn k shulz, Mathew Samimi and felix gutirrez "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" IEEE Digital Object Identifier 10.1109/ACCESS.2013.2260813.

Analysis of a 38 GHz Band Antenna for 5G wireless communication using Wave Concept Iterative Procedure (WCIP) method

Doae EL HADRI[1] , Asmaa ZUGARI[2] and Alia ZAKRITI[3]

{ [1]doae93@gmail.com, [2]asmaa.zugari@gmail.com, [3]aazakriti@hotmail.fr }

[1]Advanced Science and Technology Group, Physics Department, ENSA, Abdelmalek Essaadi University,Tetuan, Morocco

[2]Laboratory Systems of Information and Telecommunications, FS

Abdelmalek Essaadi University,Tetuan, Morocco

Abstract. The analysis of mm-Wave patch antenna for future fifth generation (5G) wireless communications is presented in this paper. The antenna is mounted on a compact Rogers substrate having dimensions 6.336 X 6.4 X 0.254 mm^3 with relative permittivity (ε_r) 2.2. The proposed antenna operates at a 38 GHz band. It is modelled by the Wave Concept Iterative Procedure (WCIP) method. The iterative procedure applies the fast modal transform (FMT) to change the analysis between spatial and spectral domains. In order to validate the present results we used the commercially available software CST MICROWAVE STUDIO. The numerical results show a good agreement compared with the CST simulation data, which illustrates the efficiency and correctness of our WCIP algorithm.

Keywords: 5G wireless communication, Antenna, CST, Wave Concept Iterative Procedure, MATLAB.

1 Introduction

The successful development of mobile wireless communications is demonstrated by a rapid rhythm of technology innovation.This growth has seen many generations from 1G, 2G, 3G, 4G and 5G. Each of these generations have various wireless technologies, data rates, modulation techniques, capacities and features compared to the other.

In initial 1980's the first generation (1G) mobile system was announced and supports only analog voice. The second generation (2G) mobile communication system debuted in 1991, it uses digital signals for voice transmission and has speed of 64 kbps. 2.5G system provides data rate up to 144 kbps [1]. The 3G system first launched in 2001, it operates at a range of 2100MHz and has a bandwidth of 15-20MHz. From 2G to 3G the wireless mobile network has been transformed from a pure telephony system to a network that can transport rich multimedia contents [2].

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284116

Fourth generation (4G) wireless networks can support data rates of up to 1 GB/s for low mobility, and up to 100 Mb/s for high mobility. The technologies driving 4G growth are LTE and WiMAX.
However, 4G still cannot accommodate some challenges like spectrum crisis, high energy consumption, poor coverage, bad interconnectivity, poor Quality of Service (QoS) and flexibility. To address all these demands 5G wireless system are expected to be deployed in the future by 2020 [3].

The fifth-generation (5G) cellular systems is more than an evolution of mobile broadband. It will be a key enabler of the future digital world, it is an opportunity to launch brand new internet services such as internet of things(IoT), mission critical services, high definition video stream, bitpipe communications at Gbps, tactile internet [4], rural access networks and autonomous car [5].

In order to meet the growing demand for more connected objects and higher data rates, the fifth generation of mobile network (5G) will be deployed. Compared to current mobile communication systems, new five technologies would be adopted like, device-centric architectures, millimeter wave, massive multiple-in multiple-out (MIMO), smarter devices, and native support for machine-to-machine (M2M) communications. These technologies may provide 10–100× higher user data rate, 1000× higher mobile data volume per area, 10–100× higher number of connected devices, 10× longer battery lifetime, and 5× reduced end-to-end latency [6].

The realization of these high level technologies brings about new challenges for the physical infrastructure designers which undoubtedly include the antenna and microwave engineers. From the electromagnetic wave perspective, due to this unprecedented growth of mobile data demand, the very limited spectrum resources available in the sub-6 GHz spectrum no longer satisfy the system needs. Consequently, the use of the vast amount of underutilized spectrum in the 6 – 300 GHz range for next-generation commercial wireless systems has recently garnered a great amount of research interest [7]. Accordingly, the mm-Wave spectrum was allocated for 5G communications [8].

The antenna is the most key components of wireless systems as it intensely affects the total receiver sensitivity, thus transceiver designs and choices of digital modulation schemes and the link budget [9]. It play an essential role in the field of wireless communication. It consists of slot antennas, patch antennas.... [10].

This paper analysed an antenna design that is inspired from recent literature [11] by modification of parameters, for mm-wave 5G applications operating in the 38-GHz using Wave Concept Iterative Procedure (WCIP) method. The advantage of this method remains in its ease of use due to the absence of test functions, its fast computation time, mainly due to the systematic use of Fast Mode Transform (FMT) and surface mesh use. In addition, there is no matrix inversion was required and the convergence was insured independently of the circuit complexity.The numerical results obtained are validated using the commercially available software CST MICROWAVE STUDIO.

2 Wave Concept Iterative Procedure

2.1 Formulation

The Wave Concept Iterative Procedure is an integral method based on wave concept [12], the waves are defined conveniently as a combination of the transverse electric field and current density, normalized in terms of root of impedance. These waves, incident A and reflected B, on the discontinuity surface (Ω) are written by the following relationships (1) [13]:

$$\begin{cases} A_i = \frac{1}{2\sqrt{Z_{0i}}}(E_{Ti} + Z_{0i}\, J_{Ti}) \\ B_i = \frac{1}{2\sqrt{Z_{0i}}}(E_{Ti} - Z_{0i}\, J_{Ti}) \end{cases} \quad (1)$$

Where Z_{0i} is the characteristic impedance of the medium i (i=1, 2).
The transverse waves Ai and Bi can be used to calculate the tangential electric fields and the current intensity of the interface as (2) :

$$\begin{cases} \overrightarrow{E_i} = \sqrt{Z_{0i}}\,(\overrightarrow{A_i} + \overrightarrow{B_i}) \\ \overrightarrow{J_i} = \frac{1}{\sqrt{Z_{0i}}}\,(\overrightarrow{A_i} - \overrightarrow{B_i}) \end{cases} \quad (2)$$

The analytical expression of the iterative process involves a system of two equations: one written in the spatial domain (3), characterizing the boundary conditions at the interface of the structure through the operator $\hat{S}$; the other expressed in the modal domain (4), describing the upper and lower media through the operator $\hat{\Gamma}$ [14]. The iterative process is defined by:

$$\begin{cases} \bar{A} = \hat{S}\,\bar{B} + \bar{A}_0 & \text{Spatial domain} \quad (3) \\ \breve{B} = \hat{\Gamma}\,\breve{A} & \text{Spectral domain} \quad (4) \end{cases}$$

A_0 is the incident wave generated by the source

The connection between the spatial domain and spectral domain is carried out via the Fast Modal Transform (FMT) (**Figure 1**) and its inverse (FMT^{-1}) [15]. In each step, the incident and reflected waves are determined.

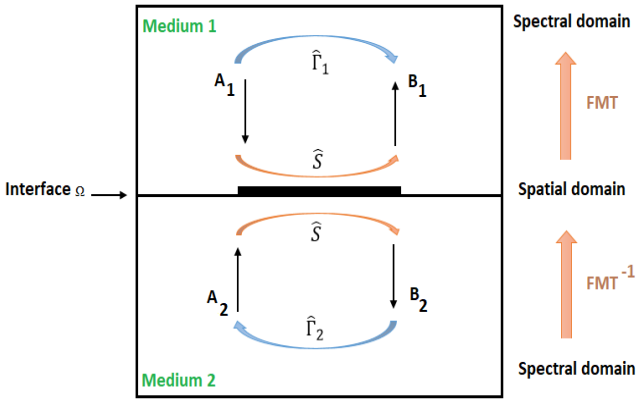


Fig. 1. Wave Concept Iterative Procedure process.

The scattering matrix of the general structure is show in equation (5) [16], it build by the physical nature of the material: if it is a source, a dielectric or a metal.

$$[\hat{S}] = \begin{bmatrix} -H_m - \frac{n^2-1}{n^2+1}H_d + \frac{-1+n_1-n_2}{1+n_1+n_2}H_S & \frac{2n}{n^2+1}H_d + \frac{n_3}{1+n_1+n_2}H_S \\ \frac{2n}{n^2+1}H_d + \frac{n_3}{1+n_1+n_2}H_S & -H_m + \frac{n^2-1}{n^2+1}H_d + \frac{-1-n_1+n_2}{1+n_1+n_2}H_S \end{bmatrix} \quad (5)$$

Where : $H_i = \begin{cases} 1 & \text{on the i domain} \\ 0 & \text{other domain} \end{cases}$ (i= Metal, Dielectric, Source)

$$n = \sqrt{\frac{Z_{01}}{Z_{02}}}\ , n_1 = \frac{Z_0}{Z_{01}}\ , n_2 = \frac{Z_0}{Z_{02}}\ , n_3 = \frac{2\,Z_0}{\sqrt{Z_{01}}\sqrt{Z_{02}}} \text{ and } \begin{bmatrix} A_{01} \\ A_{02} \end{bmatrix} = \frac{1}{1+n_1+n_2} \begin{matrix} \frac{E_0}{\sqrt{Z_{01}}}H_S \\ \frac{E_0}{\sqrt{Z_{02}}}H_S \end{matrix}$$

The expression of reflection operator in the modal domain is given by equation (6) [17]:

$$\hat{\Gamma} = \sum_{m,n} |f_{mn}^{TE}\rangle \Gamma_{mn}^{TE} \langle f_{mn}^{TE}| + \sum_{m,n} |f_{mn}^{TM}\rangle \Gamma_{mn}^{TM} \langle f_{mn}^{TM}| \quad (6)$$

Where :

$$\Gamma_{mn}^{\alpha} = \frac{1 - Z_{0i}Y_{mn}^{\alpha}}{1 + Z_{0i}Y_{mn}^{\alpha}}$$

Y_{mn}^{α} Can be calculated by equation (7) :

$$Y_{mn}^{TE}(\varepsilon_r) = \frac{\gamma_{mn}}{j\omega\mu_0} \quad \text{et} \quad Y_{mn}^{TM}(\varepsilon_r) = \frac{j\omega\varepsilon_0\varepsilon_r}{\gamma_{mn}} \quad (7)$$

γ_{mn} being the propagation constant of the medium i and it is given by (8) :

$$(\gamma_{mn}^2) = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 - K_0^2(\varepsilon_r)_i \quad (8)$$

3 Antenna Design

Figure 2 shows the different antenna structures, which are investigated during the simulation studies, and their resonance frequencies of each case are displayed :

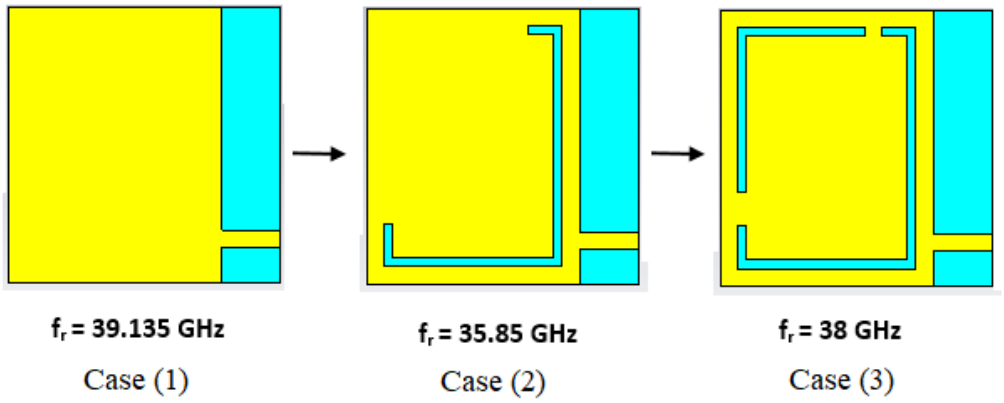


Fig. 2. Design evolution of the proposed antenna

After optimizing the antenna parameters, the geometry of the proposed antenna for future 5G wireless communications is shown in **Figure 3**. The patch is designed using conducting material with slots. The antenna is printed on a Rogers substrate with dielectric constant 2.2 and thickness h=0.254 mm. The ground plane is made up of same conducting material as the patch.

The discretization process applied by the WCIP interface shown in **Figure 4**. The dimensions of the interface are 6.336 x 6.4 mm^2.

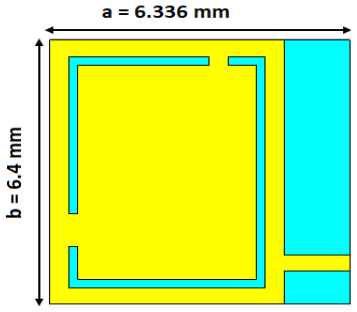


Fig. 3. Geometry of the proposed antenna.

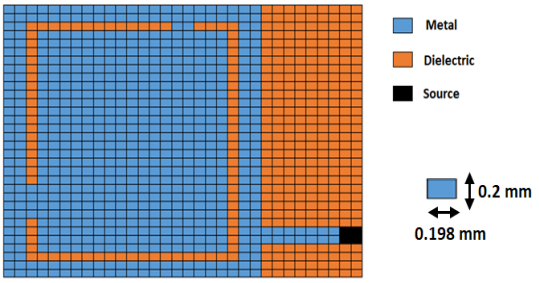


Fig. 4. Discretization of 5G antenna in the WCIP interface.

4 Numerical results

To ensure that the method WCIP has achieved convergence, a study is conducted at a frequency point at 36 GHz. The result presented in **Figure 5** shows the convergence of the S parameter module as a function of the number of iterations. This study is recommended to predefine an estimate of the number of iterations required by the method to achieve the result at each frequency point.

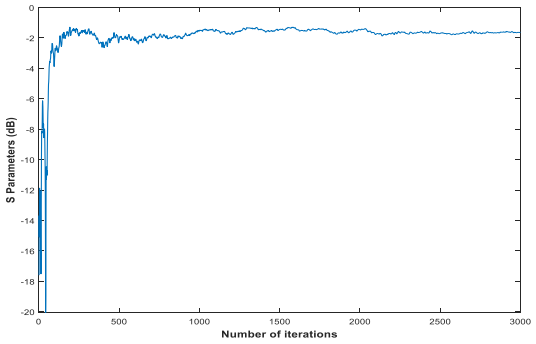


Fig. 5. Convergence of the S parameter at 36 GHz

The distribution of the electric field and current density, which is shown in **Figure 6**, conforms to the boundary conditions of the circuit interface.

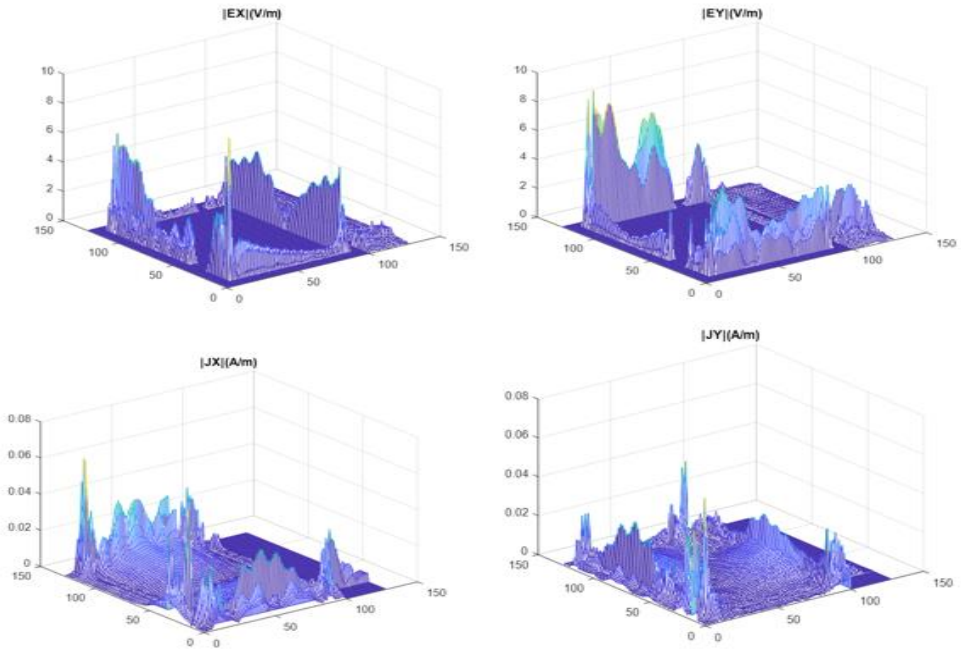


Fig. 6. Distribution of the electric field and current density

The calculated S parameters of the antenna with WCIP method show a good agreement with the CST results as presented in **Figure 7**. The resonance frequency is 38 GHz. The bandwidth at –10 dB is of order of 2 GHz (WCIP results). The phase of the reflection coefficient is presented in the **Figure 8**.

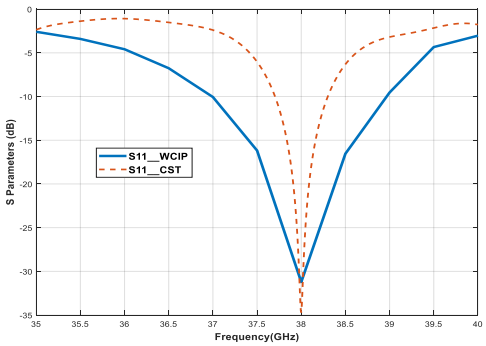


Fig. 7. The reflection coefficients (S11) for proposed antenna

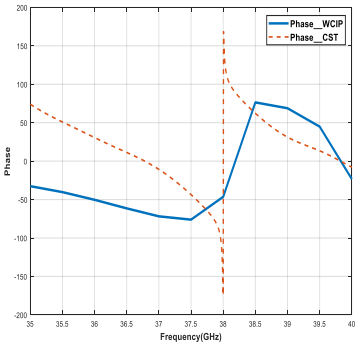


Fig. 8. The reflection coefficient phase

The figure below (**Figure 9**) shows the VSWR of the proposed antenna, the value of this parameter is compared using WCIP and CST in **Table 1** for the resonance frequency (38 GHz).

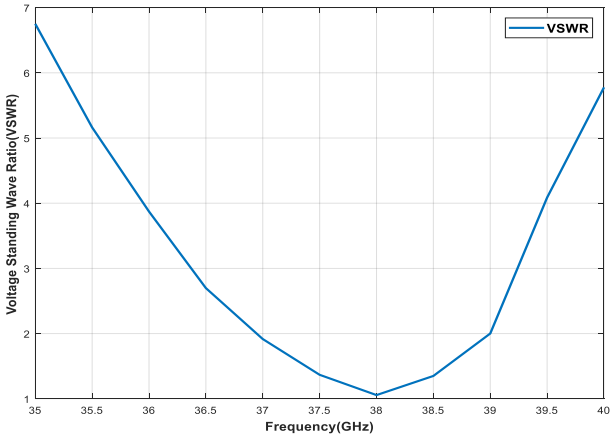


Fig. 9. The Voltage Standing Wave Ratio (VSWR)

Table 1. comparative table of the VSWR parameter at 38 GHz frequency

Parameter	WCIP	CST
VSWR	1.057	1.037

The gain of proposed 5G antenna for wireless communication at 38 GHz was 6.95 dBi.

5 Conclusion

In this paper, the analysis of 5G antenna for wireles communication using an iterative method, is carried out and discussed. The S_{11} parameters, the phases of the reflection coefficient, and VSWR for the antenna are determined. The resonance frequency is in good agreement between CST and WCIP simulation.

Acknowledgments:

The work is carried out with the support of the CNRST within the framework of the Research Excellence Scholarship Programme.

References

[1] G. R. Patil and P. S.Wankhade, "5G WIRELESS TECHNOLOGY," Int. J. Comput. Sci. Mob. Comput., vol. 3, no. 10, pp. 203–207, 2014.

[2] B. V Chaithanya and S. N. Shettar, "Smart Antenna Technology for 5g Wireless Communications," J. Android, IOS Dev. Test., vol. 2, no. 1, pp. 1–12, 2017.

[3] N. K. Reddy, A. Hazra, and V. Sukhadeve, "A Compact Elliptical Microstrip Patch Antenna for Future 5G Mobile Wireless Communication," Trans. Eng. Appl. Sci., pp. 1–4, 2017.

[4] T. S. Rappaport, Y. Xing, G. R. Maccartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks—With a Focus on Propagation Models," IEEE Trans. Antennas Propag., vol. 65, no. 12, pp. 6213–6230, 2017.

[5] A. C. S. Jr, "Antenna Development for 5G Networks," IEEE, pp. 2–6, 2017.

[6] P. Zhang, J. Lu, Y. Wang, and Q. Wang, "Cooperative localization in 5G networks : A survey," ICT Express, vol. 3, no. 1, pp. 27–32, 2017.

[7] W. Hong et al., "Multi - Beam Antenna Technologies for 5G Wireless Communications," IEEE Trans. Antennas Propag. Submitt., no. c, 2017.

[8] W. El-halwagy, R. Mirzavand, M. Jordan, H. Masum, and M. Pedram, "Investigation of Wideband Substrate-Integrated Vertically-Polarized Electric Dipole Antenna and Arrays for mm-Wave 5G Mobile Devices," IEEE Access, no. December, 2017.

[9] M. A. Matin, "Review on Millimeter Wave Antennas- Potential Candidate for 5G Enabled Applications," Adv. Electromagn., vol. 5, no. 3, 2016.

[10] R. A. Panda, N. Patnaik, C. Dash, and B. Kumari, "Dual Band Modified E-Shaped Microstrip Patch Antenna for 5g Application and Space Communication," Int. J. Res. Appl. Sci. Eng. Technol., no. September, 2017.

[11] M. El Shorbagyl, R. M. Shubair, M. I. Aihajri, and N. K. Mallat, "On the Design of Millimetre-Wave Antennas for 5G," IEEE, pp. 7–10, 2016.

[12] N. Raveu, G. Prigent, T. Callegari, and H. Baudrand, "WCIP APPLIED TO ACTIVE PLASMA CIRCUITS," Prog. Electromagn. Res. Lett., vol. 21, no. January, pp. 89–98, 2011.

[13] G. Fontgalland, P. I. L. Ferreira, T.-P. Vuong, N. Raveu, and H. Baudrand, "Analysis of asymmetric PBG ground planes for wireless communications," IEEE Xplore, pp. 352–355, 2005.

[14] T. Ziar, S. Mellal, F. Hichem, M. Zaabat, and C. Azizi, "The study of the miniaturisation effect on the characteristics of patch antenna using the WCIP method," J. New Technol. Mater., vol. 4, pp. 53–57, 2014.

[15] A. I. Alhzzoury, N. Raveu, O. Pigaglio, H. Baudrand, and K. Al-Abdullah, "WCIP APPLIED TO SUBSTRATE INTEGRATED," Prog. Electromagn. Res., vol. 33, no. August, pp. 171–184, 2012.

[16] J. Wu-sheng, L. Quan-zhen, and Y. Fan, “Analysis of H-shaped patch antenna by Wave concept iterative procedure (WCIP),” IEEE, pp. 797–800, 2010.
[17] M. Tellache, Y. Lamhene, B. Haraoubia, and H. Baudrand, “AN NUMERICAL METHOD BASED ITERATIVE PROCESS TO CHARACTERIZE MICROWAVE PLANAR CIRCUITS,” Int. J. Comput., vol. 7, no. 3, pp. 86–94, 2008.

A New Design of UHF RFID Tag Based on Split Ring Resonator For Moroccan Frequency Band

Latifa EL AHMAR[1], Ahmed ERRKIK[1], Jamal ZBITOU[1],Ilham BOUZIDA[2] and Brahim LAKSSIR[2].

elahmarlatifa@gmail.com , ahmed.errkik@uhp.ac.ma , zbitouj@gmail.com , i.bouzida@mascir.com and b.lakssir@mascir.com

[1] MEET laboratory, FST of Settat UH1 and [2]Packaging laboratory, MAScIR.

Abstract. In this paper, a new proposed design is presented for UHF RFID antenna matched to the Moroccan frequency band authorized by the ANRT. A proposed structure is inspired from circular split ring resonator theory, the geometry of the RFID antenna is double sided. Front side presenting a modified SRR while the back side is a circular ground plane. In order to achieve a good matching input impedance between the antenna and the impedance of the UHF chip, the structure is combined a modified SSR with stubs. The EM simulation of the proposed antenna done by EM Solver on an FR-4 substrate. The use of SRR permit to obtain good results of simulation with a reflection coefficient about -26,146dB at the Moroccan frequency band around 868MHz , a good read range and reduced size of RFID antenna were obtained with physical dimensions of $32\times18\times0.87mm^3$ which presents 62.9% in comparaison with diemnsions of λ/4 length.

Keywords: SRR, UHF, RFID, tag.

1 Introduction

Radio-Frequency identification is a technique which allows the exchange of the data between the RFID tag and the base station. The introduction of this technology motivated future researches to integrate RFID into various applications by using labels or transponders. These tags are the heart of RFID system that includes a base station to read the transponders. RFID system is based on two main elements, as shown in **Figure 1**.

RFID tags: composed of antenna connected to a chip activated by electromagnetic wave coming from reader antennas.

RFID base station: includes two parts, RFID reader as a principal device for communication with RFID tag via antennas and a Processing unit (PU) connected to the reader. The PU represents a software to initiate read/write operation, it can be a stationary or mobile platform to save results. [1]

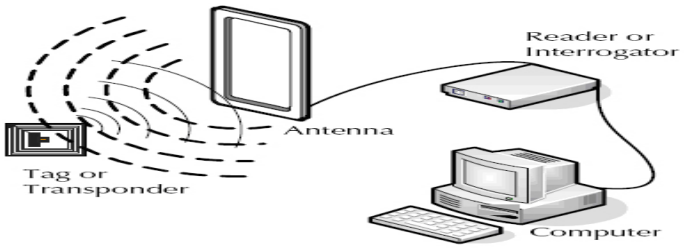


Fig. 1 Basic RFID system

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284086

The revolution of RFID is manifested by growing demand in industry. This technology represents fundamental proprieties as low cost and efficiency of RFID devices, also, it's able to add more functionalities by integrating sensors to tags and get more information (humidity, temperature…etc). In the market, there are three types of transponder: Passive tags that use electromagnetic waves to feed the chip, active tags are powered by external battery and the last technology of tags called BAP which combined active and passive tags. For RFID technology, specified frequency bands are standardized, there are Low-Frequency band (LF) (125-134 KHz), the High-Frequency band (HF) (13.56MHz), the Ultra High Frequency (UHF) (860-960 MHz) and microwave frequency band (2.4GHz and 5.8GHz).

This work presents new structure of RFID tag adapted to Moroccan UHF band based on a modified circular Split Ring Resonator (SRR). The SRR show considerable interest due to their unique properties and also their physical dimensions that are less than $\leq \lambda/10$. The flow of this paper is : the second section demonstrates the theory of the SRR structure and the design of modified SRR, the 3rd section shows results of the EM simulation of the new structure and in last section is a conclusion of work.

2 Theoretical and Design

This section is an introduction of split ring resonator theory, it describes the final geometry of the RFID tag based on circular modified SRR at Moroccan frequency (868MHz).

2.1 Electrical equation of circular SRR:

The circular Split ring resonator is commonly used for antenna design and passive circuits, thanks to their proprieties and also the simple design. The circular SRR behaviors is determined by the electrical elements of the equivalent circuit L and C which gives the resonance frequency according to the physical parameters of the structure, illustrated on **Figure 2**.

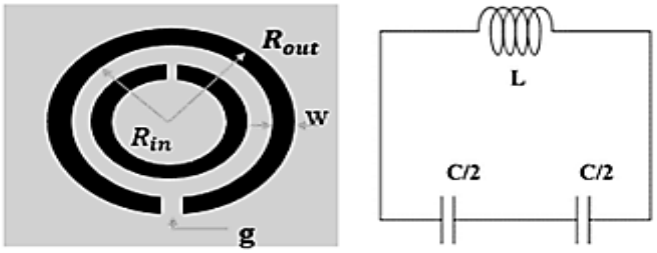


Fig.2 Circular Split Ring Resonator.

The resonance frequency is determined by (1):

$$f_r = \frac{1}{2\pi\sqrt{LC}} \quad (1)$$

L inductance of the electrical circuit is computed by (2), as referred in [5]

$$L = \mu_0 R_m \left(\ln\left(\frac{8R_m}{h+w}\right) - 0.5\right). \quad (2)$$

With μ_0 : Permeability of free space, $1.256 \times 10^{-6}[m.Kg.s^{-2}.A^{-2}]$.

R_m : Average radius $R_m = R + w/2$.

Total capacitance of the SRR structure is composed of two types of capacitance as mentioned in [5] by equation (3):

$$C = C_{gap} + C_{surf}. \quad (3)$$

With

Cgap is a capacitance due to the effect of the gap on structure computed by (4):

$$C_{gap} = \varepsilon_0 \left[\frac{wh}{g} + \frac{2\pi h}{\ln\left(\frac{2.4h}{w}\right)}\right]. \quad (4)$$

With

w is a metal width, h is the height of the susbtarte, g is width of the gap and, $\boldsymbol{\varepsilon_0}$ Permittivity of free space $\boldsymbol{\varepsilon_0} = 8.85418781710^{-12}\mathrm{F.m^{-1}}$.

And surface capacitance $\boldsymbol{C_{surf}}$ is contributed by the presence of the electric charges on metallic surface of the resonator. As denoted in equation (5):

$$C_{surf} = \int_{\theta_g}^{\pi} \frac{\sigma R d\theta}{V} = \varepsilon_0 h \int_{\theta_g}^{\pi} \frac{\cot\left(\frac{\theta}{2}\right)}{\pi - \theta} d\theta \simeq \frac{2\varepsilon_0 h}{\pi} \ln\left(\frac{4R}{g}\right). \quad (5)$$

With

$$\theta_g = \arcsin\left(\frac{g}{2R}\right).$$

σ Surface charge density, $\sigma = \frac{\varepsilon_0 V_0}{2\pi R}\cot\left(\frac{\theta}{2}\right)$ and V voltage on the gap, it's given by $V = \frac{c}{\pi}(\pi - \theta)$.

2.2 Anetnna Design :

The tag is designed as so to resonate at 868MHz and to prove a good performance, the proposed antenna design starts with a simple circular split ring resonator and two stubs are added in order to realize a good matching impedance between the antenna and the UHF chip. The ground plane is used to make the antenna more directive. The geometry is a double-sided antenna designed on a FR-4 substrate with a dielectric permittivity of 4.4, loss tangent of 0.025 and a thickness of 0.8mm, as illustrated on **Figure 3**.

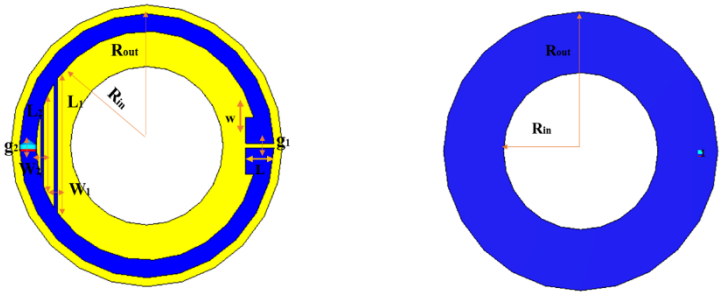


Fig.3 Proposed tag at 868 MHz.

To achieve a maximum transmitted power from the antenna to the UHF chip Alien H3 with a complex impedance at 868 MHz $Z_{chip} = 27 - j110$ .In litturature the antenna impedance should be equal to the conjugate of the impedance of UHF chip and to be strongly inductive to reduce the capacity effect of the chip.

Based on theoretical equations of circular split ring resonator, the physical parameters of the design determine the resonance frequency, for our design, several works are done to reach optimal parameters allowing good results at Moroccan UHF bands. This structure allows easy control of the resonance frequency and also more flexibility in term of the dimension of proposed antenna for RFID applications that require small antenna's dimensions. The optimal parameters of proposed structure, are presented in **table 1**.

Table 1: Optimal parameters of the antenna proposed.

Parameters	**Value (mm)**	
Ring	$R_{out} = 15$	$R_{in} = 13$
Gap	$g = g_1 = g_2 = 0.5$	
Microstrip Lines	$L_1 = 19$	$W_1 = 0.39$
	$L_2 = 15$	$W_2 = 0.32$
Stub	$L = 0.32$	$W = 1.35$
Ground Plane	$R_{out} = 16$	$R_{in} = 9$

The geometric parameters of the tag and resonance frequency are dependant as shown in [5], also the work is done to study resonance response of the structure which proves capability to reveal resonance at 868MHz. In theory the Moroccan frequency represents a wavelength about 345.6 mm, which is 10.8x of the greater diameter of the structure.

By projecting the theoretical expressions presented on [5] using optimal parameters shown on table 1 to calculate electrical elements presented in **table 2**.

Table 2: Electrical elements of the proposed tag.

Parameters	**Value(mm)**
L	60nH
C	56.03pF
C_{gap}	24.6pF
C_{gap_totale}	49.2pF
C_{surf}	6.833pF

Microstrip lines and stubs are characterized by electrical impedance presented in **table 3**.

Table 3: Characteristic impedances of the proposed structure.

	Microstrip Line 1	Microstrip Line 2	Stubs
Characteristic impedance	103.4Ω.	96.32Ω.	54,0421 Ω

3 Simulation and discussion:

According to the results of the EM simulation using CST mws , this RFID tag present a good adaptation with a reflection coefficient of -26,146dB at the Moroccan frequency (868MHz) as follows in **Figure 4** and a bandwidth of 1.94 MHz as follows in **Figure 5 ,** compared to the literature a S11 must be less than -10dB at a desired resonance frequency so the proposed tag is accepted in term of incident power that is reflected back towards the source from the antenna , in Morocco the bands is very limited 0.5MHz while the structure give us a large band of 1.94MHz.

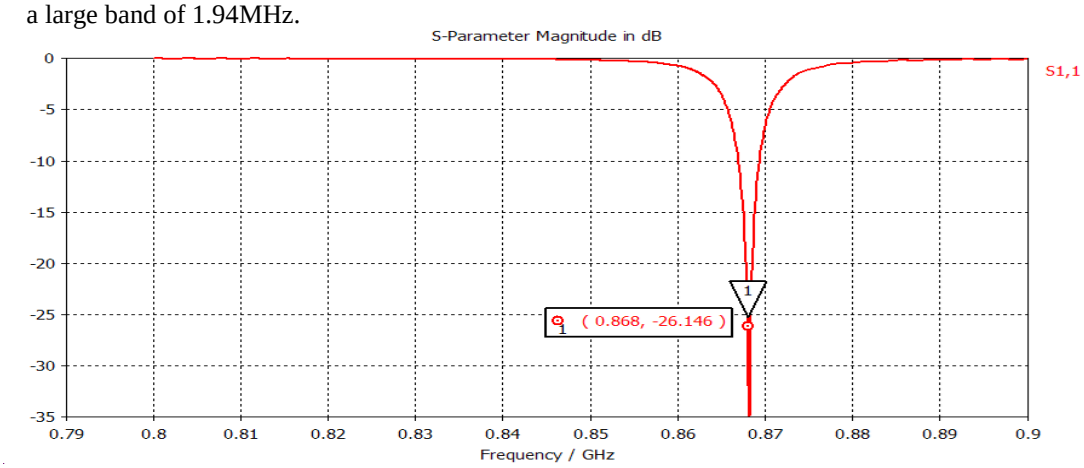


Fig.4 Simulated Reflection coefficient of the proposed RFID antenna.

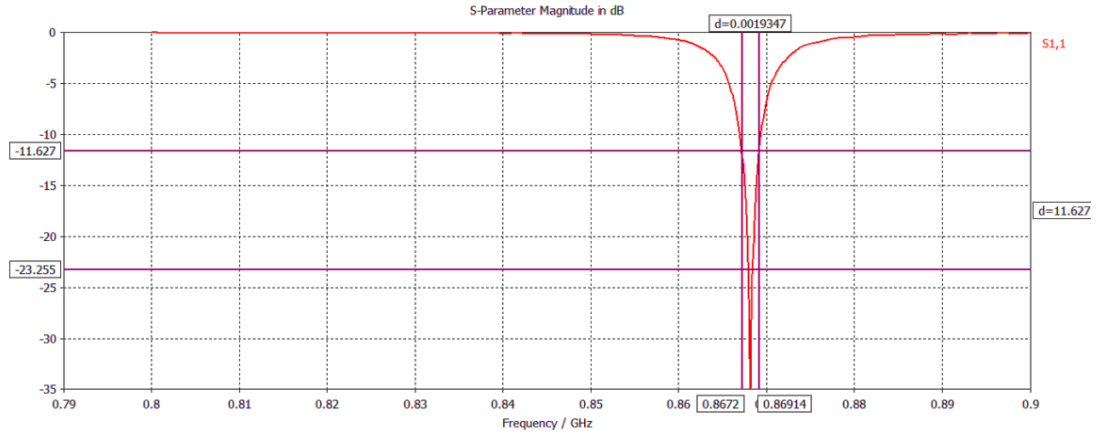


Fig.5 Bandwidth of the proposed RFID antenna.

For the selection of an antenna we evaluate also the VSWR (Voltage Standing Wave Radio) , VSWR describes the power reflection from the antenna, the ideal case is when VSWR is equal to 1 which means no reflection power. In real cases the VSWR less than 2 is acceptable for a small antenna. In case of our RFID tag , VSWR is equal to 1.1 at 868MHz as shows in **Figure 6**, that is a good VSWR , in term of power that‘s represent 99.7% of the power transmission and 0.3% of the power reflection .

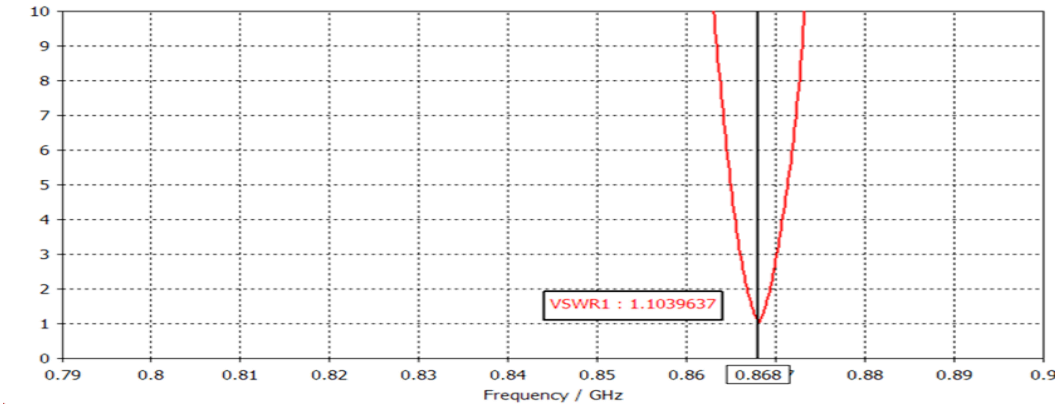


Fig.6 Simulated VSWR

The radiation pattern in the following **Figure 7** shows the RFID tag having a maximum directivity equal to 1.545dBi at 868MHz and the distribution of the radiation is more concentrated in both gaps of the structure.

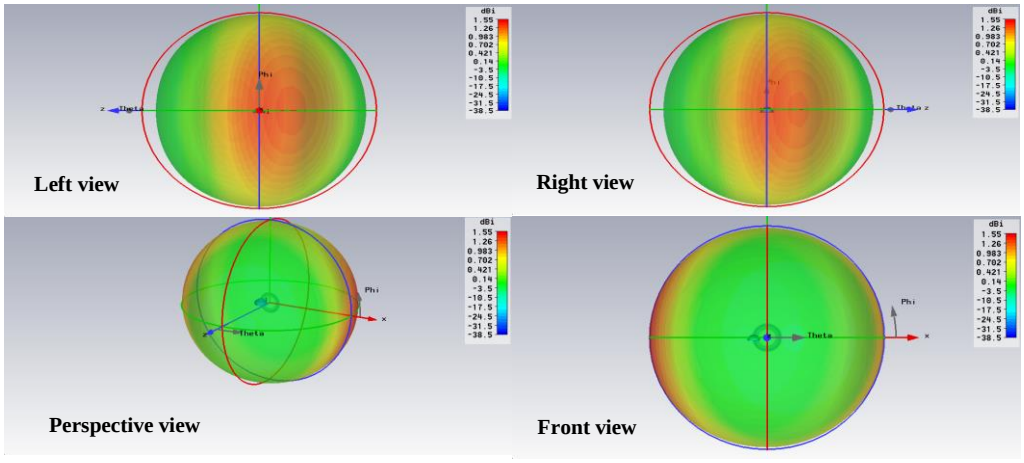


Fig.7 Radiation of the proposed RFID antenna

The read range is an important requirement of RFID tag design, it gives us an information about the maximum distance at which the tag receives the minimum threshold power required to turn on and scatter a signal back, and the maximum distance at which the reader can detect this return signal.Using the Friis free-space equation, we get to calculate the read range using

MATLAB script.The proposed antenna evaluate a read range equal to 1,6m which is the maximum distance that our proposed tag can communicate with the reader. As shown in **Figure 8.**

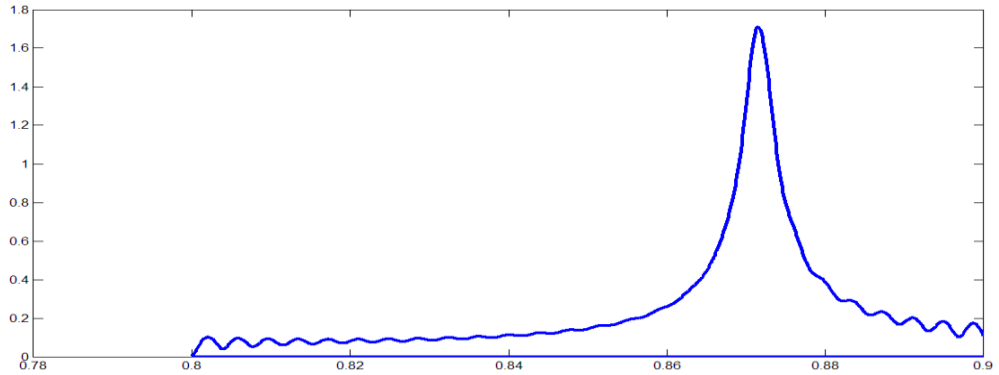


Fig.8 Read range of the proposed RFID antenna.

The proposed RFID tag having a negative gain at 868MHz, as shown in Figure 9, this result is explained according to the literature to the size of the antenna that is too smaller than $\frac{1}{2}$ or $\frac{1}{4}\lambda$,in our case the size of the antenna is 32mm less than $\frac{\lambda}{4} = 86.4$ mm , we call it an electrically small antenna that is characterized by an antenna gain less than 1.

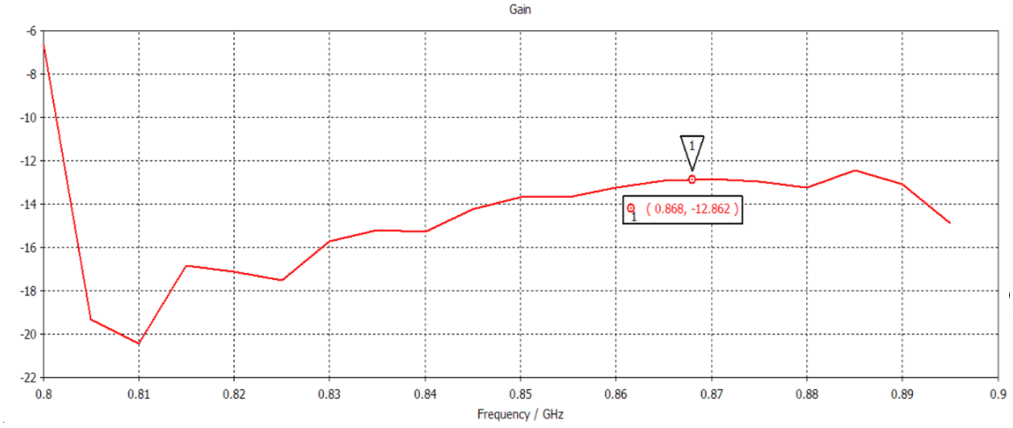


Fig.9 Gain Of the RFID tag at 868MHz.

After the validation of the structure using CST mws and fixing the dimensions and knowing the performance of the RFID tag , the tag is fabricated as shown in **Figure 10** manually in the lab. The FR-4 with εr = 4.4, tan δ = 0.025 and thickness = 0.8 mm is used as the substrate for the tag. Copper is used as the conductor.

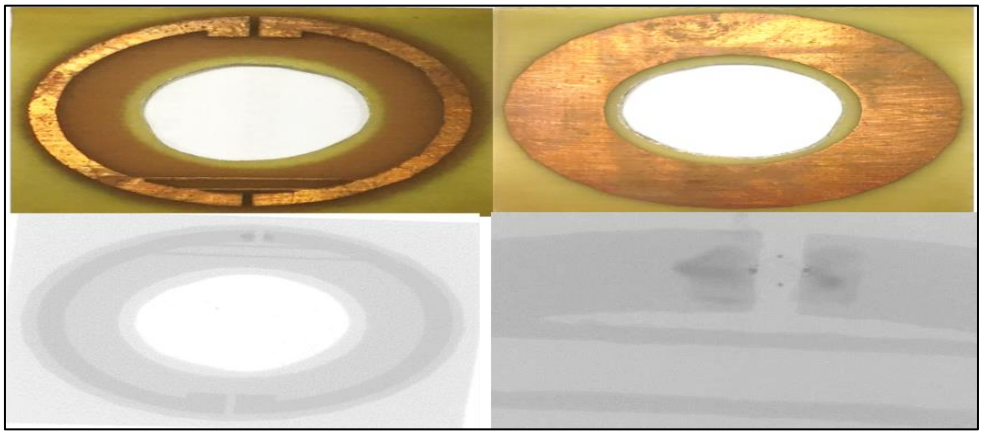


Fig.10 (i) Front view and bottom view of fabricated TAG antenna. (ii) Location of ALIEN H3 chip using X-Ray.

Conclusion

In this paper , a new contribution was given to the design of RFID tag based on metamaterials that are presented by a circular split ring resonator modified to achieve a good response at a Moroccan frequency and miniaturization of the design.Before fabricating the proposed tag , the EM simulation of the RFID antenna is demonstrated to achieve the requirements of design including matching input impedance, size and read range.

The RFID tag combined a simple split ring resonator and stubs to realize a good matching impedance. The result of the simulation presents a good matching of -26,146dB at 868 MHz with a bandwidth of 1.94 MHz.The proposed RFID antenna has a small size $32 \times 18 \times 0.87\ mm^3$ that comes to 62.9% of the theoretical size .The read range calculated using MATLAB agreed well with the test results by using a base station in the lab.

References

[1] B. Ghanem, “RFID Project Technical Report,” 1970.

[2] A. E. Abdulhadi et al., “Antenna design for UHF RFID Tags: A review and a practical application,” 2003.

[3] Advance Energy, “Impedance Matching,” 2006.

[4] I. Antennas, P. Magazine, and V. Politecnico, “The art of UHF RFID antenna design : impedance matching and size-reduction techniques,” 2008.

[5] Jabita and A.-N. Abiodun, “Design of Singly Split Single Ring Resonator for Measurement of Dielectric Constant of Materials using Resonant Method,” Master’s Thesis Electron. Telecommun. Univ. Gävle, no. June, 2013.

[6] N. H. Daud, A. Ismail, A. Reda, H. Alhawari, A. Sali, and M. F. A. Rasid, “Integration of Split Ring Resonators (SRRs) to UHF RFID Tag Antenna for Size Reduction,” 2014.

[7] S. Zuffanelli, G. Zamora, P. Aguilà, F. Paredes, F. Martín, and J. Bonache, “Analysis of the Split Ring Resonator (SRR) Antenna Applied to Passive UHF-RFID Tag Design,” IEEE Trans. Antennas Propag., vol. 64, no. 3, pp. 856–864, 2016.

[8] A. Ennajih, J. Zbitou, A. Errkik, and L. El Abdellaoui, “A New Design of UHF RFID Tag Antenna Based on Negative Index Metamaterial,” pp. 1–6, 2017.

New Design of a Multiband Microstrip Antenna

S. El Kilani[1], Larbi El Abdellaoui[1], A. Tajmouati, A. Errkik[1], Mohamed Latrach[2]

{ [1]dr.samir.2014@gmail.com }

[1] LMEET FST of Settat, Hassan 1st University, Morocco
[2]Microwave group, ESEO ANGERS France

Abstract. This paper proposes a multiband microstrip patch antenna covers a multiple operating frequency bands PCS/DCS/LTE/WIFI. The overall dimension of the proposed antenna is $65\times67.05mm^2$ printed on an FR4 substrate, having 1.6 mm as thickness with a relative dielectric permittivity of 4.4 and 0.025 for loss tangent. The proposed antenna was study and validated by using electromagnetic software's CST-MW. Three bands (1.811-1.828 GHz, 2.631-2.749 GHz and 4.554-5.457 GHz) are achieved by inserting a slot on the top patch and ground. The radiation pattern has acceptable response with low cross polarization at both E-plane and H-plane.

Keywords: Microstrip patch antenna, Multiband planar antennas, DCS, PCS, WIFI.

1 Introduction

Nowadays the multiband antenna become more and more requested, all mobiles equipments required covering all communication service includes 2G/GSM900/PCS, 3G/UMTS/HSDPA/WiFi/ Bluetooth and 4G/LTE, In various mobile communication services, long-term evolution (LTE) is one of the widely used communication systems as a fourth-generation wireless service. Because each nation or wireless carrier uses different frequency bands, a multiband antenna is desirable [1]. Since each communication protocol may operate in distinctive frequency bands, instead of using several antennas, it is highly useful to have one multiband antenna to meet the need of multiple communication systems [2].

Microstrip patch antenna is promising to be a good candidate for the future technology. Microstrip patch antenna consists of a dielectric substrate, with a ground plane on the other side. Due to its advantages such as low weight , low profile planar configuration, low fabrication costs and capability to integrate with microwave integrated circuits technology, the microstrip patch antenna is very well suited for applications such as wireless communications system, cellular phones, pagers, Radar systems and satellite communications systems [3, 4].

To achieve the multiband characteristics many efforts has been made by the researchers in recent years [5]-[6]. The methods such as notch technique [7], slots technique [8] and fractal method [9] are used to design multiband antennas. By introducing these different types of methods and techniques in the geometry of microstrip patch antenna and proper selection of feeding technique helps to achieve the multiband characteristics easily [10].

In this paper, a compact multiband design of Microstrip patch antenna is proposed. The bandwidth has improved by adding slots and by fine tuning to some parameters to reach the best possible result. The multiband antenna covers following frequency bands: DCS/PCS (1.811-1.828 MHz), LTE2500 (2.631-2.749

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284224

GHz), WIFI/Bluetooth (4.554-5.457 GHz) [11-12]. This design is fully planar, fairly compact, and low cost by using a low cost FR4 substrate material.

2 Antenna Design

Figure 1 shows the stricture of the proposed multi-band antenna for PCS/DCS/LTE/WIFI, where we can see the top and the bottom faces of the final circuit. The antenna is a planar microstrip patch antenna designed on an FR4 substrate having 1.6 mm as thickness with a relative dielectric permittivity of 4.4 and 0.025 for loss tangent. With dimensions of 65x67.05x1.6 mm^3 in FR4 substrate and 34.11x30.4 of antenna radiating part, a 50Ω microstrip line is used to feed the antenna for better impedance matching. The ground plane of the proposed antenna is modified and optimized to reach the multiband behavior in the suitable frequency bands.

With the help of the fundamental equations given in this paper [11] we find the first dimensions values of the patch antenna.

The width of the microstrip patch is finding by:

$$W = \frac{c}{2fr}\sqrt{\frac{2}{\varepsilon_r + 1}}$$

Effective dielectric constant is given by:

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{w}\right]^{-\frac{1}{2}}$$

The extension length (ΔL) is calculated by:

$$\Delta L = 0.412h\frac{(\varepsilon_r + 0.3)(\frac{w}{h} + 0.264)}{(\varepsilon_r - 0.258)(\frac{w}{h} + 0.8)}$$

Length L of the patch:

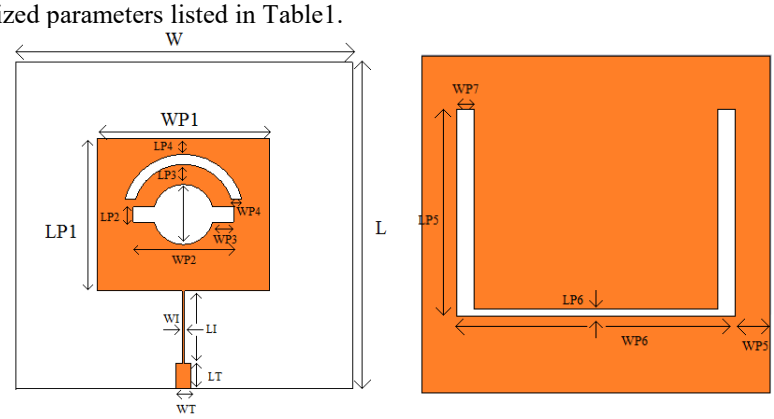
$$L = \frac{c}{2f_r\sqrt{\varepsilon_{reff}}} - 2\Delta L$$

The length and width of the ground plane is estimated by this equation:

$$w_g = 6h + w$$

$$L_g = 6h + L$$

After many series of optimization by using CST, we have obtained the different optimized parameters listed in Table1.



(a) Top view (b) Bottom view

Fig.1. Geometry of the proposed Antenna

Table1. Parameters of the proposed antenna (unit in mm)

Parameter	Value	Parameter	Value
L	65	**WP3**	4.19
W	67.05	**LP3**	4
WP1	34.11	**WP4**	2.079
LP1	30.04	**LP4**	3.5
WT	3	**WP5**	7.725
LT	5	**LP5**	40
WI	0.5	**WP6**	53.6
LI	14.5	**LP6**	1.5
WP2	20	**WP7**	3.3
LP2	3	**WS**	11.75

3 Results and Discussion

The proposed antenna is simulated by using cst microwave, whose numerical analysis is based on the Finite Element method [12]. Figure.2 shows the return loss simulation results of the proposed antenna. The simulation results ensure that the antenna covers multi-band frequencies for wireless applications. For a return loss less than 10 dB, we can deduce that the antenna operates in three frequency bands (1.811-1.828 MHz), (2.631-2.749 GHz) and (4.554-5.457 GHz), which covers DCS/ PCS/ WIFI/ Bluetooth/ WIMAX. The third band frequency is controlled by adjusting the total dimensions of the planar antenna but the first and second one are controlled by the geometry and dimensions of the ground plane.

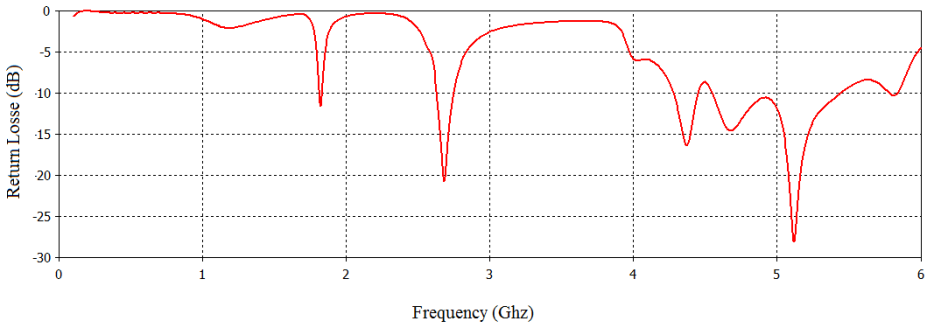


Fig.2. return losses obtained by CST-MW.

As depicted in Figure 2, The simulation results show that this antenna can be suitable for three frequency bands (1.811-1.828 GHz, 2.631-2.749 GHz and 4.554-5.457 GHz), which covers DCS/PCS/LTE/WIFI/Bluetooth for mobile phone applications.

Figure 3 shows the simulated three-dimensional (3D) radiation patterns of the proposed antenna at three resonant frequencies 1.87GHz, 2.637GHz and 5.126GHz. We can conclude that the proposed antenna can radiate unidirectional pattern at all the operating frequency bands. The simulated surface current of the proposed antenna, is presented in Figure 4.

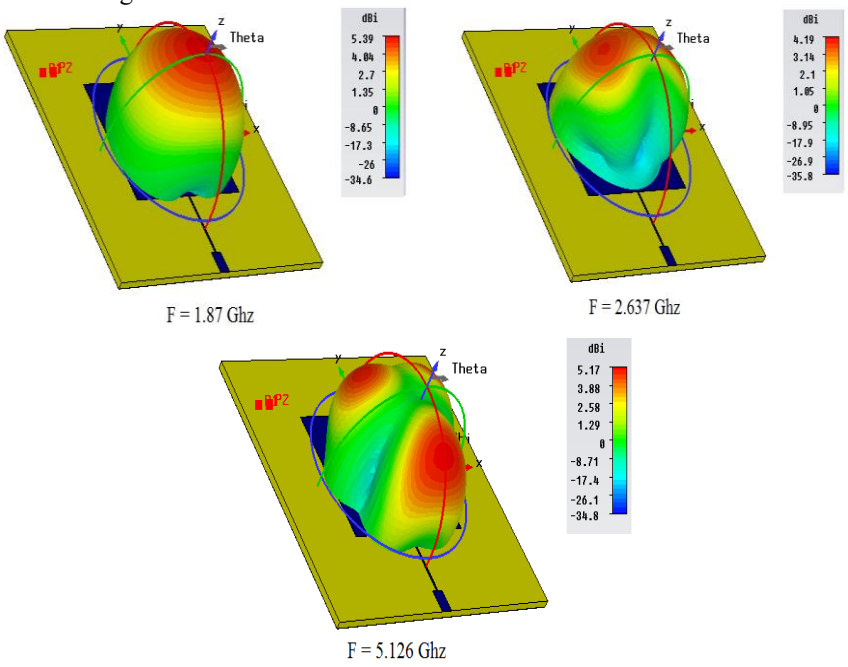


Fig.3. The simulated radiation patterns of the proposed antenna at different resonant frequencies.

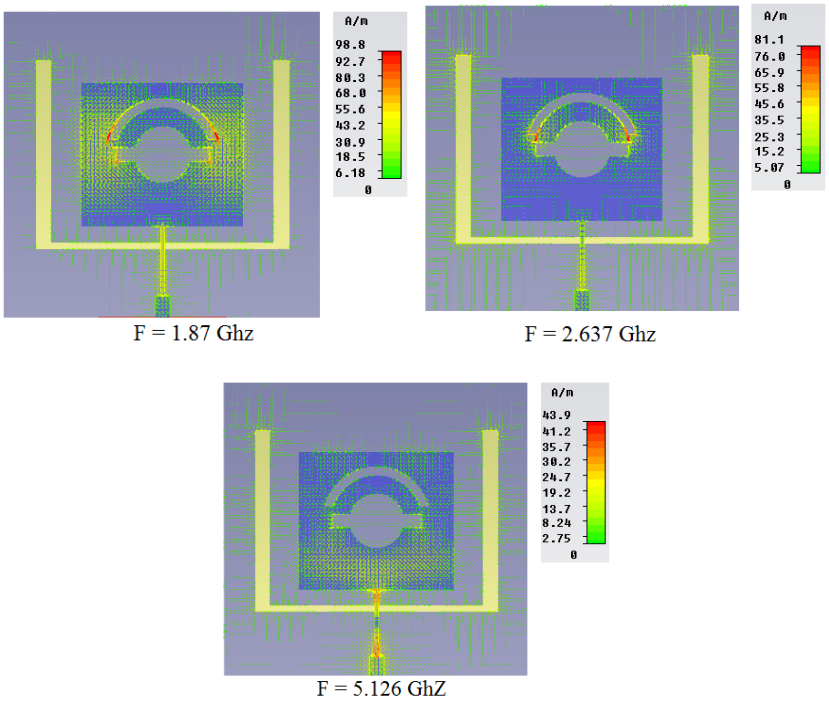


Fig.5. Simulated surface current of the proposed antenna at different resonant frequencies.

4 Conclusion

In this work, A multi-band compact patch with microstrip is presented for 1.87, 2.637, 5.126 GHz application and validated into simulation by using an electromagnetic solvers. Multiband was achieved by adding the slot in the patch and the ground. The antenna characteristic and radiation pattern is satisfactory for most of the wireless system because is unidirectional and can protect humane head. The different steps followed in this study can be used to validate an antenna for others frequency bands.

References

[1] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and Thomas, "LTE-advanced: next-generation wireless broadband technology," IEEE Wireless Communications, vol.17, no.3, pp.10–22, 2010.

[2] W.-J. Liao, S.-H. Chang, and L.-K. Li, A compact planar multiband antenna for integrated mobile devices, Progress In Electromagnetics Research 109 (2010) 1-16.

[3] FUJIMOTO T. Wideband stacked square microstrip antenna with shorting plates. IEICE Trans B: Communications. 2008 May 1;E91-B(5):1669-72.

[4] Lin S-, Row J-. Bandwidth enhancement for dual-frequency microstrip antenna with conical radiation. Electronics Letters. 2008;44(1):2-3.

[5] Jyoti Kapil, Hardyal Singh, "Compact multiband microstrip antenna for wireless communication systems," International Journal of Electrical, Electronics and Data Communication, Vol. 4, Issue 3, Mar. 2016.

[6] Hau Ran Cheong, Kim Ho Yeap, Koon Chun Lai, "A compact CPW-fed antenna with fractal s-shaped patches for multiband applications," Microwaves and Optical Technology Letter.,Vol. 59, Issue3, Mar. 2017.

[7] Mustapha El Halaoui, Abdelmoumen Kaabal, Hassan Asselman, Saida Ahyoud, and Adel Asselman,"Multiband Planar Inverted-F Antenna with Independent Operating Bands Control for Mobile Handset Applications," Inernational journal of Antennas and Propagation, Article ID 8794039, 13 pages, 2017.
[8] Manjiri Gogate, Aditya Desai, Madhuri Bhujbal, "Multiband Patch Antenna for Wireless Devices," International Conference on Emanations in Modern Technology and Engineering, Vol. 5, Issue 3, Mar 2017.
[9] Ajay Singh, G.P. Singh, Manvendra, Prinu C. Philip, Mukesh Kumar, Rohini saxena, "Analysis of V Slot Multiband Microstrip patch Antenna for S, C and X Bands," International Journal of Engineering Trends and Technology (IJETT), Vol. 48, June. 2017.
[10] Hatem Rmili,Donia Oueslati, Imen Ben, Trad Jean, Marie Floch, Abdullah Dobaie, and Raj Mittra, "Investigation of a Random-Fractal Antenna Based on a Natural Tree-Leaf Geometry," Hindawi International Journal of Antennas and Propagation, November. 2017.
[11] Balanis, C. A. Antenna Theory: Analysis and Design, 3rd ED.John Wiley & Sons, Inc., 2012.
[12] Guy A. E. Vandenbosch and Alexander Vasylchenko, A practical guide to 3D electromagnetic software tools, Microstrip Antennas, Prof. Nasimuddin Nasimuddin (Ed.), ISBN: 978-953-307-247-0, InTech (2011).

Dual-band UWB antenna for Surface Movement Radar

[1]Salaheddine Aourik, [2]Ahmed Errkik, [3]Larbi El Abdellaoui, [4]Abdelali Tajmouati, [5]Mohamed Latrach
{ [1]aouriksalaheddine@gmail.com, [2]ahmed.errkik@uhp.ac.ma, [3]lelabdellaoui@fsts.ac.ma, [4]atajmouati@fsts.ac.ma,
[5]mohamed.latrach@eseo.fr }

[1,2,3,4]Hassan 1st University, Settat, Morocco
[5]Microwave group, ESEO, Anges, France

Abstract. The proposed antenna radiation size is 9.4x6.12mm with low permittivity value 2.22, loss tangent value 0.0004 and 0.4mm thickness Duroid 5880 dielectric substrate material were used. An inset microstripline feed has been used because of its simplicity and ease of practical manufacturing. The parametric study is performed with a change of the form of patch to achieve the Ka- band UWB frequency, dual-band and a high gain. The Proposed antenna impedance bandwidth (VSWR<2) is achieved with wide band characteristics 1 GHz (23.3-24.3GHz) on the lower Ka- band frequency and 2 GHz (27.4-29.4GHz) on the upper Ka-band frequency. An extensive analysis of the return loss, gain and efficiency of the proposed antenna is discussed in this paper. The antenna is simulated by computer software technology Microwave studio (CST), the presented antenna competent for the satellite and radar applications in Ka-band.

Keywords: microstrip patch antenna, Ka-band, surface movement radar (SMR), UWB, dual-band.

1. Introduction

Aircraft traffic ground movement has a big impact on airports capacity also on the congestion in surrounding airspace. It has a great influence on the safety and quality services as well as the economy of air transport operations. The conditions of reduced visibility can cause disturbances in the terminal area of the airport, causing significant delays for both flight and ground movement aircraft.

In order to increase airport capacity and enhance the level of safety during reduced visibility operations The International Civil Aviation Organization (ICAO: United Nations Specialized Agency for the Establishment of the Global Regulatory Framework for the Safety of International Civil Aviation) has developed advanced surface motion guidance and control systems (A-SMGCS) through the use of modern technologies and a high degree of integration of the various functions (Manual A-SMGCS 9830) [1].

Following the same goal, the European organization of the air navigation safety EUROCONTROL, as part of its project to improve air traffic flows and capacity management (ATFCM) at airports by minimising delays, and by optimising the use of resources, has published specifications for the implementation of the essential requirements to the system A-SMGCS which manifests the requirement for a non-cooperative sensor in the surveillance part of airport to visualise the position of the aircraft on the aerodrome(Version 1.0 - 1 March 2018) [2].

Surface Movement Radar (SMR) plays a vital role in the overall safety management system of airports. It provides all the necessary information to establish safe and efficient flight paths for aircrafts even under high complexity conditions. (Reduced visibility, night operations).

Accurate and reliable sensors are necessary to provide detailed information to air traffic controllers and crew amount. The radar antenna or the sensor is the key word in the radar system, having a high performance antenna in terms of resolution allows the accurate and reliable detection of all projects on the movement area. Very short transmission pulses of a few nanoseconds allow a high range resolution, which permits visualisation of aircraft contours on the radar display. In the rapid development of wireless communication, a compact and discrete antenna is essential which presents a high quality price ratio, reasonable mass and size of the system. [3] [4]In the literature several studies have been presented to have a dependent radar antenna on the requirements that

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284103

imply a migration towards the band Ka following the high performance of directivity compared to the band already used.

2. Radar equation [5]

The radar equation reflects the influence of physical phenomena on the radiated power, from to the propagation of the wave until the reception of the reflected signal (echo), allows the performance estimate of the radar system. The total power captured by the radar which depends on the equivalent surface Ae of the radar antenna is:

$$\Pr = \frac{Pt \times G^2 \times \lambda^2 \times \sigma}{(4\pi)^3 \times R^4 \times (Lp)}. \qquad (1)$$

Pt , G , σ , λ , R et and Lp are respectively the power of radiation, the gain in power, the cross-section (or the radar-SER-equivalent surface), the wavelength of the radiation, the distance between the radar and the target and Lp is the loss factor due to atmospheric losses, losses of the transmission and reception chains.

This expression shows that the intensity of the echo signal is proportional to the radiation power, to the square of the gain, to the effective cross section (span) of the aircraft and to the square of the wavelength, so the maximum range of the radar depends on the square root of the Gain. This results in the performance of the antenna which has a very great influence on the performance of the radar. On the other hand the resolution of surveillance radar is divided into two parts: the distance resolution and the angular resolution (in azimuth):

- ✓ The distance resolution ΔR is the ability of a radar system to distinguish two or more targets in the same direction but at different distances. The quality of the resolution depends on the bandwidth Δf of the transmitted pulse.

$\Delta R = \frac{c}{2\Delta f}$ With ($\Delta f = \frac{\int_{-\infty}^{+\infty} |S(f)|^2}{|S(f)^2|} df$ and S (f) signal spectrum s(t)) (2)

In the case of an unmodulated pulse, $\Delta R = \frac{c}{2\Delta f}$ the distance resolution is increased by a radar antenna with a broadband based on the transmission of very short pulses, often less than the nanosecond we speak here about margin of gigahertz.

- ✓ Angular resolution is the minimum angular difference that allows the radar to distinguish two identical targets at the same distance. The angular resolution of a radar is determined by the width of its antenna lobe, itself defined by its angle at -3 dB (half-power lobe width), more the lobe is narrower, more the directivity of the antenna is important.

Current wireless communication systems and radar systems require higher speed transmission to support various services. A UWB system with high directivity and gain is considered as a promising solution to the problem of distance and angular resolution.

Various radar systems such as Synthetic Aperture Radar (SAR), Shuttle Imaging Radar, Remote Sensing Radar and other wireless communication systems operate in the L, C and X bands [6][7]. The higher the frequency is, the greater the atmospheric absorption is, and causes the attenuation of the radar beam which limits the distance resolution and range. Ka-band radars are known by their very short range, their extremely high precision and their highly fast rate of scanning, they use very short pulses of a few nanoseconds at these frequencies, they can also follow the movements of the vehicles on the tarmac and the airport runways, their resolution permits the visualization of the vehicles silhouette [4].

The micro-band antenna is the first option for this high-frequency band because of its low cost, light weight and robustness. Many researchers have worked and presented studies on micro-band antennas (Patch) which make it possible to obtain the expected solution [8]-[10]. With a narrower beam and lower output power than the X and K bands, the Ka-band antennas are the most persuasive radar type in the various research and recommendations, with a challenge of getting a wide duel band and increase Gain with a radiation pattern characterized by a narrow main lobe and eliminated the side lobes for a frequency band between 26.5 and 40 GHz.

3. Proposed antenna

The study was deeply made in order to design an antenna for surface movement Radar application; the two resonant frequencies wait respectively the reception and emission in a margin of [23-28] GHz. Many researchers have worked and presented the techniques to get multiband operation using multi stack patches [11].

Using U-slot loaded patch stacked with H-shaped parasitic elements [12]. In [13] an annular ring patch is loaded with a slot, where the loaded slot makes the design capable for dual band operation. Two asymmetric horizontal strips are used as additional resonator to produce the lower and upper resonant modes [14]. Although U-slot patch antenna is used is concerned with broadband capabilities and U slot to obtain triple band with a second slot is needed [15]. Multi frequency operation can also be achieved using spiral printed antennas, Bowtie patches loaded with slots, multiple patches, cutting rectangular patch with L-shaped slit, U-slot and inserting slits with different dimensions and adjust the length, width according to the resonance frequency and also different feeding techniques applied to match the impedance [16], [17].

- **Structure and Scope of the antenna**

The geometry of the proposed antenna is based on a rectangular patch stacked by two U-Slots used asymmetrically, to improve the antenna operation, an H-shaped slot has been stacked for the purpose of having an optimal surface current distribution, the length and width of the slots will be adjusted according to the resonant frequency to receive the High gain possible and voltage standing wave ratio (VSWR) less than 2 and to get stable radiation characteristics throughout the frequency band 23.3 GHz to 29.3GHz. . The feeding technique used is the microstrip line supported by notches in order to have a good adaptation of input impedance; a L-slot was loaded in microstrip line feed to produce multi-band operation lower and higher resonance modes.

Among the different substrates used is the FR-4 substrate which is widely used in patch antennas. The inconvenience of FR-4 is that it is difficult for a higher bandwidth; therefore the dielectric constant of the antenna substrate should be as low as possible to increase the impedance bandwidth and reduce the wave excitation surface. The use of a low dielectric constant substrate increases also the bandwidth because it is inversely proportional to the dielectric constant. The proposed antenna is printed on a Duroid 5880 substrate material which has a reduced value of the constant dielectric with relative permittivity = 2.22 and loss tangent of 0.0004, operating up to 40 GHz with good thermo-mechanical stability. Table.1 Duroid 5880 has lowest dielectric constant among the different substrates which also increases the bandwidth because bandwidth is inversely proportional to dielectric constant.

Table 1. DIELECTRIC PROPERTIES OF DIFFERENT SUBSTRATE MATERIALS

Material	Permittivity	Loss tangent
Duroid 5880	2.22	0.0004
Duroid 5870	2.33	0.0023
Duroid 6010	10.8	0.0023
FR4	4.66	0.020

In the design of a rectangular micro ribbon patch, two parameters are important (the length and width of the patch). These two parameters depend on the resonance frequency, the dielectric constant of the substrate used and its thickness. They are calculated as follow [18]:

3.1 The width of the patch:

The width W is calculated from the formula of Bahl and Bhartia :

$$W = \frac{c}{2 f_r} \left(\frac{\varepsilon_{r+1}}{2}\right)^{-0.5} \quad (3)$$

3.2 Edge effect (Fringing Effect) :

The propagation is done in two different dielectric environments: the relative permittivity air equal to 1 and the relative permittivity substrate εr with two different speeds, taking into account this inhomogeneity, an effective medium of permittivity $\varepsilon reff$ replaces the two previous environments.

For W / h> 1 ; $\varepsilon reff$ is given by the explicit formula of Hammerstad:

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-0.5} \quad (4)$$

ε_{reff} : Effective Permittivity· ε_r : Relative permittivity. h : Substrate height

3.3 The effective length

Due to the effect of the edge, electrical dimensions appear which are longer than the physical dimensions. For the rectangular patch the length L is extended by two ΔL. This extension is given by the empirical formula of Hammerstad:

$$\Delta L = 0.412h\frac{(\varepsilon_{reff} + 0.3)}{(\varepsilon_{reff} - 0.258)}\frac{(\frac{W}{h} + 0.264)}{(\frac{W}{h} + 0.8)} \quad (5)$$

- The effective length $Leff$ of the patch becomes:

$$L_{eff} = L + 2\Delta L \quad (6)$$

- For a given frequency fr the effective length is given by:

$$L_{eff} = \frac{c}{2f_r\sqrt{\varepsilon_{refff}}} = \frac{\lambda_g}{2} \quad (7)$$

, Where λg is the guided wavelength in effective medium.

According to the design calculation above, the proposed antenna was designed with 9.4x6.12mm dimensions, is powered by micro strip line feeding techniques. The geometry of the proposed antenna is illustrated in the following **Figure 1**:

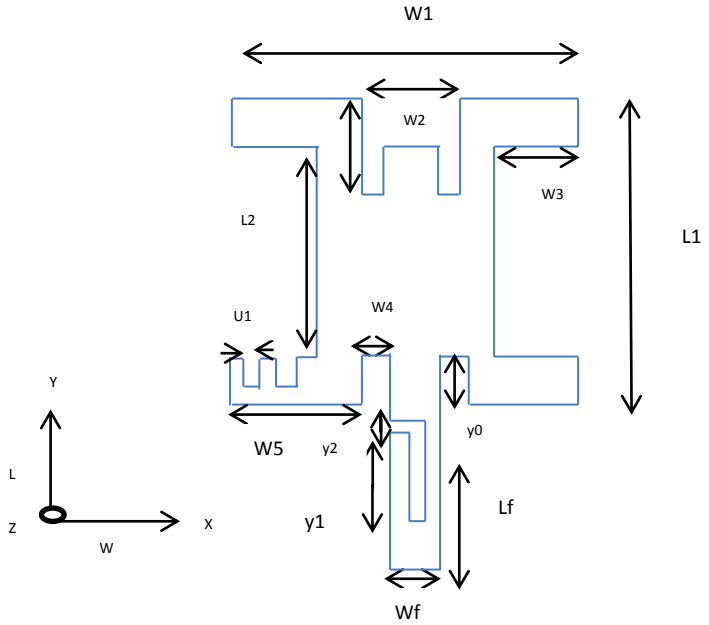


Fig.1. Geometry of the proposed antenna.

The radiating dimensions of proposed antenna are follows:

W1=9.4 mm, L1=6.12 mm, W2=1.45 mm, W3=2.8 mm, W4=0.2 mm, W5= 4.22 mm, L2= 5.12 mm, Lf=3.9 mm, y0= 1.2 mm, Wf= 0.4 mm, y1= 2.2 mm, y2=0.5 mm, u1=0.5 mm, u2= 0.25 mm, y4=1.92 mm.

4. RESULTS AND DISCUSSIONS

The antenna performances were analyzed with the Computer Simulation Technology (CST), the geometry of the proposed antenna in shown in **Figure 2**. It is printed on Duroid 5880 substrate material with low permittivity of 2.22 and thickness of 0.4mm, the ground size of proposed antenna is 18.5x3.9mm which is shown in **Figure 2,3.**

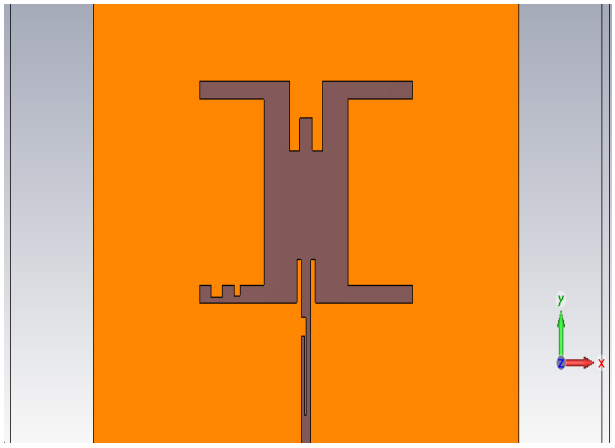


Fig.2. Design front of the proposed antenna.

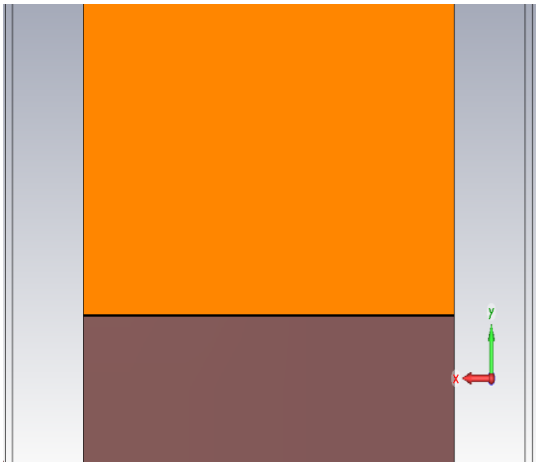


Fig.3. Back side of the proposed antenna

The proposed antenna has as a base a rectangular patch structure with dimensions calculated by transmission line model equations, in order to have a UWB in the margin of the band Ka, , the result of the design is presented in **Figure 2**.

The proposed antenna was developed by the following steps:

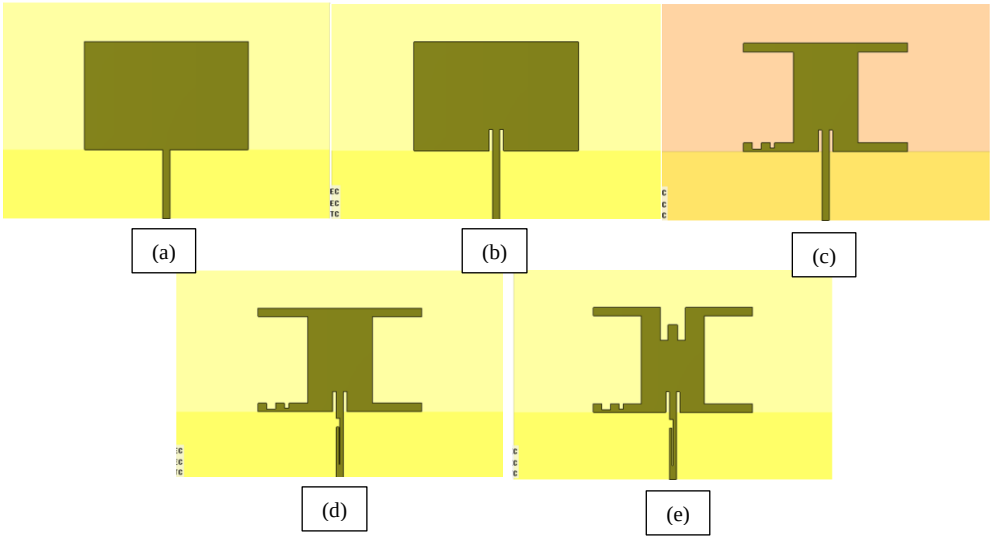


Fig.4. steps followed to optimize the proposed antenna.

(a) the proposed antenna is a basic structure of a rectangular patch antenna powered by a microstrip line, its dimensions are calculated by the transmission line analysis method to obtain resonance on the Ka frequency range. The patch etched on a substrate characterized by permittivity = 2.22 and tangent loss of 0.0004. The partial ground plane under the feeding part plays a vital role in increasing the bandwidth. This is due to the electromagnetic coupling between the split ground plane and the radiating patch. (b) in this work, we have chosen to excite the antenna with a feed line with notch because it is easy to manufacture, and simple to adapt to the antenna. The impedance of the patch is given by the relation [19] —(–) , a simple formulation is elaborated by the work carried out in [20] for the calculation of the length and the width of the notch in order to have a good adaptation.

(c) at this stage we chose to use the slot technology in the patch we made two asymmetrical U-shaped slot in such a way as to have the beginning of double resection, and to adjust on the two rectangular small slot to get a

desirable frequencies (transmission and reception), we used the experience that was in this work of M.pachiyaannan [9].

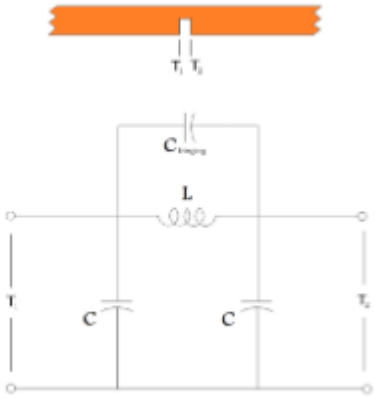


Fig.5. Rectangular small slot

Most of the studies of the U-slot patch antenna were concerned with its broadband capabilities. However, recent studies have shown that it can be also used for dual- or multiband designs. There are two approaches, depending on whether the frequency ratios are larger than or less than about 1.5. The two cases will be referred to as large frequency ratio and small frequency ratio, respectively [8]. Simple approach to designing dual-band patch antenna with frequency ratios of less than about 1.5 has been introduced in Lee and al. (2008, 2011). The method starts with a broadband antenna. When we use U-slot patch antenna in a special size, the bandwidth starts to widen towards the U-slot length frequency.

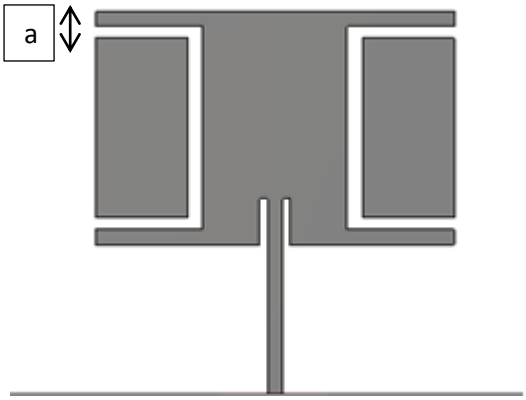


Fig.6 Two U-slot patch antenna

We have chosen to adjust parameter ‘‘a‘‘ that presents the width of the arm of U-slot :

Table 2. Table the change of frequency depending on a.

a (mm)	0.3	0.8	1.2	1.8	2	Lu/2
f_r	34-29 GHz	32-27 GHZ	31-26 GHZ	29.5-27 GHZ	28-26 GHZ	28-23 GHZ

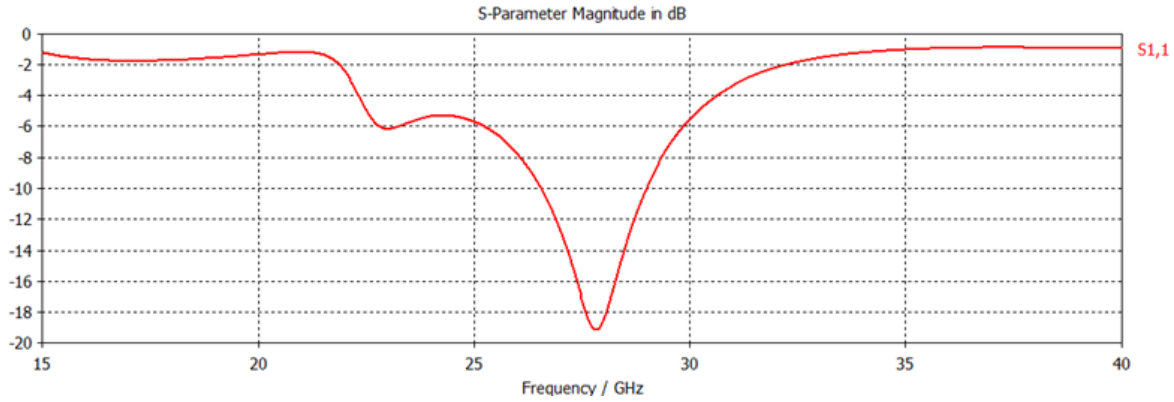


Fig.7. S-parameter S11 Patch antenna when a equal $l_u/_2$

(d) in this step an L-shaped slot used at the feed line that allows a current flow of the surface capable of adapting two resonance frequencies that was already produced by the two U-slot. (e) in the final step the antenna load by H slot in order to obtain a voltage standing wave ratio less than 2 (until 0.9). Figure .4. shows the development of the reflection's coefficient and the creation of multiband with a wide band according to the antenna design advancement a function of different stages used (a,b,c,d and e).

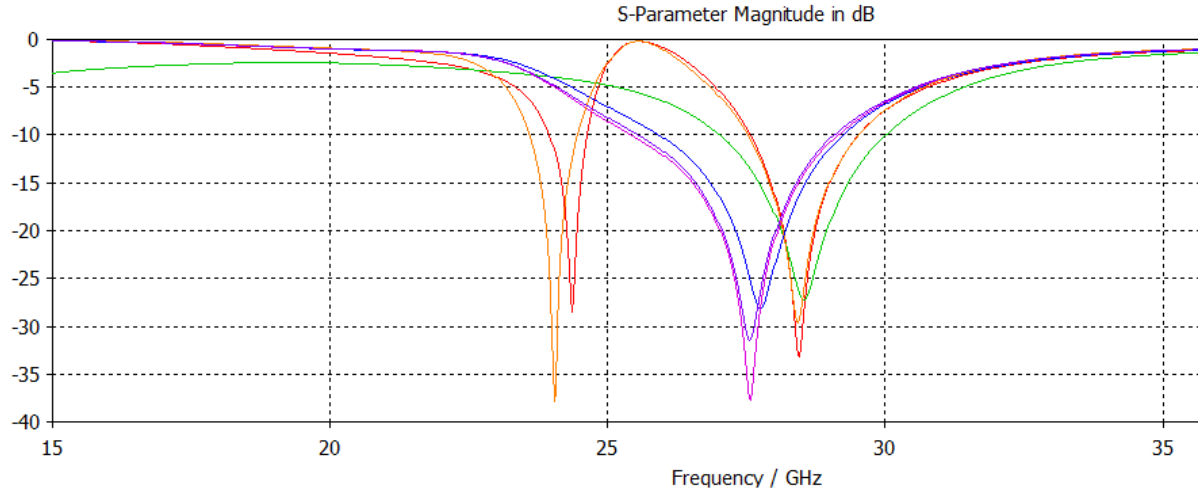


Fig.8. Simulated Return loss of different setups

Figure 5,6. Shows the return loss and VSWR characteristics of the proposed antenna which satisfying Ka-band UWB operating frequency. The antenna exhibits dual resonance at 23.8GHz and 28.36GHz and its having a return loss of -28.16dB and -40.23dB. This bandwidth is characterized by an ultra-wide band and voltage standing wave ratio <2.

- At 10 dB return loss display a bandwidth at 1 GHz in the vicinity of 23.3-24.3GHz at 23.8GHz, thus VSWR 1.09.
- At 10 dB return loss display a bandwidth at 2 GHz in the vicinity of 27.4-29.4GHz at 28.36GHz, thus VSWR is 1.006.

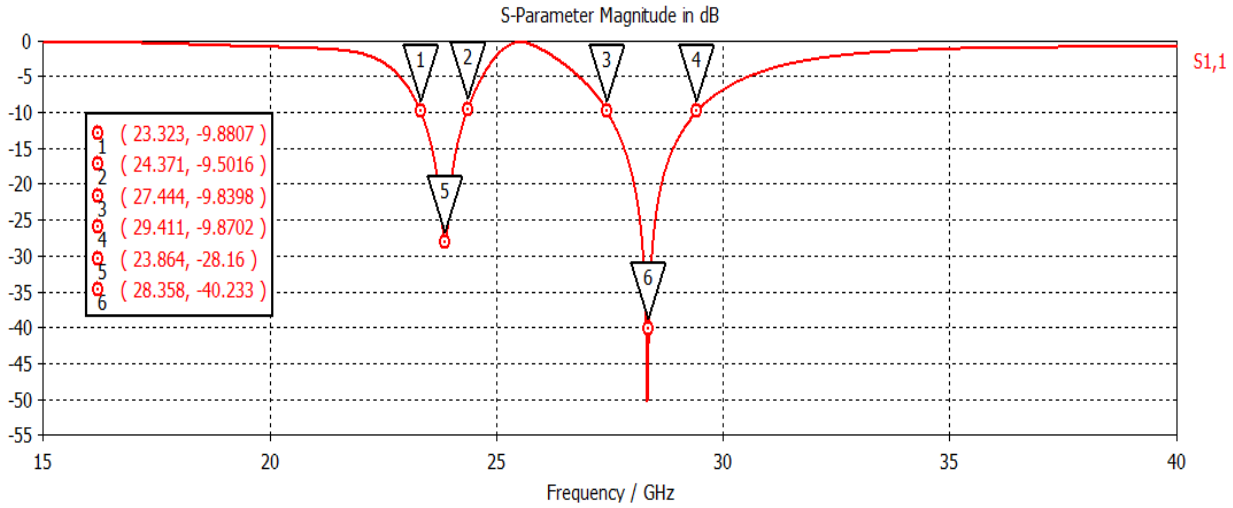


Fig.9. Simulated Return Loss of proposed antenna

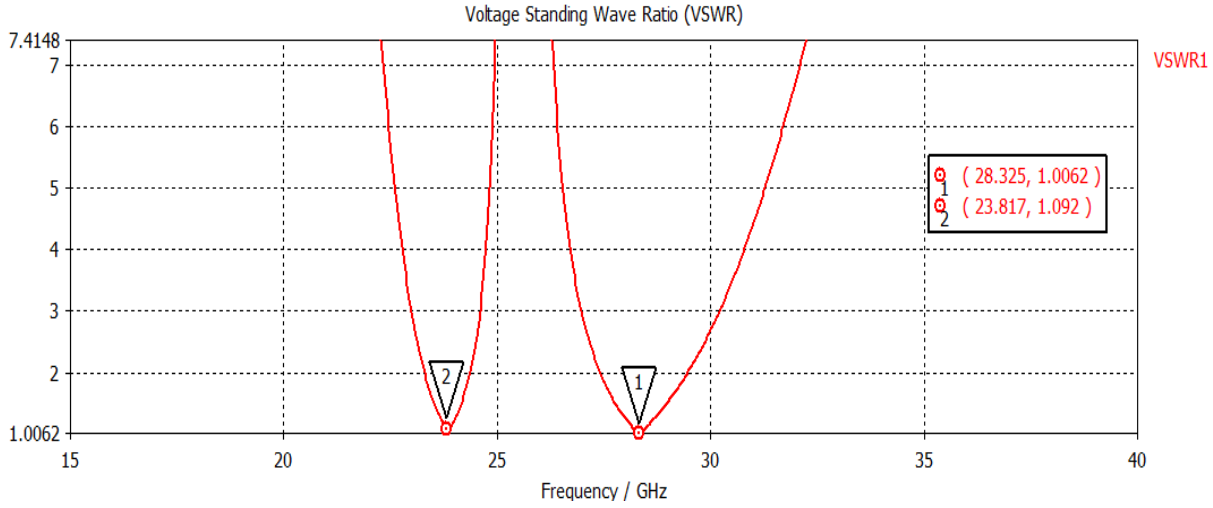


Fig.10. Simulated VSWR of proposed antenna

The simulated radiation efficiency of proposed antenna is shown in **Figure 8**. The frequency Vs radiation curve observe that the radiation efficency is greater than 92% in the both frequancy [23.8-28.3] Ghz.

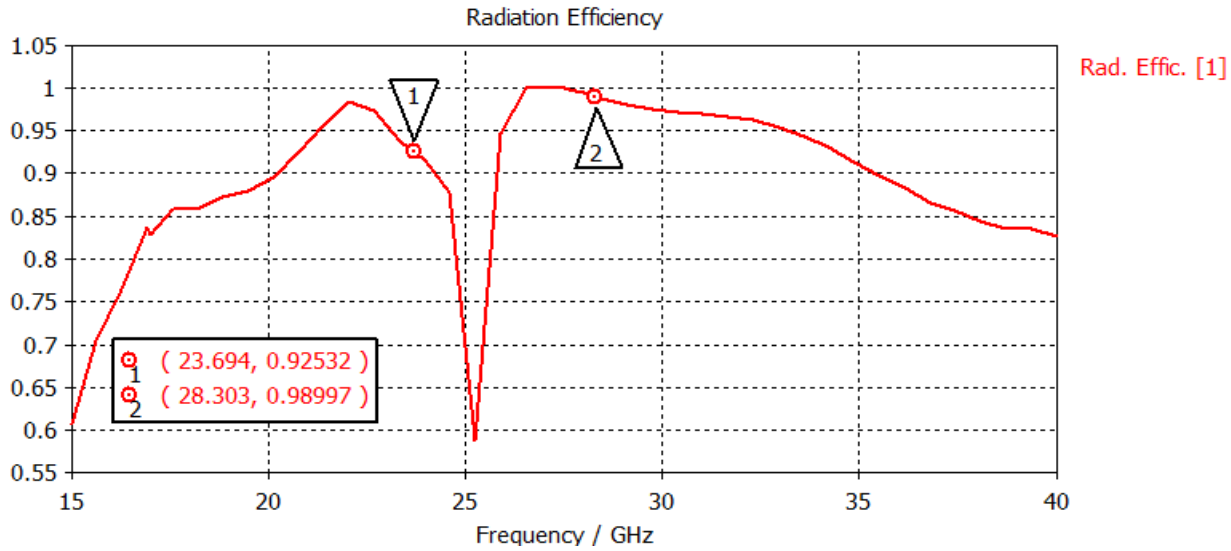


Fig.11. Simulated radiation efficiency proposed antenna

The peak gain of antenna defines that the ratio of the intensity in a given direction to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotropically. Figure 9 shows the antenna peak gain in 2D patterns. The gain of proposed antenna obtained peak gain at 3.02dB and 4.3dB for the corresponding frequency of 23.7GHz and 28.8GHz which is moderate gain for this operating band this gain is highly suitable for radar applications.

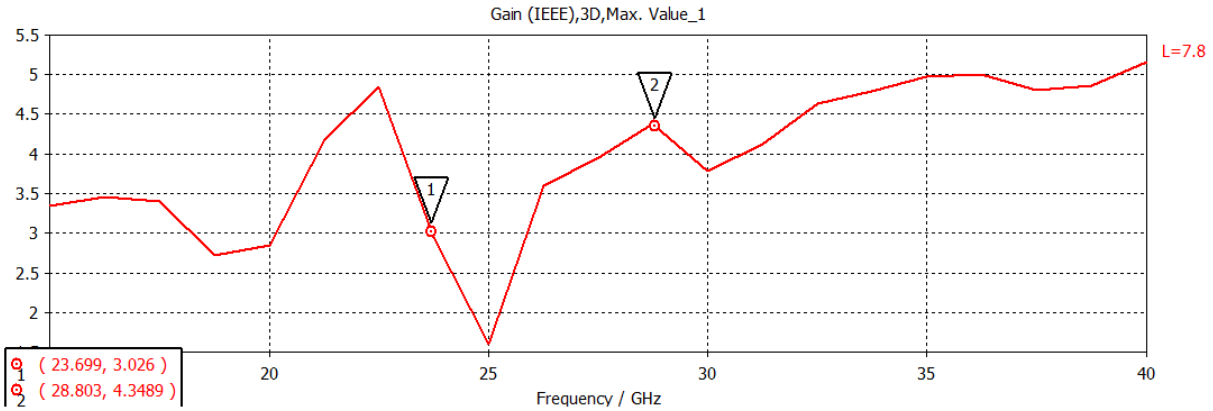


Fig.12. Simulated peak gain of proposed antenna

In literature there are more studies and researches carried out to develop the performances of the antennas in order to obtain the antenna answers as the given exigencies. In this work we compare the results obtained with very interesting work that deals with the same demands. Compact size K-band UWB antenna for surface movement radar: Design and analysis[10], Dual-band UWB antenna for Radar applications: Design and Analysis [21]. The following table shows the value of each work :

Table 3 . The comparison table between the proposed antenna and the two examples.

	Proposed antenna		Example [10]		Example [21]	
Frequency	23.3-24.3GHz	27.4-29.4GHz	19.8-20.7GHz	25.4-27.2GHz	5.7-16.5GHz	22.9-23.7GHz
Bandwidth	1 GHZ	2 GHZ	0.9 GHZ	1.7 GHZ	800 MHZ	500 MHZ
Gain	3 dB	4.4dB	2.6 dB	3.9 dB	4.6 dB	3.8 dB
S11	-25.16 dB	-40.23 dB	-20.2 dB	-26.2 dB	-24dB	-23.6 dB
Efficiency	93.1%	98.9%	94.6%	97.2%	95.1%	93.2%
VSWR	1.07	1.02	0.6	1.2	1.14	1.24

according to the comparison table, the proposed structure presents very important advantages to the adaptation, gain and bandwidth parameter which meets the need for a movement surface radar antenna [2] .the proposed antenna frequency margin to reach a very interesting range in the field of aviation and satellite communication.

5. CONCLUSIONS

In this document, we explained the study that we have conducted to obtain an elementary antenna for the movement surface radar which has important characteristics for a radar antenna in the ka-band, this result is presented as the obtain of a 23Ghz and a 28Ghz, multi band frequencies with UWB of 1 and 2Ghz. the two frequencies present the emission and the reception of the antenna. The proposed antenna presents interesting results in terms of gain and bandwidth in the Ka- band, the next studies will be dedicated to put this element in an antenna network in order to have a high gain and increase the directivity to finally have a radar antenna respecting the requirements mentioned in this paper (gain, bandwidth, radiation pattern, polarization).

References

[1] International Civil Aviation Organization: Advanced Surface Movement Guidance and Control Systems (A-SMGCS) Manual. Doc 9830 (2017)

[2] EUROCONTROL : Specification for Advanced-Surface Movement Guidance and Control System (A-SMGCS) Services. (No.171) – Version 1.0. 01 March (2018)

[3] Xinping Chen ; Zhonghui Gao ; Yunfeng Chai : The development of air traffic control surveillance radars in China" IEEE Radar Conference. pp. 1776 – 1784 (2017)

[4] P.N. Melezhik ; V.B. Razskazovskiy ; N.G. Reznichenko ; V.A. Zuykov ; A.V. Varavin ; Yu.B. Sidorenko ; S.V. Provalov ; F.J. Yanovsky : High-efficiency millimeter wave coherent radar for airport surface movement monitoring and control. 8th European Radar Conference. pp. 361 – 364 (2011)

[5] Université de Blida, Faculté des sciences de l'ingénieur. Radar course. Vol. 2, pp. 30- 40 (2010)

[6] Md. Naimur Rahman ; Mohammad Tariqul Islam ; Norbahiah Misran ; Md. Samsuzzaman: A tuning fork-shaped microstrip patch antenna for X-band satellite and radar applications. 6th International Conference on Electrical Engineering and Informatics. Pp. 1 – 2 (2017)

[7] N. Prema, : Design of multiband microstrip patch antenna for C and X band, Optik-International Journal for Light and Electron Optics, vol. 127, pp. 8812-8818, (2016)

[8] F. Oktafiani ; Y. P. Saputera ; Y. Wahyu : Microstrip patch array antenna with horizontal polarization for ISRA RADAR. International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications. pp. 169 – 173 (2017)

[9] Victor J. Marrero-Fontanez ; Rafael A. Rodriguez-Solis : Dual polarized microstrip antenna array for X-band radar applications. IEEE Antennas and Propagation Society International Symposium. pp . 2144 – 2147 (2007)

[10] M. Pachiyaanan ; G. K. D. Prasanna Venkatesan : Compact size K-band UWB antenna for surface movement radar: Design and analysis. International Conference on Applied and Theoretical Computing and Communication Technology. pp. 693 – 696 (2015)

[11] Simon Mener ; Raphael Gillard ; Langis Roy : A Dual-Band Dual-Circular-Polarization Antenna for Ka-Band Satellite Communications. IEEE Antennas and Wireless Propagation Letters. Volume: 16, pp. 274 – 277 (2017)

[12] Ansari, J.A.., Singh, P., Dubey, S.K., Khan, R.U., and Vishvakarma, B.R. : H-Shaped Stacked Patch Antenna For Dual Band Operation, Progress In Electromagnetics Research B, Vol. 5, 291–302 (2008)

[13] Jhamb, K., Li, L., and Rambabu, K.: Frequency adjustable microstrip annular ring patch antenna with multi-band characteristics", IET Microw. Antennas Propagation. Vol. 5, Iss. 12, pp. 1471–1478 (2011)

[14] Chen, S.B., Jiao, Y.C., Wang, W., and Zhang, F.S : Modified T-shaped planar monopole antennas for multiband operation. IEEE Transactions on Microwave Theory and Techniques, Vol. 54, no. 8, pp. 3267–3270, 2006.

[15] Kai-Fong Lee., and Kin-Fai Tong. : Microstrip Patch antennas-Basic characteristics and some recent advances, Proceedings of IEEE, Vol. 100, No.7,July (2012)

[16] Motinl,M.A.,Imran Hasan,Md., and Asaduz-Zaman.Md : Design and Optimisation of a Low Cost Multi Band Microstrip Patch Antenna for KBand, Ku-Band and X-Band Applications. (2012)

[14] Kin-Lu Wong : Compact and Broadband Microstrip Antennas John Wiley & Sons, Inc, Third Edition,pp.45-79,.Publication (2002)

[18] I. J. Bahl and P. Bhartia : Microstrip Antennas. Artech House, Boston, Mass, USA, 2nd edition, (1980)

[19] A. F. Alsager, "Microstrip Antennas "Design and analysis of Microstrip Patch Antenna Arrays

[20] M. Ramesh and YIP KB Design Formula for Inset Fed Microstrip patch antenna. Journal of Microwaves and optoelectronics, Vol. 3, N°. 3, December 2003.

[21] M. Pachiyaannan ; G. K. D. Prasanna Venkatesan : Dual-Band UWB Antenna for Radar Applications: Design and Analysis, 8th International Conference on Computational Intelligence and Communication Networks (CICN). Page s: 196 – 199 (2016)

New Design of a Microstrip-Fed Printed Slot UWB Antenna with Hilbert Fractal Defected Ground Structure

[1]Khadidja Krim, [2]Mehadji Abri, Hadjira Badaoui, Samira Fersiti, Fatima Zahra Mahieddine

{ [1]krim.khadidja@gmail.com, [2]abrim2002@yahoo.fr, [3]elnbh@yahoo.fr }

STIC Laboratory, Department of Telecommunication,
Faculty of Technology,University of Tlemcen,
Algeria.

Abstract. In this paper, we propose originals structures of monopole microstrip antennas for UWB applications with Hilbert curve Defected Ground Space (DGS) structure. These Antennas are made of two different shapes of slot inside a square patch, fed by a 50 Ω microstrip line on a FR4 dielectric substrate. A 3rd iteration of the Hilbert space filing curve is cut out from a partial ground plane. These antennas have a wide bandwith varying from 3.21 GHz to 13.64GHz and from 3.18 GHz to 14.32 GHz for S11< -10dB. Details of the two antenna designs approach and simulated results are presented and discussed in this paper.

Keywords: Antenna, Hilbert Fractal shape, Ultra Wide Band Antenna, Microstrip patch antenna.

1 Introduction

Defects integrated on the ground plane or slots of microwave planar circuits are colled Defected Ground Structure. DGS is used as an emerging method for improving the various parameters of microwave circuits, like: narrow bandwidth, low gain, cross-polarization ...etc.[1].

developments in wireless communication system srequired additional challenges on circuit designers and microwave antenna to ceate new designs miniaturized and multiband [2],[3].

The pionee work of Mandelbrot [4] had stimulated microwave circuits and antenna designers in 1975, who were trying to realize miniaturized circuits and components, to look for solutions by exploring different fractal geometries [5] derived from the latin word “fractus”, signifying “broken” or “fractured”. The object was to describe nature’s geometry and classify complex geometries that were designed with an iterative procedure [6], whose dimensions were not whole numbers [7].

Fractal curves have a unique property that, after an infinite number of iterations, their length becomes infinite although the entire curve fits into the finite area. This property can be used in order to minaturize microstrip antennas, resonators, and filters. Because of technology limitations [8-9], fractal curves are not physically realizable. Pre-fractals, fractal curves with finite order, are used rather [10], [11].

Hilbert, Peano, and Gosper space-filling curves have stimulated the researchers to realize miniatures antennas with multiple resonances [12].

In this paper two new types of monopole microstrip antennas are proposed, Obtained by the third iteration of the Hilbert fractal curve cut out of a partial ground plane, and two shapes of slot in a square patch.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284219

The proposed antennas are characterized by a wide bandwith with a stable direction of radiation.

2 Description of the proposed antennas

Hilbert curve is a continuous fractal space-filling curve first described by the German mathematician David Hilbert in 1891, as a variant of the space-filling Peano curves discovered by Giuseppe Peano in 1890.

Various iteration stages of Hilbert curves are shown in Fig.1 [13].

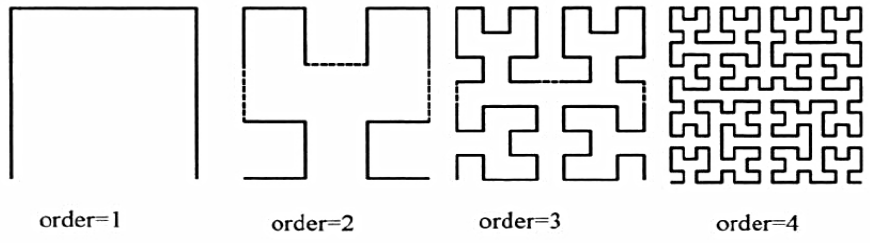


Fig. 1. First four fractal iterations for the Hilbert curve geometry.

In this paper, we propose two monopole antenna structures composed of different shapes of slots inside a square patch.

These antennas were printed and fed by a 50Ω microstrip transmission line on a FR4 dielectric substrate with thickness h=1.54mm, and a loss tangent of 0.018 and dielectric constant εr=4.3.

The third iteration of Hilbert curve slot was constructed on the ground plane side of a substrate measuring 33mm by 33 mm.

Study of these antennas was conducted using CST Microwave Studio Software.

3 Simulation Results

A "I" shape of slot inside a square patch is simulated with third iteration Hilbert curve placed on the center of the partial ground plane as shown in fig. 2 (a) and (b). The patch is modified by inserting small steps (Fig. 2 (a)).

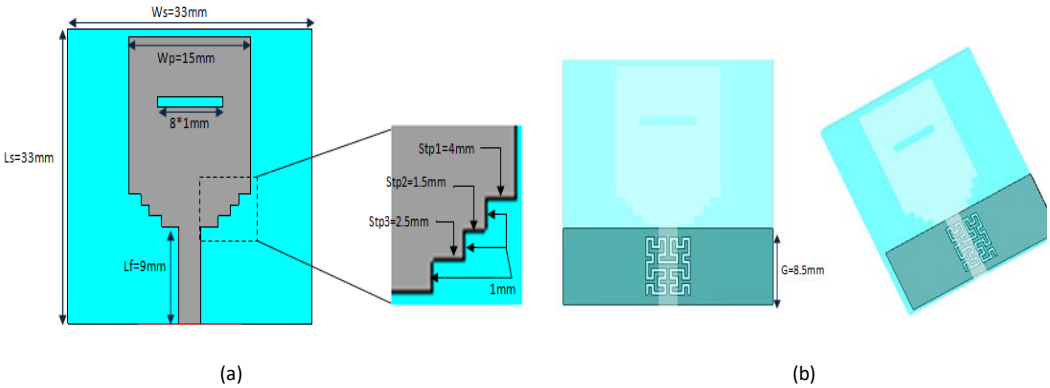


Fig. 2. (a) Rectangle Monopole Antenna geometry. (b) Ground Plane structure for the first antenna.

Fig. 3 illustrates the reflection coefficient S11 of the antecedent designed antenna from 1 GHz to 16 GHz.

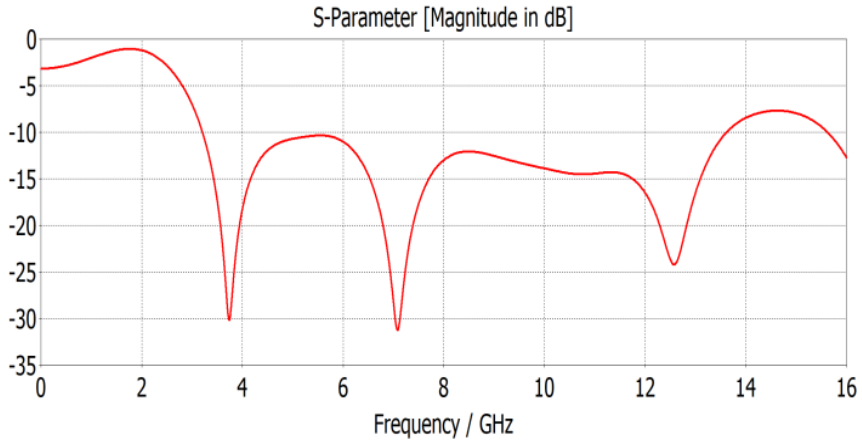
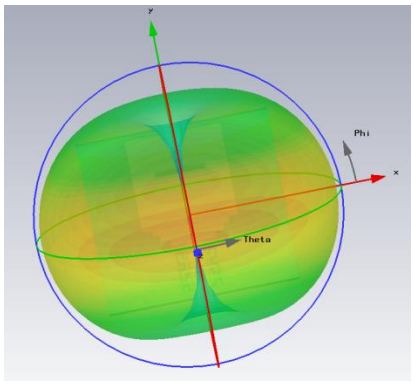


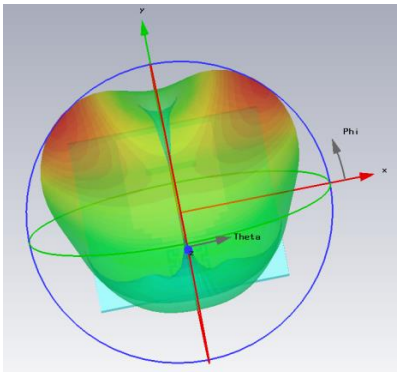
Fig. 3. Simulated Reflection coefficient versus frequency of the first designed antenna.

We observe that the antenna resonance is correctly predicted by three peaks -31.64 dB, -30.08 dB and -24.43 dB at 3.74GHz, 7.88 GHz and 12.58 GHz respectively. The band with (with the definition of -10dB) for the designed antenna is from 3.21 GHz to 13.64GHz, which may cover the UWB frequency range.

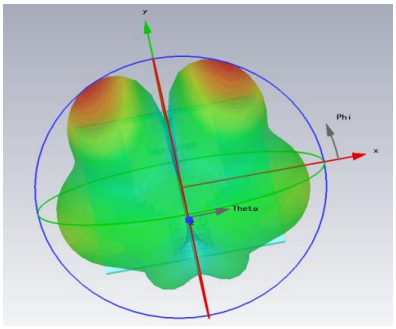
The 3D radiation patterns of the first designed antenna structure at some frequency points are illustrated in Fig.4. It shows acceptable stability of the radiation. It's shown that gain increases for higher frequencies. It's around 2.45 dBi, 3.76dBi, and 5.64 dBi at: 3.74 GHz, 7.33 GHz, and 12.33 GHz respectively.



f= 3.74GHz



f= 7.45GHz



f= 12.55GHz

Fig. 4. Simulated 3D radiation patterns of the first antenna at some frequencies.

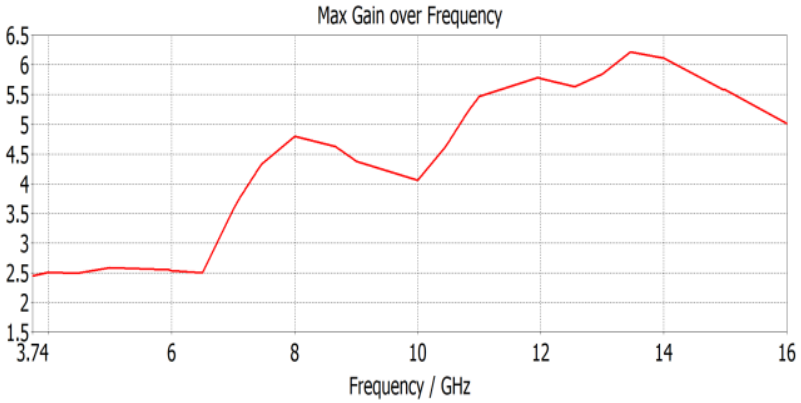


Fig. 5. Variation of simulated Gain in dBi against frequency for the first antenna.

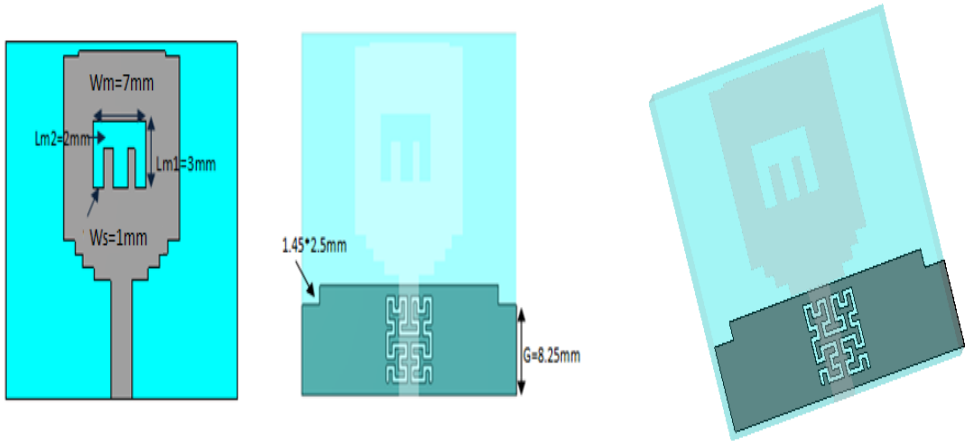


Fig. 6. Geometry of the second proposed monopole Antenna.

For this antenna we use Another shape of slot (E shape), and te ground is modified by inserting two small steps as shown in fig. 6.

Fig. 7 shows the simulated S11 plot of the second designed antenna, the band with (with the definition of -10dB) is from 3.18 GHz to 14.32 GHz, and cover the UWB frequency range.

Three peaks are remarked: -25.82 dB, -26.15 dB, and and -17.02dB at 3.71GHz, 6.97, and 13.05 GHz.

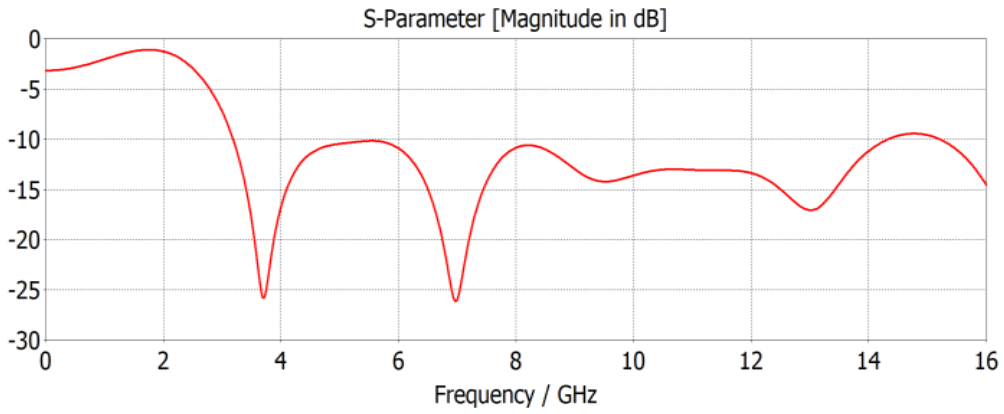
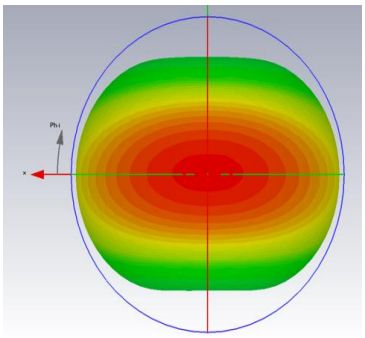


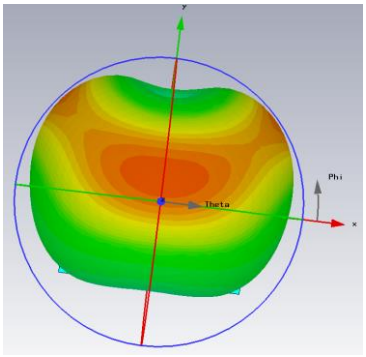
Fig. 7. Simulated Reflection coefficient versus frequency of the second designed antenna.

The 3D radiation patterns of the antenna structure at some frequency points are illustrated in Fig. 8. It shows acceptable stability of the radiation.

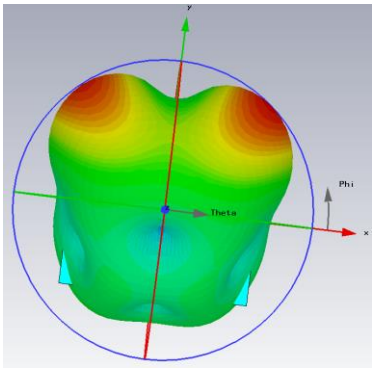
It's shown that gain increases for higher frequencies. It's around 2.5 dBi, 4.51dBi, and 5.14 dBi at : 3.91 GHz, 8.25 GHz, and 12.25 GHz respectively.

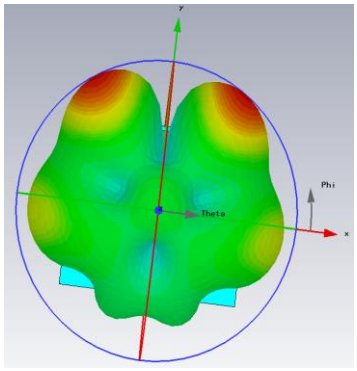


f= 3.91 GHz



f= 6.5 GHz





f= 8.25 GHz

f= 12.5 GHz

Fig. 8. Simulated 3D radiation patterns of the second antenna at some frequencies.

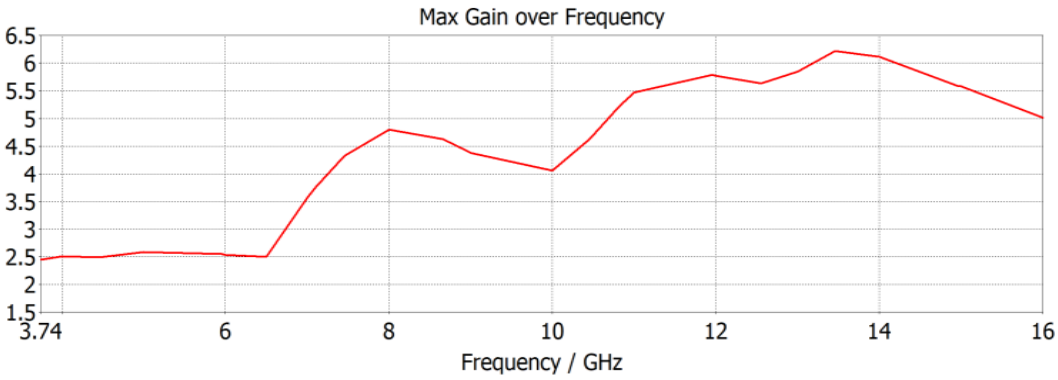


Fig. 9. Variation of simulated Gain in dBi against frequency for the second antenna.

4 Conclusion

In this paper, two new monopole microstrip antennas are proposed; composed of Hilbert curve partial defected ground plane and two different shapes of slot inside a square patch. These antennas have been simulated by SCT microwave studio software.

The simulations results show that both of the proposed antennas offer a wide band with able to be used in UWB applications.

References

[1] Mukesh Kumar Khandelwal, Binod Kumar Kanaujia, Sachin Kumar. 'Defected Ground Structure: Fundamentals, Analysis, and Applications in Modern Wireless Trends, International Journal of Antennas and Propagation' Volume 2017, Article ID 2018527, 22 pages.

[2] S. Didouh, M. Abri, and F. T. Bendimerad, 'Corporate-Feed Multilayer Bow-Tie Antenna Array Design Using a Simple Transmission Line Model', Modelling and Simulation in Engineering Volume 2012 (2012), Article ID 327901, 8 pages.

[3] Mohammed Moulay, Mehadji Abri, and Hadjira Abri Badaoui, 'Quad-Band Bowtie Antenna Design for Wireless Communication System Using an Accurate Equivalent Circuit Model', International Journal of Microwave Science and Technology, Article ID 637607, 2015.

[4] B. B. Mandelbrot, 'The Fractal Geometry of Nature' (W. H. Freeman, 1983).

[5] Jawad K. Ali , 'Microstrip-Fed Printed Slot Antennas Based on Hilbert-Type Space- Filling Curves for Wireless communication Systems' , 2013.

[6] K Falconer, 'Fractal Geometry: Mathematical foundations and applications', John Wiley and Sons, Chichester, 1990.

[7] M. ABRI, H. DIB and A. S. E. GHARNAOUT, 'Accurate Model for Single Bow-Tie Antenna Design', INTERNATIONAL JOURNAL OF MICROWAVE AND OPTICAL TECHNOLOGY, VOL.6 NO.5, pp. 294-300, 2011.

[8] M. Abri, N. Belgacem and W. Belgacem 'New GSM, DCS and GSM/DCS PIFA Antennas Designs for Wireless Networks applications', International Journal of Information & Network Security (IJINS) Vol.2, No.4, August 2013, pp. 305~310 ISSN: 2089-3299, 2013.

[9] M.Moulay, M.Abri, 'Bowtie antennas design for bluetooth/wimax/wifi application', International journalof microwave and optical technology 9 (4), 282-287,2014.

[10] J. P. Gianvittorio, 'Fractals, MEMS, and FSS Electromagnetic Devices: Miniaturization and Multiple Resonances', Ph.D. dissertation, University of California, Los Angeles, 2003.

[11] K. J. Vinoy, K. A. Jose, V. K. Varadan, and V. V. Varadan, 'Hilbert Curve Fractal Antenna: A small Resonant Antenna for VHF/UHF Applications', Microwave Opt. Technol. Lett., Vol. 29(Issue 3):215-219, May 2001.

[12] S. R. Best, 'A Comparison of the Performance Properties of theHilbert Curve Fractal and Meander Line Monopole Antennas', Microwave Opt. Technol. Lett., Vol. 35(Issue 11): 258-262, November 2002.

[13] J. Anguera, C. Puente, E. Martinez, and E. Rozan, 'The FractalHilbert Monopole: A Two-Dimensional Wire', Microwave Opt.Technol. Lett., Vol. 36(Issue 1): 102-104, January 2003.

Design of a Miniature Fractal Microstrip Antenna for RFID Tag

Mohamed Ihamji[1], El Hassan Abdelmounim[2], Hamid Bennis[3] , Mohamed Latrach[4]
{mihamji@gmail.com[1], abdelmou@hotmail.com[2], hamid.bennis@gmail.com[3]}

LASTI, FST of Settat, Hassan 1St University Settat, Morocco[1,2], TIM Research Team, EST of Meknes, Moulay Ismail University Meknes, Morocco[3], Group, ESEO Angers, France[4]

Abstract. This paper presents the miniaturization of a microstrip antenna at 915 MHz in the ISM "Industrial Scientific and Medical" band. This microstrip antenna is designed for RFID tag system, using the fractal technique, on FR4 substrate. The simulation is performed using CST Microwave Studio. The total area of the final circuit is 66.84 x 68.015 mm^2. The validated antenna has a good matching input impedance range from 910 MHz to 920 MHz with a stable radiation pattern, a loss return of -40.67 dB, a directivity of 2.35 dB, and a gain of 1.62 dBi.

Keywords: Microstrip Antenna, RFID tag, Miniature Antenna, Fractal Koch, 915 MHz, UCODE HSL IC.

1 Introduction

The RFID (Radio Frequency Identification) system is generally composed from a reader, and one or more transponders or tags. The communication between the reader and the tag is achieved by using the electromagnetic field. RFID is practiced in a variety of applications as access control, transport, bank, health, and logistic. Most RFID systems operate in either the low frequency region (30–300 kHz), the high frequency band (3–30 MHz), the ultra-high-frequency band (300 MHz–1 GHz), or in the microwave band (1GHz–40 GHz) [1-7].

The microstrip antennas consist of three parts which are radiating element, substrate and ground, and can be designed on many structure shapes such as rectangular, elliptic, circular and triangular, etc. They are famous to be low-profile, low weight, ease of fabrication, conformable to planar and non-planar surfaces and mechanically robust. Rectangular patch is the basic and the most used microstrip antenna in wireless communication [8-10].

One major consideration for handheld and portable RFID applications is the compact size. Therefore, the design of miniature antennas tag is important. Many techniques are used to miniaturize the antenna size, as the slot technique, the fractal structure, the Defected Ground Structure (DGS), or the use of Metamaterial structure. One of the techniques used to design easily a miniature antenna is the use of the fractal structures. The fractal technique is based on the repeat of an initial radiating structure in arbitrary or regular scales, in order to obtain the operational characteristics of the antenna [11-12].

The space filling by scaled replicas of the generator, leads to antenna configurations with electrically large lengths although they have small size thus. There are many popular fractal geometries, referred in the literature, as the Koch fractal, the Sierpinski fractal, the Hilbert fractal, the Minkowski and the Square Curve fractals. The Koch fractal microstrip patches are

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284079

commonly used in virtue of their attractive properties: they have small size, a single feeding port is enough, multi-frequency performance, satisfactory polarization properties and gain [13-21].

In this paper, a novel design of a miniature low cost microstrip antenna for RFID Tag is proposed at frequency of 915 MHz, this antenna is designed by using the fractal structure and the nested slot fed. This circuit is validated by optimization methods integrated in CST-MW Studio.

2 Tag antenna design

The proposed antenna is composed of a rectangular radiating patch with a FR4 substrate that is characterized by:

- ✓ Dielectric constant $\varepsilon r = 4.4$
- ✓ Loss tangent $\tan\delta=0.025$
- ✓ Thickness H = 1.58 mm
- ✓ Metal thickness t= 0.035 mm

A chip named UCODE HSL IC, SL3ICS3001 (UCODE High Frequency Smart Label) from NXP Semiconductor was chosen because it can be used in small passive smart tags. The chip can operate at 867 MHz, 915 MHz, and 2450 MHz. At the 915 MHz frequency, the chip impedance is 34.5 – j 815 Ohms.

The antenna impedance must be equal to the conjugate of the chip impedance in order to have the maximum power transfer to the tag. The nested slot method is used to conjugate the impedance matching with the tag chip, the matching chart in [21], give some idea about the nested slot dimension.

The size reduction of this microstrip antenna is based on the space filling property of fractal geometries, by designing the non radiating edges of patch with the Koch curve.

As shown in the **Figure** 1, the geometry of the first iteration is made by seven copies of the basic geometry (iteration 0), so n is equal to 7. Also, the lengths of the segments making up the geometry of the first iteration, are reduced by a factor of 4, so h = 4. The reduction reached with 2 iterations of width patch is equal to $(n/h)^2 = (7/4)^2 = 3.06$.

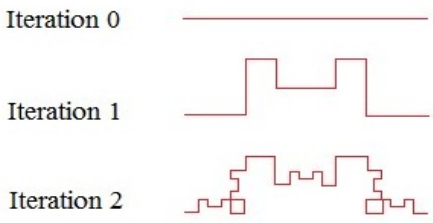


Fig. 1. The 3 iterations of the KOCH Curve.

The size reduction increases with the iteration number and the fundamental resonant frequency decreases when the number of iterations increases.

The patch antenna with the fractal side has the length L and width W. It is connected to nested slot fed. The length L and width W are related to the resonant frequency, to the permittivity and to the thickness of substrate; this length and width can be calculated, theoretically for a conventional patch antenna, by the transmission line method [5].

There is no conducting ground plane placed on the other side of the substrate. **Figure** 2 illustrates the face of the antenna:

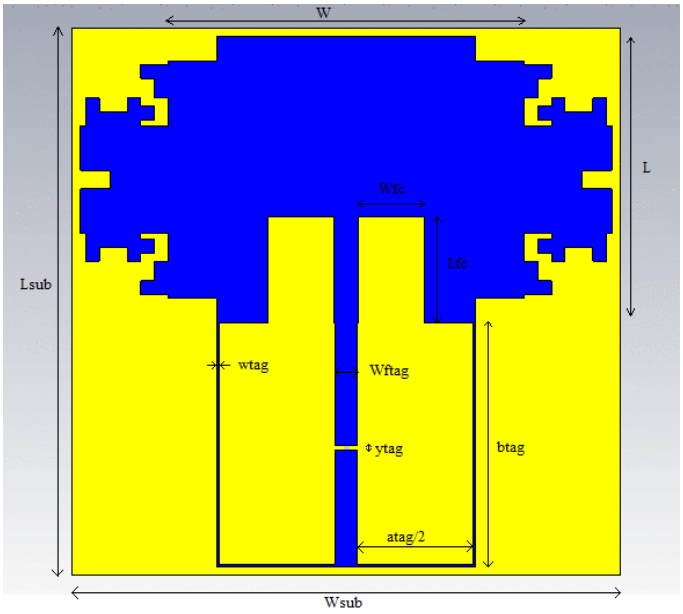


Fig. 2. Geometry of the proposed antenna: Top face.

The following Table 1 indicates the dimensions of the proposed antenna:

Table 1. Antenna Dimensions

Antenna Dimensions	Optimized Value (mm)
L	35
W	44.14
Lsub	66.84
Wsub	68.015
Lfc	13
Wfc	8.2
atag	32.14
btag	29.83
wtag	0.26
Wftag	2.76
ytag	0.5

3 Results and discussion

Using the fractal structure to minimize the patch size, a new antenna structure with fractal side and nested slot fed is proposed. The width and length value of patch antenna are optimized on CST EM solver.

The following **Figures** 3 and 4 show the real and imaginary part of antenna impedance according to the frequency. The simulated antenna impedance at 915 MHz is Zin = 38.4 + j 815 Ohms, and therefore having excellent impedance match with the chip impedance which is 34.5 – j 815. Good impedance matching could ensure a longer antenna reading distance, which is an important condition for RFID application.

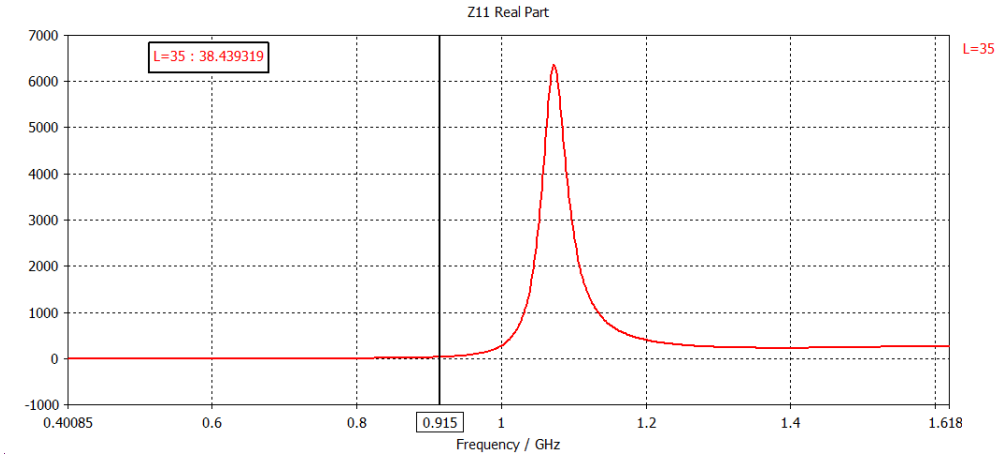


Fig. 3. Real part of simulated antenna impedance.

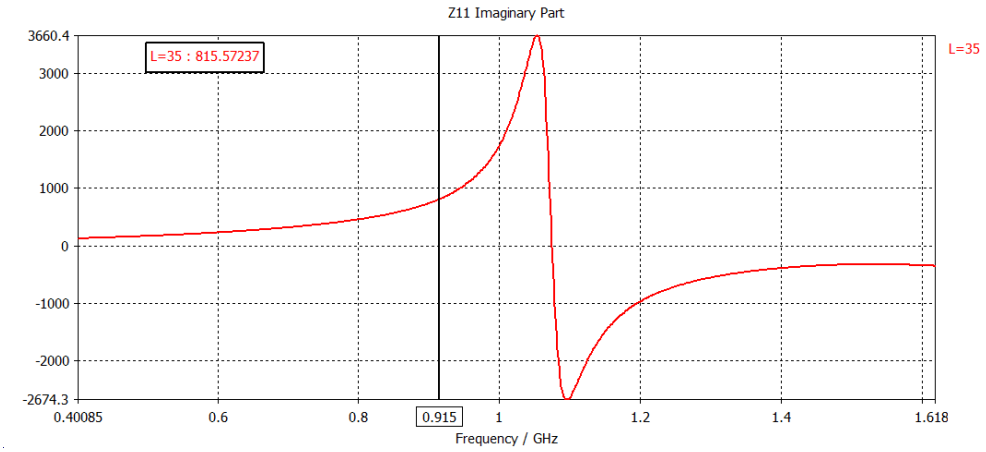


Fig. 4. Imaginary part of simulated antenna impedance.

Figure 5 shows the return loss S11 result for the antenna, it has good matching input impedance in the ISM band.

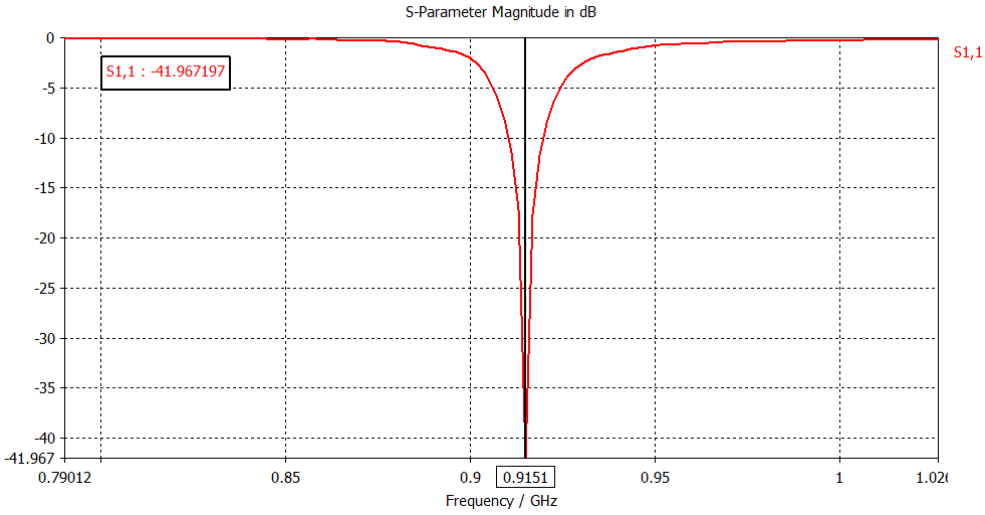


Fig. 5. Return loss (S11) value of the proposed antenna.

The return loss (S11) obyained is equal to –40.67 dB with this patch antenna at frequency 915 MHz. The bandwidth value is equal approximately to 10 MHz.

Figures 6 and 7 illustrate the 2D radiation patterns at E-plane and H-plane respectively. The proposed antenna has an omni-directional radiation pattern for both H-plane and E-plane, that means this tag antenna can be easily detected by reader antenna in any direction. The angular width is 87.4 degree at 915 MHz.

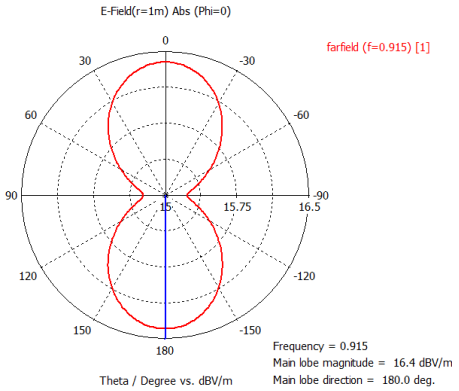


Fig. 6. 2D radiation patterns at E plane of the proposed antenna.

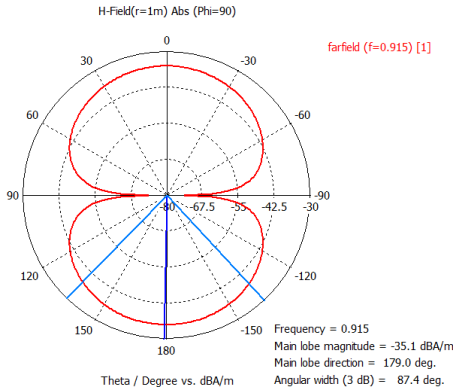


Fig. 7. 2D radiation patterns at H plane of the proposed antenna.

Figure 8 illustrates the surface current distribution of the proposed antenna at 915 MHz. A maximum current is observed around the low side, and around the feed line.

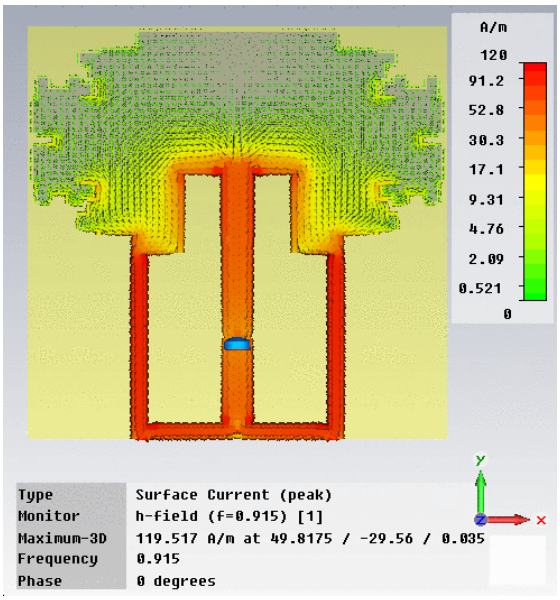


Fig. 8. Surface current distributions of the proposed antenna.

Figures 9 and 10 illustrate the directivity and the gain at the 3D radiation pattern at 915 MHz. The gain and directivity are equal respectively to 1.62 dBi and 2.35 dB, this gain is correct for RFID application in near field. The radiation pattern is omni-directional as presented in these figures.

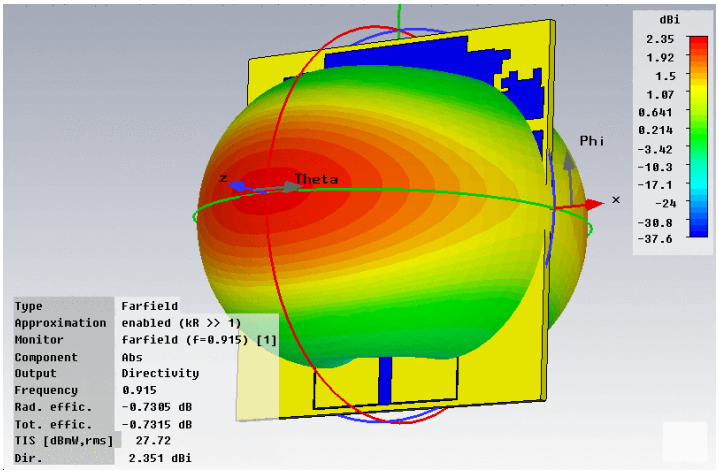


Fig. 9. 3D radiation pattern of the proposed antenna (Directivity).

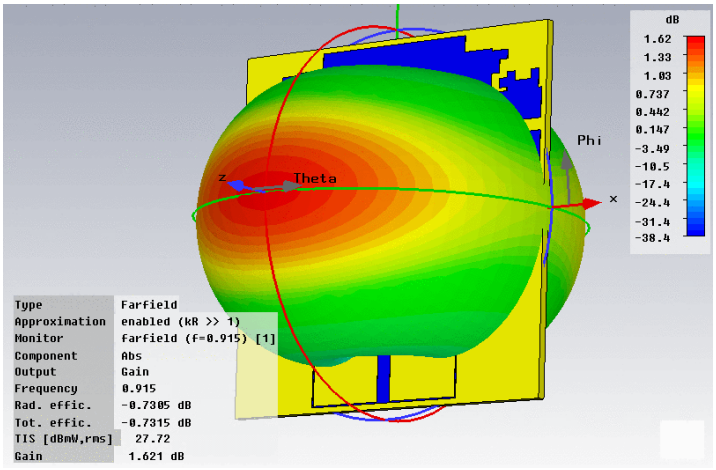


Fig. 10. 3D radiation pattern of the proposed antenna (Gain).

As shown in **Figure** 11, the size characteristics of the proposed antenna structure, Wsub x Lsub= 66.84 x 68.015 mm^2, is very small compared to the conventional rectangular radiating patch antenna at the operating frequency, Wsub x Lsub= 101 x 97.2 mm^2.

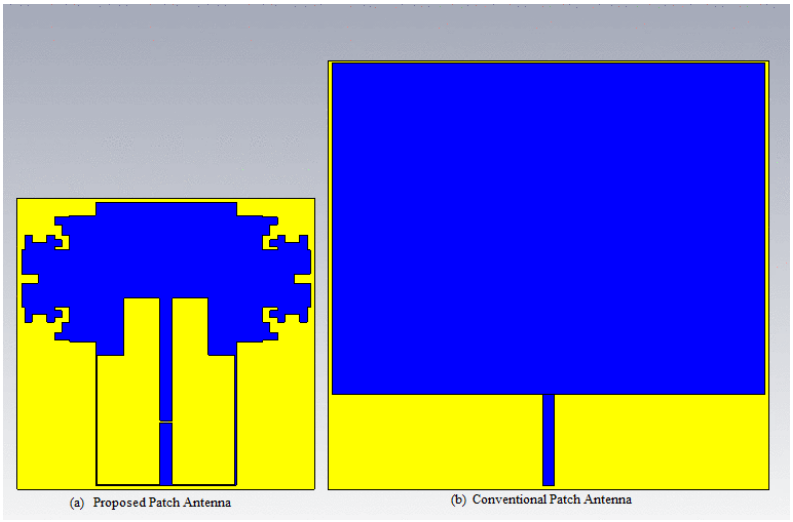


Fig. 11. (a) Proposed patch antenna, (b) Conventional patch antenna.

The results obtained of the proposed antenna, resumed in Table 2 and 3, show that a good miniaturization of the antenna dimensions is obtained with a good return loss S11, and an acceptable bandwidth for RFID applications.

As shown in Table 2, the total size reduction is for example equal to 53.7% compared to the conventional rectangular patch antenna. It is smaller than the True 3D Antenna (38%)

[22], and has approximately the same size reduction as the novel broadband UHF RFID tag antenna (56.8%) [23].

Table 2. Total Size Result Comparison between Proposed Antenna and Others Antennas

Patch Antenna 915 MHz	Wsub (mm)	Lsub (mm)	Total size reduction (Wsub x Lsub)
Conventional antenna	101	97.2	
True 3D Antenna for UHF RFID Application [22]	78	78	38 %
A Novel Broadband UHF RFID Tag Antenna [23]	106	40	56.8 %
Proposed Antenna	**66.84**	**68.015**	**53.7 %**

Table 3. Result Comparison between Proposed Antenna and Others Tag Antenna

Patch Antenna 915 MHz	S11 (dB)	Bandwidth (MHz)	Directivity (dB)
True 3D Antenna for UHF RFID Application [22]	- 28.93	244	-
A Novel Broadband UHF RFID Tag Antenna [23]	- 17.3	17	3.6
Proposed Antenna	**- 40.67**	**10**	**2.35**

4 Conclusion

This paper has presented a design of a new patch antenna with fractal structure for RFID Tag. A size reduction of antenna dimension (WsubxLsub) equal to 53.7% has been obtained compared with conventional rectangular patch. It provides appropriate characteristics, with a return loss equal to -40.67 dB, a bandwidth of 10 MHz, the directivity equal to 2.35 dB, and a gain equal to 1.62 dBi. This tag antenna has an omni-directional radiation pattern which makes RFID tags placed in any direction can be recognized by RFID readers. The antenna has been designed on a standard FR4 substrate and will be realized with conventional Printed Circuit Board (PCB) techniques in order to validate its results. This antenna can be used in inventory control or asset tracking.

Acknowledgments. We thank Professor Mohamed Latrach for allowing us to use all the equipment and EM solvers available in his laboratory.

References

[1] Nasimuddin: Microstrip antenna. Published by InTech Janeza Trdine 9, 51000 Rijeka, Croatia (2011)

[2] Debatosh Guha and Yahia M. M.: Microstrip and printed antennas: new trends, techniques and applications. Antar_ 2011 John Wiley & Sons, Ltd. (2011)

[3] Antony Ghiotto: Tag UHF RFID antenna design, application by jet of material. Thesis. Institut polytechnique Grenoble France (2008)

[4] Nathan D. Reynolds: Long range ultra-high frequency (UHF) radio frequency identification (RFID) antenna design. Thesis. Faculty of Purdue University Fort Wayne, Indiana, US (2012)
[5] D.G. Fang: Antenna theory and microstrip antennas. Taylor and Francis Group, LLC, pp. 103 (2010)
[6] Mélusine Pigeon: Design of ultra-compact antennas based on metamaterials, application to the fabrication of a miniature GNSS antenna. Thesis. Institut National Polytechnique Toulouse France (2011)
[7] K. Nithisopa, J. Nakasuwan1, N. Songthanapitak, N. Anantrasirichai, and T. Wakabayashi: Design CPW fed slot antenna for wideband applications. Piers online, Vol. 3, No. 7 (2007)
[8] T. Deleruyelle, P. Pannier, E. Bergeret, S. Bourdel. Dual band UHF and microwave RFID antenna. Proceedings of the 38th European microwave conference (2008)
[9] G. Monti, L. Catarinucci, and L. Tarricone: Compact microstrip antenna for RFID applications. Progress In Electromagnetics Research Letters, Vol. 8, pp. 191-199 (2009)
[10] Klaus Finkenzeller: RFID Handbook fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication. Third edition, Giesecke & Devrient GmbH, Munich, Germany (2010)
[11] Jithin V M.: Design and simulation of microstrip patch antenna using ADS. Department of Electrical Engineering, BITS Pilani, Pilani Campus, Rajasthan, India.
[12] Nazish Irfan, Mustapha C. E. Yagoub, and Khelifa Hettak: Design of a microstrip-line-fed inset patch antenna for RFID Applications. IACSIT International Journal of Engineering and Technology, Vol. 4, No. 5 (2012)
[13] S. Anscy: Slot microstrip antenna for 2.4GHz RFID reader application. Department of Electronics and Communication Engineering Karunya University. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), Volume 2, Issue 5 (2013)
[14] Chandrappa D. N., Ambresh P. A., Vani R. M., P. V. Hunagund: Multi-slots reconfigurable microstrip antenna with capacitive loading technique. International Journal of Innovative Research in Computer and Communication Engineering Vol. 2 , Issue 1 (2014)
[15] M. Ihamji, E. Abdelmounim, H. Bennis, M. Hefnawi, M. Latrach: Design of Compact Tri-Band Fractal Antenna for RFID Readers. International Journal of Electrical and Computer Engineering (IJECE). Vol.7, No.1 (2017)
[16] M. Ihamji, E. Abdelmounim, J. Zbitou, H. Bennis, M. Latrach: Novel Design of a Miniature Fractal Microstrip CPW Fed Antenna for RFID Reader, IEEE Conference Publications, The International Conference on Wireless Networks and Mobile Communications (WINCOM) (2016)
[17] M. Ihamji, E. Abdelmounim, J. Zbitou, H. Bennis, M. Latrach: Novel design of a miniature L-Slot microstrip CPW-fed Antenna for RFID reader. IEEE Conference Publications, The 11th International Conference on Intelligent Systems: Theories and Applications (SITA) (2016)
[18] M. Ihamji, E. Abdelmounim, H. Bennis, M. Lotfi, M. Latrach: A New Design of a Miniature L-slot Microstrip Antenna for RFID Tag. Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems (ICCWCS) (2017)
[19] M. Ihamji, E. Abdelmounim, J. Zbitou, H. Bennis, M. Latrach: A Miniature L-slot Microstrip Printed Antenna for RFID. TELKOMNIKA Journal. Vol.16, No.5, pp. 1923-1930 (2018)
[20] M. Ihamji, E. Abdelmounim, H. Bennis, M. Latrach: Design of Miniaturized Antenna for RFID Applications. Chapter in Emerging Innovations in Microwave and Antenna Engineering book, Publisher: IGI Global, pp.325-362 (2018)
[21] G. Marrocco: The Art of UHF RFID Antenna Design: Impedance-Matching and Size-Reduction Techniques. IEEE Antennas and Propagation Magazine, Vol. 50, No. 1, pp. 66-79 (2008)
[22] YS. Zhang, B.Wang, H. P. Zhang, and W. X. Kong: True 3D Antenna for UHF RFID Application. Symposium on Quality Electronic Design, pp. 83-85 (2015)
[23] L. Xu, L. B. Tian, and B. J. Hu: A Novel Broadband UHF RFID Tag Antenna Mountable on Metallic Surface. International Conference on Wireless Communications, Networking & Mobile Computing. pp. 2128-2131 (2007)

Phased array antenna for millimeter-wave application

O. Nayat-Ali[1], O. El Mrabet[2], M. Aznabet[3], F.Z. Khoutar[4]

{ Oumaima.nayatali@gmail.com[1], O.mrabet@gmail.com[2], mariamaznabet@gmail.com[3],

fs.tetouan@gmail.com[4] }

Information and Telecommunication Systems Lab[1,2,3,4], Faculty of Sciences, University of

Abdelmalek Essaadi, Tetuan, Morocco,

Abstract. Beam-steering antenna is essential and widely used in communication system.The design of 28 GHz beam steering phased array antenna for 5G millimeter-wave application presented in this communication. The objective of this research is to investigate and implement a phase shifter in a linear antenna array for angle scan. The proposed array antenna has four microstrip patch antenna elements, each feed ports are connected to phase shifter. The main mechanism of radiation is due to the excitation of element in sequence with different phase delays. The overall radiated beam can be steered to different directions depending on the initial excitation of elements. In the first, 3-D good radiation by 0° phase. In the second, good scanning performance of a range of ±10° to ±40°, acceptable side lobe level and gain of 12.1 dB are obtained.

Keywords: Linear Array; Phase Shifters; Phased Array; Radiation Pattern; Microstrip Antenna; Beam Direction.

1 Introduction

In this paper, millimeter-wave communication has become a key enabler for 5G application.

Phased array antennas can adopt a number of different configurations, including linear, planar, and circular. Phased array antennas are used mainly to steer the main beam (main lobe, broadside) of the antenna electronically. It operates by delaying signals coming out of antenna elements by using phase shifter on each antenna element. This provides constructive and destructive interference, thus giving the result that the beam is steered away from the nominal (main) direction of an individual antenna element [1],[6].

This work focuses on linear phased array antennas whose unit cell is formed by a radiating elements, each fed by a transmit and receive module that can provide amplitude and phase control. Linear phased arrays use the progressive phase excitation between the elements to scan the antenna beam electronically over one dimension. The main advantages offered by phased arrays over conventional systems are increased scanning speed, high reliability, and multifunction capability. These advantages make the use of this technology the most logical choice for the next generation weather radars. The use of linear active phased arrays as a component of future low-cost weather radars is the major motivator of this work; consequently, this paper is dedicated to explain the basic concepts related to the theory of linear phase arrays and how their characteristics can be used in the antennas [2].

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284237

The organization of the paper is as follows: Section II describes the phased array antenna principle. Section III discusses issues related to the proposed patch antenna design and simulation results. Section IV presents phased array antenna proposed structure and finally the conclusion.

2 Phased array antenna

In single element antenna, the radiation pattern is usually very broad and the directivity is relatively low. This problem can be affected by enlarging the size of the element thus increasing the directivity. Another way to enlarge the antenna without changing the size of the individual elements is to assemble the radiating elements in a geometrical configuration known as an "array" [5]. The individual elements forming the array are usually identical and they can be of any form (wire antenna, microstrip patches etc…). Many antenna system applications require that the direction of the beam's main lobe be changed with time, or scanned. This is usually done by mechanically rotating a single antenna or an array with fixed phase to the element. However, mechanical scanning requires a positioning system that can be costly and scan too slowly. For this reason, electronic scanning antennas which are known as phased array antennas are used. It can sweep the direction of the beam by varying electronically the phase of the radiating element, there by producing a moving pattern with no moving parts. Phased array antennas are known for their capability to steer the beam pattern electronically with high effectiveness, managing to get minimum side-lobe level and narrow beamwidth [7]. To achieve performance specifications such as narrow beamwidth or considerable scanning range with high angle resolution, a large number of antenna elements were needed to construct the array. Normally phase shifters are the devices in a phased array antenna that allow the radiated beam to be steered in the direction [4].

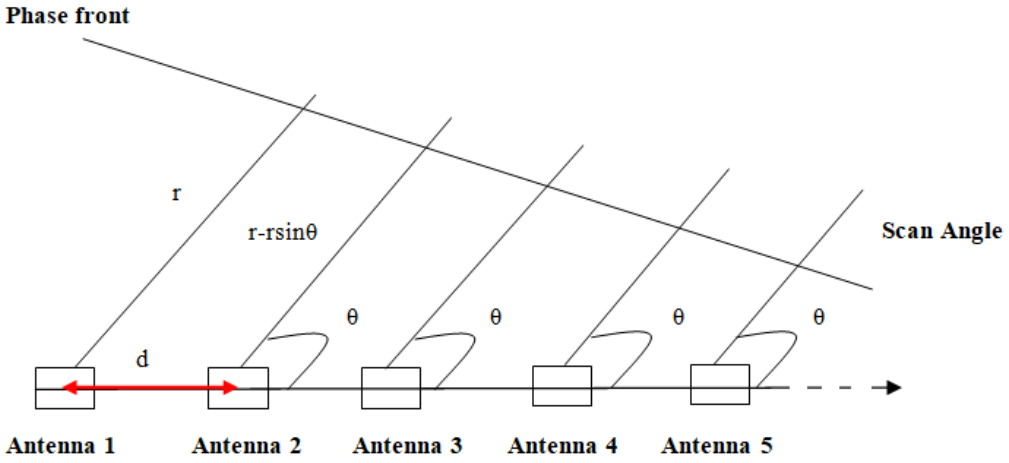


Fig. 1. Uniform linear array of N elements.

In **Figure 1.** shows a uniform linear array of N equally spaced elements having a grating step (separating two adjacent antennas). These sources are fed with the same amplitude and with a phase gradient $\Delta\varphi$ with respect to each other. r represents the maximum distance between the reference antenna and the observation plane, and θ is the beam steering angle.

In these conditions, the total electric field radiated by the array antenna is then written as (1):

$$E_{tot}=E_0*A(\psi)\ . \tag{1}$$

Where:

- E_0 represents the electric field radiated by an elementary antenna. It is named "element factor".

- A (ψ) is the "Array factor» independent of the elementary source that makes up the array.

This array factor is given by the relation (2) below:

$$A(\psi)=-\exp(j(N-1)\ \psi/2)\frac{\overline{\frac{2}{\psi}}}{}\ . \tag{2}$$

Where,

$\Psi=$ sin (θ)-$\Delta\phi$, and λ is the wavelength

The total radiation pattern of the grating antenna is defined by the product of the modulus of the electric field radiated by an elementary antenna by that of the grating factor, the latter being given by the following expression see equation (3):

$$|A\quad)|=\frac{1}{}\left|\frac{}{}\right|\ . \tag{3}$$

In these conditions, it should be noted that as soon as the number N of elementary antennas becomes large, the total radiation pattern of the array depends essentially on the array factor and very little of the radiation pattern of each antenna. Note that if all the sources are excited in phases (ψ is zero), the radiated field is maximum in the plane normal to the array.

Conversely, if the sources are excited with a non-zero phase gradient, the maximum radiation appears for $\psi= 2m\pi$. The solutions are periodic and show on the same array lobes radiation pattern for values of $m \neq 0$.

Thus, a variation of the phase gradient of the array $\Delta\varphi$ to be introduced between the adjacent elementary sources causes the orientation of the radiation pattern in the desired $\theta 0$ direction according to:

$$\Delta\varphi=\text{ sin }(\theta\quad\ . \tag{4}$$

The variation of the phase gradient of the array determines that of the beam steering angle $\theta 0$ of the array antenna: this is the principle of a "phased array antenna"[3].

3 Proposed antenna

The geometry of the proposed single-patch antenna structures is shown in **Figure 2.** consists of a the conventional square patch fed with the transmission line structure printed on a Rogers RT/Duroid5880 substrate, which has permittivity of 2.2, loss tangent of 0.001 and having a thickness of h = 0 .25mm. The area of the proposed antenna is 10× 15mm^2 and the dimensions of the antenna are calculated and optimized so that it resonates at 28GHz

frequency using patch antenna equations, the final dimensions of the proposed antenna are: L_p=W_p= 3.36mm, L_q= 2mm, W_q= 0.145mm, L_f = 3mm and W_f = 0.776mm.

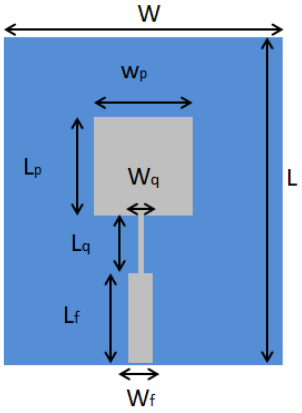


Fig. 2. Geometry of the proposed patch antenna.

3.1 Simulation results and discussions

The proposed antenna is simulated using CST Microwave StudioTM and the calculated S-parameter is shown in **Figure 3.** The proposed structure gives return loss around -19.5 dB at the resonant frequency 28GHz and has a 27.21–28.30 GHz bandwidth for |S11| < -10 dB, which is approximately 3.69% bandwidth (1.09 GHz) with a center frequency of 28 GHz.

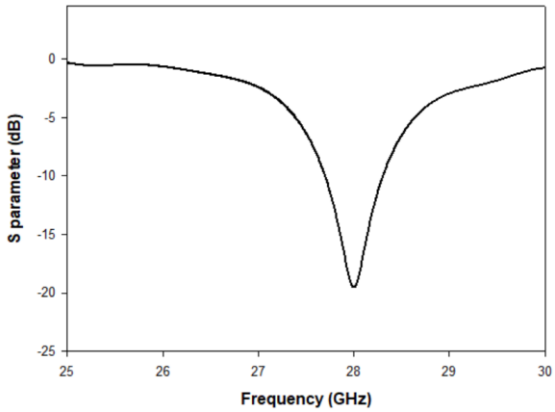


Fig. 3. Simulated result of S parameter of the initial patch.

The **Figure 4.** explains the antenna radiation pattern. The antenna exhibits a maximum gain of 7.32 dB. The Side lobe level for the E-plane and H-plane are -20.5 dB and -12.9 dB respectively. Half power beamwidth of 87.3° is obtained H-plane, whereas that of E-plane is 76°.

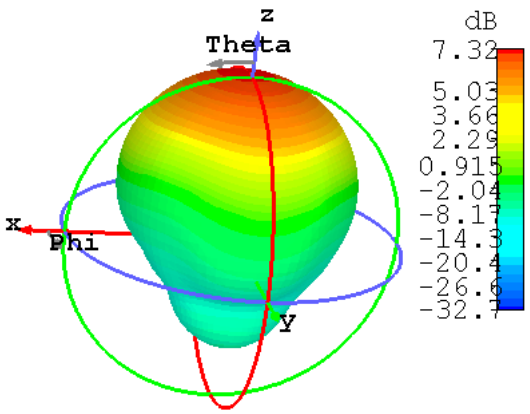


Fig. 4. Simulated 3D Far-field Radiation Pattern at 28 GHz.

3.2 Array of two element microstrip patch

Figure 5. shows an array of two element patch antenna separated by a distance d that does varies between 0.5λ and 0.8λ. The goal of this study is to optimize the separation distance between the two elements of the array in order to find a compromise between the side lobe level and the isolation between the two excitation ports. The results found are grouped in the Table 1:

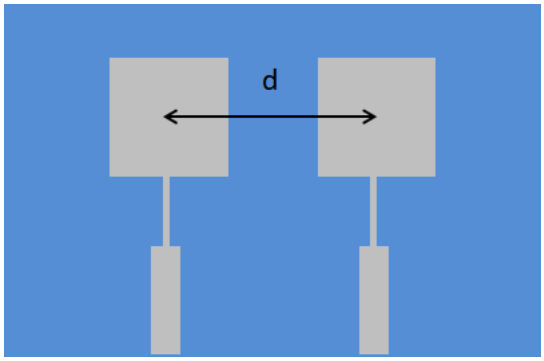


Fig. 5. Array of two element microstrip patch.

Table 1. The effect of separated disance d between the element on the antenna performance.

Separated distance d	S21 (dB)	Gain (dB)	SLL (dB)
0.5λ	-26.7	10.5	-14.7
0.6λ	-28.84	7.58	-18.4
0.7λ	-30.41	7.95	-17.7
0.8λ	-31.6	7.8	-17.5

We can see that the greater the separation distance between these two elements, the better isolation and inferior gain. However, there are a secondary lobes with high gain that are considered parasitic radiation. To reduce the side lobes level, we have to bring closer the antennas as possible. However, the closer they are the closer the field to antenna coupling is important. It was therefore necessary to find a compromise between the separation distance for the reduction of the side lobe level and the coupling between antennas. Based on this analysis, we chose as optimal separations distance 0.5λ because for this the array has good performances at the side lobes level, coupling between antennas and gain of antenna.

4 Linear array antenna

Phased array antennas are known for their capability to steer the beam pattern electronically with high effectiveness, managing to get minimum side-lobe levels and narrow beamwidths. Normally phase shifters are the devices in a phased array antenna that allow the radiated beam to be steered in the direction [4]. The objective of this part is to investigate a phase shifter in a linear antenna array for angle scan. **Figure 6.** depicts a linear array of 4 elements; all elements are excited with equal amplitude and with a progressive phase shift Δφ determined by the location of the element in the array.

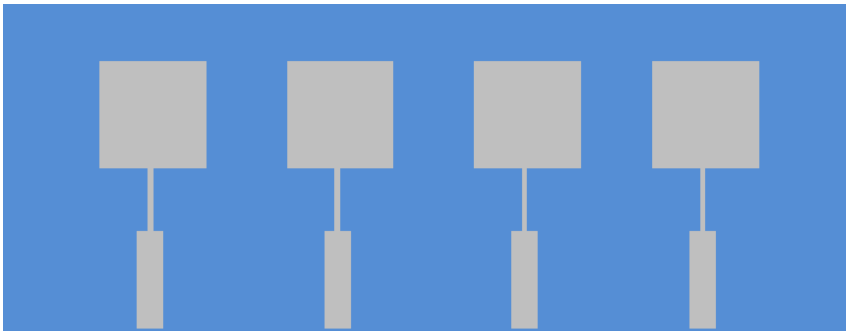


Fig. 6. Array of four element microstrip patch.

The return loss of the antenna is shown in **Figure 7**. From the result the resonance frequency of the antenna is at 28.03 GHz and S11=S22= -16.17 dB and S33=S44= -14.78 dB. The bandwidth of the antenna is 2.03 %.

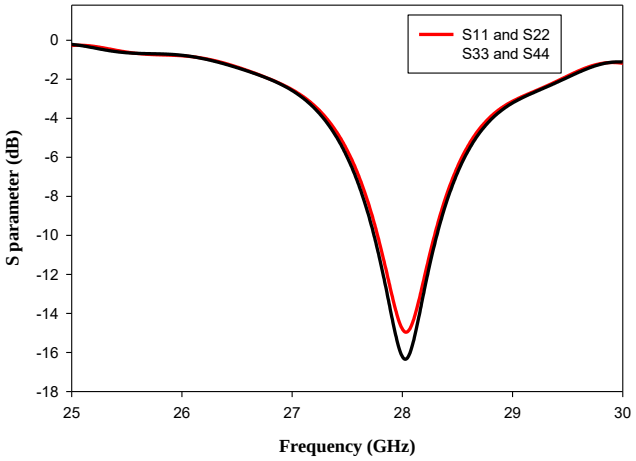


Fig. 7. Array of four element microstrip patch.

In the first, all the ports are excited by 0 phase and equal amplitude, the results of the radiation patterns in (XY) E-plane and properties such as beam position, maximum gain and SLL are shown in **Figure 8**. The maximum gain equal to 12 dB in broadside (θ=0°).Here, the SLL is equal to -13.5 dB at E-plane. In the second, to study the beam scanning capability of the array antenna all the ports are excited by a scanning over ±40° and equal amplitude. In **Figure 9**. the radiation pattern for (XY) E-plane for different values of phase from -40° to 40° at resonance frequency 28 GHz are shown. The gain for all the scanning is 12 dB, but the side lobe level is -10.8 dB when the beam is tilted by ±30° and ±40°, and its gain remains constant 12 dB. When we have the beam tilting of ±20° gets a SLL of -11.5 dB at 28 GHz and beamwidth of 26.3°.

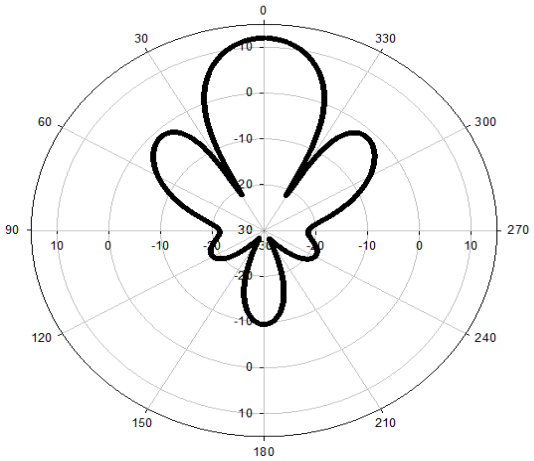


Fig. 8. Simulated radiation pattern For (XY) E-plane Φ= 0
Phase 0°.

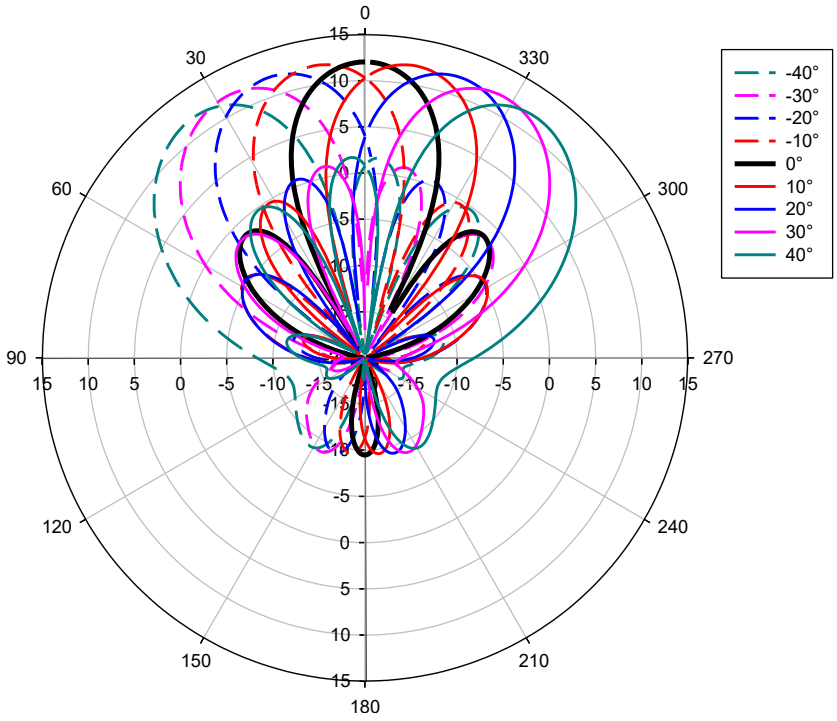


Fig. 9. Simulated radiation pattern For (XY) E-plane Φ= 0
Phase ±10 to ±40.

5 Comparison with literature

As indicated at the onset of the paper, a number of techniques have been proposed in the literature for the beam steering. Table 2 presents a comparison of the scanning angular range of the proposed antenna and the beam steering antennas in [6], [8] and [9]. We can see that the proposed antenna realizes a larger beam tilted angle.

Table 2. Comparison of techniques reported in literature for beam steering.

Antenna Model	Fr (GHz)	Beam Tilt (°)	Technology
Antenna in [6]	28	±25	Series-Fed Pach Antennas
Antenna in [8]	34.25	±30	Antenna Using a Metasurface
Antenna in [9]	9	±30	Design Of Phase-Modulated Metasurfaces
Antenna in this paper	28	±40	Uniform linear array

6 Conclusion

In this paper, a phased array structure has been proposed and analyzed. The structure is formed by uniformly arranged as an 4-element identical, which is able to achieve beam steering with using phase shifters. The phase shift are achieved by changing the phase of the element. The final proposed antenna achieves a 750 MHz bandwidth for S11 < -10 dB to cover 27.74-28.31 GHz frequency band which is intended for future millimeter wave communications. Moreover, the antenna array achieved a maximum gain of 12 dB in the broadside direction. The antenna array could be utilized as a phased array and hence could be used to steer the radiation beam. The antenna exhibited a very good radiation pattern and stable gain. The structure can have 9 kinds of reflection phase distribution, so the antenna can change its beams among 0°, ± 10°, ± 20°, ± 30° and ± 40° in xoy-plane at 28 GHz. Compare to other beam steering antenna, the proposed antenna can realize a larger beam tilted angle (40°). The next step is to fabricate the antenna and verify the simulation results which will be reported at the conference.

References

[1] D. Ehyaie: Novel Approaches to the Design of Phased Array Antennas. pp. 1-11 2011.

[2] T. Eray, H. Jurgen,W. Christoph, Z. Thomas: A novel millimeter-wave dual-fed phased array for beam steering. IEEE transaction on microwave theory and technique. Vol. 61, pp. 40-43. 2013.

[3] M. Dorra: Etude comportementale et conception d'un réseau d'oscillateurs couplés intégré en technologie silicium appliqué à la commande d'un réseau d'antennes linéaire.2013.

[4] L. Jinxin, Z. Qingsheng, L. Ruizhi, A. Tayeb: Beam-tilting antenna with negative refractive indexe metamateriam loading. IEEE antennas and wirless propagation letters. Vol. 61, pp. 40-43. 2017.

[5] Y. Yazid, G. Xun: A low-cost patch antenna phased array with analog beam steering using mutual coupling and reactive loading. IEEE antenna and wirless propagation letters. Vol. 7. 2012.

[6] M. Khalily, R. Tafazolli, T. A. Rahman, and M. R. Kamarudin: Design of phased arrays of series-fed patch antennas with reduced number of the controllers for 28 GHz mm-wave applications. IEEE Antennas and Wireless Propagation Letters, vol. 15, pp. 1305–1308, 2016.

[7] T. Eray, H. Jurgen, W. Christoph, Z. Thomas: Reconfigurable Beam Steering Using a Microstrip Patch Antenna With a U-Slot for Wearable Fabric Applications. IEEE antennas and wirless propagation letters. Vol. 15,2013.

[8] L. Teng, Z. N. Chen: Control of Beam Direction for Substrate-Integrated Waveguide Slot Array Antenna Using Metasurface. IEEE Transaction On Antennas And Propagation. 2018.

[9] R. Badreddine, L. Andre, P. Gerard, N. Shah: Modeling and design of metasurfaces for beam scanning. Appl. Phys. Lett. 50, 123,2017.

Double Layer Metamaterial Superstrate for Performance Enhancement of a Microstrip Patch Antenna

F.Z.Khoutar[1], M. Aznabet[2], O. EL Mrabet[3], O. Nayat-Ali[4]
{fatikhoutar@gmail.com[1], mariamaznabet@gmail.com[2], o.mrabet@gmail.com[3]}

Information and Telecommunication system laboratory[1,2,3,4], Abdelmalek Essaadi University,Tetuan,Morocco

Abstract. In this paper, a metamaterial antenna operating at a frequency of 5.8 GHz is presented. Metamaterial superstrate cover acts as a lens which focuses the EM radiations due to the unusual properties of the metamaterial.This leads to improvement performance of patch antenna along with its physical protection.The proposed metamaterial lens consists of a double layer of 9-by-9 matrix of Split-Ring Resonators (SRRs), and is positioned above a probe-fed rectangular patch resonating at 5.8 GHz. The structure improve the gain and directivity of a simple patch antenna up to 7 dB.Simulation results of the proposed antenna are presented and discussed in this letter.

Keywords: Metamaterial (MTM),Split Ring Resonator (SRR).

1 Introduction

An enormous growth in the field of telecommunication has generated enormous needs and has resulted in a major technological change at the antenna level, depending on the transmission frequency, data rate or transmission range.

The antenna is one of the most critical components in the wireless communication systems, since the performance of the antenna alone can seriously influence the overall performance of the entire system. The design goal of an antenna is often governed by the specifications provided by the customer. Many applications specify that the antenna must be conformal, and that it should have a high directivity along a direction when transmitting. These requirements, namely high directivity is the major developing trends among antennas and always presents a great challenge to engineers working in this field. The traditional antenna technology to achieve the high directivity needs multiple radiating elements to fabricate the antenna array [1]. However, such array design requires a complex feed network, and the demands that the multiple antenna elements be maintained properly. This is because the malfunction of one or more antenna elements may sometimes seriously affect the overall performance of the antenna system, Other interesting solutions to enhance the directivity of the patch antenna have been suggested: the first one was to make use of a DGS technique to modify the parameter of antenna and the second one proposed recently was to sandwich the antenna by dielectric layers of the same permittivity [2] [3].

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284232

More recently, another approach is proposed in order to improve the directivity of the antenna, which is based on the use of artificial materials such as left- handed metamaterials [4]. These new kinds of artificial materials, constructed from periodic unit structures that are compactly crowded into an effective material, open the way to observe some exotic properties that are impossible to achieve with natural materials.
Many research works have been achieved in obtaining the high directivity of antenna by using metamaterial structures [5–12]. MTMs can be fabricated through various ways such as photonic crystal [9], electromagnetic band gap (EBG) structure, frequency selective surface (FSS) [10–11], and other periodic artificial material which can be designed to have a low/zero refractive index [12–13].Particularly, negative refractive index zero or near zero index metamaterial, and furthermore lead to potential applications, many of which were considered difficult in the past. These include, for example, perfect lensing [14], invisibility cloaking [15] and miniaturizing microwave devices [16].
In this paper, we presented a metamaterial antenna for WLAN application operating at the frequency of 5.8 GHz. The double layer metamaterial superstrate used in this design was constructed by split ring resonator (SRRs), introduced for the first time by Pendry et al. in 1999 [17] to achieve a negative permeability in a certain frequency range. The main attractive feature of this element is its ability to exhibit a quasi-static resonant frequency at wavelengths that are much larger than its own size. We have investigated the effect of metamaterials Radom on the directivity, gain and bandwidth of the antenna by using CST Microwave studio. It is found that by optimizing the separation distance between the antenna and the double layer metamaterial superstrate the directivity of the antenna has been increased by 7 dB.

2 Metamaterial antenna design

2.1 Presentation of Metamaterial antenna

Metamaterials are typically engineered by arranging a set of small scatterers or apertures in a regular array throughout a region of space, thus obtaining some desirable bulk electromagnetic behavior. The desired property is often one that is not normally found naturally (negative refractive index, near-zero index, etc.). Over the past ten years, metamaterials have moved from being simply a theoretical concept to a field with developed and marketed applications. Three dimensional metamaterials can be extended by arranging electrically small scatterers or holes into a two-dimensional pattern at a surface or interface. This surface version of a metamaterial has been given the name metasurface For many applications, metasurfaces can be used in place of metamaterials. Metasurfaces have the advantage of taking up less physical space than do full three-dimensional metamaterial structures,Metamaterial Antenna is a class of antennas that use metamaterial to miniaturize antenna systems and improve their performance. Their purpose, as for any electromagnetic radiated element, is to propagate energy into free space. It consists of an MTM superstrate cover above a reference element.
The operation of the MTM superstrate is explained where the scattered radiations from the patch are concentrated due to the MTM superstrate cover above. This function increases the gain and directivity of the patch antenna.
Moreover, this type of antenna with metamaterial layer is called resonant cavity antenna that generally consists of one or more surfaces Metamaterial, a ground plane and a source of excitation placed inside the cavity. The different planes are separated by the cavity height (H). This cavity is also defined by its opening (Lcav). The developed antenna must have a finite

dimension in order to be easily implanted. The cavity is the first element to optimize. Its opening must be fixed in front. This structure shows a magnetic resonance at a frequency, the position of this resonance frequency can be varied by changing the different geometric parameters of the SRR, the largest dimension of the cell must be sufficiently small with respect to the wavelength ($d << \lambda$).

2.2 Antenna design

The proposed antenna geometry and configuration is depicted in **Figure 1**. Where a three-layer structure consists of a substrate, air-gap, and a superstrate of thicknesses h1, H, and h3, respectively. The antenna is designed to operate at 5.8 GHz for wireless application. The metamaterial surface is constructed on a dielectric material having a thickness of 1.6 mm and a dielectric constant of 2.2. This layer maintains a resonance distance of 0.5 λ (where λ is the operating wavelength) above the substrate, the metamaterial surface is composed of 9 × 9-unit cells. The cell unit is a rectangular split ring resonator with a compact size of X=Y= 7 mm. Each cell has a spacing of 0.5 mm in the x-direction and 0.5 mm in the y-direction, the width of the copper strip is 0.5 mm. Considering the reference element the antenna is fed by a 50-Ohm coaxial probe form backside. The antenna has a rectangular radiation patch. The operation frequency of the patch antenna is inversely proportional to the length of its main radiator and can be configured by changing the antenna parameters.

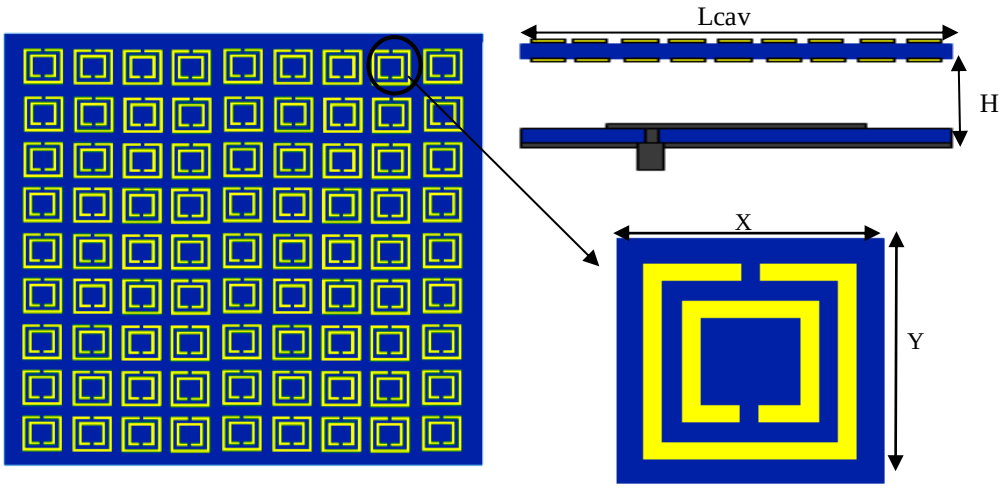


Fig. 1. Patch antenna with a double layer MMT layer. (a) top and and front view of the antenna (b)SRR unit cell.

3 Simulated results and discussion

The reflexion coefficient versus frequency of the patch antenna with and without the double layer MTM superstrate is presented in **Figure 2.**the figure show that the patch antenna itself is not in a good matching condition compared when the double layer MTM superstrate has not been used. Strong resonance in cavity improves the matching and the resonant MTM antenna successfully and it operates over a reasonable band with a center frequency of 5.8 GHz.The double layer MTM antenna is in a good matching with a return loss of about -33 dB at the resonance frequency 5.8 GHz compared with the reference antenna alone which is -12 dB.

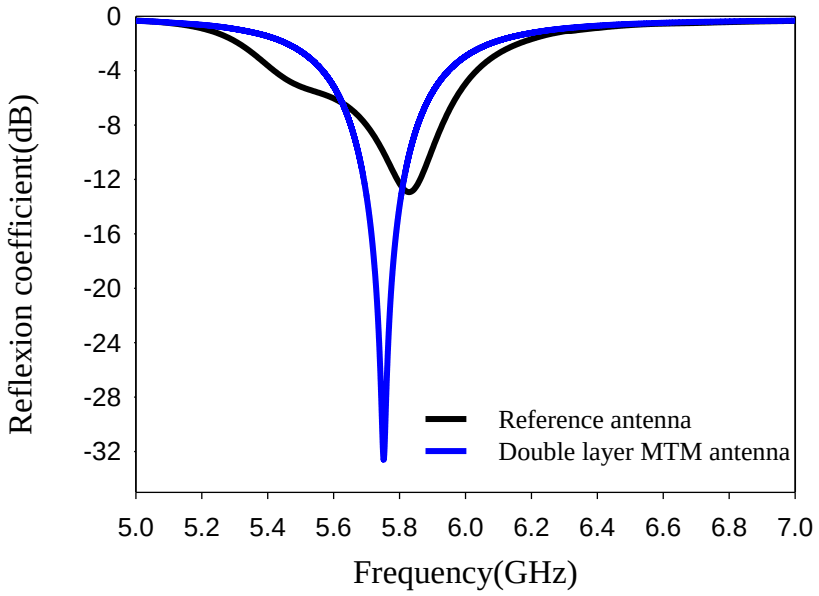


Fig. 2. reflection coefficient of the patch antenna with and without double layer MTM superstrate.

the directivity of the antenna with and without the double layer MTM superstrate is presented in **Figure 3.** It is seen that the directivity of the microstrip antenna covered by the double layer MTM superstrate is higher than that of the antenna without using the metasurface. The maximum directivity attained at the center frequency of 5.8 GHz is 7.2 dB (without double layer metamaterial superstrate) and 14.1 dB (with double layer metamaterial superstrate) which is about 7 dB improvement. This shows that the proposed double layer MTM has a remarkable effect on the enhancement of the radiation energy, at an overall cavity height of 0.5λ.

The simulated 2D E-plane and H-plane radiation pattern of the double layer MTM superstrate loaded antenna in comparison to the reference antenna at the resonant frequency 5.8 GHz are shown in **Figure 4.** We can observe that the main lobe is as expected in the direction of the

positive z-axis (the direction of propagation) and that the back-lobe level decreases and broad side level increases by the loading of double layer MTM superstrate. It's clear from the plot that loading a microstrip antenna with MTM superstrate, has an important effect on the reduction of the -3dB beam width which degrade from 28,3° to 104° in the E plan and from 30,4° to 83,5° in the H plan ,the reduction in 3-dB beam width is indication of the improved directivity in the broadside direction .

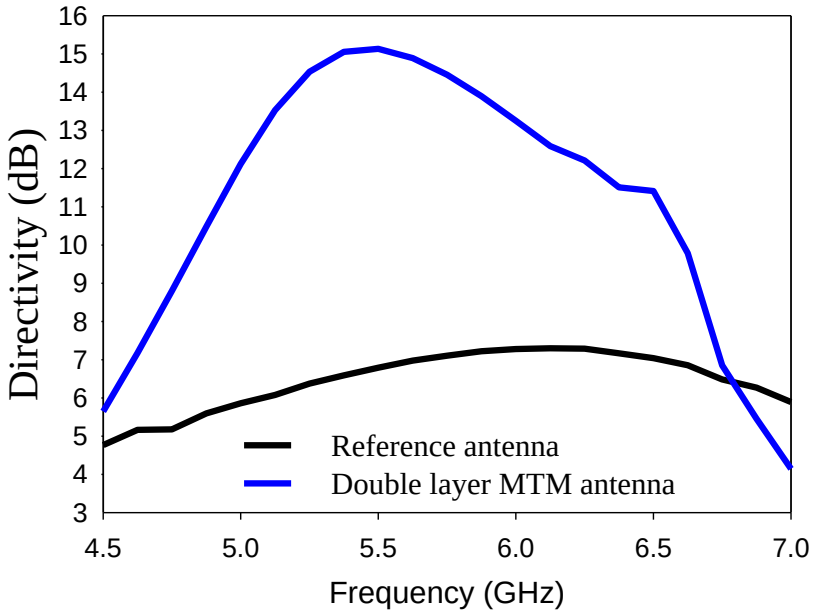
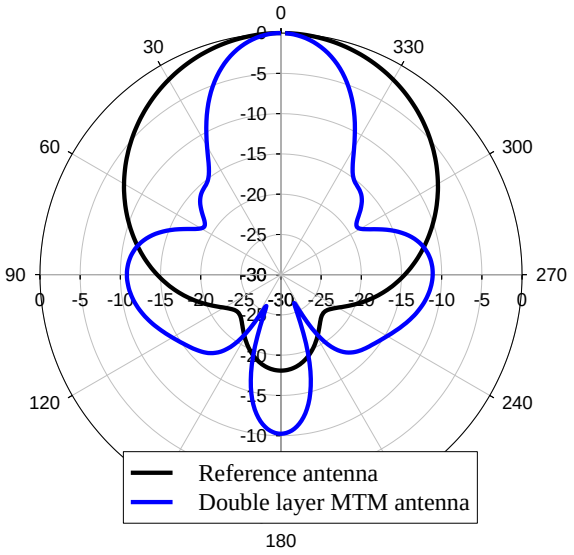
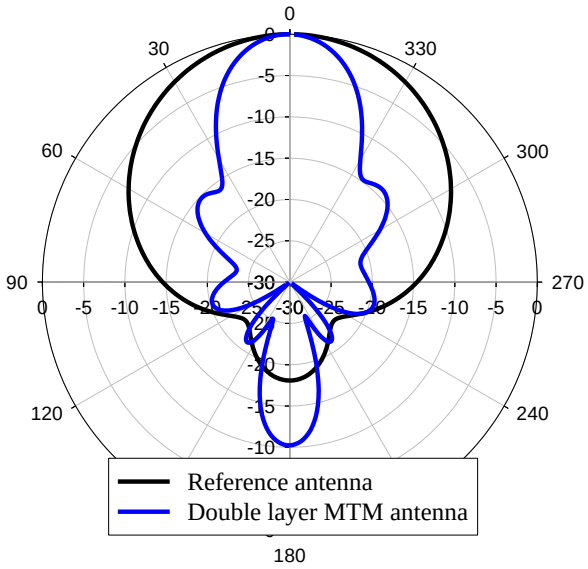


Fig. 3. Directivity of the patch antenna with and without double MTM superstrate.



(a)



(b)

Fig. 4. Radiation pattern of designed antenna at 5.8 GHz with and without the double layer MTM superstrate: (a) H-plane radiation pattern (b) E-plan radiation pattern

4 Conclusion

The results obtained are encouraging, the structure allowed us to improve the directivity of the antenna patch within the frequency range of interest by a value of 7 dB. this structure not only allows improving the directivity of the patch but also other characteristics, the double layer MTM antenna is in a good matching with a return loss of about -33 dB at the resonance frequency 5.8 GHz compared with the reference antenna alone which is -12 dB,The HPBW of the antenna with metasurface is narrowed down to 28.3° from 104° degree in the E plan and to 30.4° from 83.5° in the H plan,considering the radiation efficiency before and after using the metasurface it has been improved over the relevant range of frequencies from 70% to 83%.This indicates that the double layer metamaterial superstrate added has a very important effect on antenna patch performance. In conclusion, the double layer MTM superstrate concentrates the radiation energy of the patch antenna and therefore it makes possible to improve certain characteristics of reference element, high directivity is the major demand of telecommunication and space applications. Hence the design of such antennas is the key requirement in order to overcome the restrictive parameters of the conventional antennas, we believe that similar idea can also be applied to other types of planar antennas and arrays.

References

[1] D.Nascimento, J. da, S. Lacava: Design of arrays of linearly polarized patch antennas on an FR-4 substrate. IEEE Antennas Propag Mag.pp. 12–22 (2015).

[2] X . Tang, H. Wong, Y. Long, Q . Xue, KL . Lau:Circularly polarized shorted patch antenna on high permittivity substrate with wideband. IEEE Trans Antennas Propag. pp. 1588–92 (2012).

[3] H. Nakano, M. Ikeda, K. Hitosugi, and Yamauchi : A spiral antenna sandwiched by dielectric layers. IEEE Trans. Antennas Propag.vo1. 52.no. 6 (2004).

[4] D. Li, Z. Szabo, X. Qing, E. P. Li, and Z. N. Chen : A high gain antenna with an optimized metamaterial inspired superstrate. IEEE Trans. Antennas Propag.vol. 60, no. 12, pp. 6018–6023 (2012).

[5]U.kumar, D. K. Upadhyay, B. L. Shahu :Improvement of performance parameters of rectangular patch antenna using metamaterial. IEEE International Conference On Recent Trends In Electronics Information Communication Technology. pp. 20-21 (2016).

[6]S. Gautam, A. Agrawal : Directivity Improvement of Patch Antenna using Metamaterial, Symposium on Colossal Data Analysis and Networking (CDAN). (2016).

[7]J. W. Jayasinghe, J. Anguera, D. N. Uduwaw: A high-directivity microstrip patch antenna design by using genetic algorithm optimization. Progress In Electromagnetics Research C, Vol. 37,131–144 (2013).

[8]Z. Xiao, J .Yao,S. Yin: High-Directivity Antenna Array Based on Artificial Electromagnetic Metamaterials with Low Refractive Index,Corporation International Journal of Antennas and Propagation, Hindawi Publishing. pp. 294-598 (2015).

[9]H. D. Yang, N. G. Alexopoulos , E. Yablonovitch :Photonic band-gap materials for high-gain printed circuit antennas. IEEE Trans Antennas and Propag, vol. 45, no. 1, pp. 185-187 (1997).

[10] B. Temelkuran, M. Bayindir, E. Ozbay et al: Photonic crystalbased resonant antenna with a very high directivity. Journal of Applied Physics. vol. 87, no. 1, pp. 603–605, (2000).

[11] A. R. Weily, L. Horvath, K. P. Esselle, B. C. Sanders, and T. S. Bird : A planar resonator antenna based on a woodpile EBG material. IEEE Transactions on Antennas and Propagation. vol. 53, no. 1, pp. 216–223 (2005).
[12] S. Enoch, G. Tayeb, P. Sabouroux, N. Guerin, and P. Vincent : A metamaterial for directive emission. Physic Review Letters.vol. 89, no. 21, pp. 213902-1–213902-4 (2002).
[13]R Q. Wu, P. Pan, F.-Y. Meng, L.-W. Li, and J. Wu: A novel flat lens horn antenna designed based on zero refraction principle of metamaterials.Applied Physics : Materials Science and Processing.vol. 87, no. 2, pp. 151–156 (2007).
[14]J.B. Pendry: Negative Refraction Makes a Perfect Lens. Physical Review Lett 85, pp. 3966-3969 (2000).
[15]W. Cai, U. K. Chettiar, A. V. Kildishev, V. M. Shalaev : Optical cloaking with metamaterials .Nat.Photon 1, pp. 224 (2007).
[16]R. Marqués, F. Martín, and M. Sorolla, : Metamaterials with Negative Parameters: Theory, Design, and Microwave Application. John Wiley, New York (2008).
[17]J. B. Pendry, A.J. Holden, D.J. Robbins, and W. J. Stewart : Magnetism from conductors and enhanced nonlinear phenomena. IEEE Trans. Microwave Theory, Tech, 47 pp. 2075-2084 (1999).
[18]D. R. Smith, W. Padilla, D.C. Vier, S.C. Nemat-Nasser, S. Schultz : Composite Medium with Simultaneously Negative Permeability and Permittivity. Phys. Rev. Lett. 84 pp.4184 (2000).

Invention of an Original Tetra-Generations Patch Antenna for the New Generation of Mobile Telephony and the Study of the Thermal Effect of GSM on the Human Head and Hand

Jalal Baayer[1] and Hassan Ammor[2]
{jalal.baayer@gmail.com[1], ammor@emi.ac.ma[2]}

Smart Communications Research Team, E3S Research Center, Engineering Mohammadia School, Mohammed V University, Rabat, Morocco[1]

Smart Communications Research Team, E3S Research Center, Engineering Mohammadia School, Mohammed V University, Rabat, Morocco[2]

Abstract. The present invention relates to a miniature original tetra-generations patch antenna with rectangular slots at low cost, and dedicated to the new generation of mobile telephony by covering the frequency bands of 900MHz, 1800MHz, 2100MHz and 2600MHz. The Modeling and simulation of the basic model are performed using the Ansoft HFSS simulator employing finite element method (FEM). This study allowed us to determine the reflection coefficient, the voltage standing wave ratio (VSWR), the omnidirectional radiation and the bandwidth. With this antenna, we obtained new results from the GSM thermal effect on the human head and hand. With this in mind, a new, simple and miniaturized quadri-band antenna in patch technology meeting the standards of dosimetry and temperature for mobile phone applications is designed and proposed.

Keywords: Tetra-generations patch antenna, SAR, human head, human hand, dosimetry, temperature rise.

1 Introduction

Our study constitutes the simulation and the realization of a new tetra-generations patch antenna for the new generations of the mobile telephony operating in the bands of frequencies of 900MHz, 1800MHz, 2100MHz and 2600MHz respecting the norms of the dosimetry and the temperature.

This work is based on simulations obtained on the Ansoft HFSS software basing on two anatomical models. One concerning the human head and composed of seven layers and the other relating to the human hand and represented by 4 layers. The simulations of the SAR distribution and temperature rise are performed at four frequency bands, for seven antenna-head distances for the head and one distance for the hand during five exposure times, based on the tetra-generations patch antenna model.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284214

2 Structure of the invented antenna

2.1 Basic structure adopted for the design

The geometry of the invented antenna is represented in figure 1, the antenna is made on a FR4_epoxy substrate of dimensions L × W = 45 × 42 mm^2 with a dielectric constant εr = 4.4 [1]. The thickness of the substrate is h = 1.6mm. The antenna consists of a square patch having optimal characteristics and reduced dimensions and including technical slots.

The value of the permittivity εr is chosen such that it provides a better efficiency and effectiveness and a wider bandwidth. The antenna power is supplied via a 50Ω coaxial power supply to improve bandwidth and gain.

We performed the simulation with the HFSS software (High Frequency Structure Simulator). We chose the substrate, calculated the width, the length, the effective dielectric constant, which allowed us to have the structure proposed in figure 1.

Fig. 1. The invented and performed tetra-generations patch antenna.

The figure 2 shows the reflection coefficient of the patch antenna presented from the simulation based on the Ansoft HFSS software.

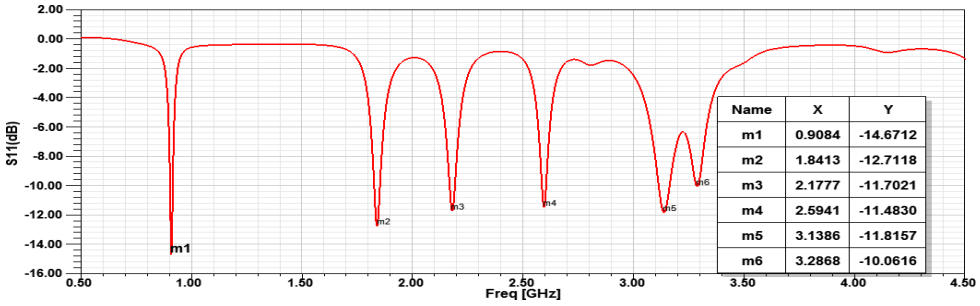


Fig. 2. The reflection coefficient of the invented antenna.

The figure 3 illustrates the comparison between the real results and the simulation of the reflection coefficient S11 of the tetra-generations patch antenna.

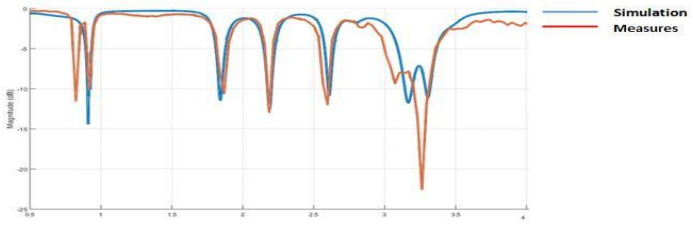


Fig. 3. Measures and simulation results of the tetra-generations patch antenna reflection coefficient.

3 SAR distribution and temperature rise

The SAR is represented by the following formula [2]:

$$SAR = \frac{\sigma}{2\rho}\left|E\right|^2 \quad (1)$$

With σ the electrical conductivity (S / m), ρ is the density of tissue (kg / m3) and E is the peak value of the electric field (V / m). The absorption of electromagnetic waves causes a temperature variation ΔT, in the case of a human body with specific heat C, moreover in thermal equilibrium with the environment, and exposed during the time Δt to these waves.

This correlation between the SAR and the temperature variation is given by the formula [3]:

$$SAR = C\frac{\Delta T}{\Delta t} \quad (2)$$

According to our recent work, the maximum SAR is observed at the level of the skin of the human head and can be represented according to the formula:

$$SAR_{max} = f(\theta_x, \theta_y) \quad (3)$$

And for a fixed ΔZ among {5,10,15,20,25,30}(mm).

With:

θx: rotation angle around the OX axis.
θy: rotation angle around the OY axis.
ΔZ: displacement along the OZ axis.

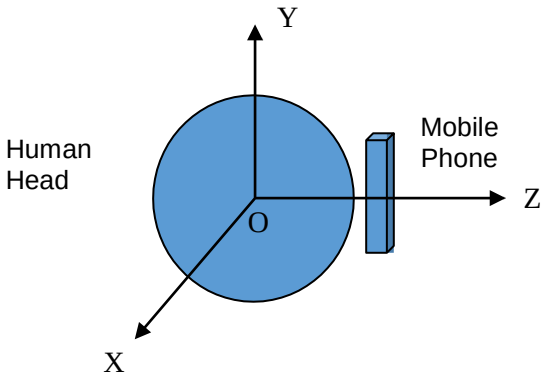


Fig. 4. Positions of the mobile phone and the human head.

4 Modeling of the human head and hand

We used in figure 5 a model of flat multilayer human head with seven types of tissues that are: skin, fat, muscle, skull, DM (dura), CSF and brain. Each layer has its intrinsic dielectric properties and its thermal characteristics.

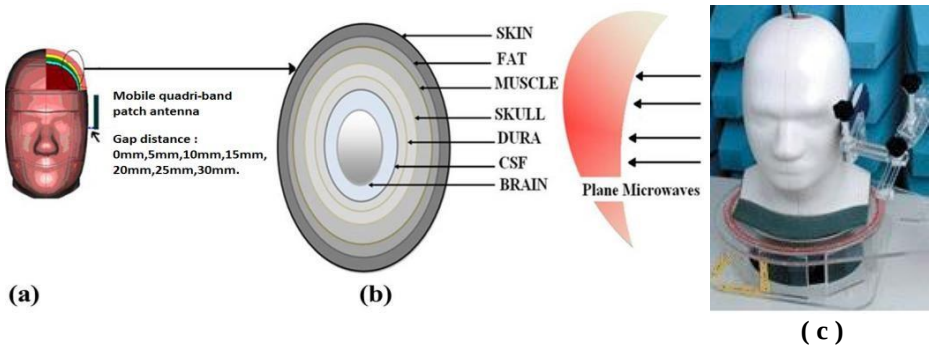


Fig. 5. The model of the human head with 7 layers and the SAM phantom.

(a) Human head model in cross section with mobile phone.
(b) Irradiation of the human head model consisting of seven layers per electromagnetic waves.
(c) SAM phantom of human head used in a SAR test bench.

We used according to figure 6 an HFSS model of human hand multilayer with four types of tissues that are: skin, fat, muscle and bone and irradiated by our tetra-generations patch antenna.

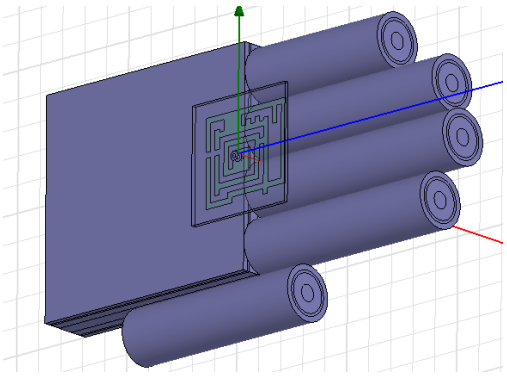


Fig. 6. HFSS model of the human hand irradiated by our tetra-generations patch antenna.

The tables 1 and 2 show the numerical values of the dielectrics, the thermal properties and the mass density of each layer. These values were used from the data of the scientific literature [4]:

Table 1. Dielectric properties of human head and hand models layers.

Layers	**900 MHz**		**1800 MHz**		**2100 MHz**		**2600 MHz**	
	Permittivity (εr)	***Conductivity σ(S/m)***	***Permittivity (εr)***	***Conductivity σ(S/m)***	***Permittivity (εr)***	***Conductivity σ(S/m)***	***Permittivity (εr)***	***Conductivity σ(S/m)***
Skin	43.8	0.86	43.9	1.23	43.9	1.23	42.9	1.59
Fat	11.3	0.11	11.0	0.19	11.0	0.19	10.82	0.26
Muscle	55.9	0.97	54.4	1.38	54.3	1.38	53.64	1.77
Bone	20.8	0.34	15.6	0.43	15.5	0.43	15.01	0.57
Dura	44.4	0.96	42.9	1.32	42.9	1.32	42.03	1.66
CSF	68.6	2.41	67.2	2.92	67.2	2.92	66.24	3.45
Brain	45.8	0.77	43.5	1.15	43.4	1.15	42.61	1.48

Table 2. Thermal properties and mass density of the layers of the human head and hand model.

Layers	Mass Density ρ (kg/m3)	Specific heat (J/g.C)
Skin	1100	3.6
Fat	1100	3.0
Muscle	1040	3.5
Bone	1850	3.1
Dura	1030	3.6
CSF	1030	4.0
Brain	1030	3.7

5 Results and discussions

5.1 Mobile thermal effect on the human head

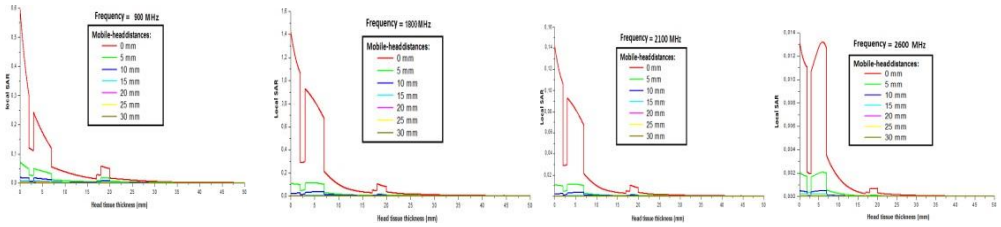


Fig. 7. Distribution of local SAR in the human head.

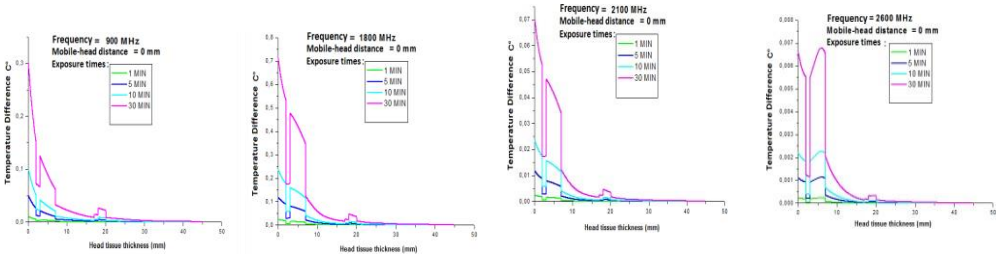


Fig. 8. Spatial distribution of temperature in the human head.

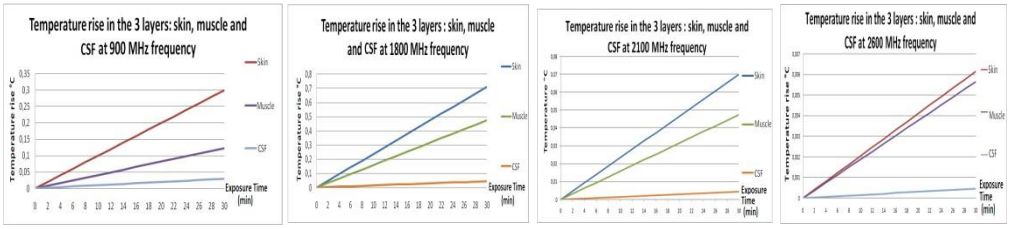


Fig. 9. The increase in temperature in the three layers: skin, muscle and CSF of the human head.

The figure 7 represents the distribution of local SAR in the human head. The figure 8 shows the spatial distribution of temperature in the human head. In the figure 9 we find the increase in temperature in the three layers: skin, muscle and CSF of the human head.

In the figure 7, we notice that:

- For each mobile-head distance, the SAR has 3 peaks in the three layers : skin, muscle and CSF.
- As we move away from the skin to the brain, the value of SAR decreases and has very low value in the brain.
- The SAR value decreases as the frequency rises from 900Mhz and 1800Mhz to 2100Mhz and 2600Mhz.

In the figure 8 and 9, we notice that:

- The temperature has 3 peaks in the three layers : skin, muscle and CSF.
- As we move away from the skin to the brain, the value of temperature decreases and has very low value in the brain.
- The temperature value decreases as the frequency rises from 900Mhz and 1800Mhz to 2100Mhz and 2600Mhz.
- The temperature value increases as the duration of telephone conversation rises.

We conclude that:

- The 3 layers skin, muscle and CSF have significant water levels so a large absorption of mobile electromagnetic waves by their biological tissues. This explains why we have 3 peaks in SAR and temperature in these 3 layers.
- The 3 layers represent a natural protection for the brain against electromagnetic waves of mobile that is why we have minimum values of SAR and temperature in the human brain.
- The migration of telecom operators from older mobile telephony generations of G1 (900Mhz) and G2 (1800Mhz) to G3 (2100Mhz) and G4 (2600 Mhz) is very beneficial for mobile phone users and this explains why we have minimum SAR and temperature values for the 2100 Mhz and 2600 Mhz frequencies.
- The warmth we feel in the superficial layer of the head during a long telephone conversation is explained by the high SAR and temperature value in the skin of the human head.

5.2 Mobile thermal effect on the human hand

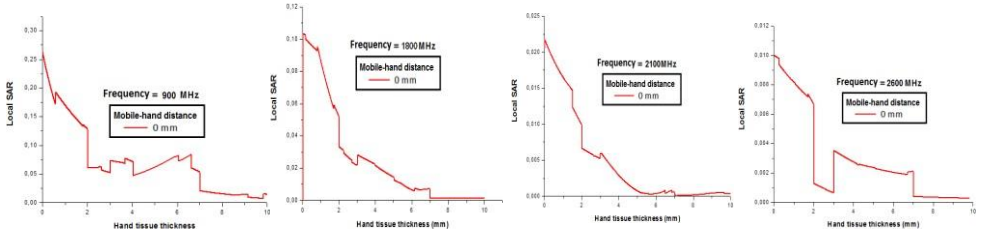


Fig. 10. Distribution of local SAR in the human hand

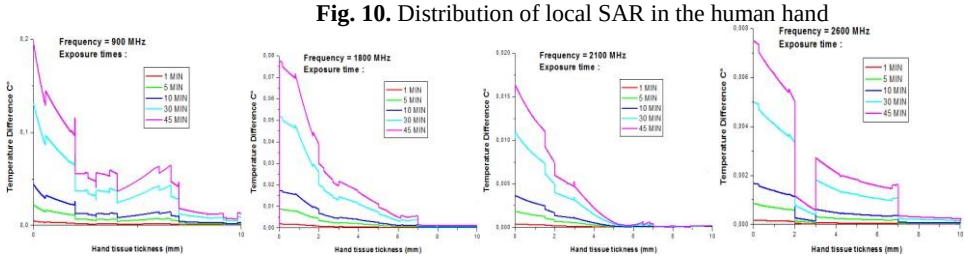


Fig. 11. Spatial distribution of temperature in the human hand.

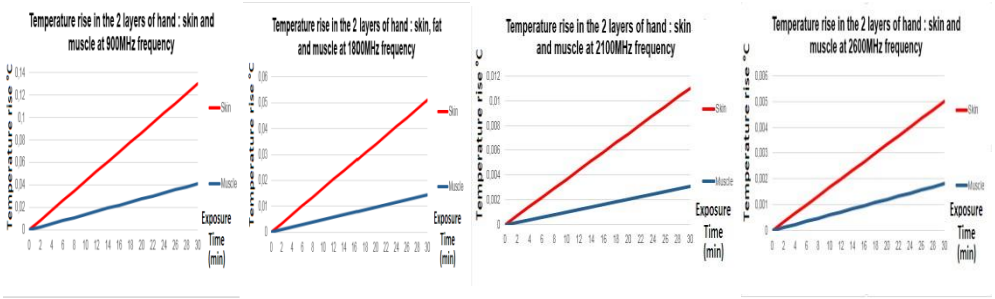


Fig. 12. The increase in temperature in both layers: skin and muscle of the human hand.

The figure 10 represents the distribution of local SAR in the human hand. The figure 11 shows the spatial distribution of temperature in the human hand. In the figure 12 we find the increase in temperature in the two layers skin and muscle of the human hand.

In the figure 10, we notice that:

- The SAR has 2 peaks in the two layers : skin and muscle .
- As we move away from the skin to the bone, the value of SAR decreases and has very low value in the bone.
- The SAR value decreases as the frequency rises from 900Mhz and 1800Mhz to 2100Mhz and 2600Mhz.

In the figure 11 and 12, we notice that:

- The temperature has 2 peaks in the two layers: skin and muscle.
- As we move away from the skin to the bone, the value of temperature decreases and has very low value in the bone.
- The temperature value decreases as the frequency rises from 900Mhz and 1800Mhz to 2100Mhz and 2600Mhz.
- The temperature value increases as the duration of telephone conversation rises.

We conclude that:

- The 2 layers skin and muscle have significant water levels so a large absorption of mobile electromagnetic waves by their biological tissues. This explains why we have 2 peaks in SAR and temperature in these 2 layers.
- The 2 layers represent a natural protection for the bone against electromagnetic waves of mobile that is why we have minimum values of SAR and temperature in the bone of human hand.
- The migration of telecom operators from older mobile telephony generations of G1 (900Mhz) and G2 (1800Mhz) to G3 (2100Mhz) and G4 (2600 Mhz) is very beneficial for mobile phone users and this explains why we have minimum SAR and temperature values for the 2100 Mhz and 2600 Mhz frequencies.
- The warmth we feel in the superficial layer of the hand during a long telephone conversation is explained by the high SAR and temperature value in the skin of the human hand.

5.3 Mobile and human head inclinations impact on the SAR values

The table 3 represents SAR values with different inclinations θxmax and θymax between the human head and the mobile phone for a distance ΔZ = 10 mm.

Table 3. SAR values with different inclinations θ_{xmax} and θ_{ymax} between the human head and the mobile phone for a distance ΔZ = 10 mm.

			SAR_{max}			
Positions	**θ_{xmax}**	**θ_{ymax}**	**900MHz**	**1800MHz**	**2100MHz**	**2600MHz**
P1	-30	0	0,01327	0,04407	0,00495	0,00157
P2	-20	0	0,00945	0,01973	0,00314	0,00107
P3	-10	0	**0,00814**	0,01926	0,00190	0,00057
P4	0	0	0,00916	0,03257	0,00193	0,00040
P5	10	0	0,01230	0,04487	**0,00119**	0,00084
P6	20	0	0,01480	0,04561	0,00256	0,00097
P7	30	0	0,02384	0,10674	0,00399	0,00144
P8	0	-30	0,01987	0,04702	0,00240	0,00171
P9	0	-20	0,01312	0,0119	0,00267	0,00121
P10	0	-10	0,01006	0,02181	0,00209	0,00082
P11	0	10	0,00916	0,05223	0,00143	0,00045
P12	0	20	0,00930	**0,01015**	0,00152	**0,00031**
P13	0	30	0,00920	0,0132	0,00212	0,00048

The figure 13 represents SAR with different head-antenna positions for 4 frequencies 900MHz, 1800MHz, 2100MHz and 2600MHz.

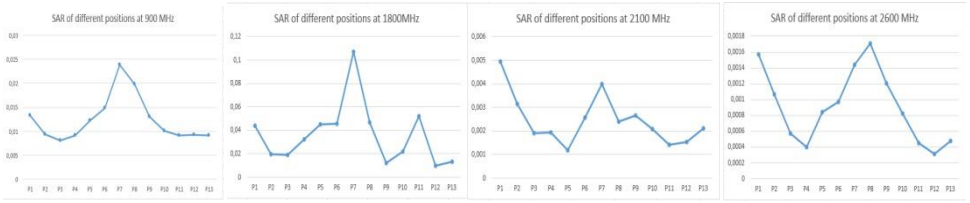


Fig. 13. SAR with different head-antenna positions for 4 frequencies 900MHz, 1800MHz, 2100MHz and 2600MHz.

The table 4 represents optimum positions with θxmax and θymax inclinations between the human head and the mobile phone for the four frequencies of 900Mhz, 1800Mhz, 2100Mhz and 2600Mhz.

Table 4. Optimum positions with θ_{xmax} and θ_{ymax} inclinations between the human head and the mobile phone for the four frequencies.

	Optimum head-mobile position	
Frequency	**θ_{xmax}**	**θ_{ymax}**
900MHz	-10°	0°
1800MHz	0°	20°
2100MHz	10°	0°
2600MHz	0°	20°

6 Conclusion

The intensity of the SAR and the increase in temperature have important values if the antenna is also close to the human head and less important when it is a little far away. The intensity of the temperature rise has small values with little exposure time and significant for a long duration of exposure. The three layers: skin, muscle and CSF protect the brain as these tissues absorb most of the power emitted by the mobile phone. The study of the thermal effect of GSM on the human head and the human hand has given results for the SAR (<1.42 W / Kg and <0.3 W / Kg) and for the temperature increase (<1 ° C) are significantly lower compared to the maximum limits (1.6 W / kg) [5] and (4.5°C) [6] set by international standards. Our study confirmed the existence of optimal inclinations between the mobile phone and the human head. This confirms that our patch tetra-generations patch is compliant and meets the standards of dosimetry and temperature when used near the head and the human hand.

Acknowledgments

We thank the members of the Institute of Electronics and Telecommunications Rennes (IETR) for their cooperation and help in the accomplishment of this work.

References

[1] J. Baayer, R. Karli, and H. Ammor : An Innovative Conception of a Compact Quadruple Frequency Microstrip Patch Antenna for Wireless and Mobile Communication. International Journal of Applied Engineering Research (IJAER). ISSN 0973-4562. Vol.10, N°.5 pp. 3979-3982(2015)

[2] P.J. Soh, G.A.E. Vandenbosch, F.H. Wee, A. Van Den Bosch, M. Martinez-Vazquez and D. Schreurs : Specific Absorption Rate (SAR) evaluation of biomedical telemetry textile antennas. Microwave Symposium Digest (IMS). IEEE MTT-S International. IEEE, pp. 1-3 (2013)

[3] Diao, Yinliang, Al, : Detailed modeling of palpebral fissure and its influence on SAR and temperature rise in human eye under GHz exposures. Bioelectromagnetics37.4 : 256-263B(2016)

[4] R. Karli, H. Ammor and J. Terhzaz : Dosimetry in the human head for two types of mobile phone antennas at GSM frequencies. Central European Journal of Engineering.Vol. 4, N. 1. pp. 39-46 (2014)

[5] A. Hirata, K. Shirai and O. Fujiwara : An Averaging Mass of SAR Correlating with Temperature Elevation Due to a Dipole Antenna. Progress in Electromagnetics Research, PIER, 84, pp. 221-237 (2008)

[6] A. C. Guyton and J. E. Hall: Textbook of Medical Physiology. Philadelphia. PA: W. B. Saunders (1996)

A New Design of a Miniaturized UHF-RFID Passive Tag Antenna Based on L-Shape Radiators with Meandered Dipole

Yassine Gmih and Abdelmajid Farchi
{yassine.gmih@gmail.com, abdelmajid.farchi1@gmail.com}

Laboratory of Mechanical Engineering, Industrial Management and Innovation
University Hassan 1st, Faculty of Sciences and Technologies, Settat- Morocco

Abstract. This article suggests a simple and new miniaturized antenna for radio-frequency identification (RFID) tags in the European and Moroccan UHF band (866–869 MHz). Our structure consists of a meandering dipole, L-shaped radiators and a changed T-Match network. The meander technique was applied to decrease the antenna size; using this method, the size of our structure has been reduced to 68% compared to the classical UHF dipole. L-shaped radiators and a modified T-matching approach are also used to easily match the antenna input impedance to that of the chip to maximize power transfer. The size of our antenna is 50 x 22 mm^2, it has been designed on a single-sided FR-4 substrate with a thickness of 0.8 mm for a relative permittivity, Ɛr of 4.4. For our design, we chose the Higgs 4 IC chip (manufactured by Alien Technology, USA) which has an impedance of 20.5-j191 Ω obtained at the centre frequency 867 MHz. The antenna performance was studied with respect to the impedance matching, realized gain, reflection coefficients and read range of the label. Therefore, the proposed design achieved a gain of 1.2 dB and a maximum reading range of 13.94 meters at 867 MHz.

Keywords: Tag antenna; meander-line; T-Match; ultra-high frequency; radio frequency identification; L-Shape radiators; read range.

1 Introduction

Radio Frequency Identification or RFID is an emerging automatic identification technology that allows remote recognition, storage and retrieval of data from a target that contains personal data to be identified such as item price, date of manufacture, and ID number, etc. At present, it is receiving considerable international interest due to the growing technological support provided by a variety of RFID applications [1]. RFID systems are intended to operate in defined frequency bands, depending on the needs of the application and local radio frequency standards. Four principal frequency ranges exist: Low Frequency (LF: 125 KHz), High Frequency (HF: 13.56 MHz), Ultra-High Frequency (860–960 MHz), and Microwave band (2.45/5.8 GHz) [2], [3]. In particular, UHF passive RFID systems are very advantageously compared to regulate passive RFID systems [4] using low frequency and high frequency bands, as they can provide a greater and faster reading range and better information storage capacity [3]. Worldwide, the UHF band for RFID systems ranges from 860 to 960 MHz where each country/region has a

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284091

unique frequency band. For example, the band allocated to RFID applications in Europe is 866–869 MHz and in North and in South and North America 902–928 MHz [5], [6].

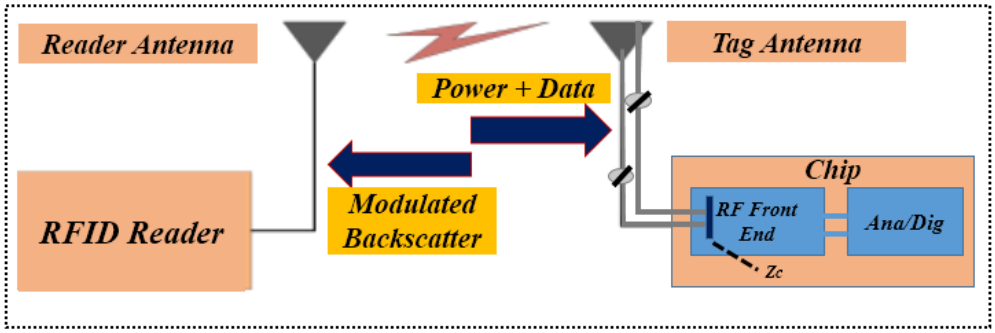


Fig. 1. RFID System Operation

A passive RFID system is composed of a tag that includes an antenna adapted to the ASIC chip (Application-Specific Integrated Circuit), a reader [7] and a personal computer to monitor the process. Figure 1 describes the functioning of a passive RFID system. In general, the chip impedance is capacitive, which requires that the antenna impedance be also inductive so as to get a proper impedance matching [8]. Several approaches are suggested in the scientific literature for impedance adaptation and miniaturization of the RFID tag antenna. The proposed feeding methods for adapting the antenna tag with the chip are T-match network, inductively coupled proximity loops, and the inclusion of nested or shaped slots [9–11]. For size reduction without significant trade-offs of tag performance, meander line structure by folding the arms of a dipole is utilized to increase the electrical length of the tag antenna [12–15], an inverted planar F antenna with a short-circuit pin used to reduce the size of a vertical wire monopole, and patch antenna with different shapes using slots on the design [16–19].

This study suggests a passive microstrip antenna with horizontal omnidirectional radiation pattern. The antenna works in the European UHF band (865–869 MHz). Its compact, low-profile structure makes it ideal for RFID applications. The antenna substrate used is an economical FR4 with a thickness of 0.8 mm. The rectangular loop of the antenna is optimized to obtain a conjugated impedance matching of the microchip. The chosen chip for the proposed tag antenna is the Alien Higgs-4 SOT-323 type chip manufactured by Alien Technology, the USA [20]. The design of our antenna is small with overall dimensions of $50 \times 22\text{mm}^2$. It is easy to design and manufacture thanks to its simple structure, without short-circuit pins or short-circuit plates.

The conjugate impedance matching and the read range of the antenna will be discussed in the following section, and Section III describes the antenna design and parametric study respectively. Section IV will present the performance of the tag and discuss the obtained results, and finally the conclusion and future work will be reported in section V.

2 Conjugate impedance matching method and read range

Read range is the activation distance d of the tag from the reader along the direction (θ, φ) [21] which is defined by the relationship (1) is generally performed in a controlled atmosphere like an anechoic chamber [12].

$$d = \frac{\lambda}{4\pi}.\sqrt{\frac{EIRP_R.\tau.Gt}{Pc}} \quad (1)$$

$EIRP_R$ is the effective power transmitted by the reader (or $EIRP_R$ =Pr x Gr, where Pr is the power transmitted by the reader, and Gr is the gain of the emitting antenna), Pc is the sensitivity of the tag transponder, it is the minimum threshold power required to activate the chip and to perform back-scattering modulations, and Gt is the tag gain.

$$\tau = 1 - [\Gamma]^2 = \frac{4.R_{ant}.R_{ch}}{|Z_{ch} + Z_{ant}|^2} \leq 1 \quad (2)$$

τ is the power transmission coefficient that is determined by the relationship (2), it is 0 to 1, if it is 1, which means that the impedance of the chip and antenna is perfectly matched, so the reading range of the labels will be maximal. Γ (formula 3) is the reflection coefficient that takes into account the impedance difference between Z_{ant} of the antenna and Z_{ch} of the chip, with $Z_{ant} = R_{ant} + jX_{ant}$ and $Z_{ch} = R_{ch} + jX_{ch}$ respectively [21].

$$\Gamma = \frac{Z_{ch} - Z_{ant}}{Z_{ch} + Z_{ant}} \quad (3)$$

Z_{ant} is the input impedance of the antenna is given [22] by (4)

$$Z_{ant} = R_{ant} + jX_{ant} = Z_{loop} + \frac{(2.\pi.f.M)^2}{Z_m} + Z_{L-rad} \quad (4)$$

Where Z_{loop}, Z_m and $Z_{L\text{-}rad}$ are the respective personal impedances of the buckle, the meander lines and the L-shaped radiator, respectively. M is the reciprocal induction between the meander dipole and the loop.

The impedance of the chip is dependent on input power and frequency, because the transponder has an energy storage level, and its impedance is capacitive is not a direct transmission line (50 Ω), which requires that the antenna impedance be inductive to get an appropriate adaptation with that of the chip (Z_{ch} = 20.5-j191 Ω).

Above d, the power absorbed by the tag decreases below the sensitivity of the chip, making it even very hard to detect the RFID tag. In order to design a compact, low-cost label, the modified t-match network is used to adjust the antenna with chip, and a meander line technique is used to decrease the geometry of the label.

3 Antenna design and parametric study

Figure 2 illustrates the structure of the proposed miniaturized tag antenna with a total dimension 50 × 22 mm^2 (Ls × Ws) it was designed on a 0.8 mm-thick (h) single layer FR-4 substrate (with relative permittivity, Ɛr = 4.4) with copper thickness of 0.035 mm and loss tangent of 0.025. Table 1 displays the parameters and the dimensions of the proposed tag

antenna. Our antenna was designed and analysed using CST Microwave Studio software, v.15. It is composed of a meander-line dipole, and L-shaped radiators, and the modified T-match used to improve the impedance matching.

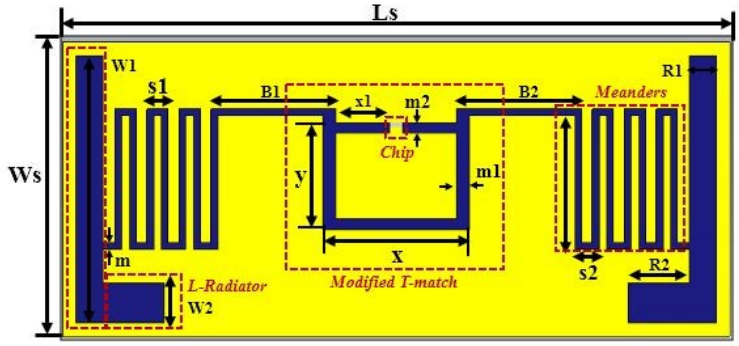


Fig.2. Geometry of the proposed antenna

Table 1. Dimensions of the proposed tag antenna

Parameters	Values (mm)	Parameters	Values (mm)
Ls	50	s2	1.8
Ws	22	m	0.5
W1	20	m1	1
W2	3	m2	0.8
R1	2	x	11
R2	4.6	x1	4
B1	9.4	y	8
s1	1.6	Gap	1

The equivalent circuit of our tag with a T-connection has been depicted in Figure 3. The entire circuit model consists of the meander antenna, the matching T-circuit and the chip model. The Rs and Cs symbols represent the microchip input resistance and capacitance respectively. Our antenna is modelized like a resistor, inductor, capacitor (RLC) in series (Ra, La, Ca), which is cascaded with the adaptation part which has an adjustable reactance. Either parameter x indicates the inductance of shunt Ls and parameter y indicates the serial inductance. By changing the length x and y, the value of the shunt and the series inductance also vary [23].

The proposed tag is designed for the Higgs-4 EPC global class 1 Gen2 UHF RFID microchip fabricated by Alien Technology [20]. The chip packaging used in the design is SOT-323. This type of packaging is easy to soldering manually. The most important advantage of using this chip is the minimum radio-frequency communication power that can reach the -18.5 dBm level. Thus, as a way of obtaining an RFID tag that can be identified at a short distance, a small antenna with a low gain can be employed.

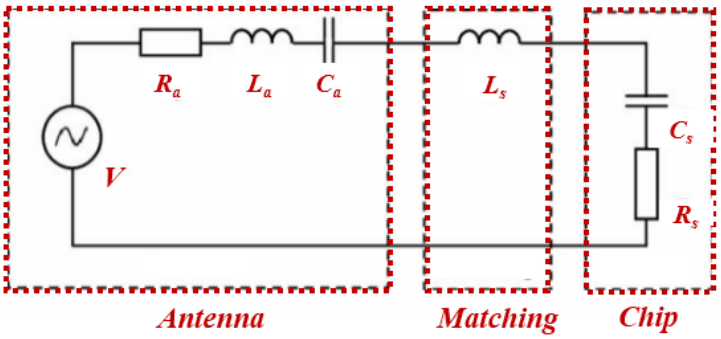
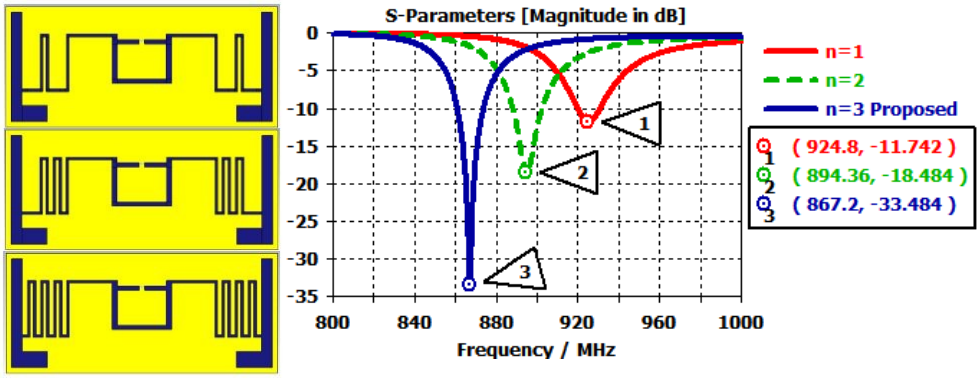


Fig.3. Equivalent circuit of tag antenna [2]

3.1 Antenna with a different number of meander lines (n)

In order to validate the design process, the tag antenna with a different number of meanders is modelled and simulated with CST-MS. n represents the number of meanders. Figure 4 (a) displays the suggested tag with a different number of meanders with the same size. Figure 4 (b) depicts the antenna reflection coefficients when n = one, two, and three, respectively. It can be seen that the -10 dB bandwidth ranges from 860–875 MHz, which covers the European RFID UHF band. When the number of meanders increases by increasing the total length of the half-wave dipole and always keeping the same total antenna size (50 x 22 mm^2), the resonance frequency shifts to the low frequencies. By using this method, the size of our antenna is decreased by up to 68% compared to the classical antenna.



(a) **(b)**

Fig.4. (a) The number of meanders n is 1, 2, and 3, respectively, **(b)** Simulated reflection coefficients of tag antenna with different meanders

3.2 Antenna with varying loop sizes

Figure 5 shows our design with different loop sizes when the number of meanders is fixed on three meanders (n = 3). The antenna in the first case has a large loop with x = 15 mm and y = 10 mm. The second has a small loop with x = 7 mm and y = 6 mm, and the optimized loop size for the proposed tag is x = 11 mm, with y = 8 mm. Figure 6 plots the input impedance characteristics ([a] real part and [b] imaginary part) of the tag. It is shown that if the horizontal x and vertical y of the T-matching rectangle (see Figure 1) are varied, the value of the shunt and inductance in series in the adaptation circuit changes. Therefore, the antenna and chip impedance can be easily adjusted both by changing the x and, y parameters, and therefore the impedance adaptation for the different chips can be easily obtained. In our case of the Higgs-4 chip, the optimal values for x and y are 11 mm and 8 mm.

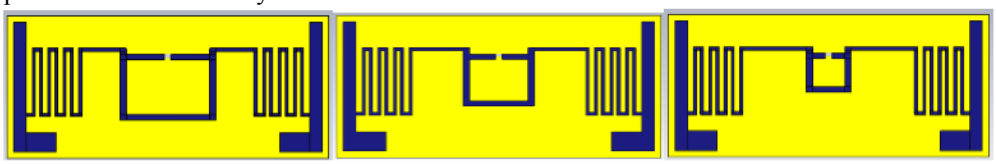
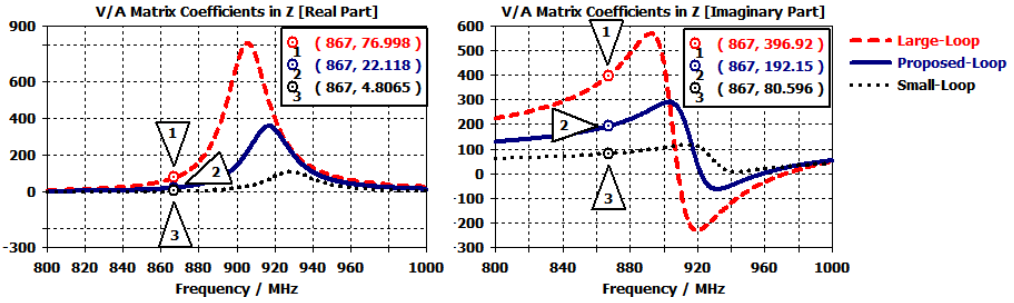


Fig.5. Tag antenna with different size of loop



(a) **(b)**

Fig.6. Simulated input impedance of tag antenna with different size of loop ((a) real and (b) imaginary part)

3.3 Antenna with and without different radiator shape

Figure 7 (b) depicts the reflection coefficients of our antenna with different shapes and without radiators, when the number of meanders is three. The first antenna has no radiator, and the second has an L-shaped radiator with a width of R1 = 2 mm, while our proposed label has an L-shaped radiator with a width of R2 = 4.6 mm in addition to R1 (see Figure 7 (a)). And since the shape of the radiator affects the impedance, the S_{11} reflection coefficients can simply be controlled by varying the width of the radiator, which can be proven by Figure 7 (b). Noting that different resonance frequencies may also be achieved with different widths of R2. By increasing its value, the resonance frequency displaces itself to the lower frequency, with an improvement in impedance matching up to -33.48dB at the 867 MHz center frequency.

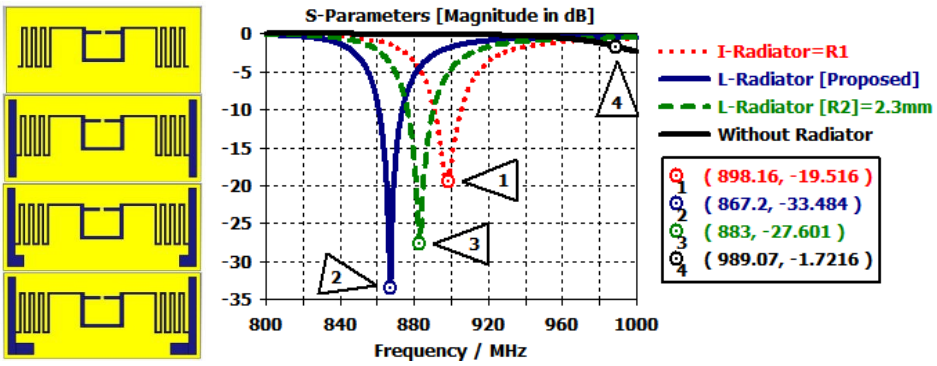


Fig.7. (a) Tag antenna with and without different radiator shapes **(b)** Simulated reflection coefficients of tag antenna with and without different radiator shapes

4 Simulated results and discussions

The proposed design is optimized using the CST-MS electromagnetic simulation software, which allows easy control of the antenna impedance matching. Figures 8 and 9 demonstrate the simulated reflection coefficient and impedance antenna of the proposed label. From the graphs, we can see that the reflection coefficient is -33.48 dB and the antenna input impedance is equal to Z_{ant} =22.1 + j192.1 Ω at 867 MHz, which is almost conjugated to the impedance of the chip Z_{ch} = 20.5-j191 Ω, this correspondence can be verified and calculated with the formula (2) of τ, we use this equation, we obtained a perfect power transfer coefficient equal to 0.98 (98% of the chip impedance). The antenna bandwidth is narrow but it covers the European band (866 - 869 MHz). After all, we can see that our antenna has a perfect adaptation response to 867 MHz, the centre frequency of the ETSI (European Telecommunications Standards Institute) band.

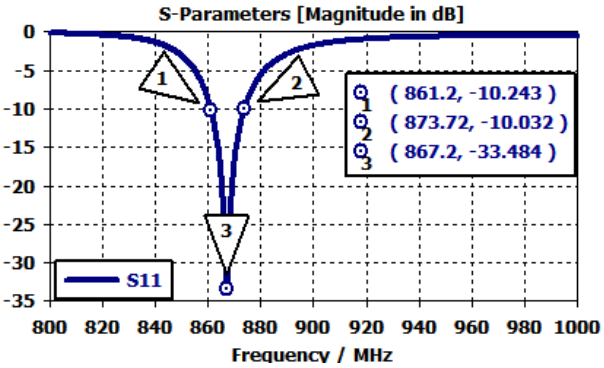


Fig.8. Simulated reflection coefficient of the proposed tag antenna

Using formula (1), it is feasible to calculate the maximum distance at which the reader can detect the tag, with a minimum output power required for chip activation of -18.5 dBm and an effective isotropic radiated power ($EIRP_R$) of the signal transmitted by the reader is 3.28 Watt in Europe. The maximum distance obtained by the suggested tag is 13.94 meters. Figure 10 (a) provides the polar radiation pattern of the label antenna. It has an omnidirectional radiation pattern in the yz- plane (phi = 90) and a bidirectional radiation pattern in the xz- plane (phi = 0) at in the 867 MHz resonant frequency, which is desirable for RFID applications. The maximum gain achieved by the proposed tag is 1.2 dB at 867 MHz, as shown in Figure 10 (b).

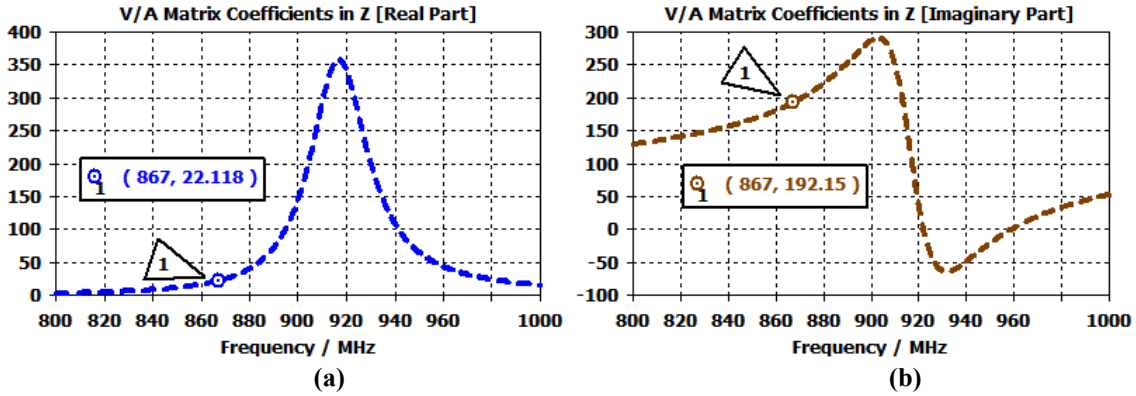
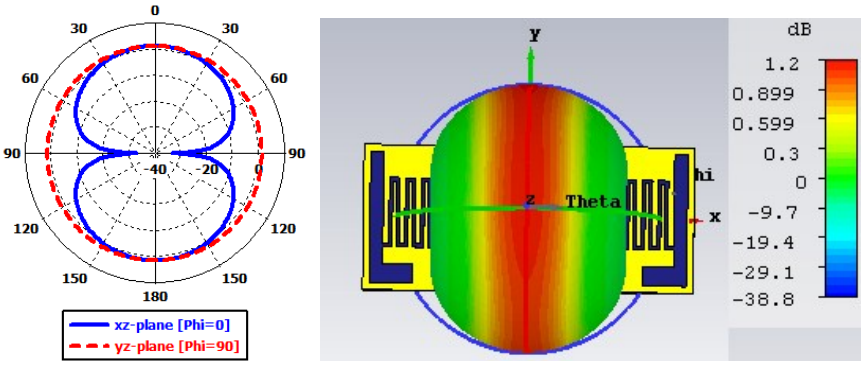


Fig.9. Simulated Input impedance of the proposed tag antenna ((a) real and (b) imaginary part)



(a) **(b)**

Fig. 10 (a) Polar radiation pattern of the proposed antenna **(b)** 3D realized gain

Table 2 compares the various antennas recently mentioned in published articles with our antenna in terms of total size, selected substrate and maximum reading range in free space.

Table 2. Comparison of previous designs with the proposed antenna

Design approach	Substrate	Frequency	Size (mm^3)	Read range in free space
Loop feed meander-line antenna [24]	Arlon foam clad (1 Er =1.1)	867 MHz	63.2 x 21.9 x 0.8	5 meters
Meander dipole tag antenna [25]	Rogers RT/duroid (Er= 2.2)	867.7 MHz	77 x 14 x 0.787	12.5 meters
Semi-circle looped with meander line antenna [26]	FR4 (Er= 4.6)	920 MHz	77.68 x 35.5 x 1.6	8.3 meters
Proposed antenna	**FR4 (Er= 4.4)**	**867 MHz**	**50 x 22 x 0.8**	**13.94 meters**

Based on the reported work on the design and implementation of the RFID tag for the UHF band, the overall antenna size in this study is the smallest of the tag antennas reported in the comparison. Additionally to its small size, our antenna offers a large maximum reading range of 13.94 meters in free space, making it a good candidate for RFID applications that require a large distance to identify it. The concept used to design compact labels, as described in the preceding section, may also be implemented on other flexible substrates for RFID identification.

5 Conclusion

A new miniaturized meandered line tag antenna in European and Moroccan band is proposed in this paper. The antenna is composed of copper layer traces and one-sided FR4 substrate with the dimension of 50 x 22 x 0.8 mm^3. The label antenna is designed to be matched to the conjugated impedance of the UHF Higgs-4 RFID chip which has an impedance of 20.5-j191 Ω. The proposed tag antenna design includes meander line, L-shaped radiators and modified T-match technique for size reduction and impedance matching respectively. The antenna does not have a short circuit plate or through hole, which makes it easier to manufacture's. The simulation results demonstrate that the impedance matching performance of the proposed tag is satisfactory at 867 MHz. In addition to this, our antenna has an omnidirectional radiation pattern with a reading range of 13.94 meters, which proves that the label antenna is a great candidate for a compact and low-power consumption system of passive RFID tags. The proposed design of the tag antenna is tested by simulation and will be manufactured and measured in the near future.

References

[1] Chawla, V., Dong-Sam, H.: An overview of passive RFID. Communications Magazine, IEEE. Vol. 45, n9, pp 7-11 (2007)

[2] Dobkin, D. M.: The RF in RFID: Passive UHF RFID in Practice. USA (2008)

[3] Finkenzeller, K.: RFID Handbook: Radio-Frequency Identification Fundamentals and Applications. Wiley, Second Ed. New York, USA (2004)

[4] El Hachimi,Y., Gmih, Y., Makroum, M., and Farchi, A.: A Miniaturized Patch Antenna Designed and Manufactured Using Slot's Technique for RFID UHF Mobile Applications. Int. J. of Electrical and Computer Engineering (IJECE). Vol. 8, No. 6, pp. 5134-5143 (2018)

[5] MingYin, L., RongLin L.: A Low Profile Broadband RFID Tag Antenna for Metallic Objects. International Conference on Microwave and Millimeter Wave Technology. pp. 1891-1893 (2010).
[6] Bhatt, H.and Grover, B.: RFID Essentials. OReilly Media, Inc. (2006)
[7] El Hachimi,Y., Gmih, Y., Makroum, M., and Farchi, A.: Miniature RFID Antenna at UHF Band using Meander-Line Technique. Int. J. of Electrical and Computer Engineering (IJECE). Vol. 8, No. 4, pp. 2280-289 (2018)
[8] Zamora, G., Zuffanelli, S., Paredes, F., Martın, F., and Bonache, J.: Design and Synthesis Methodology for UHF-RFID Tags Based on the T-Match Network. IEEE Transactions on Microwave Theory and Techniques. pp. 4090-4098, Vol. 61, n. 12, (2013)
[9] Marrocco, G.: The art of UHF-RFID antenna design: impedance matching and size-reduction techniques. IEEE Antennas Propag. Mag. Vol. 50, n. 1, pp. 66-79 (2008)
[10] Pumpoung, T.; Wongsiritorn, P.; Phongcharoenpanich, C.; Kosulvit, S.: Characteristics of a meander line tag antenna with T-Loop structure for the passive UHF RFID applications. In Proc. IEEE Int. Conf., Electron Devices and Solid State Circuit. pp. 1-4 (2012)
[11] Choo, J., J. Ryoo, and al.: T-matching networks for the efficient matching of practical RFID tags. European Microwave Conference (EuMC). Rome, Italy. pp. 5-8 (2009).
[12] Seshagiri Rao, K.V., and Nikitin, P.V.: Antenna design for UHF RFID tags: a review and practical application. IEEE Trans. Antennas Propag. Vol. 53 n. 12, pp 3870-3876 (2005)
[13] Ma, Z.L., Jiang, L.J., Xi, J., Ye, and T.T.: A compact HF/UHF dual band RFID tag antenna. In Proc. IEEE Antennas and Propagation Society Int. Symp. pp. 1122-1123 (2013)
[14] Liu, Q.; Shen, J.; Liu, H.; Wu, Y.; Su, M.; Liu, Y.: Low-cost compact circularly polarized directional antenna for universal UHF RFID handheld reader applications. IEEE Antennas Wireless Propag. Lett. Vol. 14, pp. 1326-1329 (2015)
[15] Hirvonen, M.; Pursula, P.; Jaakkola, K.; Laukkanen, K.: Planar inverted-F antenna for radio frequency identification. Electron. Lett. Vol. 40 (14), pp. 848–850 (2004)
[16] Deleruyelle, T., Pannier, P.: Multi-standard slot antenna in 860–960 MHz and 2.45 GHz RFID band. In Proc. Antennas and Propagation Society Int. Symp. IEEE. pp. 1-4 (2008)
[17] Santiago, A.G., Costa, J.R., and Fernandes C.A.: Broadband UHF RFID passive tag antenna for near-body applications. IEEE Antennas Wireless Propag. Lett. Vol. 12, pp. 136-139 (2013)
[18] Svanda, M., Polivka, M.: Matching technique for an on-body low- profile coupled patches UHF RFID tag and for sensor antennas. IEEE Trans. Antennas Propag. Vol. 63 (5), pp. 2295-2301 (2015)
[19] Mirchandani, P.; Deshpande, K.; Khan, R.: Performance analysis of miniaturized microstrip patch antennas for UHF RFID applications. In Proc. Int. IEEE Conf. Signal Processing, Comm., Computing and Networking Technologies. pp. 244-249 (2011)
[20] Higgs 4 IC Datasheet. An EPCglobal Complaint Class-1 Gen-2 ISO/IEC 18000-6C UHF RFID Integrated Circuit, Alien Technology (2017)
[21] Zhang, J., and Long, Y.: A dual-layer broadband compact UHF RFID tag antenna for platform tolerant application. IEEE Tran. On Antennas and Propagation. Vol. 61, no. 9, pp. 4447-4455 (2013)
[22] Choi, W., Son, H. W., Shin, C., Bae, J., and Choi, G.: RFID tag antenna with a meandered dipole and inductively coupled feed. IEEE Antennas and Propagation Society International Symposium. pp. 619-622 (2006)
[23] Tsai, M.-C., Chiu, C.-W., Wang, H.-C., and Wu, T.-F.: Inductively coupled loop antenna design for UHF RFID on Body applications. Progress in Elect. Research, Vol. 143, pp. 315-330 (2013).
[24] Ma, Y., Abd-Alhameed, R. A., Zhou, A., See, C.H., Abidin, Z.Z., Jin, C., and, Peng, B.: Loop feed meander-line antenna RFID tag design for UHF band. Proc. General Assembly and Scientific Symp. pp. 1-4 (2014)
[25] Dhaoudi, M., Mabrouk, M., Vuong, T.P., De Souza, A.C., and Ghazel, A.: A broadband UHF tag antenna for near-field and far-field RFID communications. Radio Eng., Vol.23 (4), 1026-1032 (2014)
[26] Rokunuzzaman, Md., Islam, M. T., Rowe, W. S. T., Kibria, S., Singh, M. J., and Misran, N.: Design of a Miniaturized Meandered Line Antenna for UHF RFID Tags. Public Library of Science (Plos One), 17 August (2016)

A Novel Compact Antenna for Passive UHF RFID Tag Using π-Match and Meander Techniques

Younes EL HACHIMI[1], Abdelmajid FARCHI[2]
{elhachimi.younes1@gmail.com[1], abdelmajid.farchi1@gmail.com[2]}

Laboratory of Engineering, Mechanics, Industrial Management and Innovation[1,2]
University Hassan 1st, Faculty of Sciences and Technologies, Settat-Morocco

Abstract. In this work, a novel passive tag antenna has been designed, investigated and proposed for UHF RFID applications. It has a small size and can operate in the North and South American band (902-928 MHz). The proposed antenna was etched on a low-cost FR4 substrate with an overall size of 54 × 24 × 0.8 mm^3 in order to be applied as UHF RFID tag antenna at the center frequency of 915 MHz. It consists of embedded Π-match, meandered arms, and an RFID chip placed at the center. The meander line technique is used to assure the miniaturization and the Π-match technique to have a good impedance matching. This antenna has a good reflection coefficient and good read range.

Keywords: Passive tag, UHF, RFID, small size, read range, meander line, Π-match.

1 Introduction

In the previous years, and until today, considerable attention has been paid to radio-frequency identification (RFID) technology among researchers everywhere throughout the world attributable to its huge applications in logistics, retail management, electronic transport payment, access control, automotive security, automated libraries, healthcare industries, livestock management, etc. [1].

Radio frequency identification (RFID) is a technology for the contactless automatic identification. By using this technology, items can be recognized and followed. A RFID device is mainly made out of a reader, a tag that affixed to objects, and implanted computer system for data processing. In a passive back-dissipated RFID system, the reader transmits a modulated RF signal consistently every which way, and the tag, arranged inside the read-zone of the reader, receives the RF signal and generates a backscatter with recognizable proof information to the reader. The latter is used to peruse and write data from the tag remotely. The data processing system consists of a host computer and software to analyze the received data. Finally, the tag. It can be divided into two categories: active (with an onboard source of power, like a battery) or passive (without battery). Passive tag usually incorporates an antenna and an RF chip which has a memory to store distinguishing data [2]. The input impedance of the RFID chip must be conjugately matched with that of the antenna impedance to accomplish maximum most extreme power exchange [3]. In the available literature, several passive RFID tag antenna structures [4–9] have already been proposed.

There are four diverse frequency bands apportioned for RFID applications, low frequency (LF: 125–134 kHz), high frequency (HF: 13.56MHz), ultra-high frequency (UHF: 860–960

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284257

MHz) [10] and microwave frequency (2.45GHz and 5.8GHz) [11]. Different nations have allocated different frequency bands for UHF-RFID applications; e.g., China (840–845 MHz), Europe (866–869 MHz), USA (902–928 MHz) and Australia (920–926 MHz). At LF and HF bands, the dimensions of the tag antenna are bigger since the wavelength is bigger, bringing about a small read range between the tag and the reader. The UHF tags have gathered more consideration as they offer longer read ranges, higher reading speeds, enable simultaneous detection of an increasing number of tags and need smaller antennas [12].

For handy applications, RFID tag antennas must be compact, suitable for low-cost and reasonable for ease large scale manufacturing [13] and must also have specific antenna impedance. Many papers presented in the literature provided different techniques for reducing the size of the antenna, such as meandering, fractal, inverted-F, or Split Ring Resonators (SRR) structures [14–17], and other for the impedance matching, such as T-match network [18], double T-matching structure [19] and inductively coupled loop [20].

In this paper, we propose a new tag antenna for UHF RFID applications which is intended to meet the requirements of reasonable compact size and good matching simultaneously. The proposed RFID antenna consists of Embedded Π-match, meandered arms, and an RFID chip placed at the center. To have a good impedance matching between the antenna and the RFID chip, the technique of loading Π-match is applied. For miniaturization, meander lines are used. This design has a simple structure with one layer of the dielectric. The performances of this RFID tag antenna supplemented with simulated results are likewise presented and discussed. It shows good read range characteristics throughout the UHF RFID band.

2 Antenna design

For the design, modeling, and simulation of the antenna, CST (Computer Simulation Technology) microwave studio [21], whose numerical analysis is based on the Finite Integration Technique was used. The geometry of the proposed tag antenna is depicted in Figure 1. It is printed on a FR4 substrate having a thickness of 0.8mm, dielectric constant of 4.4 and loss tangent of 0.025. It is completely planar in nature. The proposed design is very simple and made out of an embedded Π-match loaded at both ends with a couple of meandered arms. This tag is without a ground plane. The Π-matching structure is used to satisfy the condition of the conjugate matching so as to obtain a good adaptation. Meander line technique helps in reducing the size of the antenna. The total dimension of the tag antenna is $54 \times 24 \times 0.8$ mm3 which correspond to $(0.165\lambda) \times (0.073\lambda) \times (0.002\lambda)$, where λ is the free-space wavelength at 915 MHz. The optimized parameters are shown in Table 1.

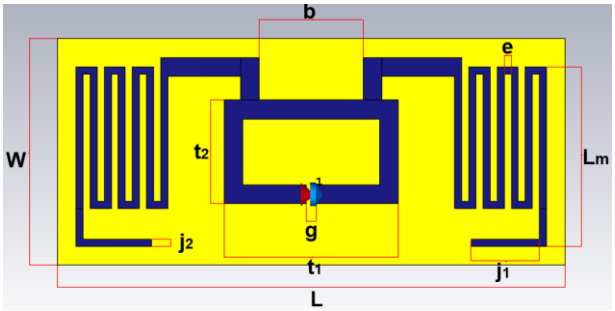


Fig. 1. Antenna structure.

Table 1. Dimensions of the proposed antenna.

Parameters	Dimensions (mm)	Parameters	Dimensions (mm)
L	54	d	11
W	24	j_1	7.25
t_1	18.5	j_2	0.75
t_2	11	Lm	19
g	1	e	0.75

3 Input impedance of the RFID chip

The microchip used in UHF RFID tags have capacitive input impedance. In order to deliver the maximum power transfer, tag antennas are designed to have an inductive impedance. The input impedance of the antenna, Za should be complex conjugate of the chip impedance, Zc (Zc = Za*). The chip used in this configuration is "Alien Higgs 4" [22]. It is a passive tag microchip having an input impedance of Zc = (18,43−j181,2) Ω at 915 MHz in North and South of America. Therefore, the antenna is designed for Za = (18,43+j181,2) Ω.

The RFID impedance can be modeled either by a parallel or series equivalent circuit Rp, Cp, Rs, Cs. These parameters can be calculated as shown in (1) (2) (3) (4) [23]. In this work, we have modeled this microchip by a series equivalent circuit (Rs=18,43 Ω , Cs= 0.955pF). Figure 2 illustrates the modeled equivalent circuit. The parameters of the equations are as follows:

$$R_p = \frac{Im_{chip}^{\ 2} + R_{chip}^{\ 2}}{R_{chip}} \quad . \tag{1}$$

$$C_p = \frac{Im_{chip}}{2\pi f . (Im_{chip}^{\ 2} + R_{chip}^{\ 2})} \quad . \tag{2}$$

$$R_s = R_{chip} \quad . \tag{3}$$

$$C_s = \frac{1}{2\pi f . Im_{chip}} \quad . \tag{4}$$

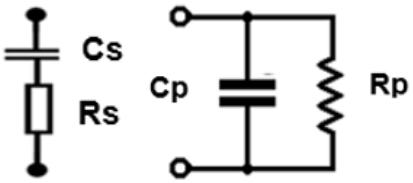


Fig. 2. Equivalent circuit of chip impedance

4 Parametric study

Since every geometrical parameter affects the antenna performance, a parametric study is being employed to optimize the final design. The radiating element has three key parameters: "t1", the length of Π-matching structure's rectangle, "b", the distance between the two arms of the Π-matching structure, and "j1", the length of the end of meandered arms. Each physical attribute of the antenna is independently varied with all other parameters being kept unchanged. In order to examine the effect of the Π-match on the resistance and reactance of the tag, t1 and b are varied and the result is presented in Figure 3 and Figure 4, respectively. It is seen from the graph that when the value of t1 is varied, the resistance values are slightly changed, and there is a significant change in the reactance values, i.e., a rise in parameter t1 causes their magnitudes to rise (important change for the imaginary part and a small one for the real part), while a fall in these parameters causes the magnitude to fall. For the distance b between the two arms of the Π-matching structure, it can be seen that the change in b causes a small change in the real and imaginary parts of the input impedance. The optimized values of t1=18.5mm and b=11mm, allowed us to have a good conjugate matching between the antenna and the chip. Figure 5 plotted the reflection coefficient of our RFID tag with different value's of j1 when t2 and b are fixed as 18.5 mm and 5 mm respectively. It is observed that the resonance frequency and the reflection coefficient decreases with the increase in j1. The best result was observed when the value of j1 was 7.25mm. We can conclude that the Π-matching structure has an effect on the impedance of the antenna and the length of meandered arms affects the resonance frequency.

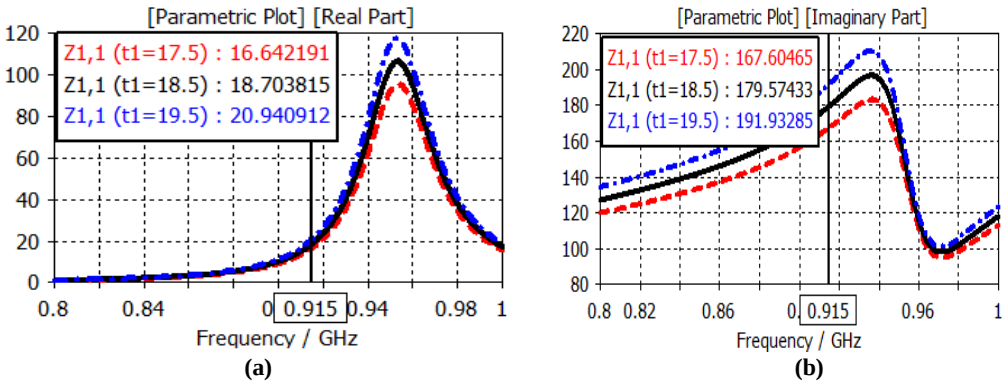


Fig. 3. Simulated input impedance of proposed antenna with various t1, (a) real part and (b) imaginary part.

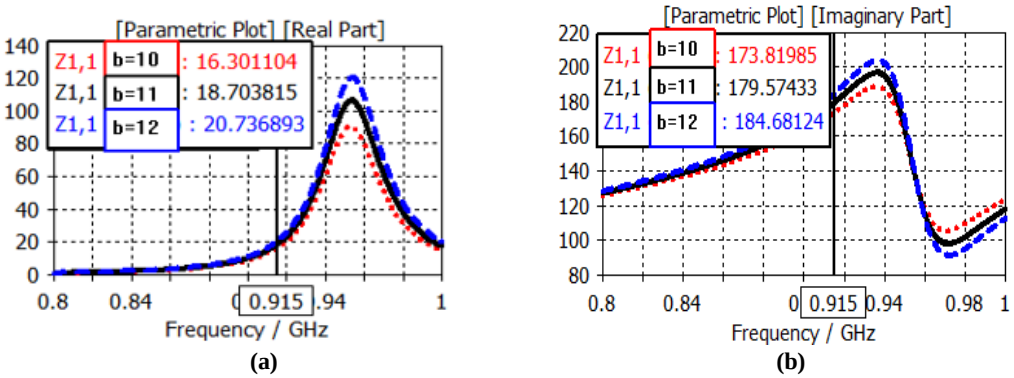


Fig. 4. Simulated input impedance of the proposed antenna, (a) real part and (b) imaginary part with various b

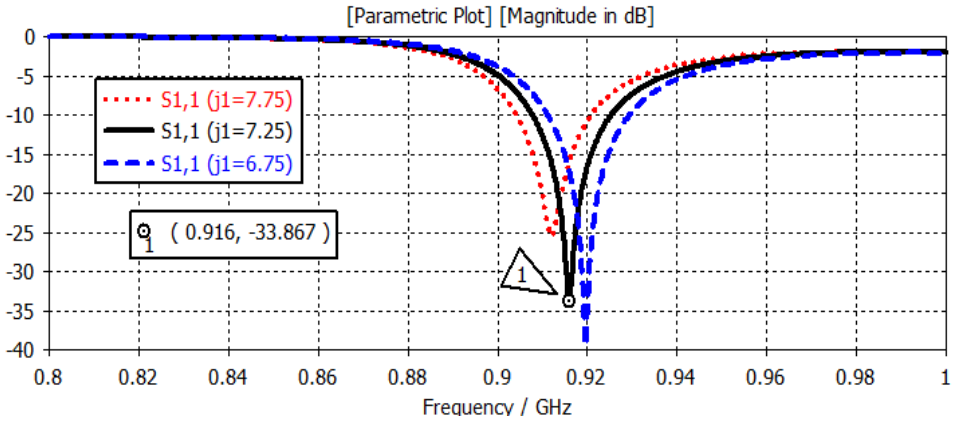


Fig. 5. Simulated return loss of the proposed antenna with various j1

5 Results and discussion

The goal of this study is to design a new tag antenna with compact and low profile structure for UHF-RFID band also a good return loss and matching between the tag antenna and the chip using the Π-matching structure and meander line technique. Figure 6 depicts the simulated reflection coefficient characteristics of the designed antenna as a function of frequency. We notice that the simulated return loss reaches a value of -33.86 dB at 916 MHz with a bandwidth of 17.61 MHz ranging from 907.79 to 925.4 MHz, almost covering the North American UHF RFID Band (902–928 MHz). This value of S11 is less than -10 dB, which reflects the antenna adaptation. The simulated real and imaginary part of the input impedance of the proposed tag antenna versus frequency is illustrated in Figure 7. As can be seen in the figure, the input impedance of the antenna is (18.7+j179.57) Ω, and it is almost the value of complex conjugate of the chip impedance (18,43−j181,2) Ω at 915 MHz. This leads to the observation that this new shape (Π) has allowed us to achieve a good conjugate matching between the antenna and the

chip, which reduces the losses due to reflections. The simulated 3D gain of the proposed tag antenna is presented in Figure 8. We can notice that the gain reached a value of 1.31 dB at 915 MHz. The simulated 2D radiation pattern at E-plane and H-plane is shown in Figure 9. It can be observed that the radiation patterns are bidirectional in the E-plane and almost omnidirectional in the H-plane. To determine the performance of the tags, it is important to get the read range [3], which can be calculated utilizing the Friis-free space formula as:

$$r = \frac{\lambda}{4\pi}\sqrt{\frac{EIRP.G_r.\tau}{P_{chip}}} \tag{5}$$

where λ is the wavelength and EIRP is the effective isotropically radiated power, determined by local country regulations (e.g., 3.8 W in Europe and 4 W in the USA). P_{chip} is the minimum threshold power necessary to activate the RFID chip, G_r is the gain of the receiving tag antenna, and τ is the power transmission coefficient. Hence, the calculated maximum read range was found to be 16.13 m at 915 MHz..

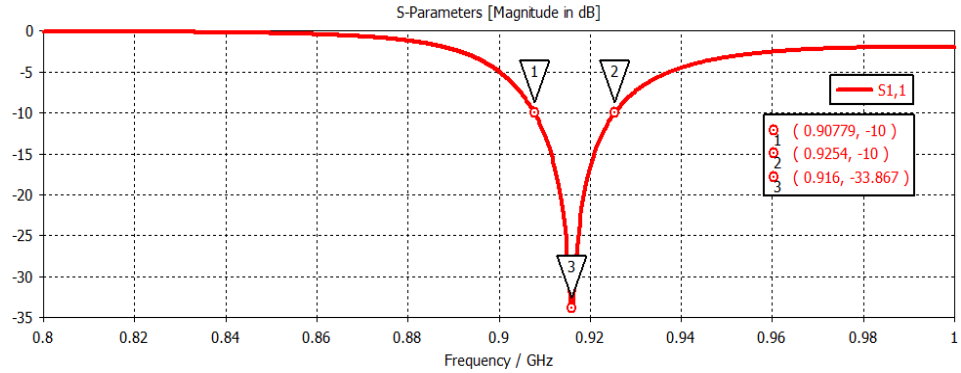
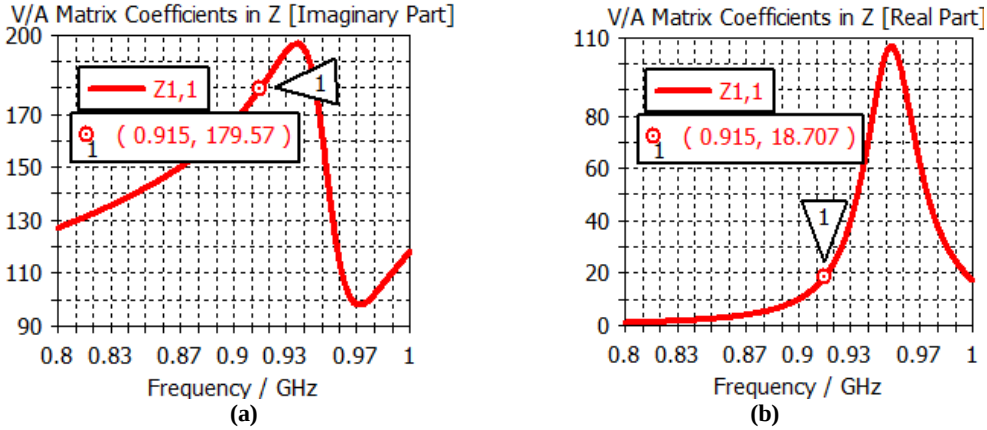


Fig. 6. Simulated return loss of the proposed tag antenna



(a) **(b)**

Fig. 7. Simulated Input impedance of the proposed antenna, (a) real part and (b) imaginary part.

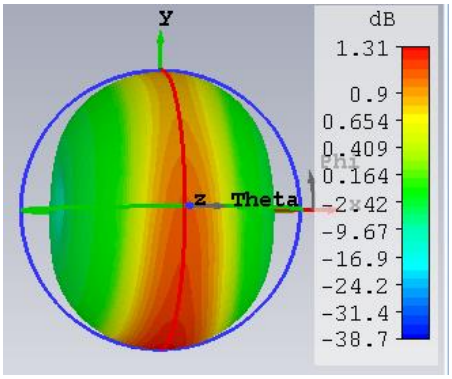
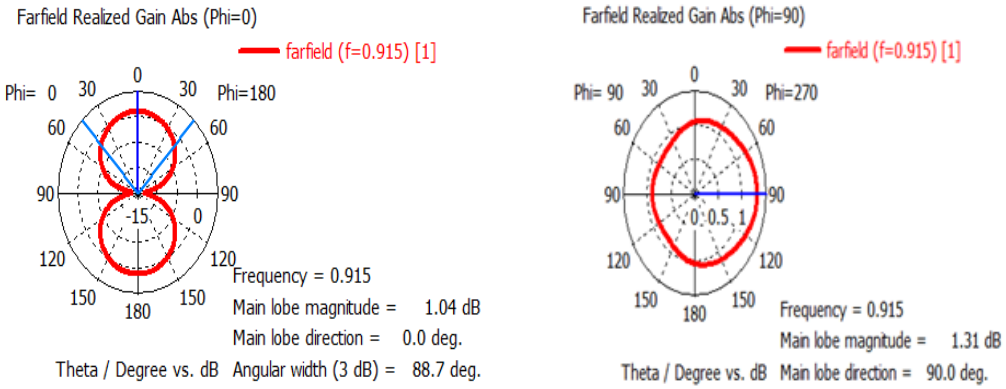


Fig. 8. Simulated 3D gain of the proposed tag antenna.



(φ=0) **(φ=90)**

(a) **(b)**

Fig. 9. Simulated 2D radiation pattern: (a) E-plane and (b) H-plane.

6 Comparison

In order to emphasize the importance of miniaturization, a comparison between the proposed tag with some other designed antennas in terms of volume, return loss, bandwidth, and read range is exhibited in Table 2, respectively. The first antenna [24] uses a semi-circle loop and two feeding strips and two triple-flipped meandered lines of symmetrical identity at both sides of the antenna. The second one [25] consists of meander short dipole antenna printed symmetrically on the same face and loaded at both ends by Split ring resonator (SRR). The proposed design is made out of an embedded Π-match loaded at both ends with a couple of meandered arms. The comparative table shows that the volume of the proposed tag antenna is smaller than the antennas in [24] and [25] which means that our tag has the compact size. It also has the best read range and return loss compared to the other ones. In fact, the results show that the bandwidth of our RFID tag is a bit small compared to the other antennas. We can conclude that the technique of "Π-match" allowed us to have the most compact structure and a good

conjugate matching between the antenna and the chip compared to the previous presented techniques employed in [24] and [25].

Table 2. Comparison of the proposed tag and other designs**.**

Antennas	Volume (mm^3)	Return loss (dB)	Bandwidth (MHz)	Read range (m)
[24]	4412.224	-23	21	8.3
[25]	2288	-30	18	4.36
This work	1036.8	-33.86	17.6	16.13

7 Conclusion

In this paper, systematic and simple design of a novel compact antenna for passive UHF RFID tag has been introduced. To get a good conjugate matching with the chip and the proposed antenna and to ensure the compactness of our structure, the Π-match network and meandered arms has been used. In order to emphasize the importance of miniaturization, a comparison between the proposed antenna with some other designed antennas in terms of volume, return loss, bandwidth, and read range is exhibited. The designed antenna has an overall volume of $54\times24\times0.8mm^3$, which is smaller than those of previously reported structures. The configuration with low cost, low profile, easy fabrication, and a miniaturized size make our proposed tag suitable for UHR RFID applications.

References

[1] M.C. Tsai, C.W. Chiu, H.C. Wang, and T.-F. Wu: Inductively coupled loop antenna design for UHF RFID on-body applications. Progress n Electromagnetics Research. Vol. 143, pp. 315-330 (2013).

[2] T. Deleruyelle, P. Pannier, M. Egels, and E. Bergeret: Dual band mono-chip HF-UHF tag antenna. In 2010 IEEE Antennas and Propagation Society International Symposium. pp. 1-4, Toronto, Canada (2010).

[3] G. Marrocco: The art of UHF RFID antenna design: Impedance-matching and size-reduction techniques. IEEE Antennas Propagation Magazine, Vol. 50, pp. 66-79 (2008).

[4] G. Marrocco: Gain-optimized self-resonant meander line antennas for RFID applications. IEEE Antennas Wireless Propagation Letters. Vol. 2, pp. 302-305 (2003).

[5] P. Kamalvand, G. K. Pandey, and M. K. Meshram: RFID tag antenna for ultra and super high frequency band applications. International Journal of RF and Microwave Computer-aided Engineering, Vol. 26, No. 7, pp. 640-650 (September 2016).

[6] H. Chen, C. Tsai, C. Sim, and C. Kuo: Circularly Polarized Loop Tag Antenna for Long Reading Range RFID Applications. IEEE Antennas and Wireless Propagation Letter. Vol. 12, pp. 1460-1463 (2013).

[7] D.L. Nguyen, K.S. Paulson, and N. G. Riley: Reduced size circular polarized square microstrip antenna for 2.45GHz RFID applications. IET Microwaves, Antennas & Propagation. Vol. 6, No.1, pp. 94-99 (2012).

[8] Y. Shi, K. Qi, and C. H. Liang: A miniaturized design of 2.45GHz RFID tag antenna. Microwave and Optical Technology Letters, Vol. 57, pp. 1905-1908 (2015).

[9] Y.Shi, C. Fang, K. Qi, and C. H. Liang: A broadband design of UHF fractal RFID tag antenna. Progress In Electromagnetics Research Letters, Vol. 58, pp. 45-51 (2016).

[10] Y. El Hachimi, Y. Gmih, E. Makroum, and A. Farchi: A compact dual-band antenna including symmetrical slots for 2.45/5.8 GHz handheld RFID Reader applications. 4th International Conference on Optimization and Applications (ICOA). Mohammedia, Morocco (April 2018).
[11] Y. Gmih, Y. El Hachimi, E. Makroum, and A. Farchi: A small printed antenna with circular slot for European UHF RFID reader devices. 4th International Conference on Optimization and Applications (ICOA). Mohammedia, Morocco (April 2018).
[12] F. Nekoogar, and F. Dowla: Ultra-Wideband Radio Frequency Identification Systems. Springer, (2011).
[13] A. Hamani, M. C. E. Yagoub, T. P. Vuong, and R. Touhami: A novel broadband antenna design for UHF RFID tags on metallic surface environments. IEEE Antennas and Wireless Propagation Letters, Vol. 16, pp. 91-94 (2017).
[14] D. K. Naji, J. S. Aziz, and R. S. Fyath: Design of Miniaturized Fractal RFID Tag Antenna with Forced Impedance Matching. International Journal of Electromagnetics and Applications. Vol. 2, no. 5, pp. 129-139 (2012).
[15] N. H. Daud, A. Ismail, A. Reda, H. Alhawari, A. Sali, and M. F. A. Rasid: Integration of Split Ring Resonators (SRRs) to UHF RFID Tag Antenna for Size Reduction. International Conference on Engineering Technology and Technopreneuship (ICE2T) pp. 204-208 (2014).
[16] Y. Shi, C. Fang, K. Qi, and C. H. Liang: A broadband design of UHF fractal RFID tag antenna. Progress In Electromagnetics Research Letters, Vol. 58, pp. 45-51 (2016).
[17] H.-T. Chou, J.-C. Chu, and Y.-Y. Kuo: Size reduction of patch antenna using a meander line feeding structure for RFID applications. Microwave Opt Technol Lett 56. pp. 918-920 (2014).
[18] G. Zamora, S. Zuffanelli, F. Paredes, F. Martin, and J. Bonache: Design and synthesis methodology for UHF-RFID tags based on the T-matched network. IEEE Trans. Microw. Theory Techn, Vol. 61, no. 12, pp. 4090-4098 (Dec. 2013).
[19] C.-W. Chiu, M.-C. Tsai, H.-C. Wang, and T.-F. Wu: Broadband Tmatching loop tag antenna design for on-body UHF RFID applications. Microwave and Optical Technology Letters, vol. 56, no. 5, pp. 1194-1200, (May 2014).
[20] Q. Liu, Y. F. Yu, and S. He: Capacitively loaded, inductively coupled fed loop antenna with an omnidirectional radiation pattern for UHF RFID tags. IEEE Antennas Wireless Propag. Lett. Vol. 12, pp. 1161-1164 (2013).
[21] CST Studio Suite version 2015 :http://www.cst.com/Content/Products/DS/Overview.aspx.
[22] Alien Higgs 4, www.alientechnology.com/products/ic/higgs-4/
[23] A. Ghiotto: Conception d'Antennes de Tags RFID UHF, Application à la Réalisation par Jet de Matière. Thèse de Doctorat. Institut Polytechnique de Grenoble (Novembre2008).
[24] Md. Rokunuzzaman, M. T. Islam, W. S. T. Rowe, S. Kibria, M. J .Singh, and N. Misran: Design of a Miniaturized Meandered Line Antenna for UHF RFID Tags. Public Library of Science (Plos One). Vol. 11, no. 8, pp. 1-14 (August 2016).
[25] I. Aznabet, M. Ennasar, O. El Mrabet, S. Tedjini, and M. Khalladi: Meander-Line UHF RFID Tag Antenna loaded with Split Ring Resonator. 5th International Conference on Multimedia Computing and Systems (ICMCS). Marrakech, Morocco (2016).

A Metamaterial inspired elliptical slotted antenna for WLAN and uplink /downlink X-bands Rejection

Mourad Elhabchi[1*], Mohamed Nabil Srifi[2], Raja Touahni[3]
{mouradelhabchi@hotmail.fr , srifimn@ieee.org , touahni-raja@univ-ibntofail.ac.ma}

[1*,3]LASTID Laboratory, Department of physics, Faculty of Sciences, Ibn Tofail University ,kenitra,morocco, mouradelhabchi@hotmail.fr and touahni-raja@univ-ibntofail.ac.ma
[2]Electronics and Telecommunication Systems Research Group, National School of Applied Science (ENSA), Ibn Tofail University, kenitra,morocco, srifimn@ieee.org

Abstract. In this communication, we propose a metamaterial inspired conventional elliptical slotted coplanar waveguide (CPW) UWB antenna for dual band-notched features. Initially, the operating frequency range is from 4GHz to 12 GHz. To filter the undesired WLAN (5-6GHz) and uplink and downlink of X-band satellite communication systems (7.25 to 8.39 GHz) frequency bands ,we are loaded respectively a symmetrical dual split ring resonator metamaterials SRR as a defected ground plane slots (DGS) and a single split ring resonator metamaterial (SRR) as a strip on the top edge of the radiating slotted elliptical antenna.The performance of the antenna is simulated both by CST Microwave Studio and Ansoft HFSS .To analyze the effects of the SRR slot and strips ,the surface current distribution is used. The antenna gain versus frequency is acceptable except the notched band regions, its value is reduced from its normal to be a negative value in (5-6 GHz) and (7.25 to 8.39 GHz) suppressed bands.

Keywords: metamaterial; UWB antenna; split ring resonator; defected ground plane; band-notched.WLAN; X-band.

1 Introduction

Since February 2002 when the Federal Communications Commission (FCC) allocated the frequency range from 3.1 GHz to 10.6 GHz to be a free band with a good effective radiated power spectral density of -41.3 dBm/MHz[1], the UWB systems become very attractive. However, there are several other excite technologies for a long time that can interfere with the UWB band systems, such as Wireless Local Area Network (WLAN) in the 5.15–5.825GHz band and 7.25-7.75 GHz band for downlink of X-band satellite communication systems.

Several published papers focus their research to filter the unwanted frequency spectrums.A.Yadav et al uses three rectangular slots loaded CPW – feed ultrawideband (UWB) to create dual band-notched [2] .B. Li et al studied a switched coplanar waveguide (CPW) dual stop-band antenna using a rectangular ring slot, two inverted S-shaped slots and four switches with a T-shaped radiating patch[3] .A microstrip UWB monopole structures with dual band-rejected characteristics based on single tri-arm resonator[4], parasitic slots[5], loading Open-Looped Resonator element in an UWB design [6] ,the π-shaped slit [7] the slotted elliptical UWB antenna with half circular ring radiator elements [8] ,T-shaped stubs and capacitive-load strips [9].Other resonator forms, such as the elliptic single complementary

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284108

split-ring resonators and rectangular split-ring resonators are etched in microstrip-fed UWB antenna [10],[11], on the hook-shaped defected ground structure (DGS) and semi-octagon-shaped resonant ring of the antenna [12].

In this letter, a conventional slotted elliptical UWB antenna has been modified to obtain a dual notched characteristic [14]. The split ring resonator(SRR) metamaterial is used as a defected ground plane(DGS) and a loaded strip to affect in the antenna behaviors and to filter the unwanted bands for WLAN and X-band satellite communication systems respectively. However, in order to realize the eliminated band in the WLAN region, a dual symmetrical DSG SRR are etched in the ground plane. The X-band elimination is achieved by loading a single SRR in the top edge of the radiating slotted element. The filtered bands can be controlled easily by tuning the length and widths of the matamaterial slots.The performances of the proposed design have been investigated, and the simulated results both in CST Microwave and Ansoft HFSS are presented.

2 Antenna design

The details and configurations of the proposed dual band notched elliptical slotted shaped are illustrated in fig. 1. The antenna is printed on a 20x24mm^2 layer of dielectric substrate FR4 with a permittivity of ε_r=4.3,loss tangent of tanδ =0.025 and height of h = 1.6mm.To achieve 50Ω impedance adaptation, the width of the feed line is fixed at wf=2.6mm. The dimensions of the proposed elliptical multiple band antenna are: Ws= 20mm, Ls=42mm, lf=7.8mm, wg=8.4mm, Lg=7mm, Wt=1mm, Lt=17mm, Wp_1=2.5mm, Lp=6mm, Ea= 3mm, Eb= 5mm,Wr_1= 5.5m ,Wr_2= 8mm ,Lr_1= 5mm,Lr_2= 7.5mm, Gr_1= 1mm ,Gr_2=1mm.

In order to excite the desired bands elimination in the frequency ranges of (5-6) GHz for WLAN rejection , a dual symmetrical split ring resonators (DGS) SRR metamaterial are etched in the CPW ground plane as shown in (fig. 2.Ant II). The unwanted bandstop for uplink and downlink of X-band satellite communication systems (7.25 to 8.39 GHz) is achieved when the researchers have been loaded the top edge of the radiating element with a single split ring resonators SRR metamaterial as a strip (fig. 2.Ant III). defining dimensions and position of introduced SRR metamaterial are one of the essential elements on band notched antenna design [13] Usually, the length of the slot is made approximately equal to half the guided wavelength at the notch frequency of the band[13].

$$L_{Total-wlan}=\frac{\lambda_g}{2}=\frac{c}{2f_{notch-wlan}\sqrt{\frac{\varepsilon_{r+1}}{2}}} \quad (1)$$

$$L_{Total-X-band}=\frac{\lambda_g}{2}=\frac{c}{2f_{notch-X-band}\sqrt{\frac{\varepsilon_{r+1}}{2}}} \quad (2)$$

Where L_{total_WLAN} is the total length of the DGS SRR metamaterial slot and $L_{total_X_band}$ is the total length of the loaded SRR metamaterial strip . by a way of CST and HFSS simulations we can determine and find the exact filter band.

The S_{11} parameters for different antenna designs are illuminated in the figure. 3. Including those of the initial reference antenna(fig.3.Ant I) [14]. As seen in figure.3, a dual symmetrical (DGS) SRR slotted on in the CPW ground plane leads to exhibits a single rejected band covering the entire WLAN band(fig. 3.Ant II), the rejected X-band is achieved when loading a single SRR on top edge of the radiating patch(fig. 3.Ant III).

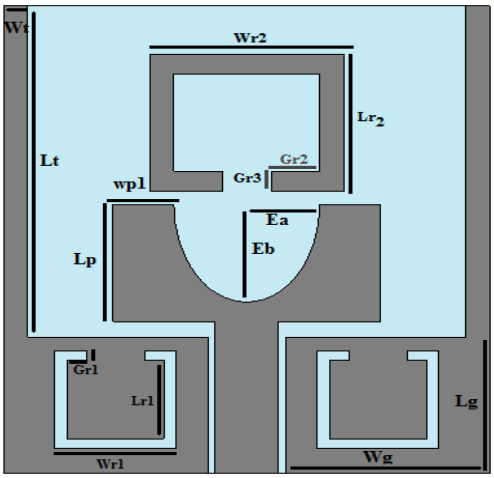


Fig. 1. The Geometry of proposed dual band notched metamaterial antenna.

Fig. 2. shows the evolution stages step by step of the proposed design.

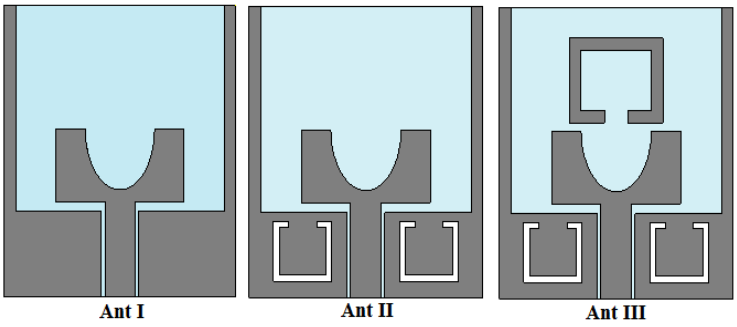


Fig. 2. The structures of the prototype and proposed antennas. Ant I: reference antenna,. Ant II: antenna with DGS SRR metamaterial. Ant III: antenna with both (DGS) SRR and loaded SRR stip.

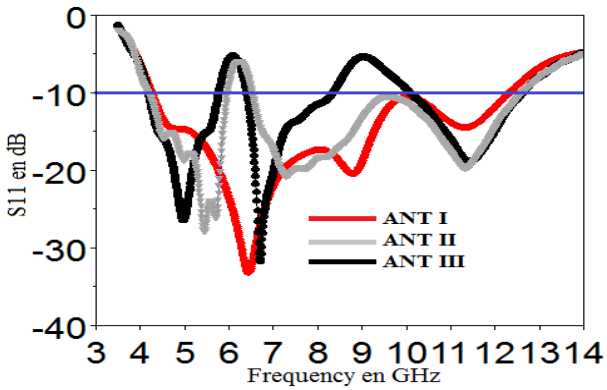


Fig. 3. The Simulated return loss for Ant I, Ant II and Ant III.

3 Parametric Study and Observation

3.1 The WLAN stop band

A parametric study is carried out to determine the influence of the essential parameters on the impedance adaptation as shown in fig. 1. Witch depicts the configuration of the proposed antenna. Figure 4 illustrates the effects of DGS SRR slot width Wr_1 dimensions. As shown, the rejected signal is shifted from 6.7 to 5.7 GHz when Wr_1 decrease from 6 to 4.5mm.

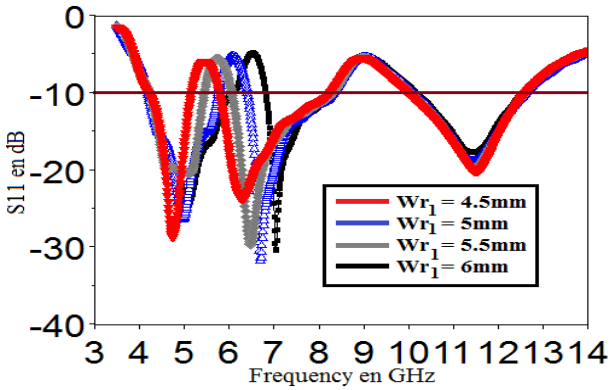


Fig. 4. Simulated reflection coefficient for different Wr_1

Moreover, the S_{11} traces increases from -7dBi to -3dBi and shifted to the higher frequency when the length of the DGS SRR Lr_1 value varies from 5.5 to 4 mm(fig. 5.)..

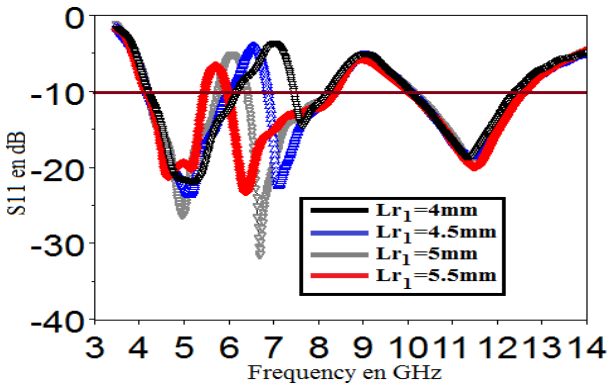


Fig. 5. Simulated reflection coefficient for different Lr_1

Figure. 6 shows the S_{11} parameters for different values of the DGS SRR slot gap Gr_1, It can be observed that the return loss S_{11} at the eliminated bands are significantly dependent on the Gr_1 values.

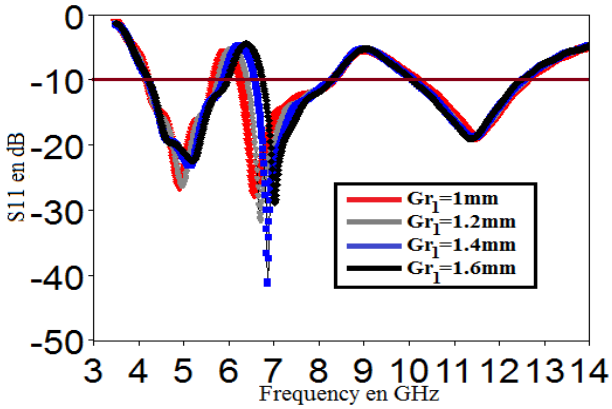


Fig. 6. Simulated reflection coefficient for different Gr_1

3.2 The uplink and downlink eliminated X-band

Figure 7 illuminates the influence of loaded SRR strip width Wr_2 values. As observed, the stopped signal to the lower frequency with desired worse impedance when Wr_2 increase from 6 to 9mm.

Thereafter, the effect of loaded SRR stub length is shown in figure 8, It can be concluded that the return loss S_{11} at the rejected bands varies when Lr_2 values increase from 6.5 to 8mm with a little effects on the low frequency due to the couplage between SRR metamaterial elements .

Figure. 9 illustrates the S_{11} parameters for different values of the SRR strip gap Gr_3, It can be observed that the S_{11} at the unwanted bands are significantly affected when the Gr_3 value changes.

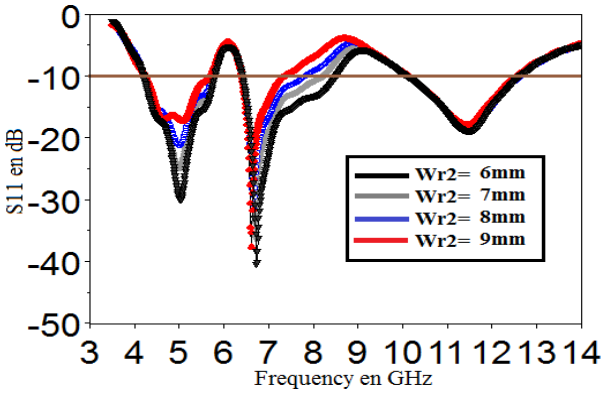


Fig. 7. Simulated S_{11} for different Wr_2

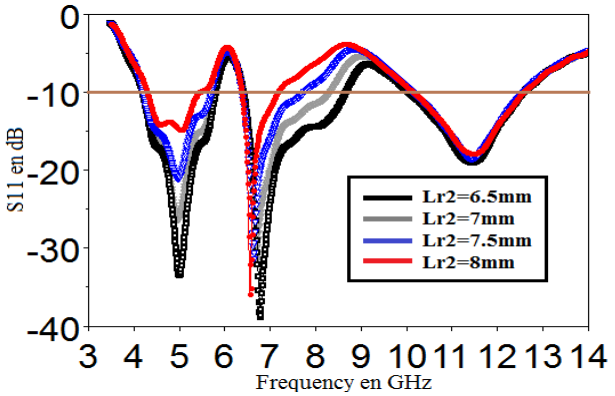


Fig. 8. Simulated S_{11} for different Lr_2

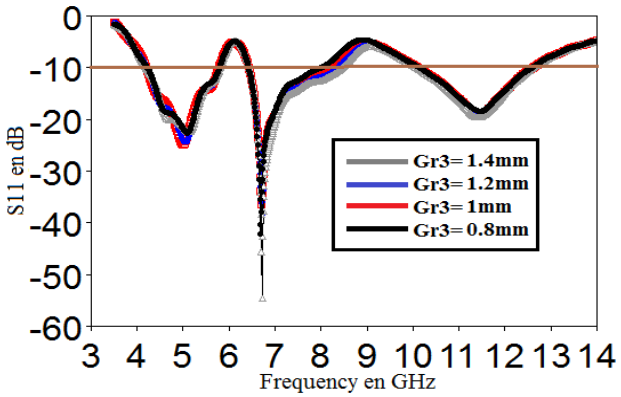


Fig. 9. Simulated S_{11} for different Gr_3

4 Current Distributions and Radiation Pattern Results

To understand the generation of the filtred bands Figure.10 depicts the current distributions density at 5.5 GHz and 8 GHz, The current is significantly important in DGS SRR slot (fig.10.a) with very current concentration in the rest part of the antenna which denotes that the loaded SRR slot metamaterials reject the WLAN band,

Figure 10 (b) illustrate the surface currents at 8 GHz, as observed in this figure the currents is important and highly distributed around the SRR loaded metamaterial strip comparing to the other part of the antenna. Therefore, the antenna relexion coefficient changed at these frequencies due to the insertion of SRR metamaterial as a slots and strip witch introduces additional paths to the surface currents at rejected signals.

Figure 11 illuminates the simulated radiation patterns both in E-plane (phi=0°) and H-plane (phi=90°) of the proposed design using CST software at sample selected frequencies of 4.5 and 10 GHz. It is observed that the proposed UWB antenna with dual stop band provide omnidirectional radiation patterns E-plane and stable bi-directional radiation patterns in H-

plane (have two main principle lobes) with a shift inaccuracy in the high frequency (10GHz) .which is obvious represents the same characteristics to those of dipole antennas.

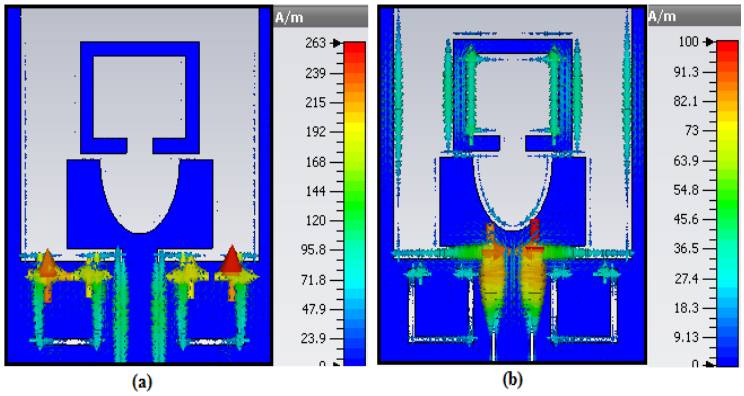


Figure 10. Current distributions at (a) 5.5 GHz, (b) 8 GHz

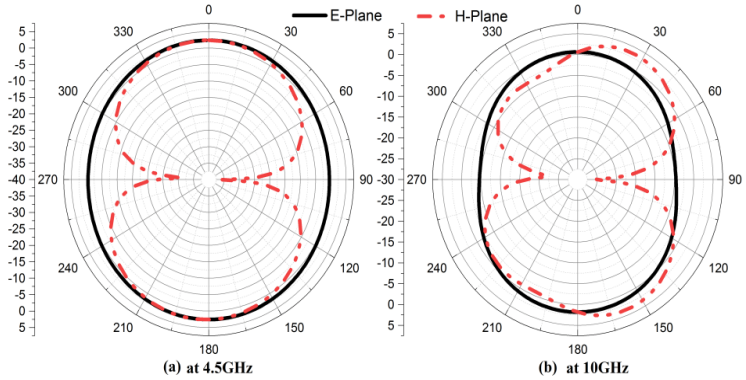


Fig. 11. Simulated radiation pattern at each selected frequencies in E-plane (solid line) and H-plane (dash dot line).

The simulated peak gain of the initial UWB structure and the proposed filtered signal antenna versus the UWB frequency range (3–14 GHz) are illustrated in figure.12. The obtained results have illuminated that substantial reductions in peak gain can be observed at the stopband frequencies.

To validate to simulated results achieved with the CST software. We designed and optimized against the proposed dual band notched antenna as HFSS to compare the results of our concept. Figure.13 present the return loss S_{11} obtained by both CST and HFSS softwares.

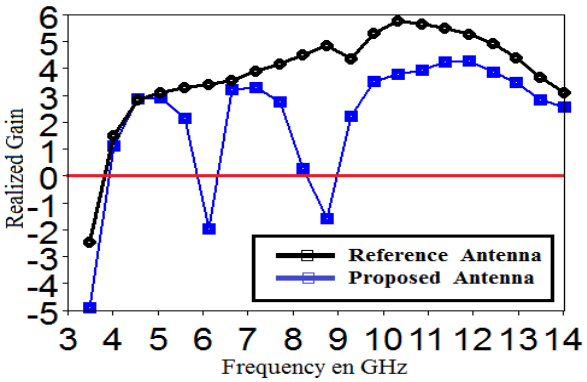


Figure. 12. The simulated gain values of the reference and modified UWB antenna.

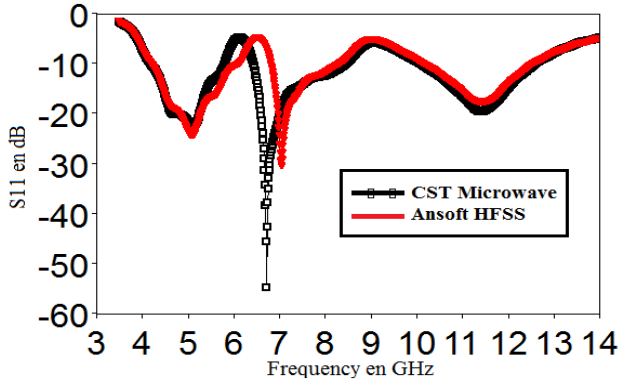


Figure 13. Comparison between the S_{11} versus frequency.

5 Conclusion

In this article, a UWB slotted elliptical with dual band notched characteristics have been presented and investigated , a dual split ring resonator metamaterials SRR as a defected ground plane slots(DGS) and a single split ring resonator metamaterials (SRR) as a stub on the top edge of the radiating antenna to avoid interference with the unwanted WLAN (5-6GHz) and uplink and downlink of X-band satellite communication systems (7.25 to 8.39 GHz) frequency bands. The bandwidths of these bands stop metamaterials can be adjusted and controlled by the dimensions of the SRR strip and slots. The prototype antenna has been simulated with CST and HFSS to validate against the obtained results.

References

[1] Federal Communications Commission Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission System from 3.1 to 10.6 GHz Federal Communications Commission. Washington, DC, ET-Docket, pp. 98–153, FCC.(2002).

[2] Yadav , A. and Sethi, D. and Khanna , R.K.: Slot Loaded UWB Antenna: Dual Band Notched Characteristics.Vol. 70(3), pp. 331-335. AEU-International Journal of Electronics and Communications (2016).
[3] Li, B. and Hong, j. and Wang, B.:Switched Band-Notched UWB/Dual-Band WLAN Slot Antenna With Inverted S-Shaped Slots .Vol. 11, pp. 572-575. IEEE Antennas and Wireless Propagation Letters (2012).
[4] Azim, R. and Islam, M.T. and Mobashsher, A.T.: Dual Band-Notch UWB Antenna With Single Tri-Arm Resonator. Vol. 13 , pp. 670-673. IEEE Antenna. Wireless Propag. Lett, (2014).
[5] Elhabchi, M. Srifi, M.N. and Touahni, R.: A Novel Dual Band Hexagonal Antenna for Bluetooth and Uwb Applications With Single Band Notched. Vol. 7, no. 5,pp. 63-67. Advanced Electromagnetics.(2018).
[6] Chu, Q. X. and Yang, Y. Y. : A compact ultrawideband antenna with 3.5/5.5 GHz dual band-notched characteristic. Vol. 56, no. 12, pp. 3637–3644, IEEE Trans. Antennas Propag. (2008).
[7] Zhao, Y. L. and Jiao, Y. C. and Zhao, G. and Zhang, L. Song, Y. and Wong, Z. B.:Compact planar monopole UWB antenna with band- notched characteristic. Vol. 50, pp. 2656-2658. Microwave Opt. Technol. Lett.(2008).
[8] Elboushi, A. Ahmed, O.M. and Sebak, A.R.:Study of elliptical slot UWB antennas with a 5.0-6.0GHz band-notch capability . Vol. 16, pp.207-222. Progress In Electromagnetics Research C.(2010).
[9] Lin, C. C. Jin, P. and Ziolkowski, R. W. :Single, dual and tri-bandnotched ultrawideband (UWB) antennas using capacitively loaded loop (CLL) resonators. Vol. 60, no. 1, pp. 102–109.IEEE Trans. Antennas Propag.(2012).
[10] Cho, Y-J. Kim, K.-H. Choi, D-H. Lee, S-S. and Park, S-O.: A miniature UWB planar monopole antenna with 5-GHz band-rejection filter and the time-domain characteristics. Vol. 54, no. 5, pp. 1453–1460. IEEE Trans. Antennas Propag. (2006).
[11] Sarkar, D. Srivastava, K.V. and Saurav, k.: A Compact Microstrip-Fed Triple Band-Notched UWB Monopole Antenna. Vol. 13, pp. 396-399. IEEE Antenna. Wireless Propag. Lett,(2014).
[12] Li, W-T. Shi, X-W. and Hei Y- Q.:Novel Planar UWB Monopole Antenna With Triple Band-Notched Characteristics. Vol. 8, pp .1094 – 1098. IEEE Antenna. Wireless Propag. Lett.(2009).
[13] Gao, G. P. Mei, Z. L. and Li, B. N.:Novel Circular Slot UWB Antenna With Dual Band-Notched characteristic. Vol. 15, pp .49-63. Progress In Electromagnetics Research C. (2010).
[14] Xu, K. Zhu, Z. Li, H. Huangfu, J. Li, C. and Ran, L. : A Printed Single-Layer UWB Monopole Antenna With Extended Ground Plane Stubs. Vol .12. IEEE Antenna. Wireless Propag. Lett (2013).

New Technologies of Power Transistors for Efficiency Increase of Power Converters: The Reliability Consideration

Pascal Dherbécourt[1], Ahmed El Oualkadi[2], Eric Joubert[1], Al Mehdi Bouchour[1,2], Wadia Jouha[2], Mohamed Masmoudi[1], Olivier Latry[1]
{pascal.dherbecourt@univ-rouen.fr, eloualkadi@gmail.com}

[1]Groupe de Physique des Matériaux UMR 6634, Normandie University, UNIROUEN, INSA Rouen, CNRS. Avenue de l'université B.P 12, Saint Etienne du Rouvray, France

[2]Laboratory of Information Technology and Communication, National School of Applied Sciences of Tangier. University Abdelmalek Essaadi, Tétouan, Maroc

Abstract. This paper proposes a methodology to study the reliability and failure analysis of new technologies of power transistors. The use of wide gap materials such as Silicon Carbide (SiC) and Gallium Nitride (GaN) is now a good alternative to meet the integration requirements of energy conversion systems. However, the reliability of these components is a crucial issue. Aging in operating conditions is considered, associated with electrical measurements to highlight nondestructive degradations of the performance of studied transistors. A step of looking for failure mechanisms in the material is made in order to identify any physical degradation. The information collected provide the user with valuable data and help to make an optimum choice of the component which can be integrated into the equipment. The presented work proposes the description of a methodology that meets these requirements and shows the study of two technologies of power transistors used in new generations of power converters.

Keywords: Electronic reliability, failure analysis, power transistor, ageing tests.

1 Introduction

The total market for Silicon Carbide (SiC) power components will account for more than one billion US dollars (US $ 1 trillion) by Yole Development organization in 2022 [1]. Applications such as PFC / power supplies and photovoltaic inverters are now contributing to the growth of this market. Photovoltaic applications seem to largely integrate SiC products. Indeed, SiC solutions offer a better system performance / cost ratio for photovoltaic inverters. In the future, applications related to electric vehicles, rail and others will also contribute to the evolution of the market.

In this context, the new power transistor technologies, now available on the market, offer increased performance compared to traditional silicon components, in terms of high voltage, high temperature and high switching frequencies.

However, these new wide gap components show random degradation problem, and reliability studies are still needed to qualify these components [2], [3], [4], [5]. The approach by long-term testing under operating conditions remains essential for integrators.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284211

The components under test are then aged thanks to dedicated benches where the electrical and environmental measurements are monitored. Degradations are observed and recorded, non-destructive and microstructural analyzes are made in order to identify the failure mechanisms responsible for performance degradations.

This set of techniques is implemented in the GPM laboratory, for applications of microwave amplifiers and switching power transistors, the work is carried out in collaboration with academic and industrial partners from the aeronautics, automotive and renewable energy sectors [6], [7], [8], [9].

In this paper, we present the general methodology developed in collaboration with the research team and partners meeting the defined objectives. Afterwards, we develop two examples of studies which are based on SiC MOSFET (Metal Oxide Semiconductor Field Effect Transistors) and GaN power transistors for switching applications.

2 Reliability and failure analysis study procedure of next-generation power transistors

The proposed methodology is summarized in **Figure 1**, which provides an overview of the various stages of the study of reliability and physical failure analysis of wide gap technology transistors. The studied components come mostly from the commercial sector and, thereby, little prior knowledge is available concerning the operation and physical structure of these devices. The first step, identified as step 1, consists in searching, from the manufacturer's documents or data sheets, the electrical, physical, and structural characteristics of the DUT.

From the data collected, failure indicators are defined, and electrical characterization based on measurements are implemented by means of the development of specific benches, this is step 2.

In step 3, the operational conditions are established, the component is aged on the dedicated bench, according to a particular mission profile, the electrical parameters are monitored, and measurements of failure indicators are regularly performed after each aging phase. A specific bench equipped with current measuring probe up to 30 A is used to record the I-V static characteristics of the transistor under test, a commercial source meter is used to measure both low gate and drain leakage currents.

According to one or more criterion of variations of the measured parameters, typically a variation of 20% of the nominal value, the aging phase is stopped, and the step 4 then consists in decapsulating the resin or ceramic packaging by laser ablation and chemical etching. This delicate operation, must be performed in order to maintain the integrity of the component. The active chip of the component is then accessible for a visual investigation, then by optical microscopy, the extrinsic elements such as metal access surfaces and bonds are controlled.

In order to precisely locate a possible defect, the PEM (Photon Emission Microscopy) and Optical Beam Induced Resistance Change (OBIRCH) techniques complete the analysis at this stage. Step 5 is implemented with a level of investigation depending on the requirements of the study. Scanning or transmission electron microscopy tools can be used, associated with a

FIB tool (Focus Ion Beam) to explore the internal structure of the component at the ultimate scale of a few tens of nanometers.

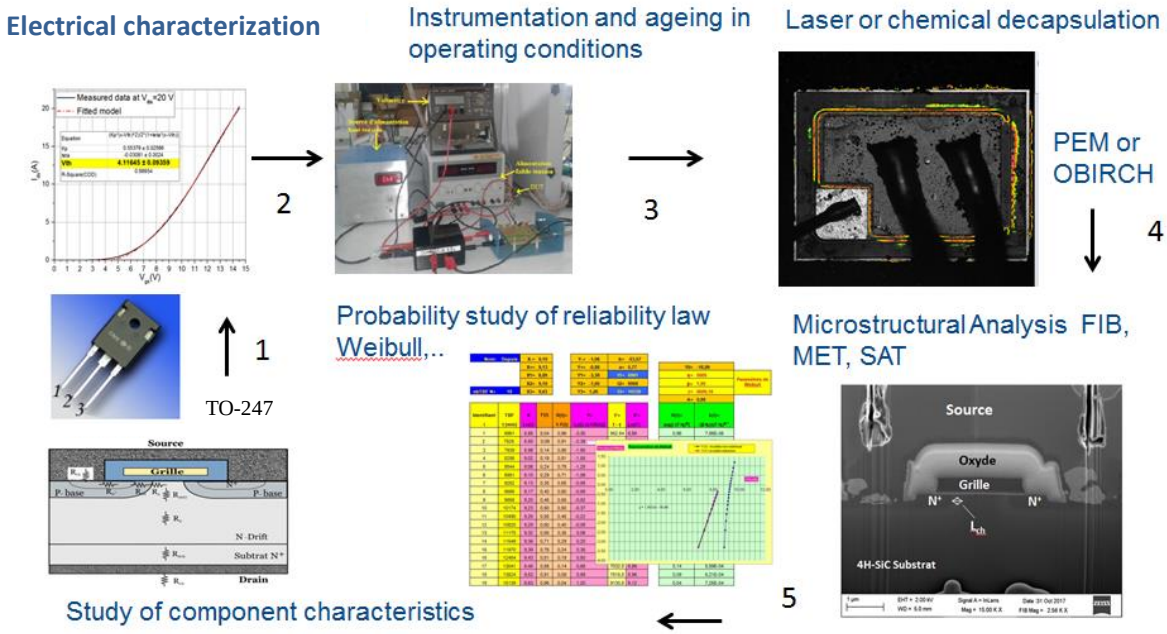


Fig. 1. Methodology to study reliability in operational conditions.

In order to meet the users requirements, study of reliability using Weibull law can be implemented in the methodology. The different parameters: Failure rate (Lambda), MTBF (Mean Time Between Failures) and MTTR (Mean Time To Repair) provide valuable information for system design and maintenance. The reference [10] describes reliability studies carried out on RFID tags, the performance of two tags of two manufacturers aged in temperature storage were compared.

These various steps make it possible to validate the study of the component in its mission profile, thus meeting the requirements of the user, but also to extract and analyze the physical defects at the origin of degradations of its characteristics. In the following paragraphs, we will develop examples of studies for new generations of SiC MOSFET and GaN power transistors, using the methodology described above.

3 Reliability and physical analysis of SiC power MOSFET failures for switching applications

Due to their interesting characteristics in thermal and in high level voltage domains, SiC MOSFETs are a promising technology. However, the feedback on their reliability is not as important as an older technology like Si, especially concerning the impact of ESD (Electro Static Discharge). This stress is important since the specific structure of SiC MOSFET may be sensitive to this kind of stress, and because electrostatic discharges are present at several places in the context of a power transistor.

The ESD appears when two objects with different electrostatic potentials get closer fast and generates a fast and important transfer of charges. This may also cause irreversible damage at manufacturing, handling or storage of an electronic device or component. The microelectronics industries are very concerned about this subject and are constantly looking for solutions, such as protection circuits, to control it and limit its consequences. An ESD test is also important to characterize the immunity of a transistor or a technology regarding this stress.

Several models of ESD have been developed and considered various contexts of stress. At first Human Body Model (HBM) simulates the electrical behavior of an operator standing in front of the table where the device is located and touching it with a finger [11]. This model is the most used since human is the main cause of ESD in industry. Usually, this simulation is achieved with an adapted couple resistor/capacitance. The Machine Model (MM) simulates the discharge produced by accumulation in a machine on a component [12]. This discharge is stronger than HBM since the electrical resistance of a machine is often very low. At least, Charged Device Model (CDM) considers a charge accumulated by the component itself during storage or manipulation confronted to a fixed potential of an external object (ground) [13].

For each discharge model, different tests protocols may be used in accordance with specific industrial standards like MIL-HDK, JEDEC, AEC, among others. For example, the ISO 10605:2008 describes tests to evaluate performances of electronic modules in road automotive context. This standard considers ESD appearing during assembly, due to maintenance staff or due to user. The tests in the ISO 10605 standard, which often involve an ESD gun comprising a R/C cell, use either a 330Ω or 2kΩ resistance, depending on the use of a metallic tools or not by user [14]. Several levels of severity are fixed by voltage from 2kV to 15kV.

When a MOSFET undergoes an ESD, a strong current is flowing during a few nanoseconds, which causes an important and local rise of temperature and various kinds of failures in the transistor structure [15], [16]:

- Filamentation: appearance of a conducting filament between source and drain
- Oxide breakdown: the gate oxide is destroyed by an overvoltage
- Hot carrier injection: nondestructive but modifies the isolating layers and the electrical characteristics of the transistor.
- Bonding breakdown: an overcurrent destroys the bondings by fusion.

The ESD stress implies also important consequences of the SiCMOSFET technology and test campaigns have to be realized to evaluate its impacts on a specific transistor family. The following parts present such a campaign on two generations of transistors of the same manufacturer, to understand evolution of ESD impact between them.

3.1 SiC MOSFETs under test

Among the leaders in SiC technology, the manufacturer Wolfspeed commercialized three generations of power transistors since 2011. Each generation improves the maximum drain current and the maximal case temperature, but valuable information is unavailable concerning ESD immunity under specific conditions.

Table 1. References of components involved in this study.

Generation	Ref.	I_{ds}(max)	V_{ds}(max)	Tc(max)
Gen2L	C2M0280120D	10 A	1200 V	150°C
Gen2H	C2M0160120D	19A	1200V	150°C
Gen3L	C3M0280090D	11,5 A	900V	150°C
Gen3H	C3M0120090D	23 A	900V	150°C

The following study involves the two last Wolfspeed generations. In table 1, letters L and H are indicating “Low” or “High” current transistors. In each generation, different references have various electrical characteristics and different power application. All these transistors are based on a VDMOS structure which is shown in **Figure 2**.

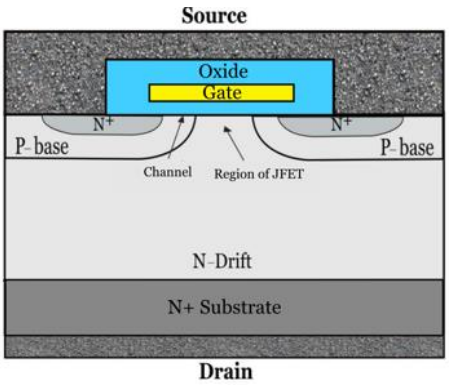


Fig. 2. Schematic structure of Wolfspeed transistors

3.2 ESD Tests

The ESD tests were realized on 8 components of 4 categories (2 of each) of Table.1 respecting ISO 10605 standard. An electrostatic gun shown in **Figure 3** was used with a couple 2kΩ/330pF simulating an HBM discharge model.

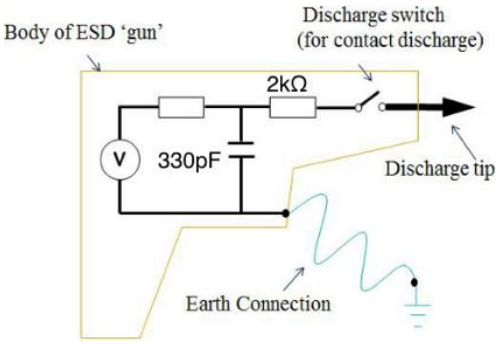


Fig.3. Electrostatic Gun schematic

Stress of 8kV and 15 kV were applied repeatedly (4 times and up to 12 times on some transistors) on gates pin to evaluate gate oxide robustness and to appreciate evolutions between generations. Drains and sources pin were leaved open.

3.3 Results and analysis

Electrical measurements have been achieved before and after stress in order to estimate discharge impact on the behavior of the transistors. This comparison yields interesting information concerning the failures that may appear in the devices.

At first, saturation current I_{dsat} and on-state resistance $R_{ds(on)}$ are determined from the output characteristics ($I_{ds} = f(V_{ds})$) of the transistors. These parameters show few variations due to the stress ($R_{ds(on)}$ decreases and I_{dsat} increases, about 1% for both) on two generations except for a transistor Gen3L that received a 15kV stress where I_{ds} decreases and $R_{ds(on)}$ increases about 10% for both.

Leakage currents I_{gss} and I_{dss} were measured on a specific bench and schematic. I_{gss} increases a lot for both generations (average values about 3000% for Gen2 and 10000% for Gen3). I_{dss} increases about 100% for Gen2L and Gen3L but doesnot change significantly for Gen2H and Gen3H.

Threshold voltage V_{th}, transconductance K_p and transverse electrical field θ were determined from output characteristics with a Levenberg-Marquardt Algorithm **[9]**. The stress induces an average decrease of 2% for V_{th} on both generations, with highest values on Gen3L transistors. K_p and θ increase slightly for both generations.

At least, C-V measurements on gate-source, drain-source and drain-gate diode were performed on a specific bench, to get more information on failure process during the stress. The measured capacitance variations globally reveal the appearance of positive traps in the oxide volume. The effect of these traps is in good accordance with the decrease of V_{th} and the increase of I_{gss}.

All the variations observed yields to a conclusion that the ESD stress mainly affects the oxide volume near the gate of the transistors and also the input parameters of the transistor.

High current transistors (Gen2H and Gen3H) are less sensitive to this stress and no significant difference seems to appear between the two generations.

4 Switching GaN transistor for energy conversion applications.

The GaN technology is another alternative to SiC for increasing the integration of energy conversion devices. Thanks to these new devices, the automobile market anticipates an integration density multiplied by 3, with a loss reduction of 50% and a cost reduction of 20% [17]. The advantage of the GaN technology compared to silicon substrate results from the ability to operate at: high frequency due to increased velocity saturation, high temperature and high voltage due to the large band gap material [18], [19]. In this paper, we use anew commercial component from GaN systems manufacturer, whose reference is GS66508P, it is a p-type gate normally-off which operates in the 650V/30A range, its main characteristics are given in reference [20].

4.1 Aging tests with switching mode

The successive aging steps and the dynamic measurements are carried out by means of a dedicated switching bench described in reference [9]. **Figure 4a.** shows a view of this setup. This includes two transistors noted DUT_C and DUT_V switching the DC source voltage to fix the desired electrical parameters in two states. The DUT_M is alternately subjected to a high current by the mean of DUT_C, then, out of conduction, subjected to an electric field by DUT_M.

The aging bench eliminates the need for significant energy consumption because, with this original operating principle, no power load is needed. To ensure this operation, the PWM_1 and PWM_2 signals are two complementary pulses whose pulse width can be modulated. The bench is instrumented using a digital oscilloscope connected to measurement probes.

Also, it places similar electrical waveforms on DUT_M, which could exist in a large class of power management products such as DC-DC power converters. Moreover, it enables dynamic measurements, like: dynamic resistance (R_{dyn}) and switching times (tr and tf).

The DUT_M current and voltage experiment waveforms are shown in **Figure. 4b**. As can be observed, V_{DS} switch from 1.14V to 24V with a rise time equals to 53.46 ns and I_{DS} switch from 0A to 14A with a rise time equals to 3.94 µA. The studied switching conditions are chosen so that the DUT_M switches in its Safe Operating Area (SSOA).

4.2 Degradation detection and measurements caused by the aging phase

The studied component is aged in operating conditions with the bench described in **Figure 4a**. At regular time intervals, measurements are made to assess the evolution of some relevant and significant electrical parameters. To meet this goal, the threshold voltage V_{th} is extracted using the Extrapolation in the Linear Region (ELR) method. The transconductance is defined as the maximum first derivative of the input characteristic in the saturation region, and

the on-state $R_{DS(ON)}$ is defined as the inverse of the slope of output characteristics in the linear region.

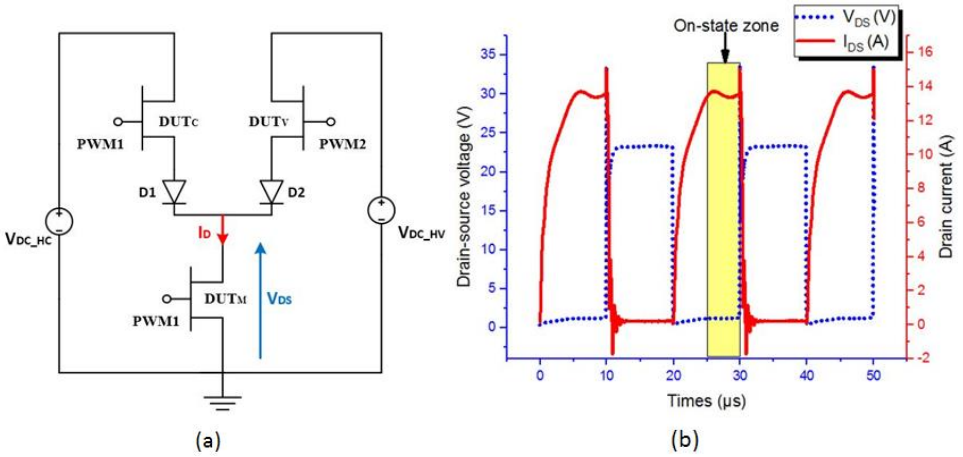


Fig. 4. (a) Setup for ageing the device in operational conditions. (b) Chronograms for I_{DS} and V_{DS} measurements.

The switching frequency is set at 50 kHz with a duty cycle of the control signal equals to 50%. The study component is thus aged at the limited Safe Operating Area (SSOA) for more than 1000 hours in operating conditions on the dedicated bench.

To ensure that the aging of the DUT_M respects the SSOA, the measurements of the junction temperature T_J and the dynamic $R_{DS(ON)}$ are performed, the temperature T_J is estimated using the thermal model and thermal Ohm's law of DUT_M package mounted on a PCB and a heat sink.

From experimental measurements chronograms, the power loss (P_{loss}) is calculated which is equal to 16.68 W, the corresponded T_J is calculated equal to 86 °C which is blow the allowable T_J=150°C from the data sheet.

Table 2 summarize the values of $R_{DS(ON)}$, g_m and I_D before and after 1008 hours of aging DUT_M under the switching conditions.

Table 2. 1008 hours aging results for the DUT.

	$R_{DS(ON)}$ mOhms	g_m(S)	I_{DS}(A)
t = 0	43.04	24.88	17.01
T=1008 hours	54.71	22.42	13.21
Degradation (%)	27.11	9.88	22.33

Drain current measurements, after each step of aging, highlight the effect of aging of the DUT operating at the limit of the SSOA. The effects are a decrease in the efficiency of DC-DC power converters, due to an increase of the power losses of the DUT_M.

Based on measurements of the aged component, considering the model provided by the constructor lacking precision, an improved SPICE model is defined taking into account the degraded performance of the component. Also, we studied the effect of aging GaN HEMT under SSOA on the efficiency of a DC-DC power converter by a simulation approach. A 12/24 V buck converter is designed which has an output power equals to 180 W in a load.

The purpose is to model the aging of the DUT_M in order to evaluate the effects of aging the GaN HEMT under SSOA on the power efficiency of the DC-DC boost converters by a SPICE simulation approach. By applying this methodology, an average decrease of the efficiency of a DC-DC boost converter equals to 4.55% after aging state for different output current. The decrease in efficiency is accompanied with an increase of conduction power loss in the DUT_M after aging, the average increase in conduction power losses equals to8.55%.

5 Conclusion

The SiC and GaN technologies allow to overcome the limits in the use of energy conversion systems by providing greater energy efficiency, including increased integration and higher switching frequencies.

The challenges are to reduce the size of the power converters and to increase energy efficiency while reducing costs. In the field of transport, the automotive sector, is addressed with the development of hybrid and electric vehicles. The aircraft industry is driven by the demand to optimize aircraft performance, the decrease of operating and maintenance costs and the reduction of gas emissions, are the most demanding sectors of energy efficiency. The field of energy conversion brings the development of the new technology for semiconductor market.

The SiC MOSFET 1200V and the GaN 650 V transistors presented in this paper are two new generation components available on the market. They are accessible to integrators and already offer an alternative to the requirements of integration of energy conversion systems. However, they continue to be the subject of reliability and failure analysis studies, in order to improve their performance and to qualify them in operational conditions for various application sectors.

Acknowledgments. This project is co-funded by European Union and Region Normandy. Europe is involved in Normandy through the European Funds for Regional Development.

References

[1] Yole Développement. Power SiC : Materials, Devices, Modules and Applications. Technical report, August 2017.

[2] A. J. Lelis, R. Green, D. B. Habersat, and M. El. : Basic mechanisms of threshold voltage instability and implications for reliability testing of sic mosfets. IEEE Transactions on Electron Devices, 62(2) : pp. 316-323, Feb 2015.

[3] V. Mulpuri and S. Choi.: Reliability of sic power mosfets under high repetitive pulse current conditions. IEEE Applied Power Electronics Conference and Exposition (APEC), pages 776-782, March 2017.

[4] Military Standard. Mil-std-750d, test methods for semiconductor devices. Technical report, 1995.

[5] Standard JEDEC. Jesd22a-108 temperature, bias, and operating life.Technical report, 2009.

[6] Pascal Dherbécourt, Olivier Latry, Karine Dehais-Mourgues, Jean-Baptiste Fonder, CédricDuperrier, Farid Temcamani, HichameMaanane et Jean-Pierre Sipma : Banc de vieillissement de transistors en conditions opérationnelles, Les systèmes mécatroniques embarqués : analyse des causes de défaillances, modélisation, simulation et optimisation, tome 2 Ed ; ISTE (ISBN : 978-1-78405-059-7) Chap 2, pp. 45-70, Mars 2015.
[7] P. Denis, P. Dherbécourt, O. Latry, C. Genevois, F. Cuvilly, M. Brault, and M. Kadi: Robustness of 4H-SiC 1200V schottky diodes under high electrostatic discharge like human body model stresses: An in-depth failure analysis. Diamond and Related Materials, 44:pp62–70, April 2014
[8] WadiaJouha, Ahmed El Oualkadi, Pascal Dherbécourt, Eric Joubert and Mohamed Masmoudi :Silicon Carbide Power MOSFET Model: An Accurate Parameter Extraction Method Based on Levenberg-Marquardt Algorithm», IEEE Transactions on Power electronics, novembre 2018,Volume: 33, Issue:11, pp: 9130-9133
[9] Al Mehdi Bouchour, Andres Echeverri, Olivier Latry, Pascal Dherbécourt, Ahmed El Oualkadi: Modeling of power GaN HEMT for switching circuits applications using Levenberg-Marquardt Algorithm, IEEE International Symposium on Advanced Electrical and Communication Technologies, ISAECT18, 21-23 november 2018, Rabat-Kenitra, Morocco.
[10] SanaeTaoufik, Pascal Dherbécourt, Ahmed El Oualkadi, Member, IEEE, and FaridTemcamani: Reliability and Failure Analysis of UHF RFID Passive Tags under Thermal Storage,IEEE Transactions on Device and Materials Reliability. Volume: 17, Issue: 3, pp. 531-538, September 2017.
[11] ESD STM5.1-2001. Standard Test Method for Electrostatic discharge Sensitivity Testing-Human Body Model (HBM) Component Level. ESD Association WG 5.1, 2001.
[12] ESD STM5.2-1999. Standard Test Method for Electrostatic discharge Sensitivity Testing-Machine Model. ESD Association WG 5.2, 1999.
[13] ESD STM5.3.1-1999. Standard Test Method for Electrostatic discharge Sensitivity Testing-Charged Device Model (CDM) Component Level. ESD Association, 1999.
[14] ISO 10605. "véhicule routiers : méthodes d'essai des perturbations électriques provenant de décharges électrostatiques". https://www.iso.org, 2001. Accessed : 2018-05-18.
[15] V.A. Vashchenko, Y.B. Martynov, and V.F. Sinkevitch.Electrical instability and filamentation in ggmos protection structures. Solid-State Electronics, 41(11) :1761 - 1767, 1997.
[16] M. Sachdev O. Semenov, H. Sarbishaei. ESD Protection Device and Circuit Design for Advanced CMOS Technologies.Springer, Dordrecht, 2008.
[17]Automotive Market Insight, GaN Systems Inc, 2018.
[18]Mitova, Radoslava, et al. "Investigations of 600-V GaN HEMT and GaN diode for power converter applications." IEEE transactions on power electronics 29.5 (2014): 2441-2452.
[19] Lu, Bin, and Tomas Palacios. "High Breakdown 1500 V AlGaN/GaN HEMTs by Substrate-Transfer Technology." IEEE Electron Device Letters 31.9 (2010): 951-953.
[20] GS66508P Preliminary Datasheet,GaN_Systems_Inc. 2018

Temperature Effects of GaN HEMTs on the Design of Power Converters

Al Mehdi Bouchour[1,2], Pascal Dherbécourt[1], Olivier Latry[1], Ahmed El Oualkadi[2]
{al-mehdi.bouchour@univ-rouen.fr, pascal.dherbecourt@univ-rouen.fr, OLivier.latry@univ-rouen.fr}

[1]GPM UMR CNRS 6634 Normandie Univ, INSA Rouen, CNRS, Rouen, France
[2]Laboratory of Information and Communication Technologies, National School of Applied Sciences of Tangier, Abdelmalek Essaadi University, Tanger, Morocco

Abstract. This paper proposes an experimental study of temperature effects on Gallium Nitride (GaN) High Electron Mobility Transistors (HEMTs). The output and transfer characteristics are monitored at temperatures ranging from 5°C to 105°C. The temperature dependency on static parameters of GaN HEMT is examined, such as: drain current (I_{DS}), on-state resistance ($R_{DS(ON)}$), transconductance (g_m), threshold voltage (V_{TH}) and the gate leakage current (I_{GSS}). The decreases of I_{DS} and g_m accompanied with the increase of $R_{DS(ON)}$ and I_{GSS} when increasing temperature have been observed. Moreover, the decrease in electron mobility with increasing temperatures is considered to be one of the causes of the reduction in the drain current and transconductance. In order to study the impact of temperature on power converters with GaN HEMTs by simulation approach, the thermal characteristics of a 650V, 30A GaN HEMT have been modelled. The used model is a non-segmented Electro-thermal SPICE model of Motorola. The model parameters are extracted using Levenberg-Marquardt Algorithm.

Keywords: characterization; modeling; temperature; GaN; HEMT; Levenberg-Marquardt.

1 Introduction

The global renewable energy consumption is estimated to increase by 78% within 2040 [1]. The use of Gallium Nitride (GaN) power transistors in the renewable energy industry, enable 20% power density increase, 50% power loss reduction, 10% lower system cost, and more efficient energy storage systems, along with bi-directional power flow architectures that are not achievable with silicon (Si) [2].

Furthermore, GaN HEMTs power transistors can operate at higher power levels, frequencies and temperatures, and with an improved energy efficiency with respect to that guaranteed by Si devices [3]. These excellent features are due to the excellent physical properties of GaN semiconductor [4]. In fact, they can allow a reduction of the on-state resistance and parasitic capacitances with respect to Si devices, with an overall reduction of the power losses [5]. The above advantages of the GaN semiconductor make GaN power transistors attractive for renewable energy domain in particular for use in photovoltaic and wind energy systems [6].

In this context, DC-DC power converters are the most inevitable electronic units in power electronic systems, by providing the control and management of the electric power [7]. The investigation of the temperature dependency of GaN HEMT power transistors is vital to the use

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284227

of the GaN HEMTs in power converter applications. Such an investigation is important particularly to understand the cause of deterioration of their electrical performance at elevated temperatures.

In the literature, several studies on the temperature dependency of GaN HEMTs were performed. References [8] and [9] indicate the evolution of GaN HEMTs electrical parameters during thermal stresses depending on the different temperature ranges.

In this paper, current–voltage characteristics of a GaN HEMT power transistor at operating temperatures are studied experimentally. Moreover, the physics underlying various high-temperature operations of current–voltage characteristics is discussed. Finally, the effect of temperature on power converters with GaN HEMT is verified by a simulation approach.

2 Static characterization

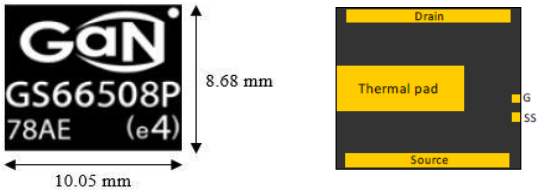
The tested device is a GS66508P from GaN Systems [10]. It is a p-type gate normally-off AlGaN/GaN power transistor which operates in the range of 650V/30A. **Figure 1** provides a bottom and a top view of the tested GaN transistor package. The device package allows to study the temperature dependency of the GaN HEMT characteristics from low temperature (5˚C) to high temperature (105˚C). **Figure 2** illustrates the p-AlGaN gate formed over the undoped AlGaN/GaN heterostructure [11]. The p-AlGaN lifts up the potential at the channel, which enables normally-off operation.

The static characterization is performed using AMCAD pulsed I-V system. The pulses setting is fixed at a pulse frequency of 100 Hz with a gate pulse duty cycle of 0.04%. The pulse width of 4 µs is short enough to ensure iso-thermal measurement of pulsed I-V GaN HEMT characteristics. The used I-V pulsed probes have a high measurement accuracy, which is equal to ±100mA.

The tested GaN HEMT (DUT) is characterized at temperature ranging from 5 ˚C to 105 ˚C in order to study the effect of temperature on GaN power transistors. The temperature is fixed by a Peltier heater plate. The I-V pulsed bench with associated instruments are shown in **Figure 3**.

The measured output characteristic curves for a gate-to-source voltage (V_{GS}) equals to 6V under operating temperatures equal to 5°C, 25°C, 75°C and 105 °C are shown in **Figure 4 (a)**. As observed, the drain current I_D is modulated by the voltage V_{GS}. The transfer characteristic is obtained for a drain-source voltage (V_{DS}) equals to 2V and for various temperatures equal to 5°C, 25°C, 75°C and 105 °C see **Figure 4 (b)**. The tested device has the advantage of safe normally-off operation for temperature up to 105˚C.

According to the output characteristic in **Figure 4 (a)**, the evolution of the following electrical parameters are extracted: the on-state resistance, the threshold voltage. From the transfer characteristic in **Figure 4 (b)**, the evolution of the transconductance g_m versus temperature is determined.



(a) (b)

Fig. 1. GS66508P package [10]: (a) Top view; (b) bottom view

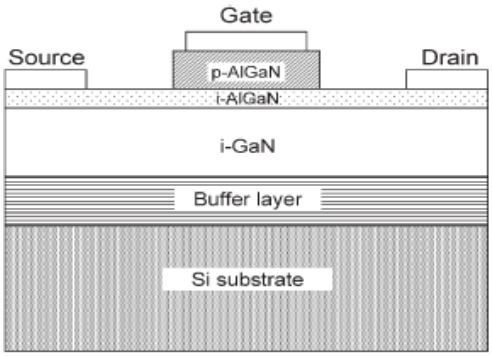


Fig. 2. Schematic illustration of the tested GaN HEMT power transistor structure [11]

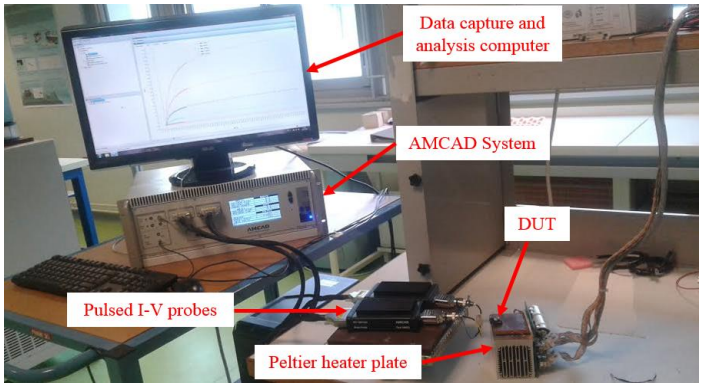


Fig. 3. I-V pulsed bench with associated instruments

2.1 On-state resistance ($R_{DS(ON)}$)

The R_{DSON} is defined as the inverse of the slope of output characteristics in the linear region and it calculated by the following equation:

$$R_{DS(ON)} = \left(\frac{\Delta I_D}{\Delta V_{DS}} \right)^{-1}_{V_{GS}=6V, V_{DS}\to 0V} . \quad (1)$$

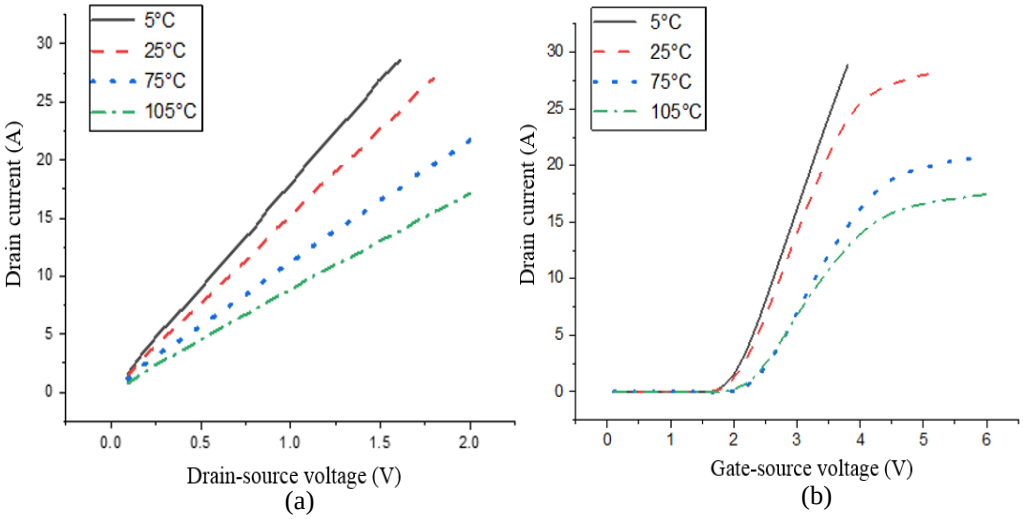


Fig. 4. (a) Output characteristic of GaN HEMT at V_{GS} = 6V; (b) Transfer characteristic of GaN HEMT at V_{DS} = 2V for various temperatures: 5°C, 25°C, 75°C and 105 °C.

The measurement conditions of normalized $R_{DS(ON)}$ are: V_{GS} = 6V and varying temperature from 5°C to 105°C. **Figure 5** shows the variation of the normalized $R_{DS(ON)}$ versus temperature. An increasing of 180% in on-state resistance at 105°C can be observed compared to its value at 5°C.

The temperature dependency of $R_{DS(ON)}$ is due to the impact of temperature on the channel between the source and gate contacts [12]. In fact, if the device temperature is raised, the low-field electron mobility in the channel is decreased [13], and a proportional increase of the channel resistance related to the drain-source resistance (R_{DS}) can be expected. This result is also confirmed in [14] which is suggested to use GaN HEMT transistors with higher 2-D electron gas (2DEG) channel density for enhanced performances.

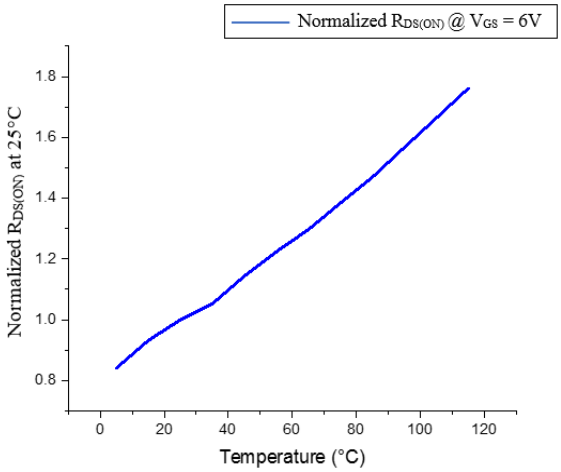


Fig. 5. Normalized $R_{DS(ON)}$ as a function of temperature at V_{GS} = 6V

2.2 Threshold voltage (V_{TH})

The V_{TH} is extracted using Extrapolation in the Linear Region (ELR) method [15]. The temperature dependency of the extracted threshold voltage is given in **Figure 6 (a)**. As shown in this figure, the threshold voltage of GaN HEMT at 105 °C decreases about 12% compared to their values at 5°C.

The dependency of V_{TH} on temperature is due to holes injected from the metal to the p-GaN layer. The injected holes are accumulated at the p-GaN/AlGaN interface. This lead to a negative shift in threshold voltage, which can also be observed in [16]. Recent studies [17] shows that using hydrogen plasma treatment instead of etching technology may compensate holes in the p-GaN layer above the two dimensional electron gas (2DEG) channel to release electrons in the 2DEG channel and form high resistivity area to reduce leakage current and increase threshold voltage stability.

2.3 Transconductance (g_m)

The g_m quantifies the drain current variation with a gate-source voltage variation while keeping the drain-source voltage constant (bias voltage). It is defined by the following equation, while V_{DS} is the bias voltage.

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}=2V} . \tag{2}$$

Figure 6 (b) shows the variations of the transconductance g_m as a function of temperature. As shown in this figure, the transconductance g_m at 105 °C decreases about 58.38% compared to their values at 5°C. The decreasing of g_m with the increase of temperature is due to the decrease of both the electron mobility in the channel and the electron velocity [18]. The decrease of the transconductance when increasing temperature is observed in [19] which is confirmed that the channel mobility decreases with the increase of temperature [20].

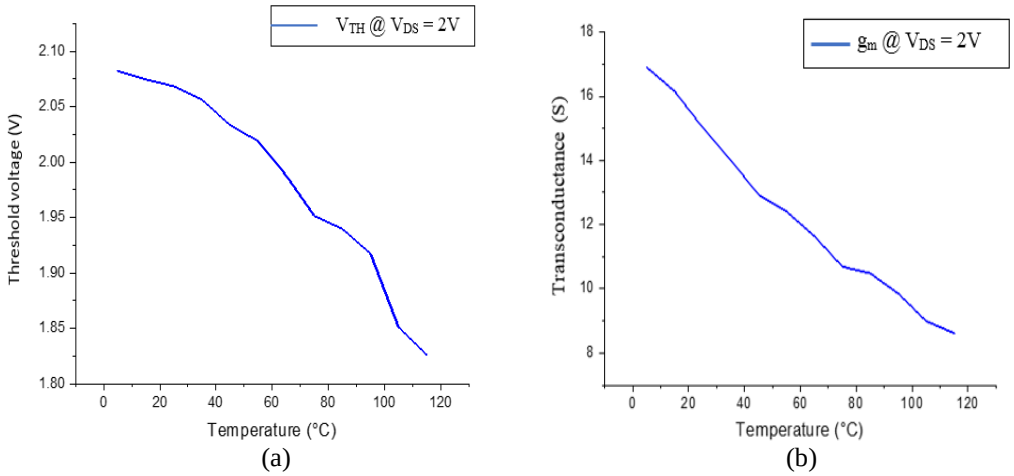


Fig. 6. (a) Threshold voltage; (b) Transconductance as a function of temperature at V_{DS} = 2V

2.4 Gate leakage current

The gate leakage current is defined as the leakage that occurs when the specified voltage is applied across the gate and source with drain and source short-circuited. The leakage current measurements were performed by using Keithly 2636B SourceMeter.

Figure 7 shows the evolution of the gate leakage current as a function of temperature. As observed, the current I_{GSS} is increased by 77.77% when varying temperature from 5°C to 105°C. It was demonstrated in reference [21] that increasing temperature may reduce the schottky barrier height at the metal/p-GaN contact and, consequently, increase the gate leakage current. Moreover, in p-GaN/AlGaN/ GaN heterostructure, the reduction of the Schottky barrier height at the metal/p-GaN gate enhances the tunneling of holes through the barrier [22]. Then, as highlighted by Hwang et al. [23], the enhancement of holes injection leads not only to a negative shift of the V_{TH} but also to an increase in the leakage current in p-GaN HEMT. The increase in the leakage gate current when increasing temperature is also observed in [24], which is attributed to the lowering of the Schottky barrier height of the metal/p-GaN gate.

3 Modeling temperature dependency

The purpose of this part is to model the temperature dependency of the GaN HEMT in order to evaluate the effects of GaN HEMT on the power efficiency of the DC-DC converters by a SPICE simulation approach. **Figure 8** shows the studied DC-DC boost converter, which is a 12/24 V boost converter with an output power equals to 180 W. The tested device (DUT) is the SPICE Electro-thermal model of the GaN HEMT.

The proposed drain current model of the DUT is a non-segmented, smooth and continuous equation inspired from the Motorola Electro-thermal Model (MET) developed by Curtice et al. [25] and described in [26].

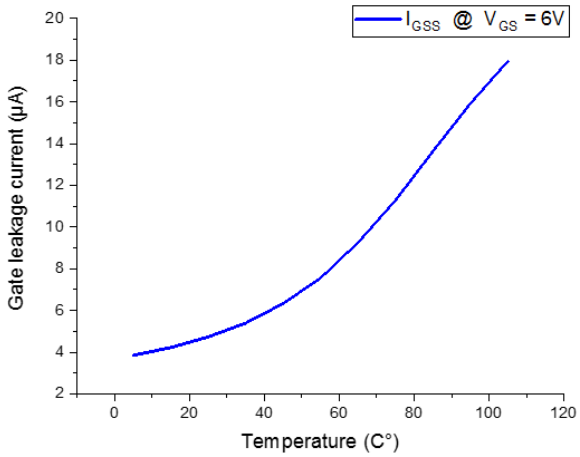


Fig. 7. Gate leakage current as a function of temperature at V_{GS} = 6V

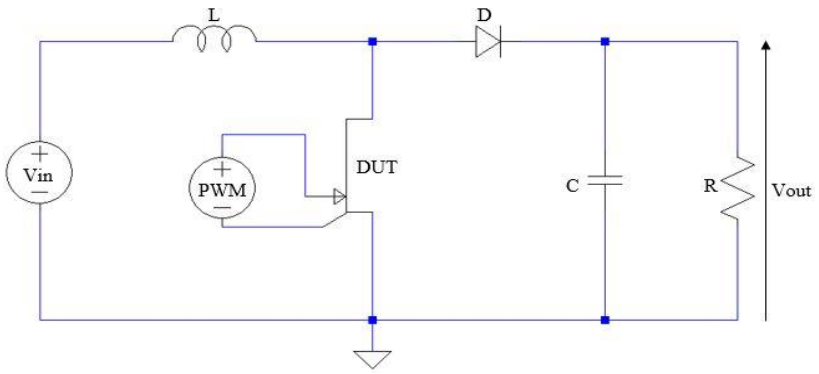


Fig. 8. DC-DC boost converter with L = 21.41 µH, Vin = 12V, Vout = 24V, C = 90.9 µF and R = 1.71 Ω. The switching conditions of the DUT transistor are: f = 50 kHz and duty-cycle = 50 %.

The proposed drain current equation is shown as follows:

$$\begin{cases} I_{DS} = K(T)\cdot\log\left[1+\exp\left(\dfrac{V_{GS}-b}{c}\right)\right]\cdot\dfrac{(m+n\cdot V_{GS})V_{DS}}{1+P(T)\cdot(d+e\cdot V_{GS})V_{DS}} \;,\; V_{DS}\geq 0 \\ K(T) = K.\left[1+T_{C1}.(T-25)+T_{C2}\cdot(T-25)^2\right] \\ P(T) = P.\left[1+T_{C3}.(T-25)+T_{C4}\cdot(T-25)^2\right] \end{cases} . \quad (3)$$

Where K is the device forward transconductance ($A.V^{-1}$) parameter at 25°C, P is the output conductance (V^{-1}) at 25°C, b and c are related parameters of the transfer characteristic, while m, n, d and e are related parameters of the output characteristic. Tc1, T_{C2}, T_{C3} and Tc4 are temperature coefficients. These parameters are extracted by using the Levenberg-Marquardt algorithm [27] as shown in Table 1.

The conduction power loss (P_C) is defined by:

$$P_c = \frac{1}{T_{on}}\int_0^{T_{ON}} i_{ds}\cdot v_{ds}\cdot dt \;. \quad (4)$$

Where, i_{ds} is instantaneous drain current, v_{ds} is the instantaneous drain voltage and T_{on} is the conduction time which is given by:

$$T_{on} = \frac{d_c}{f} \;. \quad (5)$$

Where, d_c is the duty cycle and f is the frequency.

Figure 9 shows the conduction power losses of the DUT used in the DC-DC converter of **Figure 8** at various junction temperatures: 25°C, 70°C and 105°C. From **Figure 9**, an increase in conduction power losses is observed, which is mainly due to the increase of the dynamic on-state resistance of the DUT when increasing temperature.

Table 1. Extracted Electrothermal SPICE model parameters

Parameter	K	b	c	m	n	P
Value	1.42	1.65	0.14	8.12	-1.15	0.75
Parameter	d	e	T_{C1}	T_{C2}	T_{C3}	T_{C4}
Value	1.94	-0.35	-8.02E-3	3.22E-5	-3.75E-3	4.01E-6

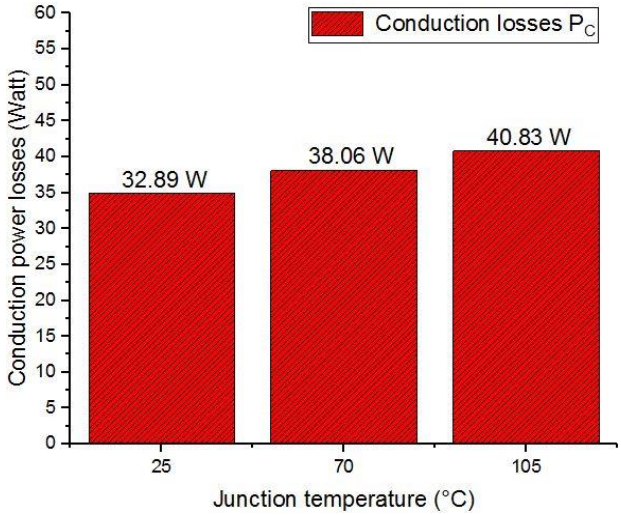


Fig. 9. Conduction power losses of GaN HEMT at T_J = 25°C, 70°C and 105°C

4 Conclusion

In conclusion, the temperature dependency of GaN HEMT static characteristics has been studied based on experimental measurements. The decreases in $R_{DS(ON)}$, I_{DS} and g_m with the increase of temperature are due to the decrease in electron mobility. The increase in I_{GSS} and the negative shift of V_{TH} when increasing temperature is attributed to the lowering of the Schottky barrier height of the metal/p-GaN gate and to hole injection at the metal/p-GaN. It is suggested to use GaN HEMT transistors with hydrogen plasma treatment instead of etching technology to reduce leakage current and increase threshold voltage stability. Upon cooling after high temperature characterization at 105 °C, GaN HEMT transistors showed recovered output and transfer characteristics. The effect of increasing the operational temperature of the GaN HEMT transistors on a DC power converter is verified by modeling the thermal characteristics of the GaN HEMT. Such an investigation is important particularly to understand the cause of deterioration of GaN HEMT electrical performances at high temperatures. The temperature dependency of GaN HEMT dynamic characteristics will be studied in the next work in order to take into account the switching losses. This enable an accurate design of power converters with GaN HEMTs for renewable energy applications by a simulation approach.

Acknowledgments

This work is financially supported by the PHC Toubkal project (code TBK/17/41).

References

[1] Global EV Outlook 2017, International Energy Agency, 2017.
[2] Renewable Energy Market Insight, GaN Systems Inc, 2018.
[3] Chen, Feida, Xiaofeng Ding, and Ehtisham Lodhi. "Efficiency and Current Harmonics Comparison Between SiC and Si Based Inverters for Microgrids." Energy Procedia 103 (2016): 298-303.
[4] Bhat, Ishwara. "Physical properties of gallium nitride and related III–V nitrides." Wide Bandgap Semiconductor Power Devices. Woodhead Publishing, 2019. 43-77.
[5] Perkins, S., et al. "A comprehensive comparison of the static performance of commercial GaN-on-Si devices." Wide Bandgap Power Devices and Applications (WiPDA), 2017 IEEE 5th Workshop on. IEEE, 2017.
[6] Kaplar, R. J., et al. "Characterization and reliability of SiC-and GaN-based power transistors for renewable energy applications." Energytech, 2012 IEEE. IEEE, 2012.
[7] Ioinovici, Adrian. Power Electronics and Energy Conversion Systems: Fundamentals and Hard-switching Converters. Volume 1. Wiley, 2013.
[8] Chihani, Omar, et al. "Temperature and voltage effects on HTRB and HTGB stresses for AlGaN/GaN HEMTs." 2018 IEEE International Reliability Physics Symposium (IRPS). IEEE, 2018.
[9] Chihani, Omar, et al. "Effect of HTRB lifetest on AlGaN/GaN HEMTs under different voltages and temperatures stresses." Microelectronics Reliability 88 (2018): 402-405.
[10] GS66508P Preliminary Datasheet, GaN_Systems_Inc. 2018.
[11] Uemoto, Yasuhiro, et al. "Gate injection transistor (GIT)—A normally-off AlGaN/GaN power transistor using conductivity modulation." IEEE Transactions on Electron Devices 54.12 (2007): 3393-3399.
[12] Kuzmik, J., et al. "Determination of channel temperature in AlGaN/GaN HEMTs grown on sapphire and silicon substrates using DC characterization method." IEEE transactions on electron devices 49.8 (2002): 1496-1498.
[13] J. D. Albrecht, R. P. Wang, R. P. Ruden, M. Farahmand, and K. F. Brennan, “Electron transport characteristics of GaN for high temperature device modeling,” J. Appl. Phys., vol. 83, pp. 4777–4781, 1998.
[14] Wong, King-Yuen, Wanjun Chen, and Kevin J. Chen. "Characterization and analysis of the temperature-dependent on-resistance in AlGaN/GaN lateral field-effect rectifiers." IEEE Transactions on Electron Devices 57.8 (2010): 1924-1929.
[15] Ortiz-Conde, Adelmo, et al. "A review of recent MOSFET threshold voltage extraction methods." Microelectronics Reliability”, 42.4-5 (2002): 583-596.
[16] Matteo Meneghini, Oliver Hilt, Joachim Wuerfl and Gaudenzio Meneghesso, “Technology and Reliability of Normally-Off GaN HEMTs with p-Type Gate “, mdpi,energie 2017, pp. 15-17, 2017
[17] Hao, Ronghui, et al. "Normally-off p-GaN/AlGaN/GaN high electron mobility transistors using hydrogen plasma treatment." Applied Physics Letters 109.15 (2016): 152106.
[18] Arulkumaran, S., et al. "High-temperature effects of AlGaN/GaN high-electron-mobility transistors on sapphire and semi-insulating SiC substrates." Applied physics letters 80.12 (2002): 2186-218
[19] Kühn, Jutta. AlGaN-GaN-HEMT Power Amplifiers With Optimized Power-Added Efficiency for X-Band Applications. Vol. 62. KIT Scientific Publishing, 2011.
[20] Maeda, Narihiko, et al. "High-temperature electron transport properties in AlGaN/GaN heterostructures." Applied Physics Letters 79.11 (2001): 1634-1636.
[21] Aktas, Ozgur, et al. "High temperature characteristics of AlGaN/GaN modulation doped field-effect transistors." Applied Physics Letters 69.25 (1996): 3872-3874.
[22] Greco, Giuseppe, et al. "Effects of annealing treatments on the properties of Al/Ti/p-GaN interfaces for normally off p-GaN HEMTs." IEEE Transactions on Electron Devices 63.7 (2016): 2735-2741.

[23] Giuseppe Greco, Ferdinando Iucolano, Fabrizio Roccaforte, "Review of technology for normally-off HEMTs with p-GaN gate", Materials Science in Semiconductor Processing, pp. , 2017, ISSN 13698001.
[24] I. Hwang et al., “p-GaN gate HEMTs with tungsten gate metal for high threshold voltage and low gate current,” IEEE Electron Device Lett., vol. 34, no. 2, pp. 202–204, Feb. 2013
[25] W. Curtice et al., “A New Dynamic Electro-Thermal Nonlinear Model for Silicon RF LDMOS FETs,” IEEE MTT-S International Microwave Symposium Digest, 1999, p. 419.
[26] Bouchour, Al Mehdi, et al. “Modeling of power GaN HEMT for switching circuits applications using Levenberg-Marquardt Algorithm,” In : International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 2018 International Conference on. IEEE, 2018.
[27] Jouha, Wadia, et al. "Silicon Carbide Power MOSFET Model: An Accurate Parameter Extraction Method Based on Levenberg- Marquardt Algorithm." IEEE Transactions on Power Electronics (2018).

Square Complementary Split Ring Resonator (CSRR) Low Pass Filter

Mohammed Bendaoued[1], Rachid Mandry[1], Larbi El Abdellaoui[1], Aytouna Fouad[2] and Mohamed Latrach[3]

{ mohammed.bendoued@gmail.com }

[1] FST, Hassan II, University, Settat, Morocco
[2] ENSA, Abdelmalek Essaadi, University, Tetouan, 93000, Morocco
[3] RF& Microwave Group, ESEO, Angers, France

Abstract. A novel CPW Low pass filter structure using square complementary split ring resonator (CSRRs) is presented in this work. The CSRRs are etched periodically in each ground plane of the CPW line which permit to suppress the spurious response. The low pass filter shows a good rejection in the stop band. The cut-off frequency at -3dB is equal to f_c = 5.28GHz. This filter has total area of 35.48x21.16 mm^2.

Keywords: CPW, LPF, CSRR, Resonator, Metamaterial.

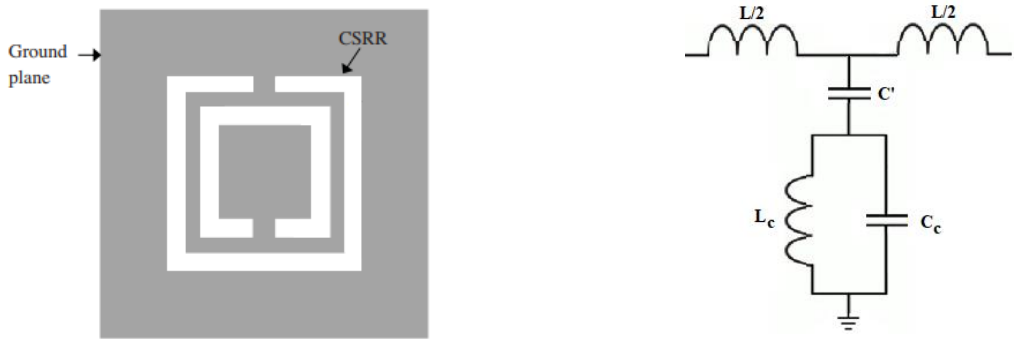
1 Introduction

Metamaterials are artificial materials that exhibit singular electromagnetic properties that we do not find in nature or in their constituents taken separately. The most interesting feature is the possibility to control or modify the permittivity and permeability of the material to obtain a behaviour adapted to a specific application [**1**]. Metamaterials are used to improve the performance of antennas, filters and couplers. Their main advantage is the miniaturization of the microwave circuits. Thanks to these left-handed materials as Calls them, which are used the split-ring resonators (SRR) and Complementary Split Ring Resonators (CSRRS), it may be possible to build compact microstrip low pass filters, stop band filters and band pass filters for microwave frequency uses, etc... These resonators have a structure (cell of base) very small in front of the guided wavelength. It is accepted that the limit of homogeneity is fixed at sizes less than λg/4.

The SRR is one of the essential structures of metamaterials. It produces a negative Permeability effect close to the resonance frequency While the metal wires behave like a two-dimensional plasma, with negative Permittivity to the plasma frequency. The use of CSRRs metamaterials is to promote the miniaturization of band-stop and UWB filters [2-3]. The CSRRs are formed by parallel combinations of inductances (s) and Capacitors (C), with the electromagnetically coupled LC circuit to the host transmission line. The equivalent circuit model for CSRRs loaded by transmission lines, and its relevant values of inductances and capacitors can be calculated using the methods described in [4]. At resonance, the CSRR provides

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284083

negative effective permittivity and produces a sharp rejection stop band. **Figure 1.** shows the geometry of the CSRR cell structure with it's equivalent lumped elements.



(a) (b)

Fig. 1. (a) Geometry of a CSRR unit cell (b) Lumped-element equivalent circuit of the CSR.

The zero transmission frequency of each CSRR coupled to the transmission line is given by:

$$f_z = \frac{1}{(2\pi\sqrt{LC[C_c+C]})} \quad (1)$$

The resonance frequency of CSRR can be tuned for the desire application.

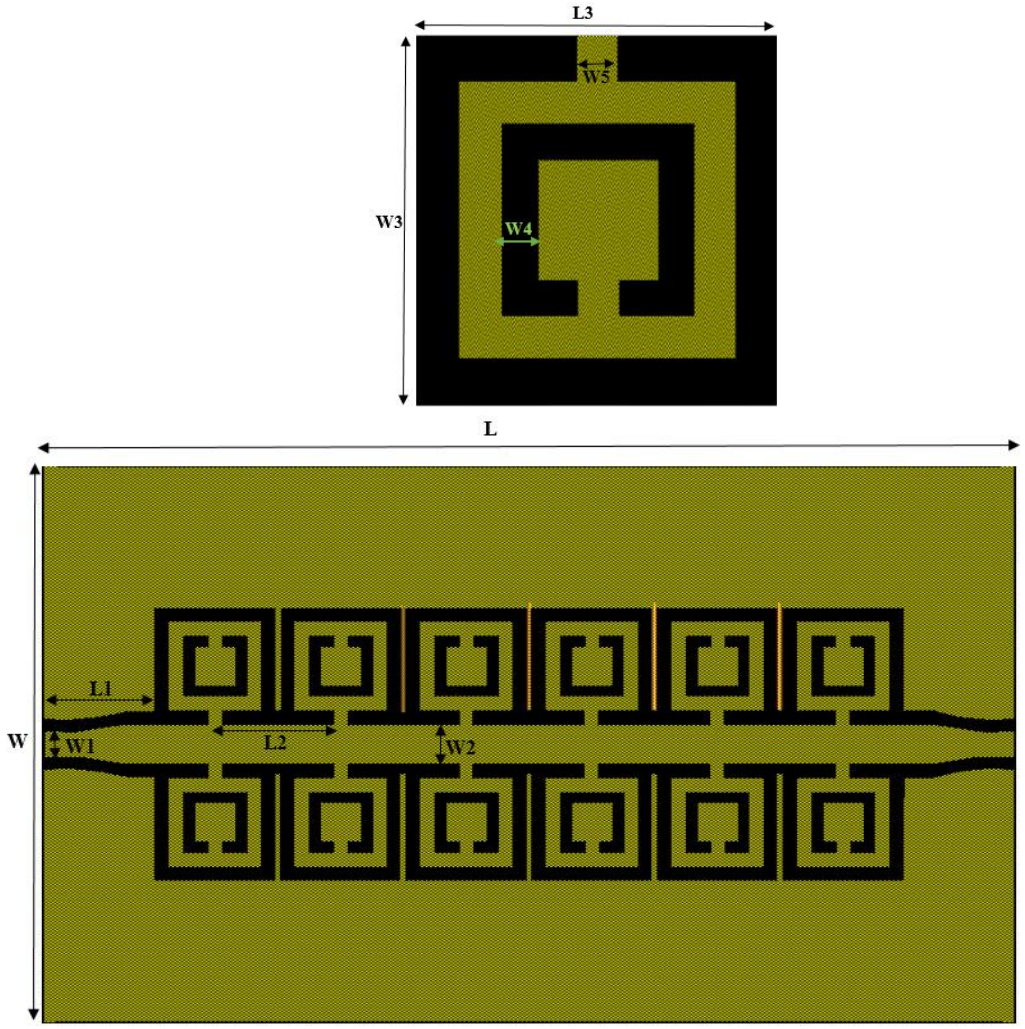
It is shown by many publications [5-9] that for a band stop filter, CSRR can greatly reduce the size of the filter while retaining very high stop band attenuation. This is because the evanescent modes can propagate in CSRR structures. In other side the periodic integration of these complementary split ring resonator on coplanar waveguide (CPW) provide some different design advantages in comparison with the conventional microstrip technology, e.g. the easy realization of shunts and the possibility of mounting active and passive lumped components [10-11]. in this paper we will present a CPW low pass filter loaded with six CSRR etched periodically in the two sides of the ground plane.

2 Design procedure

2.1 Theoretical study of the CSRR

2.2 Design procedure by ADS Solver

We have started the design of this filter by studying the unit cell of CSRR structure, which permit to adapt our filter structure to the desire applications, then we have inserted six complementary split ring resonator shape periodically in each ground of CPW line. After many series of optimization using methods integrated in ADS (Advanced Design System), we have validated into simulation the proposed CPW LPF. **Figure 2** presents the geometry of the proposed low pass filter (a) and CSRR unit cell (b).



(b)

Fig. 2. (a) Geometry of a CSRR unit cell (b) Proposed CPW LPF structure.

The optimized dimensions are presented in Table 1.

Table 1. Dimensions of the proposed LPF structure.

Parameters	Values (mm)
L	35.48
W	21.16
L1	4
L2	4.11
L3	4.4
W1	1
W2	1.5
W3	3.9
W4	0.44
W5	0.5

The final CPW LPF is mounted on FR4 substrate having a thickness of 1.6mm, a dielectric permittivity εr =4.4 and loss tangent tanδ=0.025. The proposed CPW low pass filter using metamaterial structures is simulated by using electromagnetic solver ADS. The **Figure 3** illustrates the S parameters of the filter structure.

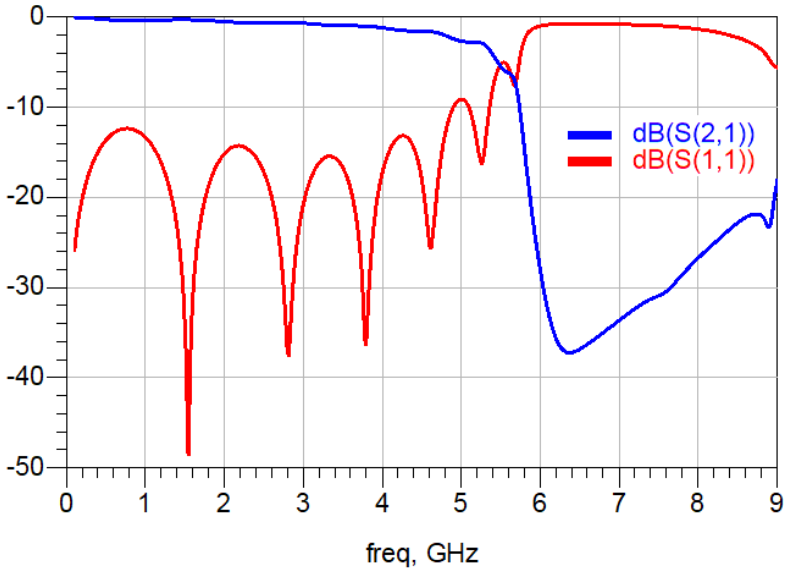


Fig. 3. The S parameters of the proposed LPF structure.

As depicted in **Figure 3**. We have obtained a low pass filter having a large pass band with a cut-off frequency around 5.28GHz at -3dB. This frequency can be tuned to achieve any low pass filter by controlling the dimensions and values of capacitance and inductance when we pass from lumped elements to distributed elements. The filter presents a good result in term of insertion loss around -0.5dB, a good rejection band until 9 GHz which is about -38dB. This rejection is due to the insertion of periodic square complementary split ring resonator shape into the CPW line ground plane.

2.3 Design procedure by another Solver

After validation of the proposed filter using ADS electromagnetic solver, we will proceed to design the same filter structure by using another electromagnetic solver in order to compare the simulation results and to confirm the validation of the final CPW low pass filter integrating CSRR structures. The topology of this filter is depicted in the **Figure 4**.

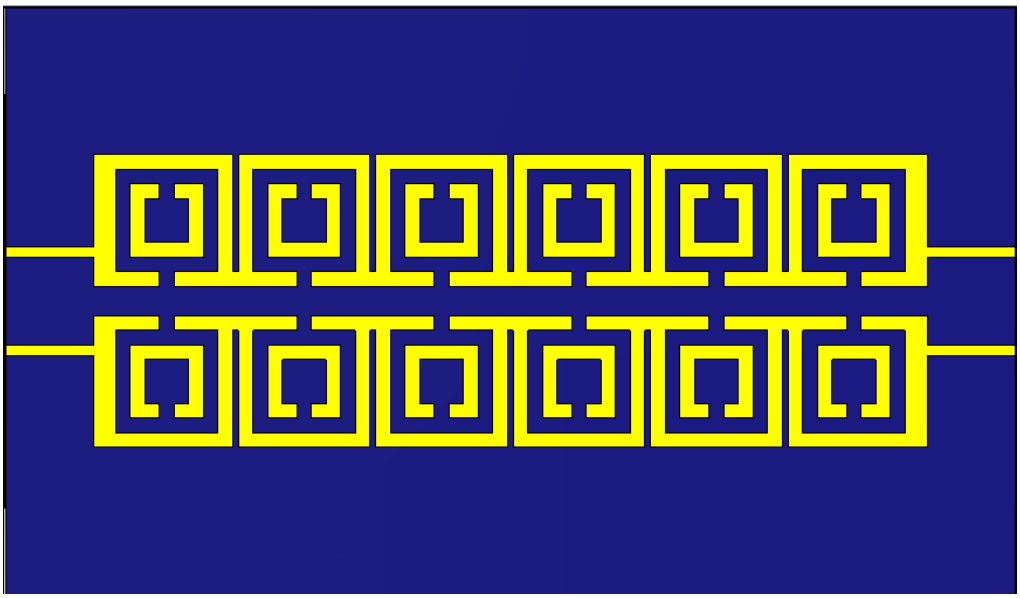


Fig. 4. Geometry of the proposed CPW LPF structure by ANOTHER SOLVER Solver.

The **Figure 5** illustrate a comparison of simulation results obtained by using two electromagnetic solvers.

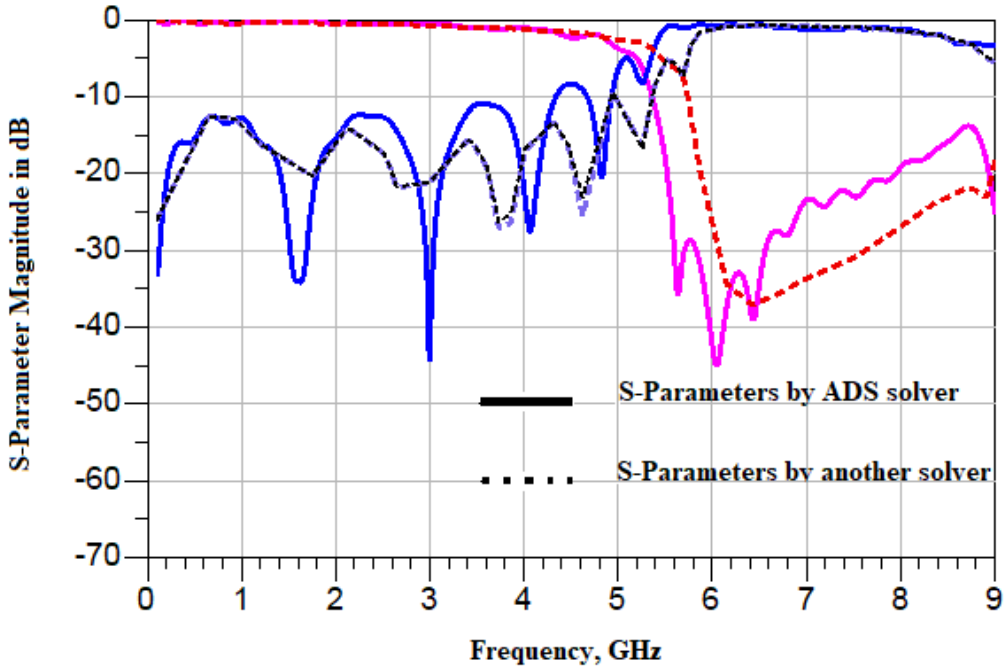
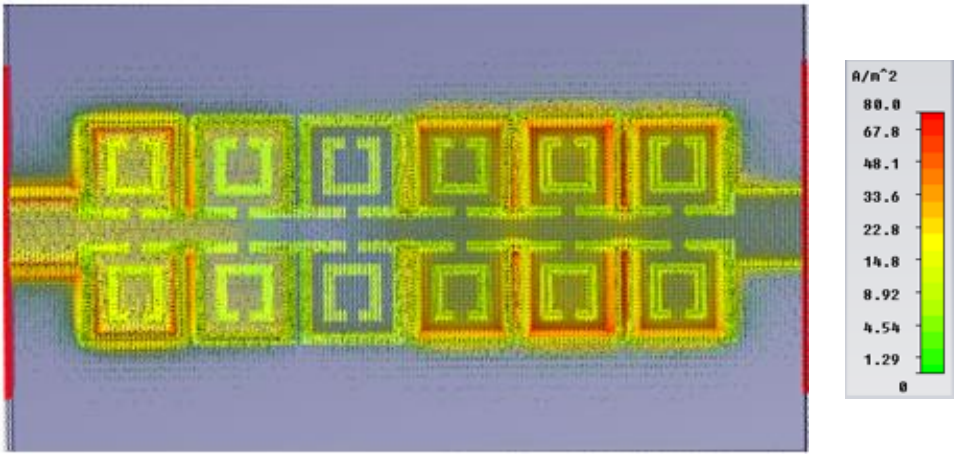


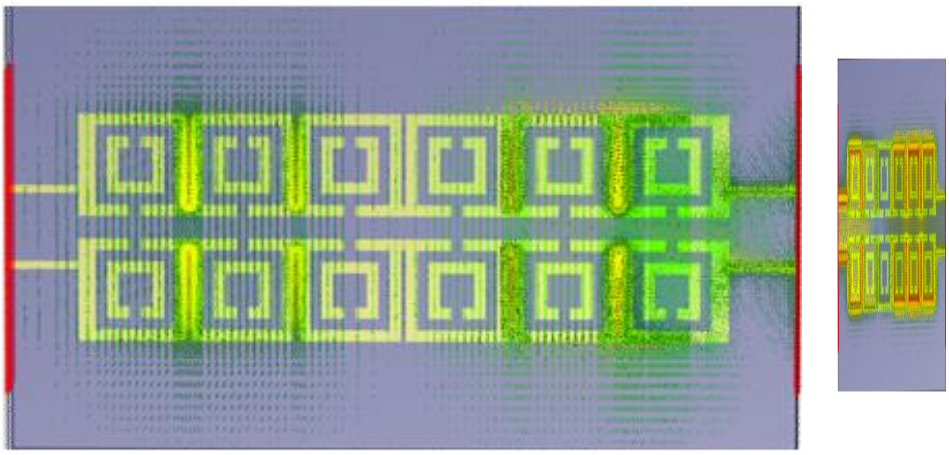
Fig. 5. The S parameters ADS and another solver Comparison results of the proposed CPW LPF

As depicted in **Figure 5**, we have obtained a good agreement between simulation results using ADS and another electromagnetic solver which confirm the operation of the proposed filter.

To validate this filter in term of current density we have launched a simulation at two frequencies one at 2 GHz in the pass band and another one at 7 GHz in the rejection band. As shown in **Figure 6**, we have the current density presented in the both ports at the two frequencies which confirm the operation frequency bands of the proposed filter.



(a) At 2 GHz



(b) At 7 GHz

Fig. 6. Presents the current distribution (a) at 2 GHz and (b) at 7 GHz

3 Conclusion

In this work we have presented a novel CPW low pass filter structure based on the use of metamaterial structures placed periodically in the ground plane of coplanar waveguide line. The circuit is designed, simulated and validated by using two electromagnetic solvers. As results we have developed a low pass filter structure with a good rejection which is due to the use of CSRR structures. The LPF can be used for several microwave applications.

Acknowledgment

We have to thank Pr. M. Latrach and his research Team RF& Microwave in ESEO Engineering institute in Angers in France for all the support brought to our research studies by permitting us to use Electromagnetic solvers.

References

[1] Veselago, V.: The electrodynamics of substances with simultaneously negative values of and. Soviet Physics Uspekhi. (1968)

[2] Smith, D. R., Padilla, W. J., Vier, D. C, Nemat-Nasser, S. C. and Schultz, S.: Composite medium with simultaneously negative permeability and permittivity. Physical review letters. pp. 4184-4187 (2000)

[3] Zeng, H. –Y, Wang, G. –M, Zhang, C. –X. and Zhu, L. : Microwave & Optical Technology Letters. (2010)
[4] Gil, M., Bonache, J. and Martin, F.: Metamaterial Filters: A Review. Science Direct. pp. 186-197 (2008)
[5] Kehn, M.N.M., Quevedo, O.T. and Rajo, E.I. :Split-ringresonator loaded waveguides with multiple stopbands. Electronics Lett., pp.714–716 (2008)
[6] Kim, J., Cho, C.S. and Lee, J.W. :CPW bandstop filter using slot-type SRRs. Electronics Lett., pp.1333–1334 (2005)
[7] Lin, X. and Cui, T. :Controlling the bandwidth of split ring resonators. IEEE Microwave and Wireless Components Lett., pp.245–247 (2008)
[8] Xu, Y. and Alphones, A. : Propagation characteristics of complimentary split ring resonator (CSRR) based EBG structure. Microwave and Optical Tech. Lett.,pp.409–412 (2005)
[9] Baena, J.D., Bonache, J., Martin, F., Sillero, R.M., Falcone, F., Lopetegi, T., Laso, M.A.G., Garcia-Garcia, J., Gil, I., Portillo, M.F. and Sorolla, M.: Equivalent-circuit models for split-ring resonators and complementary splitring resonators coupled to planar transmission lines. IEEE Microwave Theory and Tech.,. pp.1451–1461 (2005)
[10] Simons, R. N.: Coplanar Waveguide Circuits, Components and Systems. Wiley-IEEE, New York (2001
[11] Wolff, I. :Coplanar Microwave Integrated Circuits. John Wiley & Sons, (2006)

A Novel Periodic CPW UWB Low Pass Filter With Large Rejected Band

Fouad Aytouna[1], Otmane Chakkor[1], Mohamed Latrach [2]

{ [1]aytouna2002@gmail.com }

[1] ENSA, Abdelmalek Essaadi, University, Tetouan, 93000, Morocco
[2] RF& Microwave Group, ESEO, Angers, France

Abstract. In this paper, we present a new UWB CPW low pass filter structure by using periodic elements. The originality of this work is to develop a neFigure 4 showsw UWB LPF structure which is compact, miniature, and easy for fabrication. The validated LPF is a compact planar filter easy to associate with others microwave planar circuits, having a large band pass and a large rejection band. The final circuit is simulated and optimized by using the electromagnetic solver ADS (Advanced Design System). After many series of optimizations, we have validated the final circuit into simulation by using different optimization methods taking into account a high density of meshing in order to cover the whole circuit. The simulated LPF circuit shows good results in term of matching input impedance and insertion loss with a cut-off frequency of 6.5GHz. The entire area of the proposed LPF is 16x40 mm^2.

.**Keywords:** CPW, UWB, LPF, Periodic Structure.

1 Introduction

Periodic structures of different types have always been a field of interest for researchers. They are actually used in many applications in the microwave and millimeter-wave regime [1]. The periodic structures have been used to fabricate high-performance filters, to perform harmonic tuning in power amplifiers, to suppress superior's harmonic and to enlarge the band pass [2-5]. Coplanar Waveguide (CPW) structures are one of the important subjects for the miniaturization of microwave integrated circuits (MICs). Due to uniplanar structure, low dispersion effect and simple fabrication process, CPW are used more than microstrip structures in the planar technology filter applications [6]. Harmonic-suppressed and wide stop-band properties have been highly demanded in millimeter systems to reject the spurious responses caused by nonlinear devices, such as high power amplifiers, mixers, and oscillators, etc Different approaches can be used to reach this goal like employing a periodic structure [7-8] or a Defected Ground Structure (DGS) [9], many research activities have been performed in order to apply it to microwave LPF circuits [10].

Many techniques can be used to implement such filters like stepped-impedance, coupled line, using coupled resonators, periodic structures and others. In order to design a planar filter many steps can be followed. Firstly, we can start by designing a low pass filter prototype using Butterworth, Chebyshev approximations [11]. The prototype g-values are scaled (demoralized) to the desire filter termination resistance and cut-off frequency as follow:

$$L = \frac{gR}{\omega} \quad (1)$$

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284082

$$C = \frac{g}{\omega R'} \tag{2}$$

High pass, band pass and stop-band can be obtained by applying an L-C transformation on low pass filter [12].

The next step is use the Richards' transformation, which allows the inductors and capacitors of a lumped-element filter to be replaced with short-circuited and open-circuited transmission line stubs as depicted in **Figure 1.**

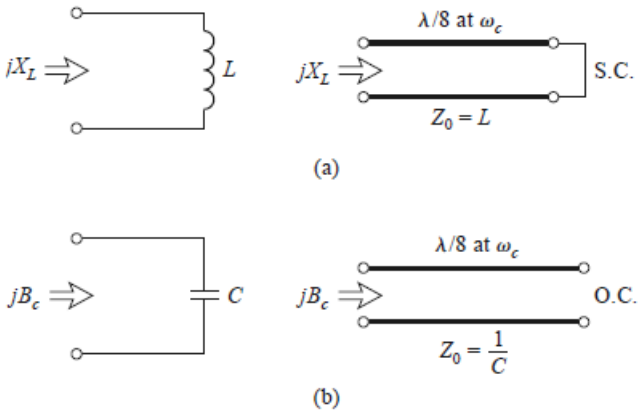


Fig.1. Richards' transformation. (a) For an inductor to a short-circuited stub. (b) For a capacitor to an open-circuited stub.

The transformation

$$\Omega = \tan \beta l = \tan \frac{\omega l}{\vartheta_p} \tag{3}$$

Map ω plane to Ω plane. if we replace the frequency variable with ω, we can write the reactance of an inductor as:

$$jX_L = j\Omega L = jL \tan \beta l \tag{4}$$

$$jB_c = j\Omega C = jC \tan \beta l \tag{5}$$

In addition to these, kuroda identities can be very useful in making the implementation of Richard's transformations more practicable. We can use Kuroda's Identities [13] to:

- Physically separate transmission line stubs.
- Transform series stubs into shunt stubs.
- Change impractical characteristic impedances into more realizable ones.

In general, the filter synthesis isn't always possible; therefore we can achieve new filter structures by using optimization methods. In this paper, we have used the optimization methods to achieve a novel LPF based on the coplanar waveguide (CPW) with improving the insertion loss and stop-band performances.

2 Design procedure

To develop this novel structure, we have done a bibliography study, in order to have an idea about the different shapes and techniques used for such design. By consequent, we have started our work taking into account the study done in [14]. The filter design is started by passing from a simple CPW line, in which we have inserted octagonal shape as shown in **Figure 2.**

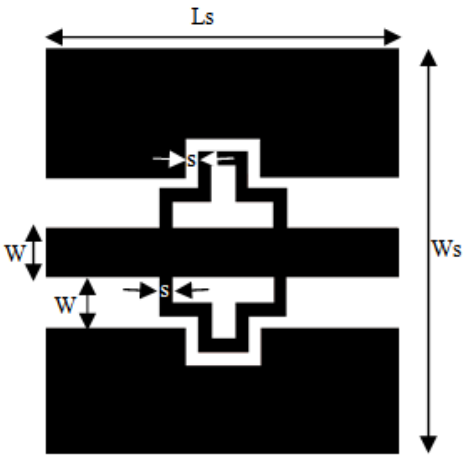


Fig.2. The geometry of the CPW (Ws=16mm, Ls=14mm, s=0.5mm et W =2mm).

The simulation results of this filter structure are depicted in **Figure 3**.

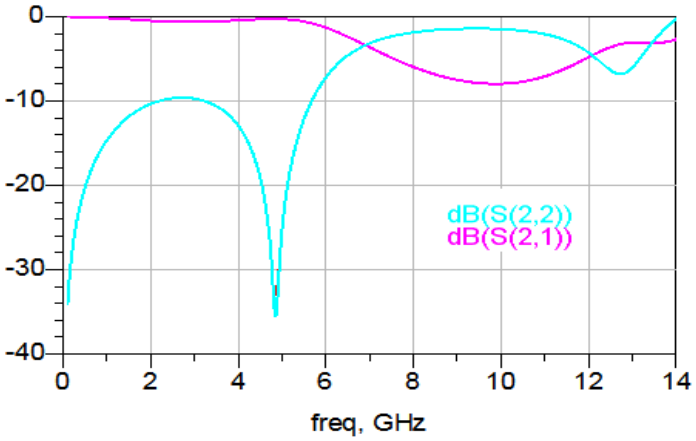


Fig.3. The S-parameters of the simulated CPW line versus frequency.

The Figure 3 shows the simulation results of the structure with one cell, as we can conclude, we have obtained a low pass filter behavior with a cutoff frequency equal to 6.77 GHz corresponding to a reflection coeffecient around -10dB in the whole bandwidth with an insertion loss less than -0.5dB. In order to enhance the rejection band characteristics we have done another study based on the technique of the use of the periodic structure. The number of unit cell and the distance between cells influence the rejection band characteristics (Aytouna et al., 2015).

After many series of simulation using optimization methods integrated in ADS, we have validated into simulation the proposed CPW LPF presented in **Figure 4**. The final proposed CPW LPF structure contains 4 cells which gave the good performances.

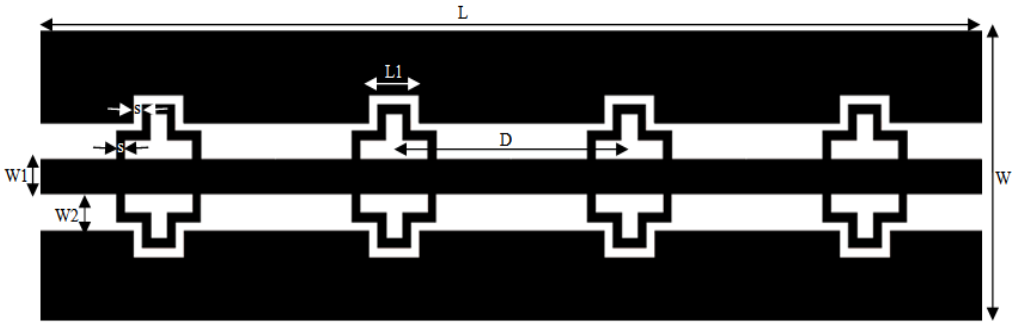


Fig.4. The topology of the proposed LPF.

The optimized parameters are presented in **Table 1**.

Table 1. Dimensions of the proposed CPW LPF structure.

Parameters	Values (mm)
W	16
W1	2
W2	2
L	40
L1	3
D	9
S	0.5

This CPW LPF is printed on an FR4 substrate with a thickness of 1.6mm, a dielectric permittivity $\varepsilon_r = 4.4$ and loss tangent $\tan\delta = 0.025$.

Figure 5 illustrates the S parameters of the proposed filter.

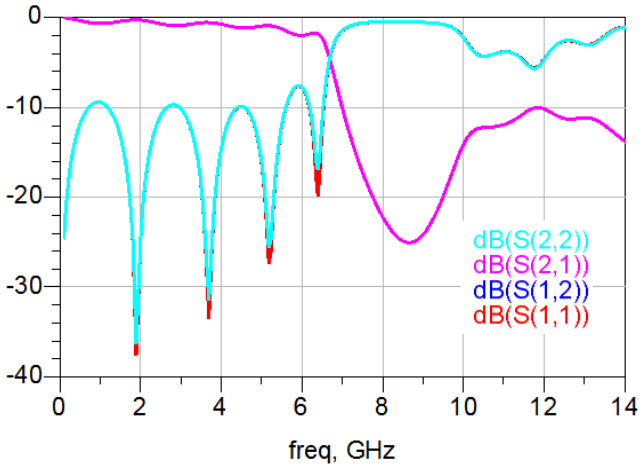


Fig.5. S-parameters versus frequency of the designed CPW LPF

As depicted in **Figure 5** we have obtained good results in term of insertion loss around -0.8dB, a cutoff frequency of 6.77 GHz and a good rejection until 15GHz. These results are due to the insertion of periodic structures.

To show the effect of the variation of the distance and the number of cells, we have conducted a series of simulations based on two parameters:

2.1 Effect of the number of cells

The first study is done by fixing the distance between cells at 9 mm. **Figure 6** and **Figure 7** illustrate the S-parameters versus frequency of the designed LPF by changing the number of cells.

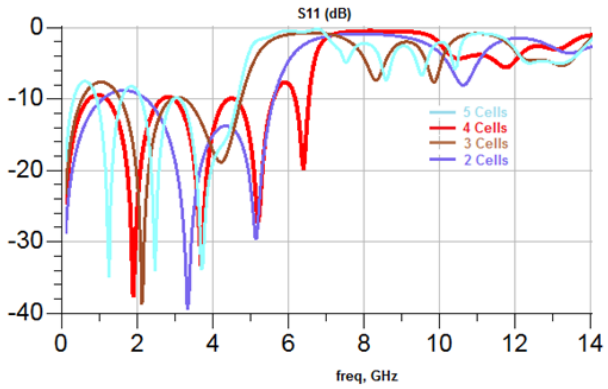


Fig.6. S11 parameter versus frequency for different number of cells of the proposed LPF.

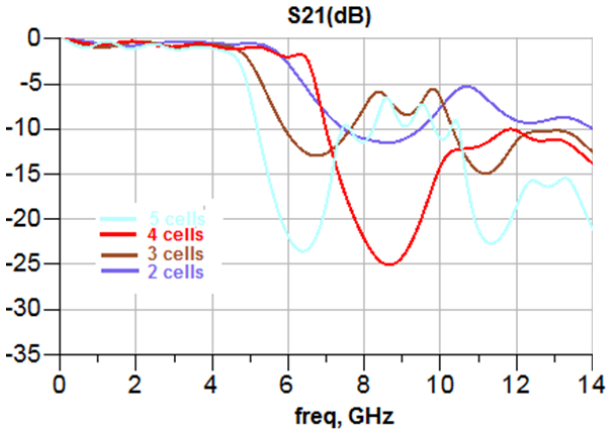


Fig.7. S21 parameter versus frequency for different number of cells of the proposed LPF.

As shown in **Figure 6** and **Figure 7** the use of adequate number of cells in the proposed CPW LPF permits to enlarge the rejection band.

2.2 Effect of the distance between cells D

The second study is done by fixing the number of cells; in this case we have used 4 cells and we have changed the distance D between cells, the different results obtained are presented in **Figure 8** and **Figure 9**.

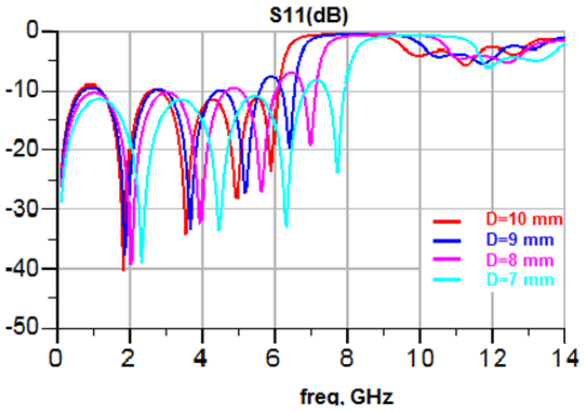


Fig.8. S11 versus frequency for different value of distance between cells of the proposed LPF.

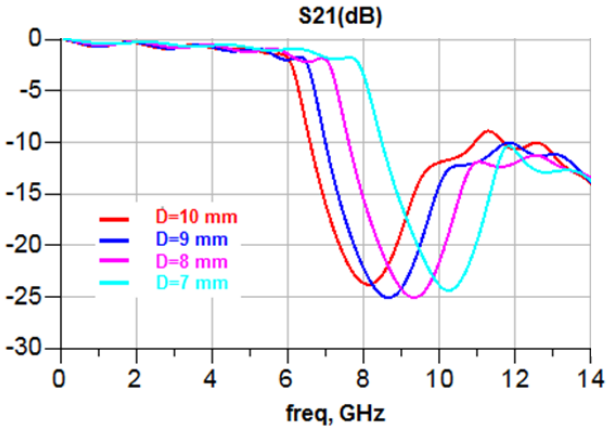
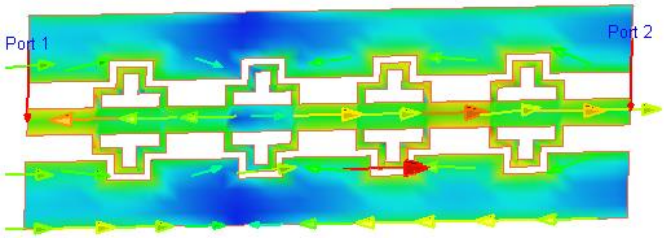
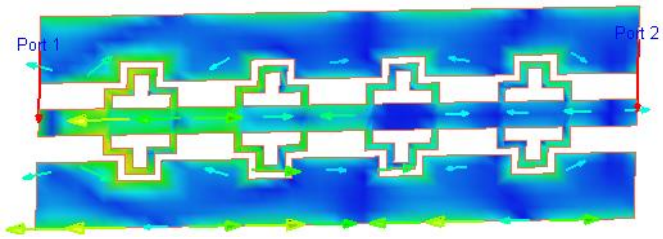


Fig. 9. S21 versus frequency for different distance between cells of the proposed LPF

To show the flow of energy in the band pass and rejected band, we have launched a simulation which confirms the function of the CPW LPF filter. **Figure 10** presents surface current distribution , one in the bandwidth at 2.58 GHz and another one at 10 GHz in the rejected band which validate the frequency bands of the proposed filter.



(a) at 2.58 GHz



(b) at 10 GHz

Fig.10. Presents the current distribution at 2.58 GHz and at 10 GHz

3 Conclusion

This work comes with a new configuration of LPF structure based on the use of CPW technology and periodic structures formed by octagonal shape slots permitting to adjust the bandwidth and to enlarge the rejection band. The design procedures were based on many optimization methods integrated into ADS. The simulation of the LPF was done taking into account a high mesh density to cover the whole circuit. The final proposed filter structure can be used for wireless communications and many other microwave applications. The methodology followed in this study can be used to match the LPF filter to another frequency band.

Acknowledgement

We have to thank Pr. M. Latrach and his research Team RF& Microwave in ESEO Engineering institute in Angers in France for all the support brought to our research studies by permitting us to use Electromagnetic solvers and instruments of test.

References

[1] Yablonovitch, E.: Photonic band-gap structures. J. Optical Soc. America B. pp. 283-295 (1993)
[2] Yang, F. R., Coccioli, R., Qian, Y. Itoh, T.: Analysis and application of coupled microstrips on periodically patterned ground plane. IEEE MTT-S Int. Microwave Symp. Dig. pp. 1529-1532 (2000)
[3] Hang, Radisic, V., Qian, Y.,Itoh, T.: High efficiency power amplifier with novel PBG ground plane for harmonic tuning. IEEE MTT-S Int. Microwave Symp. Dig. pp. 807-810 (1999)
[4] Yang, F. R., Ma, K. P., Qian, Y., Itoh, T.:A novel TEM waveguide using compact photonic-bandgap (UC-PBG) structure. IEEE Trans. Microwave Theory Tech. pp. 2092-2098 (1999)
[5] Ma, K. P., Kim, J., Yang, F. R., Qian, Y. Itoh, T. :Leakage suppression in stripline circuits using a 2-Dphotonic bandgap lattice. IEEE MTT-S Int.Microwave Symp. Dig. pp. 73-76 (1999)
[6] Meng, Q., Zhang, Pu, X., J., 2009. :A Novel CPW Low-Pass Filter Using Electromagnetic Band-gap Structure. In: Proc. IEEE. Int. Conf. Appl. Supercond. Electromagn. Devices. pp. 267-269 (2009)
[7] Yang, F.-R., Qian, Y., Itoh, T.: A novel uniplanar compact PBG structure for filter and mixer applications. IEEE MTT-S Int. Dig. pp. 919–922 (1999)
[8] Aytouna, F., Zbitou, J., Aghoutane, M. , Touhami, N.A. , Tribak, Latrach, A. M.:A Novel CPW BandPass Filter Integrating Periodic Rectangular Slot Cells. Journal of Engineering Science and Technology Review. pp. 131–134 (2015)
[9] Ahn, D., Park, J.-S., Kim, C.-S., Kim, Qian, J., Y., Itoh, T.: A design of the low-pass filter using the novel microstrip defected ground structure. IEEE Trans. Microw. Theory Tech.pp.86–93 (2001)
[10] Chen, H.-J., Huang, T.-H., Chang, C.-S., Chen, L.-S., Wang, N.-F., Wang, Y. [10] H., Houng, M.-P.:A novel cross-shaped DGS applied to design ultra-wide stopband low-pass filters. IEEE Microw. Wireless Compon. Lett. pp. 252–254 (2006)
[11] Hettak, K., Delisle, G., 1989. :A new miniature uniplanar lowpass filter using series resonators. IEEE MTT-S. pp. 1193–1196 (1989)
[12] Ishii, T. K.: Hand book of Microwave Technology, components and devices. by academic press, Inc., USA (1995)
[13] Pozar, D. M.: Microwave Engineering. John Wiley& Sons, Inc., USA (2012)
[14] Sor, J., Qian, Y.: Miniature low-loss CPW periodic structures for filter applications. IEEE Transactions on Microwave Theory and Techniques .pp. 2336–2341 (2001)

A SPICE IGBT Model with easy parameters extraction

Hadini Yassine[1], Galadi Abdelghafour[2] , Echchelh Adil[3]
{yassine.hadini@uit.ac.ma[1], agaladi@yahoo.fr[2], adilechel@gmail.com[3]}

[1,3] Laboratory of Electrical Engineering and Energy System, Department of physics Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco
[2]Team: Electronic, Electrical engineering, Instrumentation and Imaging, National School of Applied Sciences, Cadi Ayyad University, Safi, Morocco

Abstract: An accurate IGBT model is presented in this paper. It is based on the IGBT physical structure. The model uses an equivalent circuit which combines a MOSFET, controlled by a voltage-controlled voltage source, in series with a diode. the required parameters can be extracted with a simple procedure from the datasheet dc characteristics given by the manufacturers. The results of the comparison between the simulation and the datasheet information shows an average error less than 9% for the output characteristics and less than 7% for the transfer characteristic. Moreover, the model is simple and can be implemented in any SPICE-based circuit simulator. A description of the model and the parameter extraction procedure will be provided along.

Keywords:IGBT, SPICE, MODEL, MOSFET, DIODE.

1 Introduction

In power electronic circuit, the power transistors are used as a switch to control the power electric energy. To sustain high voltage in the off-state, the power transistor structure contains a thick and lowly doped region called the drift region [1]. In case of the power metal-oxide-semiconductor field-effect transistor (MOSFET), the drift region turns into high resistance restricting the power electronic applications of this type of devices. In bipolar power transistor, the resistance of the drift region is reduced considerably thanks to the modulation of the conductivity of this drift region [1]. The insulated gate bipolar transistor (IGBT) transistor structure (Figure 1), is a hybrid MOSFET-bipolar configuration which is equivalent to an bipolar current controlled by an MOSFET. Thanks to the high input impedance of the MOSFET, the IGBT is controlled by a low gate drive power. As IGBT voltage and current ratings increase, its application range is extending to high power applications; that is why, in the very near future, the IGBT will be a serious competitor to the GTO [2], [3].

On the other hand, the peculiar IGBT structure makes the SPICE modeling of the IGBT a very complex task. In literature, the IGBT model has been represented by a couple of a simplified equivalent circuit. The first presents the IGBT as a N-channel MOSFET driving a PNP bipolar transistor in a Darlington configuration [4]–[8] , whereas the second consists of an N channel MOSFET in series with a diode [9]. The proposed models cannot describe the IGBT behaviour accurately, because the wide base in the IGBT differs from the existing discrete power BJT model. Furthermore, these models cannot extract the BJT current gain BF parameter from

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284073

the datasheet information like all the other parameters used in the model. While Mihaliç et al [7] chose the maximum possible value for BF parameter as BF=100, Kim et al [5] assumed the current gain BF equal to 1.

In this paper, an IGBT equivalent circuit, based on an N-channel MOSFET in series with diode and effective voltage controlled voltage source (VCVS) is proposed. All needed parameters of the model are extracted easily by using a simple method from the measurements. The MOSFET model parameters are extracted from the transfer characteristic, while the diode parameters are extracted from the output characteristics. Moreover, our model takes into consideration the IGBT transconductance flattening out [10] which occurs in high gate voltage, by VCVS in the MOSFET gate. The validity of this model has been confirmed by the comparison between the simulation and the measurements given in the manufacturer's datsheet for the IGBT CM600HA-24H. The comparison shows that the proposed model gives 9 % as an average error for the output characteristics and 7% for the transfer characteristic, the DC model description will be given in section 2, the parameters extraction procedure will be explained in section 3, the results of the comparison between the measurements and simulation will be presented in section 4 and the conclusion of this paper is given at the end.

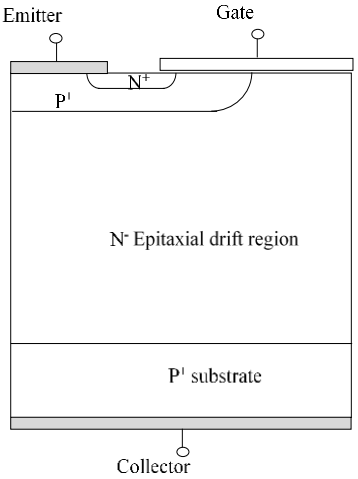


Fig 1. Cross-sectional view of NPT IGBT structure

2 Model description

As in the MOSFET, the IGBT collector current is controlled by the applied gate voltage VGE. When this applied voltage exceeds the threshold voltage (Vth), a channel is created underneath the gate allows the current to flow. The output characteristics show the liaison between collector voltage, collector current and the gate voltage. They are different from those of the MOSFET as they have a cut-off region where the output current is zero [11] .Between the output voltage and the saturation voltage, the IGBT operates in the linear region, where the collector voltage is determined by the collector current and the gate voltage. Above the saturation voltage up to breakdown voltage, IGBT operates in the saturation mode. The transfer characteristic and the gate voltage determine the collector current at any collector voltage [12].

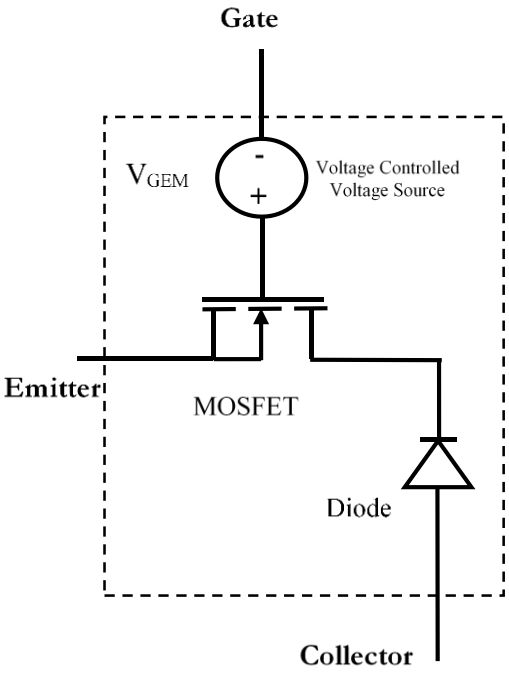


Fig 2. Schematic of the proposed model for the IGBT

Based on the description presented before, and on the physical structure, the proposed model uses a channel intrinsic MOSFET controlled by a Voltage Controlled Voltage Source (VCVS), in series with a diode as shown in Figure 2 to model the IGBT behaviour in its different states (cut-off, linear and saturation).

The intrinsic MOSFET is used to represent the effect of the channel created under the gate as described above. Therefore the MOSFET model describes the IGBT behaviour in saturation region and also in a part of the linear region [13]. The SPICE MOSFET model level 1 has chosen for this task, because of its simplicity and the reduced number of its parameters. The channel transconductance (Kp) and the threshold voltage (Vth) are the essentials parameters to model accurately the IGBT behaviour in these regions. Then the current in these regions has the following equations:

$$I\begin{cases} = 0 \\ if \quad V_{GE} < V_{th} \quad or \quad V_{CE} < V_D \\ = K_P\left[(V_{GE} - V_{th})(V_{CE} - V_D) - \dfrac{(V_{CE} - V_D)}{2}\right] \\ if \quad V_{GE} \geq V_{th} \quad and \quad V_{CE} < V_{GE} + V_D - V_{th} \\ = \dfrac{K_p}{2}(V_{GE} - V_{th})^2 \\ if \quad V_{CE} \geq V_{GE} + V_D - V_{th} \end{cases} \tag{1}$$

The diode is used to model the IGBT behaviour in the cut off region and the beginning of the linear region. And the current is modelled by the Diode SPICE model equation:

$$I = I_S\left[\exp\left(\frac{V_D}{NV_T}\right)-1\right] \quad (2)$$

The two basic required parameters for the diode model are the saturation current Is and the emission coefficient N.

The sub-circuit composed, just, by a MOSFET in series with a diode, could not describe accurately the IGBT behaviour in the saturation region, especially in the high VGE. Because the MOSFET SPICE model level 1 could not model the drop of the IGBT transconductance due to the channel saturation in the high VGE. So, to eliminate the difference between the simulation and the measurement in the saturation region, the VGE applied in the model should be less than the VGE used in the measurements for the same IC value. Therefore, to overcome this disagreement, a voltage-controlled voltage source (VCVS) is added. The main role of the VCVS is the control of the gate voltage, in the purpose to compensate the difference existed between the real VGE value and this applied in the model. the VCVS depends linearly on the VGE as described in the following equations:

$$V_{GEM}\begin{cases} = V_{GE} \\ \quad if \quad V_{GE} < V_{GEC} \\ = V_{GEC} + (a \times V_{GE}) + b \\ \quad if \quad V_{GE} \geq V_{GEC} \end{cases} \quad (3)$$

When

$$f(V_{GE}) = a \times V_{GE} + b \quad (4)$$

VGEC is a critical gate-emitter voltage at the onset of the transconductance drop phenomenon happened in the high VGE. a and b are the VCVS parameters.

Our proposed model uses the sub-circuit approach based on the physical structure to describe the IGBT behaviour. The new proposed IGBT model is easy to implement in the PSPICE either as sub-circuit or as a compact model using a netlist. Also, all required parameters of the model are simply extracted from the experimental results given in the manufacturer datasheet.

3 Parameter extraction

3.1 MOSFET parameters

The basic parameters of this model are the transconductance Kp and the threshold voltage Vth. The two parameters are extracted from the transfer characteristic in high VCE.

The MOSFET channel controls the IGBT behaviour in saturation region. the saturation region of the MOSFET and the IGBT can be assumed to be the same, therefore they have the same transfer characteristic in high VCE. For this etude the VCE chosen is 10 V. the level 1 SPICE model is used to model the MOSFET, then the MOSFET drain current in the saturation region is given by the following equation:

$$I = \frac{K_p}{2}(V_{GE} - V_{TH})^2 \quad (5)$$

From the linear portion of the transfer characteristic, we trace a straight line whose equation has the following form:

$$\sqrt{I} = \sqrt{\frac{K_P}{2}}\, V_{GE} - \sqrt{\frac{K_P}{2}}\, V_{TH} \qquad (6)$$

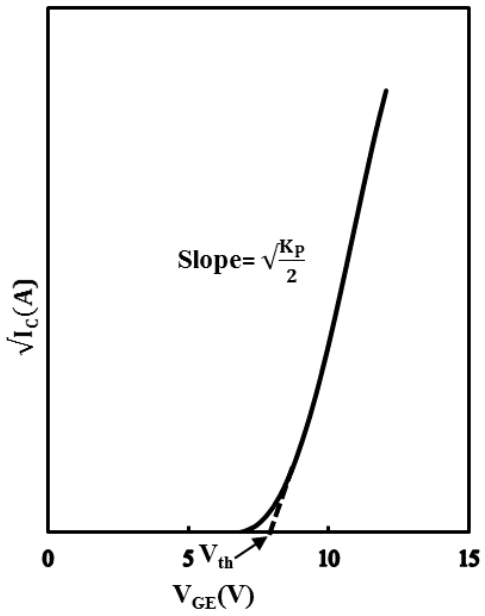


Fig 3. Square root of collector current vs gate voltage defines V_{th} and K_P.

KP can be determined from the slope factor of the √IC – VGE curve, while Vth is the intersection between the extension of the linear portion and the VGE-axis as shown in Figure 3.

3.2 . Diode parameters

The saturation current Is and the ideality factor N are considered as the essential parameters required to model the diode. They are extracted from the output characteristics in low collector voltage and high gate voltage. The diode controls the IGBT behaviour when VCE is in vicinity of the junction PN cut-in voltage.

$\exp\left(\frac{V_{CE}}{NV_t}\right) \gg 1$ then the equation (2) may be simplified

$$\ln I = \ln I_S + \frac{V_{CE}}{NV_t} \qquad (7)$$

Is and N can be extracted from the equation (7).

3.3 The voltage controlled voltage source parameters

3.3.1 Critical voltage V_{GEC}

In order to determine the critical VGE (VGEC). We trace the IC-VGE curve. from it we can define the VGEC as shown in the figure 4.

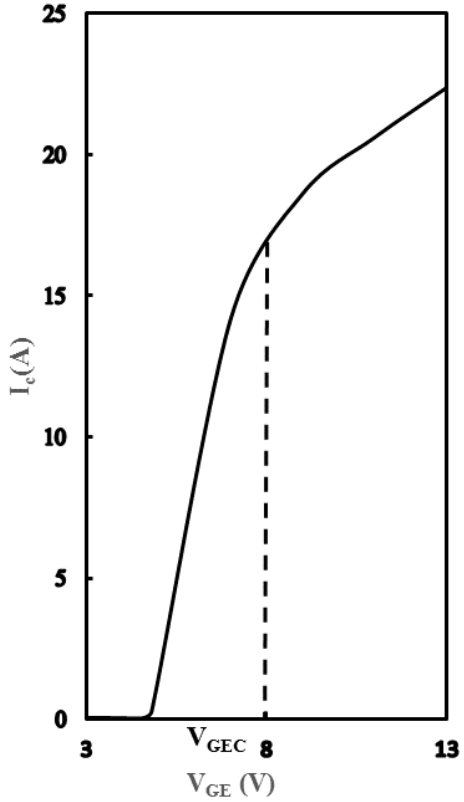


Fig 4. Collector current vs gate voltage defines V_{GEC}

3.3.2 a and b VCVS parameters

When VGE > VGEC, the VGEM can be given by:

$$V_{GEM} = \frac{I_C}{K_P(V_{CE} - V_D)} + \frac{(V_{CE} - V_D)}{2} + V_{th} \qquad (8)$$

Where VD is the voltage drop in the diode calculated by the following equation

$$V_D = NV_T \times \ln(\frac{I_C}{I_S}) \qquad (9)$$

Then from the equations (3), (4) and (8) a and b can be determined.

All the proposed methods in this section are verified by comparison with the experimental results given in datasheet. The comparison between simulated and experimental results will be shown and discussed in the following section.

4 Results and discussion

After having determined the parameters of the model, using the methods described above. The sub-circuit is simulated using PSPICE in order to compare the simulation results of the dc analysis with the dc characteristics given in the datasheet.

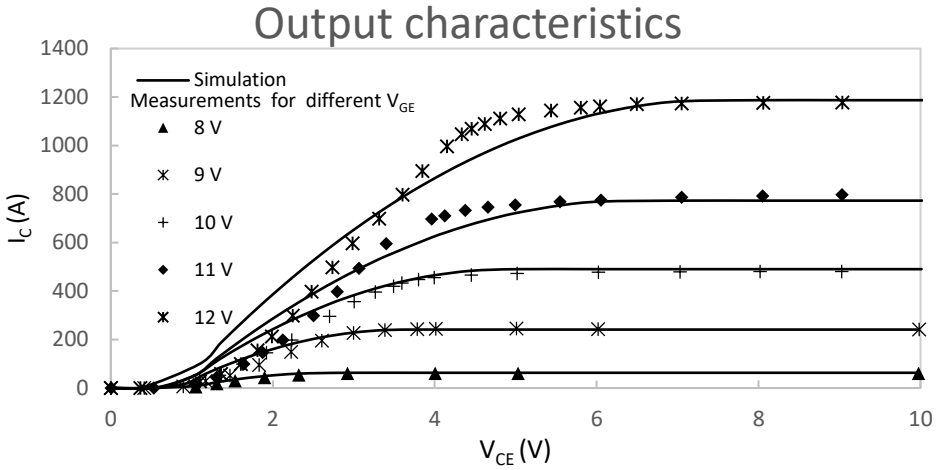


Fig 5. Measured output characteristics in comparison to the proposed model applied to the IGBT CM600HA-

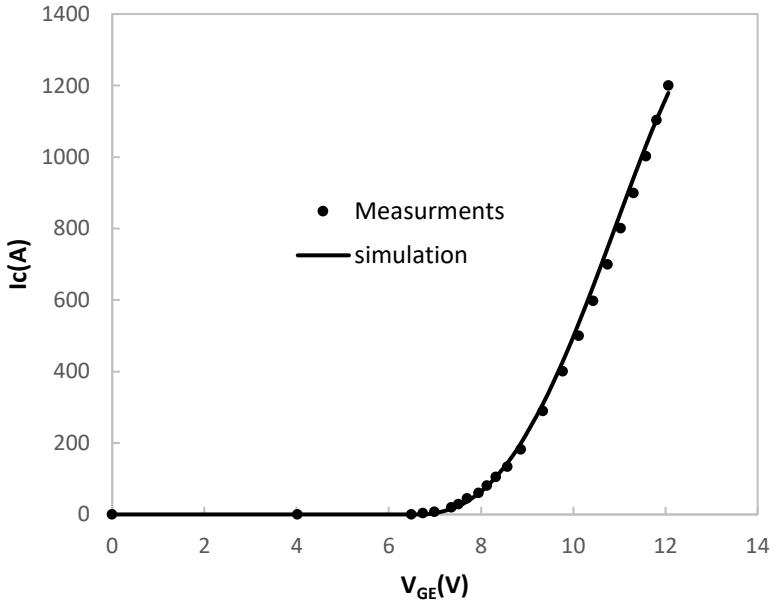


Fig. 6. Measured transfer characteristic in comparison to the proposed model applied to the IGBT CM600HA-24H for V_{CE} = 10 V.

While the Figure 5 shows the measured output characteristics in comparison with the simulation result of the proposed model applied to the IGBT CM600HA-24H, the Figure 6 shows the measured transfert characteristic in comparison to the proposed model applied to the IGBT CM600HA-24H for VCE = 10 V.

The proposed model presents a good agreement with the experimental results. The average error is less than 9% for in output characteristics and less than 7% in transfer characteristic.

It can be observed that in the linear region the measurement is a little higher than the simulation, this difference is due to two main causes, the first is that the model of the diode is unable to precisely describe the fort level injection presents in the IGBT N- base region [14], the second is that the transition from the linear region to the saturation region in the IGBT differs from which in the MOSFET SPICE level 1 model used in this work [11]. For this reason, the model proposed in this work can be used only for the NPT IGBT, where the level of injection in the base region is less important than the other IGBT types.

5 Conclusion

In this paper, the IGBT SPICE macromodel has been described. The static I-V characteristics offered by the macromodel are shown and the model efficiency is demonstrated by the comparison between the simulated and experimental results of two IGBT devices. The parameters extraction method is represented and all parameters such as Vth, KP, Is and N are successfully obtained from the datasheet characteristics. Our model is simple and easy to be implemented in any SPICE-based circuit simulator.

The future research direction could have as purpose, the improvement of this model to describe accurately all kind of IGBT.

References

[1] B. J. Baliga, *The IGBT device: physics, design and applications of the insulated gate bipolar transistor*. Kidlington, Oxford ; Waltham, MA: William Andrew/Elsevier, 2015.

[2] K. Sheng, B. W. Williams, and S. J. Finney, “A review of IGBT models,” *Power Electron. IEEE Trans. On*, vol. 15, no. 6, pp. 1250–1266, 2000.

[3] S. Yuan, S. Yuan, Y. Hong, and Y. Liu, “SPICE Model for IGBT Temperature Effects Compared with Measurements,” in *2010 International Conference on Electrical and Control Engineering*, 2010, pp. 3037–3040.

[4] K. Asparuhova and T. Grigorova, “IGBT high accuracy behavioral macromodel,” in *Microelectronics, 2008. MIEL 2008. 26th International Conference on*, 2008, pp. 185–188.

[5] H. S. Kim, Y. H. Cho, S. D. Kim, and Y. I. Choi, “Parameter extraction for the static and dynamic model of IGBT,” in *, 24th Annual IEEE Power Electronics Specialists Conference, 1993. PESC ’93 Record*, 1993, pp. 71–74.

[6] L. Michel, A. Chériti, and P. Sicard, “Development of an efficient IGBT simulation model,” in *Electrical and Computer Engineering, 2009. CCECE’09. Canadian Conference on*, 2009, pp. 252–256.

[7] F. Mihalič, K. Jezernik, K. Krischan, and M. Rentmeister, “IGBT spice model,” *Ind. Electron. IEEE Trans. On*, vol. 42, no. 1, pp. 98–105, 1995.

[8] H.-S. Oh and M. El Nokali, “A new IGBT behavioral model,” *Solid-State Electron.*, vol. 45, no. 12, pp. 2069–2075, Dec. 2001.

[9] “IGBT tutorial: Part 1 - Selection | EE Times,” *EETimes*. [Online]. Available: http://www.eetimes.com/document.asp?doc_id=1273173. [Accessed: 23-Jul-2016].

[10] V. K. Khanna, *Insulated Gate Bipolar Transistor IGBT Theory and Design*. John Wiley & Sons, 2004.
[11] K. Asparuhova and T. Grigorova, "IGBT behavioral PSpice model," in *Microelectronics, 2006 25th International Conference on*, 2006, pp. 203–206.
[12] I. Baraia, J. Galarza, J. A. Barrena, and J. M. Canales, "An IGBT behavioural model based on curve fitting methods," in *Power Electronics Specialists Conference, 2008. PESC 2008. IEEE*, 2008, pp. 1971–1977.
[13] P. Igić and N. Janković, "Review of advanced IGBT compact models dedicated to circuit simulation," *Facta Univ.-Ser. Electron. Energ.*, vol. 27, no. 1, pp. 13–23, 2014.
[14] B. J. Baliga, "Analytical Modeling of IGBTs: Challenges and Solutions," *IEEE Trans. Electron Devices*, vol. 60, no. 2, pp. 535–543, Feb. 2013.

Design of a Microstrip Balanced Amplifier Using the Wilkinson Power Divider

Amine Rachakh[1], Larbi El Abdellaoui[2], Ahmed Errkik[3], Abdelali Tajmouati[4], Mohamed Latrach[5]

{rachakh.amine1@gmail.com[1], elabdellaoui20@hotmail.com[2], ahmed.errkik@uhp.ac.ma[3]}

[1,2,3,4]MEET Laboratory, FST of Settat, Hassan 1st University, Morocco
[5]Microwave Group, Ecole Supérieure d'Electronique de l'Ouest, Angers, France

Abstract. The broadband bandwidth of power amplifiers is one of the most important features that restrict its wide usage. This work presents a new structure of a broadband power amplifier suitable for ISM band. The proposed amplifier, which is realized with two identical single-stage amplifiers and two power dividers on low loss Epoxy (FR4) substrate with a substrate thickness of 1.6mm and a relative permittivity equal to 4.4. The proposed PA demonstrates a power gain of 16 dB at 2.45GHz. The input and output matching impedances are better than -15dB throughout the band. the power amplifier delivers a maximum output power equal to 16dBm with a flat gain varied between 15-16dB.

Keywords: Power Amplifier (PA), Gallium Arsenide (GaAs) , Field-effect transistor (FET), Microstrip technology , Advanced Design System (ADS).

.

1 Introduction

Broadband power amplifiers are critical circuit components for wireless communication systems, as well as cellular telephone applications. The broadband power amplifier should be designed to trade-off the power gain, output power, return loss and bandwidth. The broadband amplifier can be designed by utilizing the compensation network, but the mismatching will raise the VSWR **[1]**. When VSWR improve, the power gain and output power will deteriorate. The balanced power amplifier can equilibrium these performance parameters for its coupling circuit at the input and output ports.

Conventionally, balanced power amplifiers are made up by associating two identical amplifiers with branch line couplers at the input and output as illustrated in **Figure 1.** Based on the circuit configuration illustrated in **Figure 1.** Balanced power amplifiers have been realized with various types of couplers, such as broadside couplers **[2–3]**, branch-line couplers **[4–5]**, and Lange couplers **[6– 7]**.

Various technologies are utilized in designing PA for several applications in wireless communications. For microwave frequencies, the GaAs MESFET technology appears to be prominent. This is owing to it has superior electrical characteristics at high frequencies, which helps to transmit high output power with a minimum parasitic loss. Moreover, its inherent characteristic of low resistance and low drain-source offset voltage **[8-10]**.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284098

Following these studies, this paper presents a new design of a balanced power amplifier by using a Wilkinson power divider. To improve circuit performance, the microstrip lines technology is used. While in the design process, the advanced design system (ADS) software is utilized, and it can also give simulation result of amplifier final circuit such as power gain, reflection coefficient, isolation coefficient, and output power.

2 Broadband Power amplifier design methodology

The balanced power amplifier is a convenient method for realizing a broadband PA that has good performances especially in input/output return loss and flat gain. The balanced amplifier is constructed with two single-stage power amplifiers and two hybrid couplers as shown in **Figure 1.**.

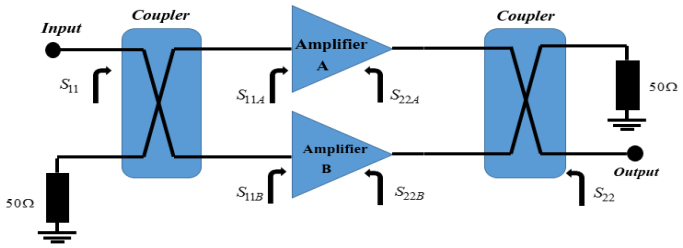


Fig. 1. Conventional topology of balanced amplifiers with hybrid couplers.

Considering the balanced amplifier architecture in **Figure 1.** the coupler on the left is working as the power divider and the one on the right is work as the power combiner, while the two amplifiers (Amplifier-A, Amlifier-B) are identical and have S-parameters [S_{-A}] and [S_{-B}] respectively. Thanks to the phase difference of the couplers, the two PA work in quadrature. it can be demonstrated that [S] parameters of the overall circuit (the balanced power amplifier are **[11]** :

$$|S_{11}| = \frac{1}{2}|S_{11A} - S_{11B}| = 0 \qquad (1)$$

$$|S_{12}| = \frac{1}{2}|S_{12A} + S_{12B}| = S_{12A} \qquad (2)$$

$$|S_{21}| = \frac{1}{2}|S_{21A} + S_{21B}| = S_{21A} \qquad (3)$$

$$|S_{22}| = \frac{1}{2}|S_{22A} - S_{22B}| = 0 \qquad (4)$$

It is clear from equations (1 - 4), that the magnitude of S_{11} and S_{22} will turn zero. By adopting the balanced configuration, excellent input and output matching conditions can be acquired simultaneously, while the reverse isolation and gain are kept unchanged.

In order to design a balanced power amplifier. As a first step, a single-stage power amplifier is designed and then transformed it into a balanced amplifier configuration by incorporating the hybrid couplers.

2.1 Design of the single-stage power amplifier

The important step that should take into consideration in designing a power amplifier is the choice of an appropriate active device (transistor). The choice of an active device for a microwave power amplifier is very complicated. It implicates choosing an active device having an acceptable current rating with output power and power gain capability that meets the requirements of the intended application. It is also important that the chosen an active device has a breakdown voltages which will not be exceeded by the RF and DC voltages that appear through the various junctions of the transistor and that permit the gain at frequency objectives to be met by the active device.

According to the design target, we select ATF-21170 transistor **[12]**. ATF-21170 is a low noise enhancement-mode GaAs FET designed to be employed for low cost commercial applications in the 500MHz to 6GHz frequency range [12]. It has high linearity, low noise and high gain performance and so on, therefore, that it can satisfy our request perfectly.

The bias circuit design is very important during an amplifier design process, The aim of the bias circuit is to set the quiescent point that is the Ids and Vgs for an active device that causes it to work in the preferred region. From a general perspective, there are different kinds of biasing networks although in power amplifier applications. Low complexity is required. One commonly and probably used method is to utilize a radial stub immediately after λ/4 high impedance bias line that will facilitate to realize proper isolation at desired RF/microwave frequency, whatever the component to use after λ/4 long bias line **[13].**

A stable and strong bias circuit to supply an appropriate quiescent operating point is important since the bias circuit will also affect the output power, gain, stability and so on. A class-A biasing point was chosen to obtain a good performance in power gain with good matching impedances and high output power. Depending on the datasheet **[12]**. We select the typical quiescent operating point, Vds = 4 V, Vgs= 0.5 V, Ids = 60 mA.

The power amplifier may fail if the design circuit is unstable because of oscillation. There are two natures of stability namely conditional stability and unconditional stability. Conditional stability in PA presents when the stability factor (k) less than 1 and this condition rely on the source and the load termination. It can keep the system stable for a certain range of source and load impedances **[14]**.To contrariwise, the unconditional stability in PA happens when stability factor more than 1, ensures the circuit to be stable for all source/load impedances. For the proper running of the power amplifier, the unconditional stability must be ensured. For this, we add a taper line to the source of the transistor **[15]** as shown in **Figure 2** .

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |S_{11}.S_{22} - S_{12}S_{21}|^2}{2|S_{12}S_{21}|} \qquad (5)$$

Where k is the stability factor **[14]**.

In the design of power amplifier, the matching network makes the input/output impedance transform to 50Ω impedance in the condition of good power gain and low return loss .The schematic of the proposed single-stage power amplifier is shown **Figure 2**; it comprises a stability network, bias network, the input and output matching networks, For machining, a "T" type matching is used at the input and the output sides of the transistor.

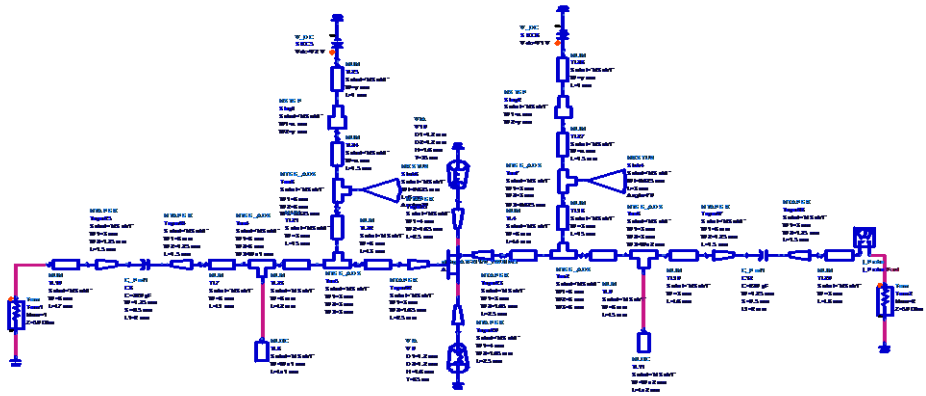


Fig. 2. The schematic of the proposed single-stage power amplifier

The simulation results of the proposed single-stage amplifier circuit are illustrated in **Figures 3–5**. **Figure 3** presents the simulated S-parameters of the performance of the proposed single-stage power amplifier. the power gain (S_{21}) of this power amplifier is about 19dB with an isolation coefficient (S_{12}) lower than -24dB at 2.45GHz, and the reflection coefficients are lesser than -10dB over the interested band; it can be observed that the input reflection coefficient (S_{11}) is less than -38dB and the output reflection coefficient (S_{22}) is less than -33.5dB at 2.45GHz. Whereas the circuit achieves a maximum saturated output power of 14dBm as shown in **Figure 4**. As depicted in **Figure 5**, it can be seen that k >1 over operating frequency band. Consequently, the condition for unconditional stability of the proposed single-stage power amplifier is confirmed on the interested band, which means there is no risk of oscillations. Despite this performance but the single-stage amplifier it has a narrow band.

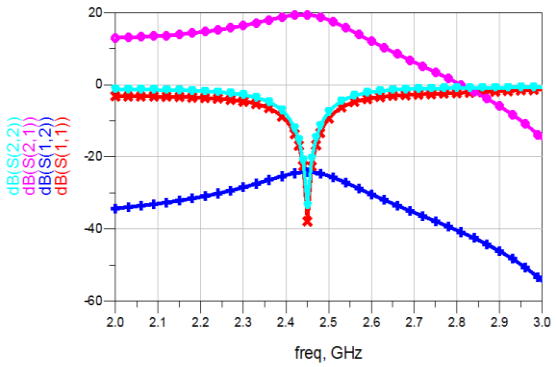


Fig. 3. S-parameters versus frequency of the proposed single-stage amplifier.

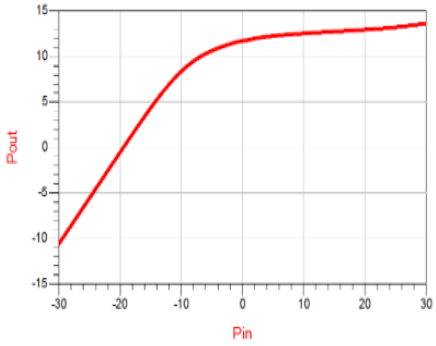


Fig. 4. Input power versus output power for 2.45GHz.

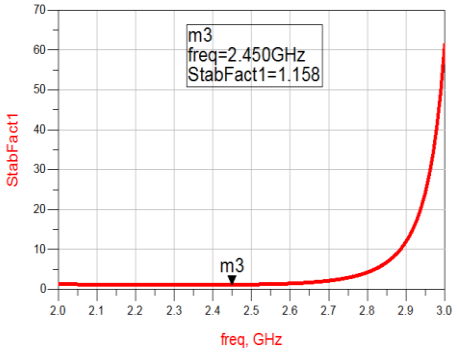


Fig. 5. Stability factor versus frequency characteristics.

2.2 Design of the balanced power ampifier configuration

After we have finished designing the structure of the single-stage PA we move on to the next step, which is the design of the hybrid coupler.

Due to the branch-line coupler has low-performance and large-coupler is difficult to implement, the Wilkinson power divider has been utilized in this work. The Wilkinson power divider is a type of classical one with a straightforward structure and easy to implement. The circuit has been designed on Epoxy substrate (ε_r = 4.4, tanδ = 0.025) with 1.6 mm thickness. The selection of the substrate should consider its permittivity and its thickness. These characteristics impact on bandwidth. Epoxy (FR4) combines good electrical features, availability and price. In order to verify the accuracy of the design, the simulation has been carried out on an Agilent ADS. The characteristic impedance is 50 Ω and the center frequency is 2.45GHz, then the length and width of the microstrip can be determined using LineCalc tool. The schematic circuit of the Wilkinson divider and simulation results are illustrated in **Figure 6** and **Figure 7**, respectively. The results indicate that the isolation (S_{32}) between port 2 and port 3 is lower than -27dB. The transportation gain (S_{21}) is -4.8dB, the reflection coefficient (S_{11}) of each port is also less than -38dB.

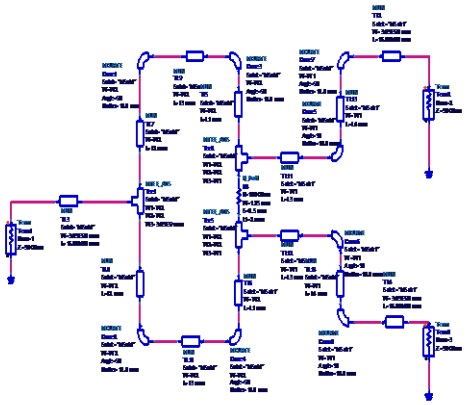


Fig. 6. The schematic circuit diagram of the Wilkinson power divider

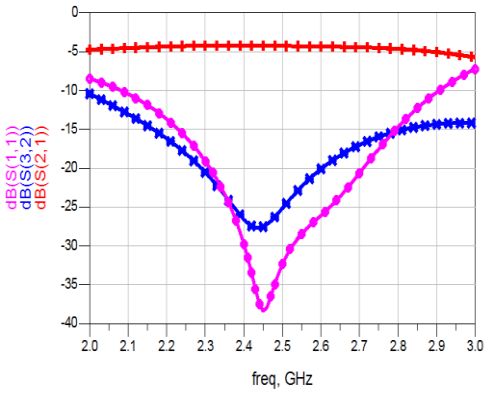


Fig. 7. The simulated results of the Wilkinson power divider

Finally, we mix two parallel single-stage amplifiers and two power dividers. The printed circuit of the balanced power amplifier was implemented in microstrip technology using the Epoxy substrate (FR4) with a relative permittivity of 4.4, a thickness of 1.6mm, a metallization thickness t=0.035mm and a tangential loss of 0.025. For more realizable results, Momentum tool is performed. Layout architecture utilized to show the location of the components and the real circuit size. The layout of the final proposed balanced power amplifier structure is presented in **Figure 8**, where the overall size of this balanced PA is 260 mm (L) * 93 mm (W). The simulation results of the balanced power amplifier circuit are illustrated in **Figures 9–11**.

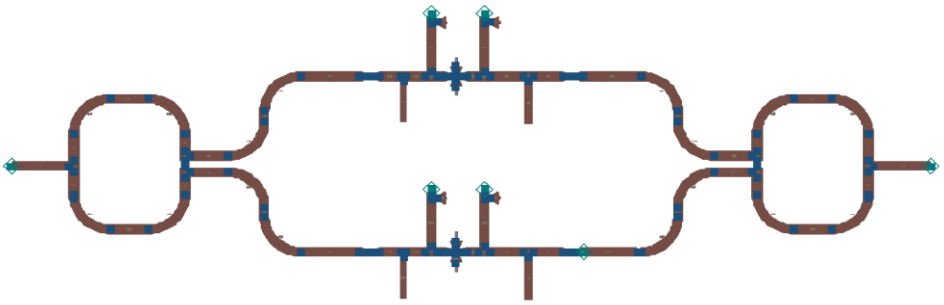


Fig. 8. The complete structure of the balanced amplifier

The S-parameter of the proposed balanced power amplifier is presented in **Figure 9**. For this structure, the power gain is more than 16dB in the required frequency range of 2.40–2.50GHz with an isolation coefficient (S_{12}) is less than -28dB in the operating frequency band. Note that the power gain is dropped by 3dB at the peak-gain frequency for the balanced PA due to the additional loss of the Wilkinson power divider. Whereas the input and output reflection coefficient (S_{11} and S_{22}) of this amplifier are less than −15dB with the input reflection coefficient S_{11} is between -26dB and -38dB, while the output reflection coefficient S_{22} ranges between -15dB and -19dB. It is clear from the result that the proposed balanced amplifier provides good matching impedances.

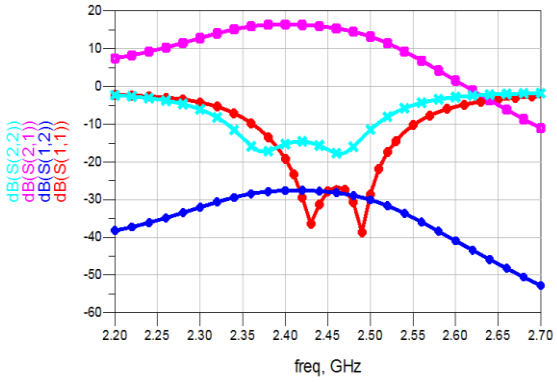


Fig. 9. S-parameters versus frequency of the proposed balanced amplifier.

Thereby this balanced amplifier has a bandwidth equal to 100MHz from 2.40GHz to 2.50GHz. From **Figures 10** and **11**, we can see that the max value of the output power rise up to 16dBm and the stability factor of the amplifier is bigger than 2.1 that signifies the circuit is unconditionally stable.

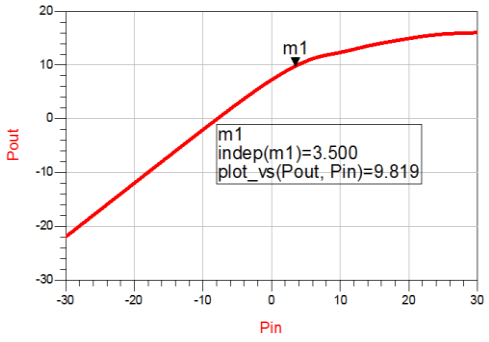


Fig. 10. Input power versus output power for 2.45GHz.

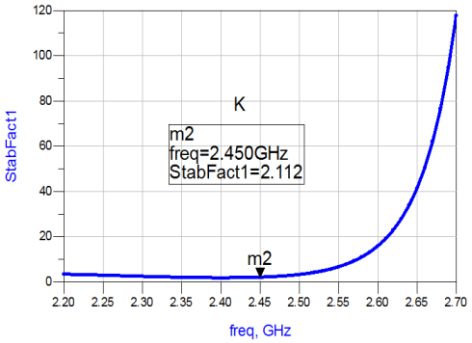


Fig. 11. Stability factor versus frequency characteristics.

After the completed the design of the balanced amplifier, a performance comparison between the proposed amplifier and other designed amplifiers in literature in terms of power gain, reflection coefficients and supply voltage are introduced in **Table 1**. As mentioned in **Table 1**, we can clearly observe that the proposed amplifier achieves a better performance compared to the cited works.

Table 1. Comparison between the proposed amplifier with published literature.

Ref	Technology	Frequency (GHz)	Supply Voltage (V)	Power Gain (dB)	Output Power (dBm)	Input Return Loss (dB)	Output Return Loss (dB)
[16]	GaAs FET	1.10-3.00	3	12	14.8	-25	-17
[17]	CMOS	1.65-2.00	10	10	18.5	-21	-8
[18]	Si BJT	1.75 - 2.15	12	11 ± 0.5	8@1dB	-22	-18
This Work	GaAs FET	2.40-2.50	4	16	16	-38	-19

3 Conclusion

In this study, we have developed a new balanced power amplifier by using a Wilkinson power divider. This balanced amplifier was designed using the microwave simulation Advanced Design System (ADS) software. The proposed amplifier was designed by a microstrip technology and packaged GaAs FET ATF-21170. The final circuit has achieved a wide bandwidth of 100 MHz, with The S_{11} and S_{22} are all less than -15dB, output power up to 16dBm, and 16dB power gain. Furthermore, the proposed balanced amplifier can keep more than 1dB

gain flatness at the whole operation band which shows much better performance than the conventional narrowband amplifier (single-stage amplifier).

Acknowledgments. The authors are thankful to Mr. Mohamed Latrach Professor in ESEO, engineering institute in Angers, France, for allowing us to use electromagnetic solvers available in his laboratory.

References

[1] Ng, H. T., Ziazadeh, R. M., et Allstot: A multistage amplifier technique with embedded frequency compensation. IEEE Journal of Solid-state circuits, vol. 34, no 3, p. 339-347 (1999).

[2] Gillick, M., Robertson, I. D., Et Joshi, J. S: Coplanar waveguide two-stage balanced MMIC amplifier using impedance-transforming lumped-distributed branch line couplers. IEE Proceedings-Microwaves, Antennas and Propagation, vol. 141, no 4, p. 241-245 (1994).

[3] Akkaraekthalin, P., Jongjitaree, S., et Vivek, V: Coplanar waveguide balanced amplifier using bipolar junction transistors and backed ground-plane hybrids. In: TENCON 2001. Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology. p. 732-735 (2001).

[4] Nelson, B. L., Umemoto, D. K., Perry, C. B., et al: High-linearity, low DC power monolithic GaAs HBT broadband amplifiers to 11 GHz. In: Microwave and Millimeter-Wave Monolithic Circuits Symposium, Digest of Papers. p. 15-18 (1990).

[5] Seo, S., Pavlidis, D., et Moon, J. S: A wideband balanced AlGaN/GaN HEMT MMIC low noise amplifier for transceiver front-ends. In: Gallium Arsenide and Other Semiconductor Application Symposium, EGAAS 2005. European. IEEE. p. 225-228 (2005).

[6] Engelbrecht, R. S., et Kurokawa, K: A wide-band low noise L-band balanced transistor amplifier. Proceedings of the IEEE, vol. 53, no 3, p. 237-247 (1965).

[7] Imaoka, T., Banba, S., Minakawa, A., et Imai, N: Millimeter-wave wide-band amplifiers using multilayer MMIC technology. IEEE transactions on microwave theory and techniques, vol. 45, no 1, p. 95-101 (1997).

[8] Duvanaud, C., Dietsche, S., Pataut, G., et Obregon, J: High-efficient class F GaAs FET amplifiers operating with very low bias voltages for use in mobile telephones at 1.75 GHz. IEEE Microwave and Guided Wave Letters, vol. 3, no 8, p. 268-270 (1993).

[9] Rachakh, A., El Abdellaoui, L., et Latrach, M: Design and Analysis of a New Configuration of Microwave Power Amplifier. Emerging Innovations in Microwave and Antenna Engineering, p. 364 (2018).

[10] Tajima, Y., et Miller, P. D: Design of broad-band power GaAs FET amplifiers. IEEE transactions on microwave theory and techniques, vol. 32, no 3, p. 261-267 (1984).

[11] Gonzalez, G: Microwave transistor amplifiers: analysis and design. New Jersey: Prentice hall (1997).

[12] Avago Technologies Datasheet. ATF-21170 0.5–6 GHz Low Noise Gallium Arsenide FET.

[13] Rachakh, A., El Abdellaoui, L., Zbitou, J., Errkik, A., Tajmouati, A., et Latrach, M: A Novel Design of a Microstrip Microwave Power Amplifier for DCS Application using Collector-feedback Bias. International Journal of Electrical and Computer Engineering (IJECE), vol. 8, no 3, p. 1647-1653 (2018).

[14] Kassim, S., et Malek, F: Microwave FET amplifier stability analysis using geometrically-derived stability factors. In: Intelligent and Advanced Systems (ICIAS), 2010 International Conference on. IEEE, p. 1-5 (2010).

[15] Rachakh, A., El Abdellaoui, L., Errkik, A., Tajmouati, A., et Latrach, M: A New Configuration of a Microstrip Power Amplifier Using Si-BJT in Class-A Mode for DCS Applications. In: Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems. ACM. p. 58 (2017).

[16] Ribate, M., Zbitou, J., Mandy, R., Erkik, A., et Latrach, M: Broadband GaAs FET Power Amplifier for L and S Bands Applications. The International Journal of Intelligent Engineering and Systems (IJIES), vol. 11, no 5, p. 96-104 (2018).
[17] Aniktar, H., Sjoland, H., Mikkelsen, J. H., et Larsen, T: A Class-AB 1.65 GHz-2GHz broadband CMOS medium power amplifier. In: NORCHIP Conference. 23rd. IEEE. p. 269-272 (2005).
[18] Rachakh, A., El Abdellaoui, L., Zbitou, J., Errkik, A., Tajmouati, A., et Latrach, M: A Novel Configuration of a Microstrip Microwave Wideband Power Amplifier for Wireless Application. Telkomnika, vol. 16, no 1, p. 224-231 (2018).

A Novel Zero Bias Microstrip MESFET Power Limiter

Khalifa Echchakhaoui[1], Elhassane Abdelmounim[1], Hamid Bennis[2], Mohamed Latrach[3]
{k.echakhaoui@gmail.com, abdelmou@hotmail.fr, hamid.bennis@gmail.com}

ASTI Laboratory, FST of Settat, Hassan 1St University Settat, Morocco[1], TIM Research Team, EST of Meknes, Moulay Ismail University, Meknes, Morocco [2], RF-EMC Group, ESEO, Angers, France[3]

Abstract. In this article, a novel limiters power is designed and validated by ADS software. The new limiter is based on a mictrostrip circuit and uses MSFET transistors and Schottky diode as actives components. This Power Limiter have been optimized in two steeps: the first circuit is composed of one stage. Simulation of this circuit presents some limitation in termes of limitation rate. To improve this performance, a new circuit composed of two stages is simulated and optimized. The final circuit exhibits 25 dB of limitation rate while insertion loss is -1 dB with a threshold input power level of 0 dBm until a maximum input power level of 30 dBm.

Keywords: Microstrip, Power limiter, MESFET transistor, Schottky Diode.

1 Introduction

Currently, various types of semiconductor diode power limiters are used to improve the reliability and protection of communication and radar systems implementing high power microwave sensitive electronic components [1]–[4].

Generally, Power Limiter designs incorporate a planar-mounted PIN diodes network or Schottky diodes [5]–[8]. Figure 1 presents the classical topologies found in several scientific literatures on circuit protectors [5]–[7], [9]–[11]. These limiters can be optimized for operating frequency ranges and power levels.

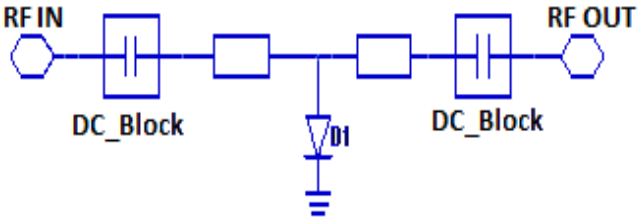


Fig. 1. Classic Power Limiter Topology

In this article, a new limiter power is designed and validated using a transistor and two Schottky diodes. The transistor operates as variable impedance that absorbs the extra power over the threshold supported by the load. While the Schottky diode functions as a signal detector and generates the DC electric current needed to bias the transistor and reducing the power threshold limit to activate the protective mode of the power limiter.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284090

Depending on the mode of supply of the power limiters, there are two types: the active power limiters that require an external current to operate and the passive power limiters that are zero biais. In this article, the proposed design is a new passive power limiter that does not require DC current.

In the next section, the theory approach and the description of the proposed circuit will be exposed, in the validation section, the s-parameters and Harmonic simulation results will be exposed and discussed.A comparison between the proposed circuit and other circuits in literature will be presented and commented at end.

2 Research Method

2.1 MESFET transistor operation

The structure of the Schottky contact field effect transistor is based on an active layer (channel) directly implanted in the semi-insulating substrate [12]–[14]. Figure 2 shows the cross section of a MESFET transistor. This device is composed of an insulating layer of GaAs above which, a layer of N-doped semiconductors above, a metal port called gate is connected directly to this layer. The source and drain ports are two negatively doped semiconductor layers. The source and drain contacts have ohmic contacts in contrast to the Schottky gate contact. The length of the metal gate is typically 0.5 to 1.0 µm for discrete transistors but it may be 0.2 µm for integrated circuits [10].

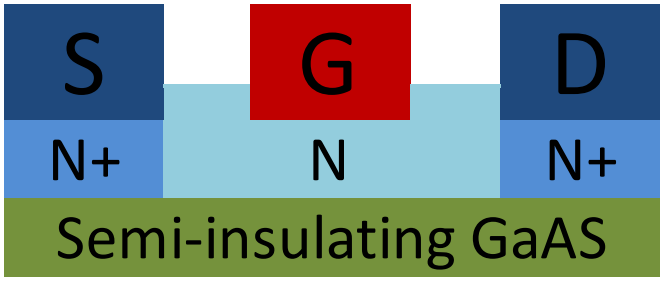


Fig. 2. Structure of MESFET transistor

In addition to their use as active elements in amplifiers, MESFET Transistors are used as active devices in variable attenuators, switches and phase shifters [10], [12], [15], [16]. The use of MESFET transistors in solid state passive power limiters is another application of this device [1], [6], [17], [18].

For most control applications, the MESFET is switched between a low impedance state, or a high impedance state, or pinchoff state corresponding respectively to the linear, no-linear and cutoff regions of this component. This switching is achieved by applying a negative voltage on its Gate port.

Russell Gaspari and Harold Yee [19] report four reasons to prefer FET technology over PIN diode in microwave control circuit:

1. The FET biasing is isolated from RF line and needs minimal RF Choke
2. Possibility to achieve insertion gain rather than insertion loss

3. Switching from On to Off consume very low energy
4. The switching speed is better in the FET than the PIN diodes

The basis of operation of a MESFET is the ability to modulate the thickness of the channel via the gate [12]. A depopulated layer of free electrons, called the space charge zone (SCZ), is created under the gate. No current can cross this layer. The region where the current can flow is therefore reduced to the fraction of the active layer that is not depopulated. In normal operating mode the drain is positively polarized with respect to the source, while the gate is always biased negatively, with respect to the source.

The operation of a MESFET is summarized by the following steps [12], [14]:

At a fixed drain voltage, the negative bias of the gate has the effect of increasing the penetration of the space charge zone (SCZ) in the active layer, thereby decreasing the flow of current. When the gate voltage is sufficiently negative, the SCZ completely close the channel, no longer allowing the current to flow. The transistor is then said "pinched off" and the voltage applied to the gate is then called clamping voltage or pinchoff voltage Vp.

In high power, Schottky diode will generate a DC rectified current. The orientation of the schottky diode relative to the gate port of the MESFET transistor, will allow in one direction to bias the gate by a negative voltage.

2.2 Use of MESFET transistor as limiter devices

The realization of MESFET-based power limiters can be achieved in several ways [7]: The first way is presented in Fig. 3. The MESFET transistor is connected parallel to load with an amplitude detector such as a schottky diode [7]. The rectified DC voltage is generated when the amplitude of incident signal is detected. This voltage is applied to the MESFET gate and its impedance is modified. As a result, the output RF amplitude is decreased.

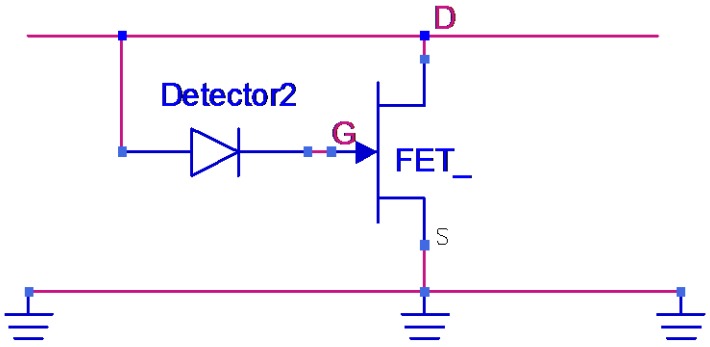


Fig. 3. Use of MESFET parallel to the load

The second way is to connect the MESFET transistor in series [18] with the transmission line as shown in Fig. 4. When the incident power is small, the MESFET has low impedance and therefore the insertion loss of the limiter is negligible. Increasing the power of the input signal provides the voltage of the Schottky barrier junction created by the gate of a MESFET. When this voltage exceeds the static threshold of the Schottky barrier junction, the direct current begins to flow. This provides the negative gate bias voltage across the gate ballast resistor.

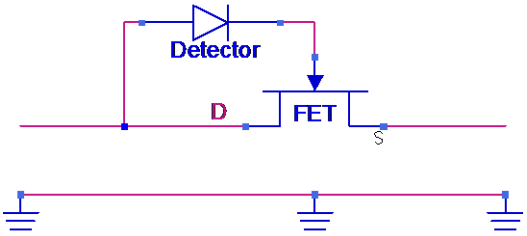


Fig. 4. Use of MESFET in series

In this proposed work, The MESFET Transistors is used in a new way as depicted in Figure 5. The MESFET is implanted as variable resistance R_{GS} and R_{GD} controlled by the voltage applied in the Gate port.

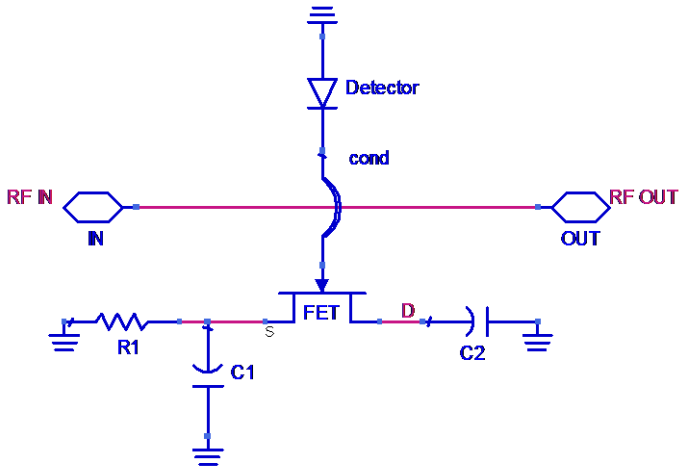


Fig. 5. Use of MESFET as variable resistance

3. Design Description

The proposed limiter is a passive power limiter based on microstrip line. One of them is a linear line that transmits the main signal to the load and the second line is derivated and connected at its end to Schottky diode and MESFET Transistor. The second line is used to divert a portion of incident signal when the amplitude of the signal reaches the threshold of the diodes. The Schottky diode operates as voltage controlled attenuator and current detector. Figure6 presents the proposed architecture. Commercial devices are used in this circuit:

- Schottky diode HSMS286B made by Avago technologies [20]
- MESFET ATF21170 made also by Avago technologies [21]

The commercial diode Avago HSMS286B has used in this circuit is a Surface Mount RF Schottky diode presented in SOT23/143 package[20] and suitable for incident power above -20 dBm. The HSMS286x family has been designed for RF applications [17], such as:

- DC biased small signal detectors to 1.5 GHz.
- Biased or unbiased large signal detectors (AGC or power monitors) to 4 GHz.
- Mixers and frequency multipliers to 6 GHz

The equivalent impedance of this Schottky diode, as depicted in Figure7, is modeled as a combination of the variable junction resistance (Rj), the parasitic capacitance and parasitic series resistance (Rs).

The ATF-21170 is a high performance MESFET GaAS transistor designed for use in low noise or medium power amplifier applications in the frequency range 0.5-6 GHz [21].

This GaAs FET device is housed in an hermetic package with a gold-based metalization. This device has a nominal gate length of 0.3 micron with a total gate periphery of 750 microns. AT21170 can dissipate a power of 27dBm at 25°C and its channel temperature can reach 175°C [21].

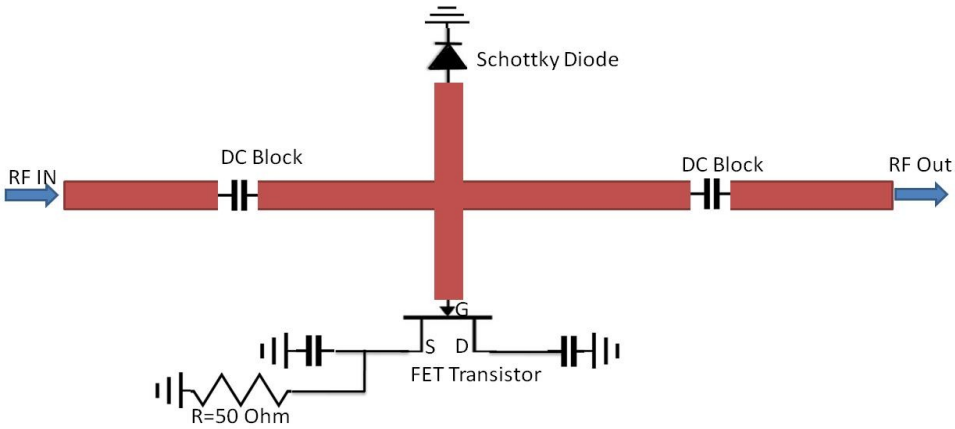


Fig. 6. Power Limiter circuit

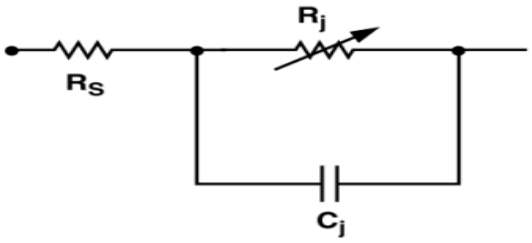


Fig. 7. Linear Equivalent Circuit Model of Schottky Diode

The designed circuit has been optimized in two steps by improving in each step the limiting rate while maintaining the insertion loss as minimum as possible.

The first circuit consists of a single stage identical to Figure 3: a stage is composed of an anti-parallel line to the main line with MESFET transistors and a Schottky diode at its ends.

In order to improve the limiting rate of the power limiter, the final circuit is obtained by joining in series two circuits. This final circuit has been optimized to keep the insertion loss as low as possible while doubling the limit rate.

4. Circuit validation by simulation and discussion

The designed circuit was validated in simulation by ADS software to check s-parameters response and harmonic simulation response of the circuit. For the simulation we use a standard FR4 substrate with a relative dielectric constant value of 4.4 and a dielectric Loss tangent value of 0.025 and a thickness of 1.6 mm. The choice of FR4 as substrate is dictated by its availability in the laboratory and its cheap cost.
As the circuit has been optimized in two steps, the first ciruit is constituted of one stage. The second circuit is formed by two stages. For each steps, the layaout schematic, the s-parameters results and harmonic simulation results will be presented and discussed.

4.1 One stage circuit

The layout of the first circuit studied is presented in Figure 8.

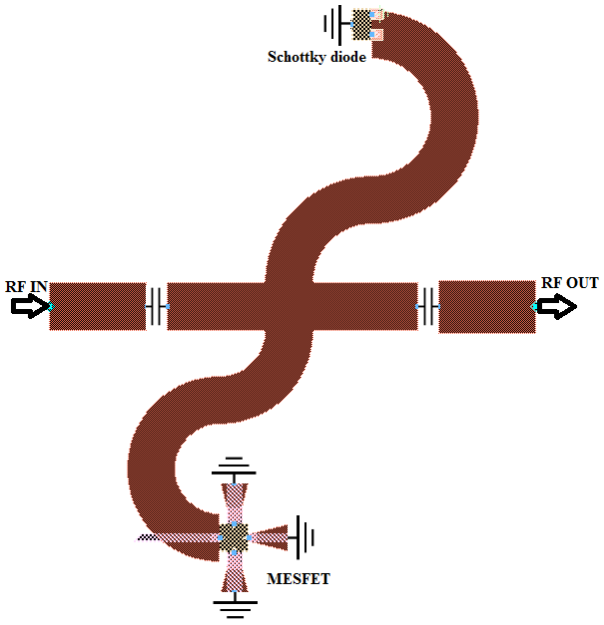


Fig. 8. Layout of one stage circuit

Figure 9 presents the s-parameters simulation results. As shown in the figure 9, the circuit presents good matching input impedance in the ISM band. The insertion loss is around -0.9dB over [3, 3.6 Ghz].

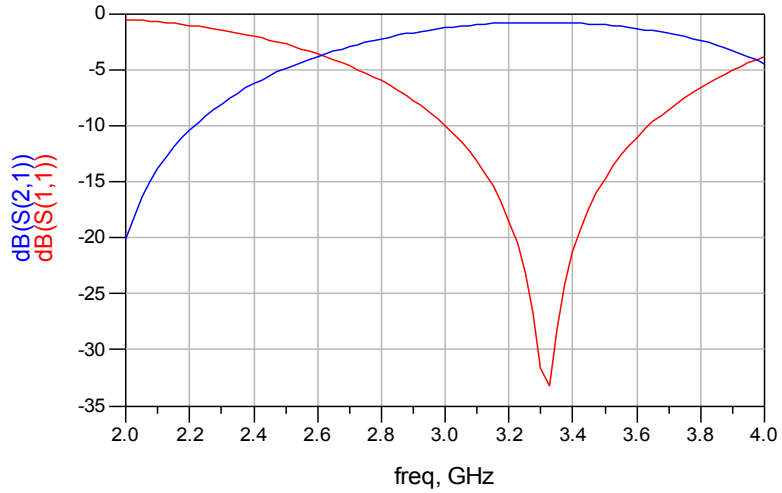


Fig. 9. S-parameters simulation results for S11 and S21

Figure 10 shows the output power versus input power. This circuit presents a low isolation at large signal (12 dB at Pin=30 dBm). These Values of protection will not be considered sufficient in many applications. In order to improve the performances of this circuit, we have simulated a second circuit with two stages. The layout of the second optimized circuit is presented in Figure 11:

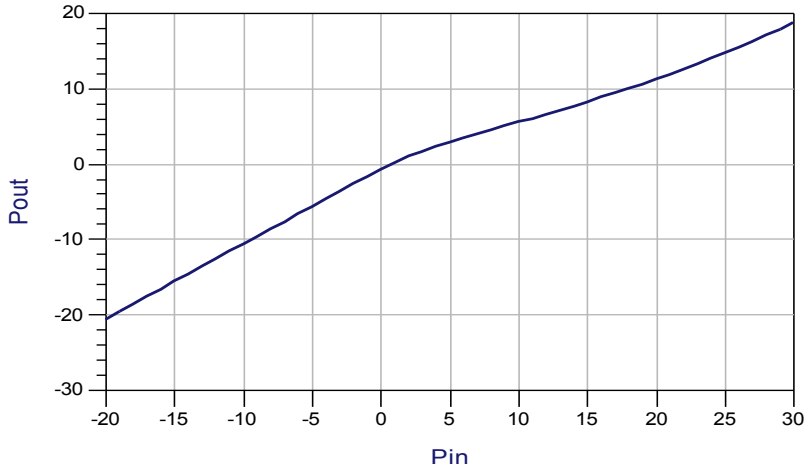


Fig. 10. Output Power versus input power

4.2 Circuit with two stages

The second circuit is obtained by adding a second stage. However, in order to reduce circuit dimension, the antidopal line is replaced by a half circle line. The final circuit is presented in the Figure 11:

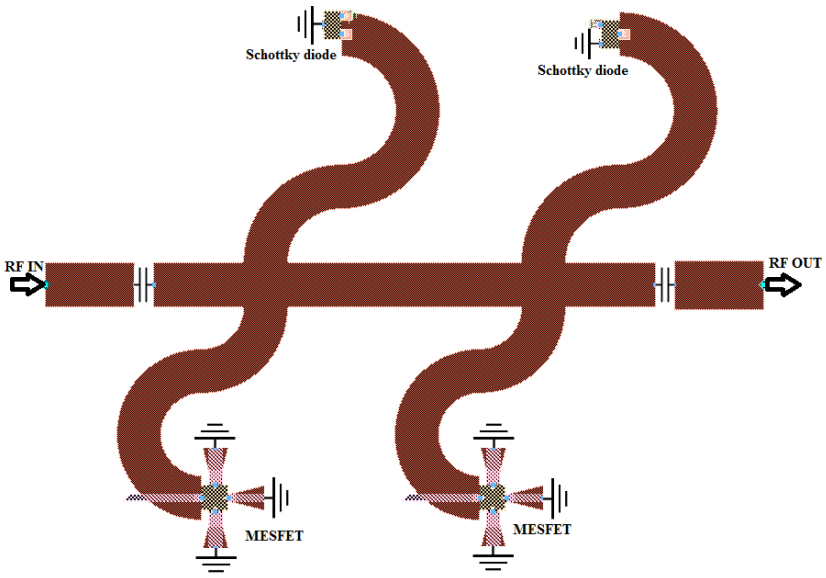


Fig. 11. Layout of two stages circuit

The simulation results of this circuit are presented in Figure 12 (S-parameters results) and in Figure 13 (Output Power Versus Input Power).

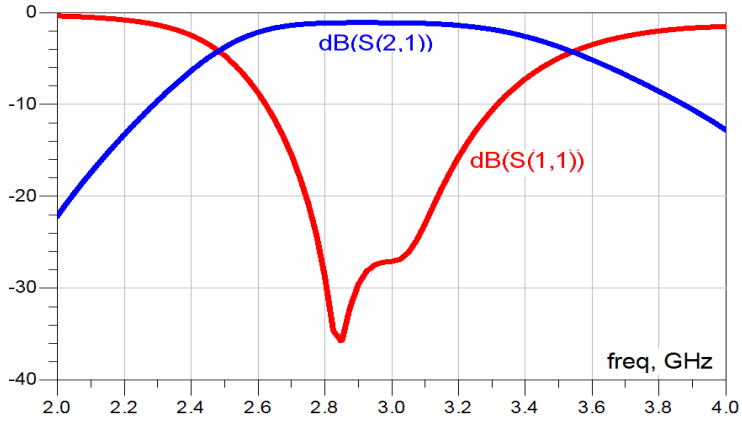


Fig. 12. S-parameters simulation for S11 and S21

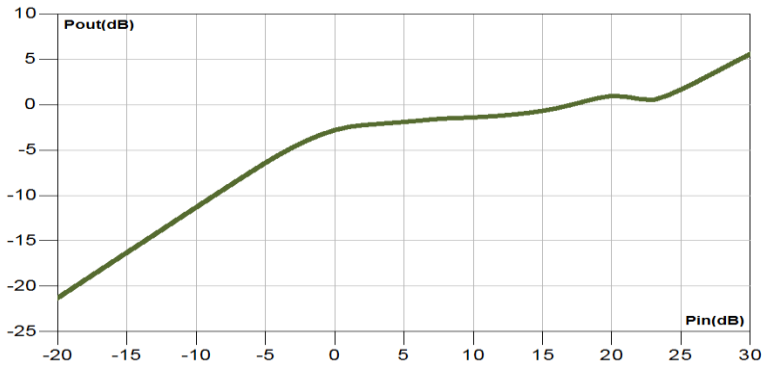


Fig. 13. Output power versus input power

Compared to one stage circuit, the two stages circuit shows better performance in isolation rate:

- Broadband frequency range of 1Ghz from 2,6 to 3,4Ghz
- Low insertion loss (-1 dB)
- The isolation rate reaches 25 dB at 2.7 Ghz.

Also the circuit presents a good stability of limitation rate across the broadband range.

5. Comparaison between the designed circuit and similar circuit in letterature

By comparing the proposed circuit against power limiter topologies constructed on a microstrip line published in the literature, the designed circuit presents an efficient and simple solution for a passive power limiter based on MESFET transistors and Schottky diodes. The table below relates different technical characteristics of the proposed limiter with those of the scientific literature.

Table 1. Comparison of microwave power limiter.

Design	Insertion loss (S21)	Frequency range	Limiting rate	References
Microstrip Passive limiters using discrete MESFET and Schottky diodes	1 dB	3 Ghz	25 dB	The proposed limiter
Planar Schottky diode	1 dB	1 Ghz	20 dB	[6]
Planar Schottky diode and MESFET based limiter	1 dB	1 Ghz	15 dB	[6]
Pin diode limiter (3 stages)	1.3 dB	Ku band	29.7 dB	[11]
Passive limiters using Discrete MESFET and Schottky diode	0.9 dB	7 Ghz	15 dB	[18]

In addition, the designed circuit shows important advantages in comparison to the state of art presented in literature:

- The circuit presents simple construction by avoiding the need of external DC current.
- The use of zero bias diodes reduces Johnson noise caused by DC current and improves reliability of the circuit.
- MESFET transistors present faster switching and RF power detections thus enhance circuit response to high power pulses.

6. Conclusion

In this paper, we have designed and validated by simulation a power limiter circuit based on zero bias MESFET transistors and Schootky diodes.
The first circuit based on one limiter stage achieves 12dB of isolation. The final optimized circuit achieves 25 dB of isolation while the insertion loss at low signal remains acceptable (around -1 dB). The circuit dimensions for the proposed frequency are: 48x38 mm². In comparaison to other power limiters in literatures, the proposed circuit presents the best performance in terms of limitation rate.

References

[1] Bahl I. J.: 10W CW broadband balanced limiter/LNA fabricated using MSAG MESFET process.International Journal of RF and Microwave Computer-Aided Engineering, vol. 13, no. 2, pp. 118–127, (Mar. 2003).

[2] Echchakhaoui K., Abdelmounim E., and Bennis H.: Microwave Power Protectors: Attenuators and Limiters. in Handbook of Research on Advanced Trends in Microwave and Communication Engineering. IGI Global. pp. 244–282 (2017).

[3] Echchakhaoui K., Abdelmounim E., Zbitou J., Tajmouati A., Bennis H., and Angel M.: An evolved design of a Zero Bias broadband microstrip Power Limiter. in 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), Fez, Morocco, pp. 1–4(2017).

[4] Xin Z. Y., Han S. X., Luo Y. L., Xu L., and He J.: A compact broadband high-power PIN limiter using LC network. in 2013 International Workshop on Microwave and Millimeter Wave Circuits and System Technology, Chengdu, China. pp. 336–338(2013).

[5] Maloratsky L. G.: Passive RF&Microwave integrated circuits, Newnes. Elsevier Inc.pp. 385(2004).

[6] Chaturvedi S., Saravanan G. S., Bhat K. M., and Bhalke S.: MESFET process based planar schottky diode and its application to passive power limiters. in IEEE MTT-S International Microwave and RF Conference, New Delhi. pp. 1–4. india(2013)

[7] Zbitou J., Latrach M., and Toutain S.: Wide band rectenna with high sensitivity detection. 30 June-2 July 2004, Granada, Spain (ESA SP-567, December 2004). in Proc. Of the 4th Int. conf. on Solar Power from Space- SPS'04, together with the 5th Int. Conf. on Wireless Power transmission-WPT 5, Granada, Spain(2004).

[8] Coaker B. M. ,Dowthwaite D. M.: Planar limiters and receiver protectors. in 2007 European Radar Conference. pp. 405–408(2007).

[9] Sundararajn R., Peterson E., Nowlin R.: Attenuator Design Reference Manuals. in ENCYCLOPEDIA OF RF AND MICROWAVE ENGINEERING, Kai Chang, Editor., vol. 1, 6 vols., John Wiley & Sons, Inc., Publication, pp. 452–480.

[10] Bahl I. J., Control components using Si, GaAs, and GaN technologies. Boston: Artech House, (2014).

[11] Yang S., Kim T., Kong D., Kim S., Yeom K.: A Novel Analysis of a Ku-Band Planar p-i-n Diode Limiter.IEEE Transactions on Microwave Theory and Techniques, vol. 57, no. 6, pp. 1447–1460, (Jun. 2009).
[12] Golio J. M.:RF and microwave semiconductor device handbook. Boca Raton: CRC Press, 2003.
[13] Noll A.,Stuble W.: A broadband receiver protection limiter for FET based integrated circuits. in 2015 10th European Microwave Integrated Circuits Conference (EuMIC), Paris, France. pp. 164–167(2015).
[14] Gutmann R. J.,Fryklund D., and Menzer D.: Characterization and Design of GaAs MESFETs for Broadband Control Applications. in 1986 IEEE MTT-S International Microwave Symposium Digest. pp. 389–392(1986).
[15] White J. F., Microwave Semiconductor Engineering. Dordrecht: Springer Netherlands, 1982.
[16] Zagorodny A. S.,Voronin N. N., Yunusov I. V., Goshin G. G.,Fateev A. V.,Popkov A. Y..: Microwave microstrip attenuators for GaAs monolithic integrated circuits. in International Conference and Seminar of Young Specialists on Micro/Nanotechnologies and Electron Devices, Altai, Erlagol, pp. 67–71. Russia (2012)
[17] Katz A., Dorval R..: Characteristics and modeling of passive MESFETs for non-linear control applications. in 1994 IEEE MTT-S International Microwave Symposium Digest (Cat. No.94CH3389-4), 1994, pp. 437–440 vol.1 (1994).
[18] Drozdovski N. V.: Microwave Passive Power Limiters Based On Mesfets.Journal of Microwaves and Optoelectronics, Vol. 1, No. 2, pp. 11(April 1998).
[19] Gaspari R. A., Yee H. H.: Microwave GaAs FET Switching. in 1978 IEEE-MTT-S International Microwave Symposium Digest, 1978, pp. 58–60.
[20] Avago Technologies, HSMS-282x, Surface Mount RF Schottky Barrier Diodes. Avago Technologies, US(2014).
[21] Hewlett-Packard: Data Sheet of MESFET AT21170 0.5–6 GHz Low Noise Gallium Arsenide FET. Hewlett-Packard Co.US(1998).

Broadband Solid State GaAs Power Amplifier for L and S Bands Applications.

Mohamed Ribate[1], Rachid Mandry[1], Larbi El Abdellaoui[1], Ahmed Errkik[1], Mohamed Latrach[2]

{ [1]ribate.mohamed@gmail.com }

[1]LMTEE, FST of Settat, Hassan 1st University, Settat, Morocco.
[2] Microwave Group ESEO Angers, France

Abstract. This paper proposes the research of a single stage Broadband Solid State Power Amplifier (BPA) based on ATF13786 transistor, using GaAs process. This BPA operates in the frequency band ranging from 1.35 GHz to 3 GHz which covers the mainstream communication standards running in L and S Bands. The design approach is based on the real frequency technique. The simulated results are obtained by using ADS circuit simulator. In order to improve the broadband response, a multi-section quarter wave impedance transformer combined with an approximate transformation of a previously designed lumped elements into a transmission lines are used. The simulation results show a saturated output power of 17.85 dBm, with PAE of the 11% and a maximum power gain of 16.68 dB.

Keywords: Broadband Power Amplifier, Real Frequency Technique, Solid State, GaAs.

1 Introduction

At the present, almost every communication system has some sort of transceiver, and inherently a Radio Frequency Power Amplifier (RFPA), running in one of the following bands (defined by IEEE standard 521-1984): L – S – C – X or Ku bands, from 1 GHz to 18 GHz, operating at Ultra-High Frequencies (UHF) and Super-High Frequencies (SHF) [1-7].

Nowadays, power amplifiers are used everywhere in our life, we are driving cars that are smarter, equipped with new technologies such as parking radar sensor system and GPS. We receive signals from satellites in our home with satellite TV receivers and in our cars or phones from Global Positioning System (GPS) [8-11].

Essentially, Power Amplifiers are critical components for many applications, among which communication systems, Magnetic Resonance Imaging (MRI), semiconductor manufacturing, microwave heating, radars, avionics, medical microwave imaging, electronic warfare and navigation systems represent just a few examples [12-17]. Given such enormously spread fields, power amplifiers sets may greatly differ in technology, design and operating requirements [18]. As a consequence, a widespread architectures of PA have been realized in the past 2 decades, from travelling-wave tube power amplifiers for very high power systems such as radars or satellite payloads, to solid state power amplifiers for medium and low power systems like wireless communication handsets [19].

On the other hand, the improvement in broad bandwidth demands is primarily driven by the next generation of communication systems, which require significantly wide bandwidth. As a result, the interest in the power amplifiers with broad bandwidths and with good performance

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284102

has grown immensely over recent years. Such systems involve stringent requirements from broadband amplifiers in terms of high gain, high stability, high Power Added Efficiency (PAE) and high output power over a wide frequency ranges [20].

The design of broadband power amplifiers is basically equivalent to build an appropriate circuit able to convert DC power supplied to microwave transistor, into Microwave/Radio Frequency power, successively transferred to an external load, in a given band or a range of frequencies of interest [21].

Various approaches, configurations and techniques have been developed and explored for this purpose, trying to meet at the same time several contrasting requirements in terms of gain flatness *vs* broadband matching, efficiency *vs* linearity or low distortion *vs* high output power. Among such configurations of BPA, balanced structures, distributed structures and Doherty configurations that use various matching network schemes as well as reactive filter synthesis, shunt or series feedback and resistive/reactive matching [22].

The reactive/resistive matching technique gives a virtuous power performance and a moderate bandwidth. Thus, the wideband design is difficult. This scheme also has poor VSWR and gain flatness. The shunt/series feedback topology results in lower power density and higher noise figure. However, this configuration is good for VSWR, multi-octave bandwidth and gain flatness. The distributed power amplifiers have the capability of broad bandwidth and achieves excellent gain, but shows weak points in limitation of output power, poor PAE and requiring large area for circuit realization [23].

The work presented in this paper is basically based on the Real Frequency Technique (RFT), in combination with other broadband matching techniques such as multi-section quarter wave impedance transformer, approximate transformation of previously designed lumped element matching circuits into the transmission line matching networks and both single and double stub tuners.

Comparing the matching techniques adopted in this work, with similar contemporary state of the art broadband matching techniques, the RFT is probably the best circuit design available for the design of broadband power amplifiers. The RFT has being pertinent for all matching problems and permits optimizing many performance factors of single or multi-stage RF power amplifiers [24].

In this paper we present the design of a broadband solid state power amplifier, operation in the frequency range from 1.35 GHz to 3 GHz which compromise the mainstream applications and communication standards running in L and S bands. The proposed BPA reaches a saturated output power of 17.56 dBm and a maximum power gain of 16.85 dB. The design architecture achieves a simple configuration with an excellent matching performance over 1650 MHz bandwidth.

2 Broadband Power Amplifier circuit design

Regardless of its characterization as narrowband, wideband, high power, low noise or else, all power amplifiers have in common feature of converting the DC power supplied to their active devices, into RF power as well as providing a positive power gain at the requested frequency or rang of frequencies.

The design of Broadband Power Amplifier requires an accurate active device, appropriate bias and matching networks, stability during operation and easiness of circuit implementation. The effectiveness of the design approach is evaluated by its aptitude to achieves high output

power, maximum power gain across a broad bandwidth with minimum stages of PA and under stable operating conditions. The exigency for linearity or high efficiency can be reviewed where it is needed. Keep in mind that, for the stability requirement, it is indispensable to examine the frequency ranges in which the RF transistor may be unstable.

The simplest possible single stage power amplifier consists of a single RF transistor –be it a BJT, FET or otherwise – DC supplied through a DC bias network, and connected through an input and output matching networks, to a source and load respectively, both having same impedance Z_0, as shown in Figure 1.

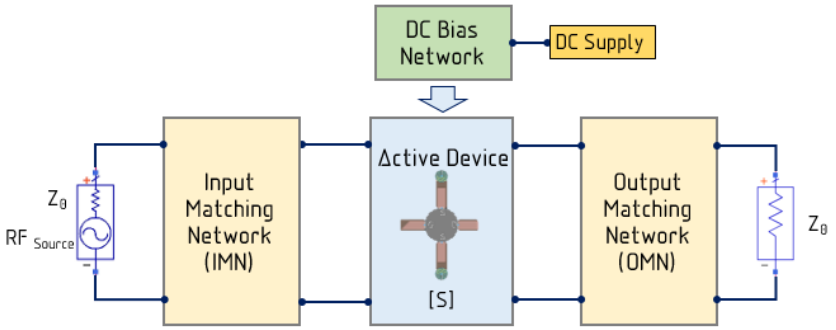


Fig. 1. Single Stage Power Amplifier Block Diagram

At high frequencies, to provide maximum power transfer of the RF power available from the RF source to the load, an accurate input and output matching networks are required as well as a specific DC bias.

2.1 Biasing circuit of the BPA

At high frequencies, biasing networks takes an important parts of power amplifiers design, and they must be designed carefully to avoid low frequency oscillations due to high device gain, high capacitance and low substrate resistivity. Despite the biasing parameters depends on the active device technology, the design approaches of biasing circuit are basically the same [25].

In practice, there are many basing schemes available for RF power amplifiers. According to the designer choice, the biasing circuits could be integrated in the matching circuits or they stand-alone.

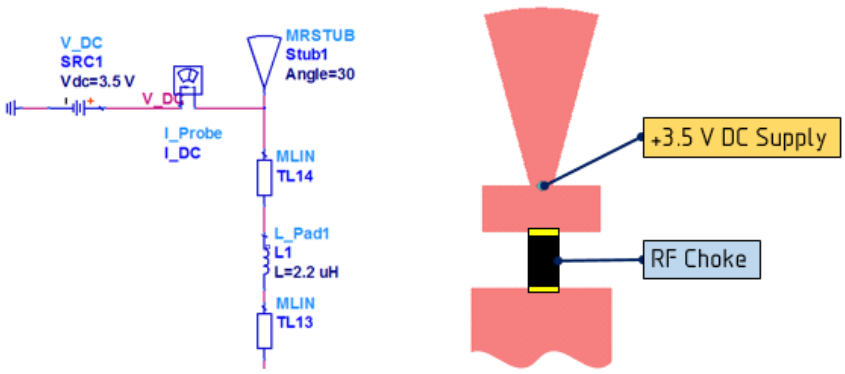


Fig. 2. The proposed microwave biasing circuit.

However, when the DC bias voltage supplied to the active device, the RF energy must not break through the bias port, otherwise degradation of RF performance in terms of PAE, noise figure and power gain takes place.

Basically, a biasing circuit consists of RF choke and a DC block. The DC block can be either a 3-dB backward –wave coupler or chip capacitors, and must be present a short circuit at the operating frequency. The value of the capacitor must be selected in such a way that $\omega C = 2\pi f C$ is very large.

The configuration of the proposed biasing network is illustrated in Figure 2. The active device -ATF786- is supplied with a single supply voltage VCC of +3.5V. the biasing network is synthesized by using a radial line section terminated by a RF choke. This configuration is smaller in size and provides better bandwidth than λ/4 open circuited stub.

2.2 Broadband matching networks design

The impedance matching networks takes a major part in the design process of the BPA. Basically, the active device is mismatched with the source and the load. Consequently, the RF energy available from the source is not delivered to the load, and this makes a loss of RF energy. In other words, to provide a maximum power transfer of the RF energy from the input port to the active device, from the active device to the output port or between devices, an accurate matching circuit should be designed at the interface between two sections having mismatched impedances, otherwise, a portion of the electrical signal propagated in the circuit will be reflected. In a broad sense, the matching network can be defined as a circuit that transforms a given impedance into a desired impedance.

Essentially, the matching network design depends on the operating frequency of the PA. According to the broadband design approach, there are various impedance matching schemes and techniques that can be deployed to match two given mismatched impedances, regardless of whether the matching is accomplished with transmission lines or lumped elements.

The selection of an appropriate matching technique might not always be clear, however, it is basically determined by the following factors: bandwidth, complexity and implementation.

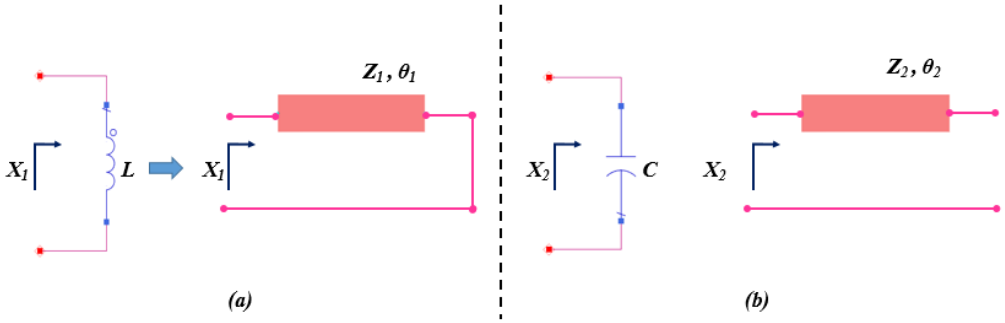


Fig. 3. Equivalence between lumped elements and transmission lines.

In this work, a combination of matching techniques is deployed to achieve the broadband bandwidth of the proposed BPA. The first one is an approximate transformation of lumped elements into transmission lines. At high frequency, excepting an MMIC (Monolithic Microwave Integrated Circuit) design, the implementation of lumped elements such as capacitors and inductors is very difficult where they can be considered as distributed elements. Thus, in this work, the design of the input and output matching networks is based on an

approximate transformation of lumped elements into transmission lines, which can be established by applying Richard's transformation.

Figure 3 shows the equivalence between lumped elements and transmission lines. The value of Z_1, $\theta 1$, Z_2, $\theta 2$ can be obtained by using the following equations:

$$X_1 = Z_1 \tan(\theta_1) \approx \omega L \quad (1)$$

$$X_2 = \frac{Z_2}{2\tan(\theta_2)} = \frac{1}{\omega C} \quad (2)$$

From the equation (1) and (2), the variables are Z_1, Z_2, $\theta 1$, $\theta 2$. As a results, there are two degrees of freedom in implementation of a capacitor or inductor respectively.

The second technique used in the design of the matching networks is the Multi-Section Quarter Wave Impedance Transformer (MSQWIT). Typically, there are two widely known MSQWIT: Chebyshev and Binomial. In this paper we adopt the second transformer. A schematic diagram of a binomial transformer is shown in Figure 4.

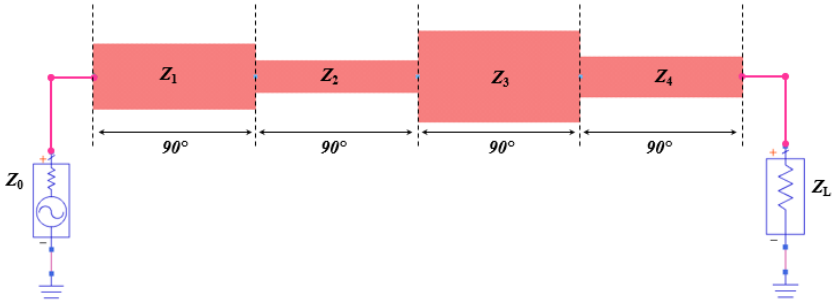


Fig. 4. four-section quarter-wave impedance transformer.

The sections impedances can be calculated approximately by using the following equations:

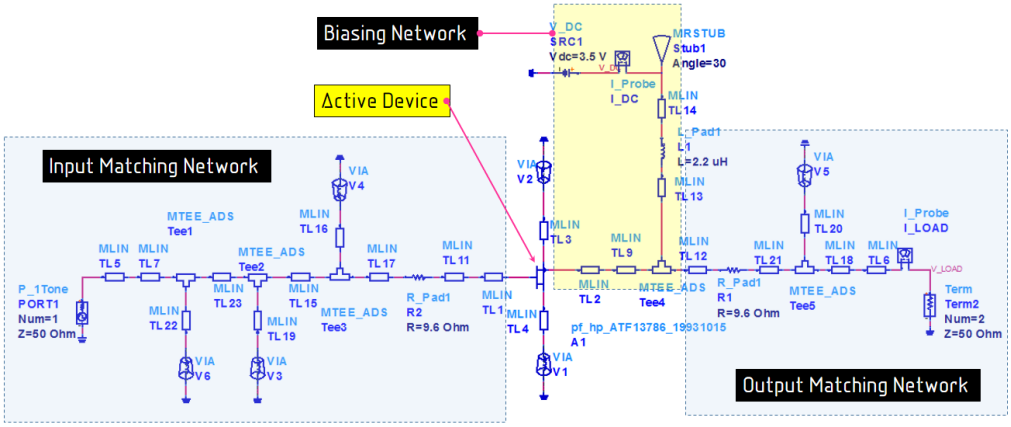


Fig. 5. The proposed single stage BPA.

$$ln\frac{Z_{n+1}}{Z_n} = 2\rho_n = 2^{-N} C_n^N ln\frac{Z_L}{Z_0} \quad (3)$$

$$C_n^N = \frac{N!}{(N-n)!\,n!} \tag{4}$$

Where C_n^N are the binomial coefficients. Z_n and Z_{n+1} are the impedances of the *nth* and *(n+1)th* sections and ρn is the reflection coefficient of the junction between Z_n and Z_{n+1}.

By using the concepts mentioned above, the proposed BPA schematic is illustrated in Figure 5.

3 Simulation results and discuss

3.1 Stability analysis

According to the stability criterion of PA, the unconditionally stability is achieved when the following Rollett and Bodway equations are simultaneously validated:

$$k = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{12}S_{21}|} > 1 \tag{5}$$

$$B_1 = 1 + |S_{11}|^2 + |S_{22}|^2 - |\Delta|^2 > 0 \tag{6}$$

Where

$$\Delta = S_{11}S_{22} - S_{12}S_{21} \tag{7}$$

K is the Rollett stability factor and B_1 is the Bodway stability factor. As shown in Figure 6, the stability requirement of the proposed BPA is established over the operating bandwidth.

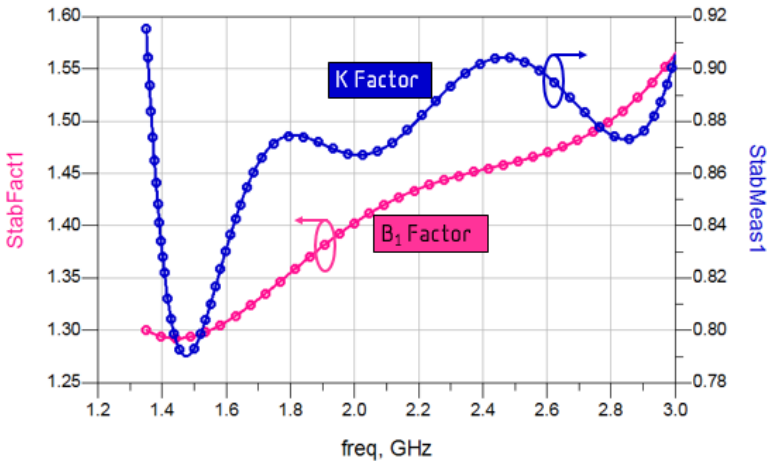


Fig. 6. Curves of K and B_1 Factors versus frequency.

3.2 Small Signal Simulation

The RF performance of the proposed BPA is examined using small signal two port scattering parameters simulations which consists of measuring the reflected, transmitted and

incident waves. We note that the S-parameters simulations are done in 50 Ω source and load impedances. Figure 7 shows the curves of input and output return loss over the bandwidth ranging from 1.35 GHz to 3 GHz. As we can see, the maximum value that can be reached by the input return loss S_{11} is -13dB, where the output return loss S_{22} reaches a maximum value of -15dB over the operating frequencies. As a result, the good broadband matching of the proposed BPA is performed.

The internal feedback of the active device is represented by the parameter S_{12}. In order to improve the degree of isolation between the output and input as well as the stability of the proposed BPA, it is more required to have S_{12} as small as possible.

The simulated revers isolation S_{12} and small signal gain S_{21} are illustrated in Figure 8. As we can observe, the revers isolation S_{12} reaches a maximum value of -20 dB. The small signal gain S_{21} varies between a maximum value of 16.78 dB and a minimum value of 10 dB. Then, for the proposed BPA, we can clearly assume that the unconditionally stability is assured, the broadband impedance matching is accomplished with a good power gain performance over a broad bandwidth of 1650 MHz.

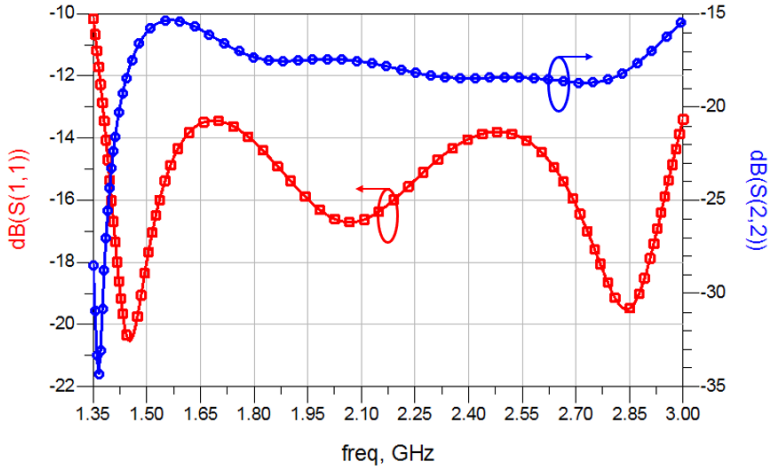


Fig. 7. Curves of **S_{11}** and **S_{22}** from 1.35 GHz to 3 GHz.

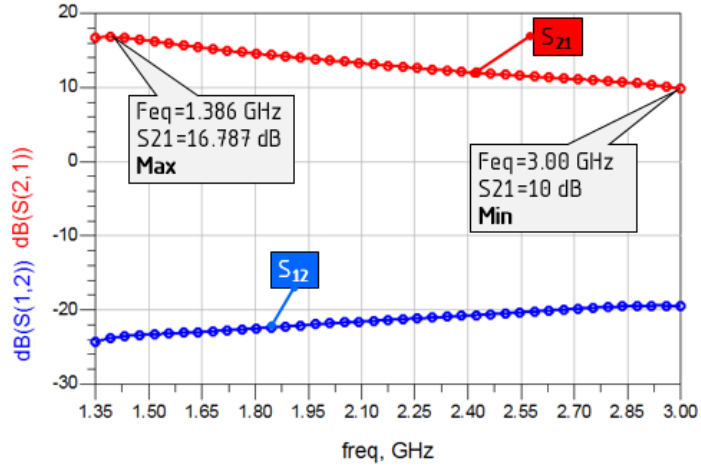


Fig. 8. Curves of S_{12} and S_{21} versus frequency.

3.3 Large signal simulation

From energetic standpoint, a power amplifier can be considered as a device able to convert DC power supplied to the active device, into an RF energy. The efficiency of this conversion is performed by means of Power Added Efficiency (PAE), defined as the ratio between the added power and the supplied DC power, which the added power is the subtraction of the input power from the output power.

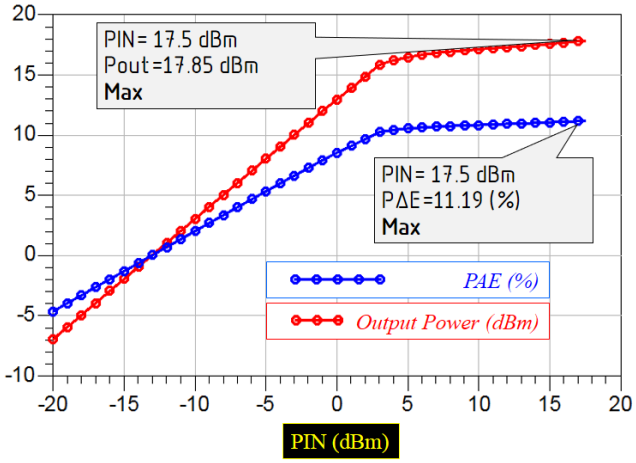


Fig. 9. PAE and Output Power versus frequency.

As shown in Figure 9, at 2.15 GHz, the proposed BPA achieves a maximum value of PAE of 11.19 % with a saturated output power of 17.85 dBm, thus corresponding to 60.95 mW of output power.

Table 1. Performance comparison between the proposed PBA and other BPAs

Ref.	Freq [GHz]	Psat [dBm]	Gain [dB]	PAE [%]	S11 [dB]	S22 [dB]	Supply [V]
[26] 2015	1.9-2.7	28.1	11	13.7	-	-	2.5
[27] 2016	1.8-2.8	28	25	6.1	-12 | -25	-09 | -19	5
[7] 2018	1.75 -2.15	8	-	11.7	-11 | -22	-13 | -19	12
[1] 2018	1.1 - 3	17.14	14.9	14.9	-10 | -35	-10 | -25	3
This Work	1.35 - 3	17.85	16.78	11.19	-10 | -24	-15 | -34	3.5

By comparing the proposed BPA with the similar contemporary state-of-the-art broadband PAs, we can clearly see that the matching techniques deployed in this work are very sufficient and permits to achieve a broad bandwidth. From Table 1, the proposed BPA exhibits excellent bandwidth, high gain, good output power and excellent matching across the operating bandwidth as well as simple circuitry when it is compared with the other BPAs.

4 Conclusion

This paper introduces a Broadband Power Amplifier operating in the frequency ranges from 1.35 GHz to 3 GHz with simple circuit configuration. The effectiveness of the designed matching networks was verified by the simulation results.

The large signal simulation shows that the proposed BPA reaches a maximum output of 17.85 dBm, with a PAE higher than 11%. This BPA is unconditionally stable over the whole operating bandwidth. A power amplifier with those features is more suitable for various microwave applications running in L and S bands such as cellular communication standards, medical microwave imaging, digital television broadcasting, avionics and navigation applications represent just a few examples.

References

[1] Ribate, M. Zbitou, J. Mandry, R. Errkik, A. Latrach, M.: Broadband GaAs FET Power Amplifier for L and S Bands Applications. International Journal of Intelligent Engineering and Systems. Vol. 11 (5), pp. 96-105 (2018).

[2] Ribate, M. Mandry, R. Latrach, M. Errkik, A. El Abdelaoui, L.: GaAs FET Broadband Power Amplifier for L and S Bands Applications, In Proceeding of The 2nd International Conference on Computing and Wireless Communication Systems (ICCWCS, Morocco (2017).

[3] Watkins, G. Wang, S.: The impact of power amplifier turn-on characteristics in cognitive radio networks. Microwave Journal. Vol. 57 (2), pp. 86-92 (2014).

[4] Fang, J. Moreno, J. Quaglia, R. Camarchia, V. Pirola, M. Guerrieri, SD. Ramella, C. Ghione, G.: 3.5 GHz WiMAX GaN Doherty power amplifier with second harmonic tuning. Microwave and Optical Technology Letters. Vol. 54 (11), pp. 2601-2605 (2012).

[5] IEEE Std 521-1984, IEEE Standard Letter Designations for Radar-Frequency Bands, (1984).

[6] Wagner, R.: 3G power amplifiers: moving from GaAs to CMOS. Microwave Journal. Vol. 55(4), pp. 44–48 (2012).

[7] Rachakh, A. El Abdellaoui, L. Zbitou, J. Errkik, A. Tajmouati, A. Latrach, A.: A Novel Configuration of a Microstrip Microwave Wideband Power Amplifier for Wireless Application. Telkomnika. Vol. 16 (1), pp. 2014-2031 (2018).

[8] Khannur, PB. Chen, X. Yan, DL. Shen, D. Zhao, B. Raja, MK. Sindunata, R. Yeoh, GW. Singh, R.: Universal UHF RFID reader IC in 0.18-µm CMOS technology. IEEE Journal Solid-State Circuits. Vol. 43(5), pp. 1146–1155 (2008).

[9] Osepchuck, JM.: Microwave Power Applications. IEEE Transactions on Microwave Theory and Techniques. Vol. 50 (3), pp. 975-985 (2002).

[10] Sobol, H. Tomiyasu, K. Milestones of microwaves.: IEEE Transactions on Microwave Theory and Techniques. Vol. 50 (3), pp. 594–611 (2002).

[11] Kopp, B.A. Borkowski, M. Jerinic, G.: Transmit/receive modules. IEEE Transactions on Microwave Theory and Techniques. Vol. 50 (3), pp. 827–834 (2002).

[12] Willert-Porada, M.: Advances in Microwave and Radio Frequency Processing. Report from the 8th International Conference on Microwave and High Frequency Heating. Springer-Verlag, Germany (2001).

[13] Osepchuk, J.M.: A history of microwave heating applications. IEEE Transactions on Microwave Theory and Techniques. Vol. 32 (9), pp. 1200–1224 (1984).
[14] Rosen, A. Stuchly, M. A. Vorst, A. V.: Applications of RF/Microwaves in Medicine. IEEE Transactions on Microwave Theory and Techniques. Vol. 50 (3), pp. 963-974 (2002).
[15] Rekanos, I. T. Räisänen, A.: Microwave Imaging in the Time Domain of Buried Multiple Scatterers by Using an FDTD-Based Optimization Technique. IEEE Transactions on Magnetics. Vol. 39 (3), pp. 1381-1384 (2003).
[16] Kaynak, M. Tekin, I. Gurbuz, Y.: Fully integrated low-power SiGe power amplifier for biomedical applications. IET Microwaves Antennas & Propagation. Vol. 5 (2), pp. 214–219 (2011).
[17] Fear, E.C. Li, X. Hagness, S.C. Stuchly, M.A.: Confocal microwave imaging for breast cancer detection: localization of tumors in three dimensions. IEEE Transaction on Biomedical Engineering. Vol. 49 (8), pp. 812–822 (2002).
[18] Poole, C. Darwazeh, I.: Microwave Active Circuit Analysis and Design. Elsevier Ltd, UK (2016).
[19] Colantonio, P. Giannini, F. Limiti, E.: High Efficiency RF and Microwave Solid State Power Amplifiers. John Wiley & Sons, Ltd, UK (2009).
[20] Virdee, B.S. Virdee, A.S. Banyamin, B.Y.: Broadband Microwave Amplifiers. Artech House, Inc, US (2004).
[21] Eroglu, A: Introduction to RF Power Amplifier Design and Simulation. CRC Press, US (2016).
[22] Wu, H. Wang, L. Zhou, P. Ma, J.: A 0.1-1.2 GHz CMOS Ultra-Broadband Power Amplifier. In Proceeding of 2014 IEEE MTT-S International Microwave Symposium, USA (2014).
[23] Inder Babel, Prakash Bhartia.: Microwave Solid State Circuit Design. 2nd Edition, John Wiley & Sons, Inc, (2003)
[24] Jarry, P. Beneat, J.N.: Microwave Amplifier and Active Device Design Using the Real Frequency Technique. John Wiley & Sons, Inc, Hoboken, New Jersey (2016).
[25] Raghavan, A. Srirattana, N. Laskar, J.: Modeling and Design Techniques for RF Power Amplifiers. John Wiley & Sons, Hoboken, New Jersey (2008).
[26] Francois, B. Reynaert, P.: Highly linear fully integrated wideband RF PA for LTE-advanced in 180-nm SOI. IEEE Transactions on Microwave Theory and Techniques. Vol. 63 (3), pp. 649–658 (2015).
[27] Chen, C.Q. Hao, M.L. Li, Z.Q. Du, Z.B. Yang, H.: A 1.8–2.8 GHz Highly Linear Broadband Power Amplifier for LTE-A Application. Progress in Electromagnetics Research C. Vol. 66, pp. 47–54 (2016).

Microwave Absorption Properties of Garnet Ferrite Filled Polymer Nanocomposites in X-Band

S.M.Alami , H.Ammor
{ [1]souhail.alami.m@gmail.com, [2]ammor@emi.ac.ma }

Engineering 3S Research Center,Smart Communications Research Team
Mohammadia School of Engineering, Mohammed V University in Rabat-Morocco

Abstract. The electromagnetic properties of ferrite filled polymer nano-composites can be determined by a classical Nicolson-Ross and Weir technique. The measurement of S-parameters makes it possible to calculate the complex permittivity ε and permeability μ of the Lanthanum Iron Garnet (LIG)-loaded polyvinylidene fluoride (PVDF)-polymer nanocomposites over the X-band waveguide. The reflection S_{11} and transmission S_{21} coefficients were calculated by the transmission/reflection rectangular waveguide technique. The LIG/PVDF sample, placed in the rectangular waveguide over X-band frequencies. The average values of the relative complex permittivity and permeability of LIG/PVDF over X-band frequencies [8 - 12 GHz] are (4.34 - j0.09) and (1.27 - j0.15) respectively.

Keywords: Complex permittivity and permeability, Ferrite, Garnet, LIG, PVDF, waveguide W-90, X-band.

1 Introduction

Ferrites are ferrimagnetic oxides materials have been heavily used in passive microwave devices such as circulators, isolators, phase shifters, resonators , filters, , and miniature antennas [1]. The garnets are soft ferrites having the general chemical formula $R_3Fe_5O_{12}$ where R is yttrium or trivalent rare earth ions from lanthanum to ytterbium[2]. Garnet ferrites have a cubic crystal structure and include three types of crystallographic lattice of (a,b and c), the $3R^{3+}$ ions are located the dodecahedra (c) site, $2Fe^{2+}$ ions are distributed to the octahedral (a) site and $3Fe^{3+}$ ions occuping the tetrahedral (d) site, where as O^{2-} ions are located to the interstitial sites [3].

Lanthanum Iron Garnet (LIG) has a chemical formula $La_3Fe_5O_{12}$, it is strongly used in electronic devices, due to the efficient absorption of electromagnetic waves, low permeability, low losses at high frequencies, extremely easy to magnetize and demagnetize, low saturation flux, and high resistivity. In recent years, many works have been made on polymer matrix - based nanocomposites filled with ferrite particles, such as Co-ferrite [4], Ni-Zn-ferrite [5], and Mn-Zn-ferrite [6], using different characterization methods [7].

In this paper, the transmission/reflection (TR) caracterization method based on Nicolson–Ross–Weir approach allows to calculate simultaneously the complex relative permittivity and complex relative permeability of the lanthanum iron garnet filled -PVDF polymer nanocomposite, from the S-parameters, the sample was placed in a rectangular waveguide over the X-band frequencies [8 – 12 GHz] , and with the ANSYS HFSS simulation, we can

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284217

measure the reflection and transmission coefficients and calculate the complex relative permittivity and complex relative permeability of LIG/PVDF composites.

2 Materials and methods

Many papers for the physico-chemical synthesis of lanthanum iron garnet have been described in [8,9] by sol-gel procedure.

2.1 Preparation of the LIG powder

The procedure for preparing LIG magnetic powders is shown in the diagram below. **Figure 1.**

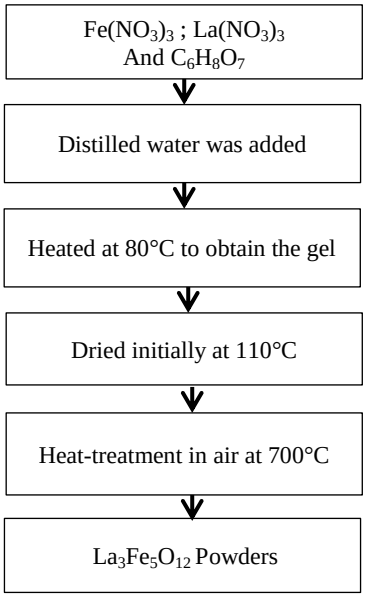


Fig. 1. Procedure for preparing Lanthanum Iiron Garnet powders

The LIG/PVDF as a nanocomposite was prepared by the solvent method , with 13% LIG nano-powders and 87% of PVDF plymer in the shape of a rectangular sheet 3 mm thickness.

2.2 Simulation of the rectangular waveguide loaded with the sample LIG/PVDF

The PVDF-LIG nano-composite sample of dimensions (22.86*11.43*3mm^3), was well adjusted into a rectangular waveguide cross section WR-90. The S-parameters then were measured in the frequency range from 8 to 12 GHz by using the HFSS simulator. The simulation can be shown in **Figure 2.**

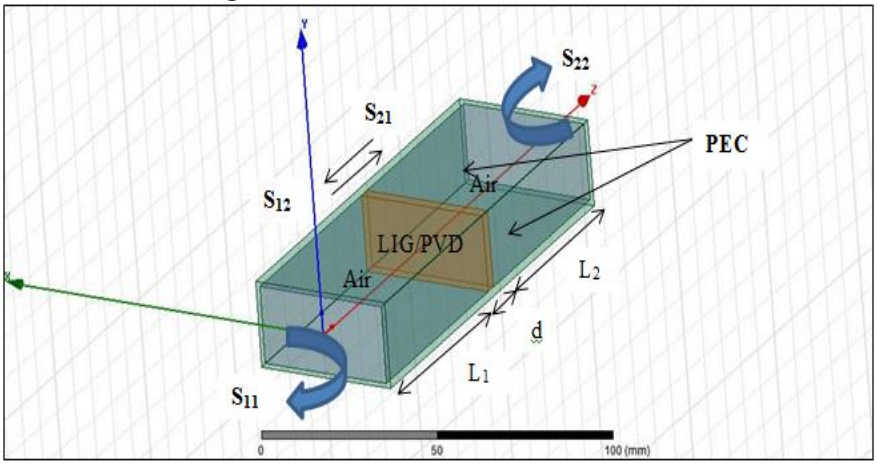


Fig. 2. Simulation of the rectangular waveguide

2.3 Nicholson-Ross-Weir (NRW) method

Determination of the electromagnetic properties of PVDF-LIG composite sample using transmission/reflection rectangular waveguide technique[10-12]. The organigram shown in the **Figure 3** describes the procedure for extracting the S-parameters and estimation complex permeability and permittivity of the sample under study.

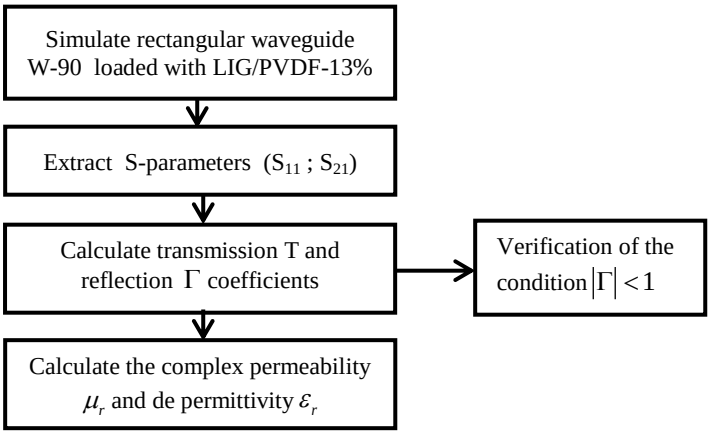


Fig. 3. Algorithm of Nicolson-Ross and Weir

The proposed classical procedure by Nicolson-Ross and Weir is based on the following formulas, these equations are applicable for the rectangular waveguide with the transverse electric TE_{10} mode.

$$S_{11} = R_1^2 \frac{\Gamma\left(1-T^2\right)}{1-\Gamma^2 T^2} \tag{1}$$

$$S_{21} = R_1 \times R_2 \frac{T\left(1-\Gamma^2\right)}{1-\Gamma^2 T^2} \tag{2}$$

$$R_1 = \exp(-\gamma_0 L_1) \quad \text{and} \quad R_2 = \exp(-\gamma_0 L_2) \tag{3}$$

$$T = \exp(-\gamma d) \tag{4}$$

$$\gamma = j\frac{2\pi}{\lambda_0}\sqrt{\varepsilon_r \mu_r - \left(\frac{\lambda_0}{\lambda_C}\right)^2} \tag{5}$$

$$\gamma_0 = j\frac{2\pi}{\lambda_0}\sqrt{1-\left(\frac{\lambda_0}{\lambda_C}\right)^2} \tag{6}$$

$$\Gamma = K \pm \sqrt{K^2 - 1} \tag{7}$$

$$K = \frac{S_{11}^2 - S_{21}^2 + 1}{2S_{11}} \tag{8}$$

$$T = \frac{S_{11} + S_{21} - \Gamma}{1-\left(S_{11}+S_{21}\right)\Gamma} \tag{9}$$

$$\mu_r = j\left(\frac{1+\Gamma}{1-\Gamma}\right) \times \frac{c}{2\pi fd}\ln\left(T\right) \tag{10}$$

$$\varepsilon_r = -\left(\frac{c}{2\pi fd}\ln(T)\right)^2 \times \frac{1}{\mu_r} \tag{11}$$

- Γ and T are respectively the reflection and transmission coefficients;

- γ_0 and γ are respectively the propagation constants in the vacuum and sample;

- λ_0 and λ_C represent respectively the free space and the cut-off wavelengths;

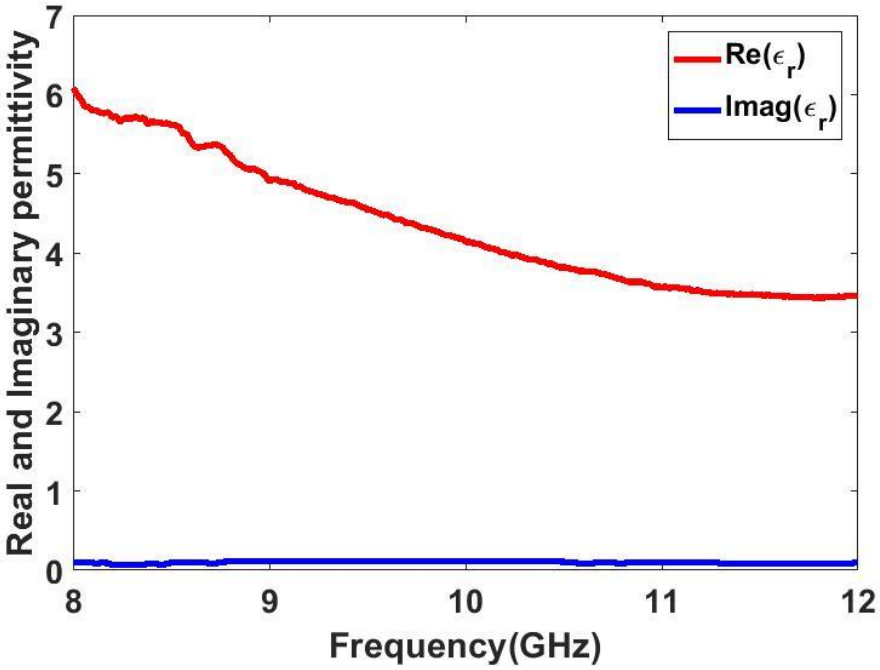
- ε_r et μ_r are the relative permittivity and permeability respectively.

-R_1 and R_2 correspond the expression of reference planes respectively.
- L_1 and L_2 are the respective distances from calibration reference planes to sample ends.
-d is wavelenght of sample.

3 Results and discussion

The Nicolson-Ross and Weir (NRW) technique was used to calculate the complex permittivity and permeability of LIG/PVDF nanocomposite, from reflection S_{11} and transmission S_{21} coefficients using the transmission/reflection (TR) measurement technique. **Figure 4a.** below shows two response the curves of LIG/PVDF sample of the real permittivity and imaginary permittivity as function of frequency. The real part of permittivity is decreasing from 8GHz to 12GHz, this real part has a high value nearly 6.08 in frequency at 8GHz, and then decreases to 3.45 at 12GHz when the frequency increases. This figure also represents the imaginary part of the permittivity curve. The imaginary part has a small variation with frequency increases from 8GHz to 12GHz. The observed response of the permittivity constant is the result of adding the portion of the LIG magnetic in PVDF polymer matrix [13].
The **Figure 4b.** below, presents two response the curves of LIG/PVDF nanocomposite of the real permeability and imaginary permeability as function of frequency. it illustrates that the real part of permeability is 3.53 at 8GHz, from this point, the curve of the real permeability shows a gradual decrease with increasing frequency to reach 0.23 at 12 GHz.
Similarly, the imaginary part decreases from 0.75 to 0.01 with increasing frequency from 8GHz to 12GHz, the decreasing is the product of the domain wall motion and the spine-lattice interaction.
The average values obtained in this work of the complex permittivity and permeability respectively of mentioned above nanocomposite in X-band frequencies were (4.34-j0.09) and (1.27-j0.15). The results achieved in this project are in excellent agreement with the literature[14].

(a)



(b)

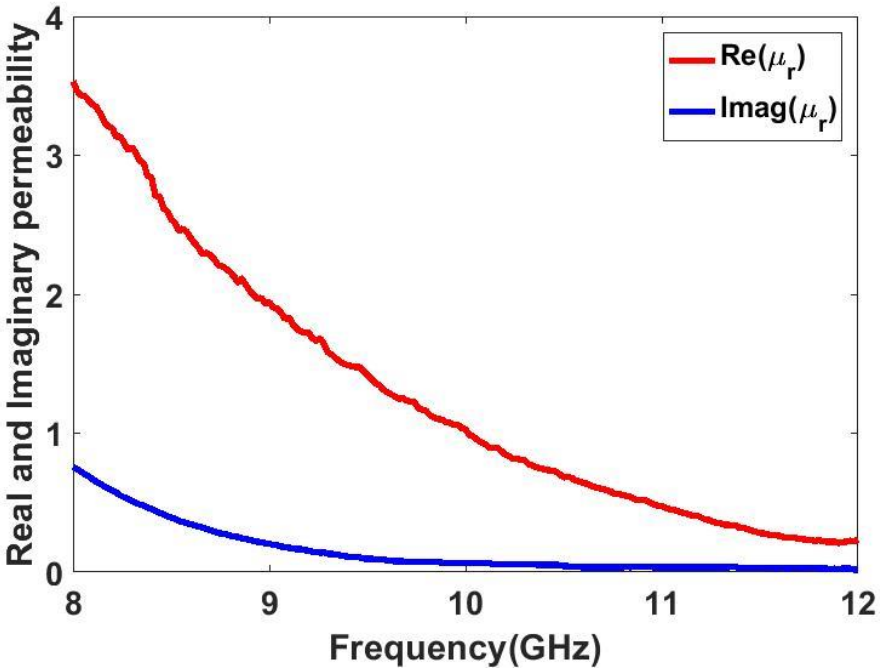


Fig. 4. Evolution of (a) Complex permittivity and (b) Complex permeability of LIG/PVDF nanocomposite versus frequency

Table 1. below represents the comparison between the values calculated and measured [14].

Table 1. Comparison table between the calculated parameters and the measured parameters

LIG/PVDF-13%	ε_r	μ_r	Error $\frac{\Delta\varepsilon_r^{'}}{\varepsilon_r^{'}}$	Error $\frac{\Delta\mu_r^{'}}{\mu_r^{'}}$
Calculed	4.34-j0.09	1.27-j0.15	$\leq$ 0.23%	$\leq$ 2.3%
Measured	4.33-j0.09	1.24-j0.15		

Conclusions and perspectives

In this paper, the LIG/PVDF nanocomposite has been fabricated in the shape of rectangular sheets with thickness of 3mm. The S-parameters were measured by the cross section rectangular waveguide at the X-band. The Nicholson-Ross-Weir (NRW) procedure was employed to calculate the complex permittivity and complex permeability of the LIG/PVDF sample. The results indicated at X-band frequencies, show the both real part of permittivity, real part of permeability and imaginary part of complex permeability are decreased when increasing the frequency in X-band. But, the imaginary part of the dielectric permittivity probably remains constant in this frequency range. The average values obtained in this work of the complex permittivity and permeability respectively of mentioned above nanocomposite in X-band frequencies were (4.34-j0.09) and (1.27-j0.15). The results achieved in this project are in excellent agreement with the literature[14].

Acknowledgement

The authors would like to thank the members of Institute of Electronics and Telecommunications of Rennes, France for allowing us to use the commercial solvers available in their laboratory.

References

[1] Y.J. Wu, H.P. Fu, R.Y. Hong, Y. Zheng, D.G. Wei.: Influence of surfactants on co-precipitation synthesis of Bi–YIG particles. Vol. 470, N°1-2, pp. 497–501. Journal of Alloys and Compounds (2009)
[2] Ümit Özgüri, Yahya Alivov, and Hadis Morkoç.: Microwave Ferrites, Part 1: Fundamental properties. Vol. 20, N°9, pp. 789–834. Journal of Materials Science: Materials in Electronics (2009)

[3] Makiyyu Abdullahi Musa, Raba'ah Syahidah Azis, Nurul Huda Osman, Jumiah Hassan, Tasiu Zangina.: Structural and magnetic properties of yttrium iron garnet (YIG) and yttrium aluminum iron garnet (YAlG) nanoferrite via sol-gel synthesis. Vol. 7, pp. 1135-1142. Results in Physics (2017)
[4] Vaishnava PP, Senaratne U, Buc E, Naik R, Naik VM, Tsoi G, et al.: Magnetic properties of cobalt-ferrite nanoparticles embedded in polystyrene resin. Vol. 99, N° 08G702, pp. 1-3. J. Appl Phys (2006)
[5] Ting, T. H.; Yu, R. P.; Jau, Y. N.: Synthesis and Microwave Absorption Characteristics of Polyaniline/NiZn Ferrite Composites in 2−40 GHz. Vol. 126, pp. 364−368. Mater. Chem. Phys (2011)
[6] Amiri, Gh. R.; Yousefi, M. H.; Abolhassani, M. R.; Manouchehri, S.; Keshavarz, M. H.; Fatahian, S. : Magnetic Properties and Microwave Absorption in Ni-Zn and Mn-Zn Ferrite Nanoparticles Synthesized by Low-Temperature Solid-State Reaction. Vol. 323, N°6 pp. 730−734. J. Magn. Magn. Mater (2011)
[7] H. Elmajid, J. Terhzaz, H. Ammor.: Implementation of the Mode Matching Technique to Estimate the Complex Permittivity of Dielectric Material at Microwave Frequency Using Rectangular Waveguide Measurements. Vo1. 108, pp. 123-138. European Journal of Scientific Research (2013)
[8] P. Vaqueiro, M. P. Crosnier-Lopez, and M. A. Lopez-Quintela.: Synthesis and Characterization of Yttrium Iron Garnet Nanoparticles. Vol. 126, N° 0324 pp. 161-168. Journal of solid state chemistry (1996)
[9] Majid Niaz Akhtar , Noorhana Yahya, and Nadeem Nasir.: New EM Transmitter with Y3Fe5O12 based Magnetic Feeders Potentially Used for Seabed Logging Application. Vol. 667, pp 10-23. Advanced Materials Research (2013)
[10] A. M. Nicolson and G. F. Ross.: Measurement of the Intrinsic Properties of Materials by Time-domain Techniques. Vol. 19, N°4, pp. 377-383. IEEE Trans. Instrum. Meas (1970)
[11] A. M. Nicolson.: Broad-band Microwave Transmission Characteristics From a Single Measurement of the Transient Response. Vol. 17, N°4, pp. 395402. IEEE Trans.Instrum. Meas (1968)
[12] W. B. Weir.: Automatic Measurements of Complex Dielectric Constant and Permeability at Microwave Frequencies. Vol. 62, N°1, pp. 3339. Proc.IEEE (1974)
[13] Christian Brosseau, Jamal Ben Youssef, Philippe Talbot, and Anne-Marie Konn.: Electromagnetic and magnetic properties of multicomponent metal oxides heterostructures: Nanometer versus micrometer-sized particles. Vol. 93, N° 11, 1, pp. 9243- 9256. J. Appl. Phys (2003)
[14] Hassan Soleimani , Zulkifly Abbas, Noorhana Yahya, Hojjatollah Soleimani,Mohammad Yeganeh Ghotbi.: Determination of complex permittivity and permeability of lanthanum iron garnet filled PVDF-polymer composite using rectangular waveguide and Nicholson–Ross–Weir (NRW) method at X-band frequencies. Vol. 45, N°6, pp. 1621–1625. Measurement (2012)

Contribution to the numerical study of the thermal gradient during the fabrication of composite materials by the RTM process

Soukayna ELYOUSSFI[1], Abdellatif EL OUJDI[2], Aouatif SAAD[3], Mohamed HATTABI[4], Adil ECHCHELH[5]

{ [1]ssoukaynaelyou@gmail.com, [2]abdellatif.eloujdi@gmail.com, [3]saad_aouatif@yahoo.fr, [4]mohamdhattabi@hotmail.com, [5]adilechel@gmail.com}

[1, 2,3,5] Laboratory of Materials and Energy, faculty of sciences, University of kenitra, Morocco.
[4]Department of Mechanical Engineering, ENSEM, University of Casablanca, Morocco.

Abstract. Resin Transfer Molding (RTM) has become one of the most efficient processes in the field of composite materials, among the steps in composite part processing is the curing reaction. In the majority of cases this reaction is of exothermic nature and accompanied by a rise in temperature in the center of the laminate, this leads to the appearance of a thermal gradient. This research work aims to highlight the thermal gradient generated as a result of a cross linking reaction. The objective is to reduce the temperature excess in the Polyester / Glass laminate. By means of a one-dimensional numerical study using the finite- difference method we have shown that the energy balance depends not only on the temperature and the degree of curing but also on several other factors namely: the volume fraction of the fibers and the temperature cycle. The comparison of the results with literature showed a good agreement.

Keywords. Resin transfer molding (RTM), thermal gradient, Polyester/Glass, one-dimensional, finite difference, temperature, degree of curing, volume fraction of the fibers.

1. Introduction

Since their advent in the thirties and until today composite materials have strongly contributed to the development of our techniques. Among the industrial processes used in the production phase is resin transfer molding RTM. The latter is considered one of the most promising techniques available today; it is able to make a great deal complex shapes with a great mechanical performance, tolerance, dimensional, narrow and high surface finish [1]. The manufacturing of a part by the RTM process can be freely divided into four main stages [2]: fiber preform manufacturing, mold filling, curing and remolding the part.

After the filling of the whole mold, a curing reaction is initiated and is intended to bind the liquid resin to the solid reinforcement. In most cases, this reaction is initiated by a heating of the mold. The exothermic curing reaction is accompanied by an increase in the external temperature. This released heat can thus, if not properly distributed, further promote the curing reaction which once again generates heat; therefore, this gives rise to a temperature gradient in the composite which increases as a result of the low thermal conductivity of the polymer matrices. This thermal gradient in turn causes residual stresses, defects in the final part.

A numerical study of the curing process remains a good way to predict the degree of cure and the temperature distribution. Several attempts have been made to reduce this thermal gradient by optimizing the energy cycle. Non-exhaustive, the work of Choi and al. [2] who manipulated with a temperature profile that represents a constant temperature interval, we observed as a consequence the disappearance of this gradient which is argued by the weak transfer heat in this time interval, so the temperature will not exceed its value at the wall of the mold. Antonucci and al. [3] presents a methodology based on the scale analysis of the energy balance equation to manage the heat generated by the curing reaction. Their approach allows to minimize temperature gradients before the resin solidifies. Loos and al. [4] involved the development of a one-dimensional simulation of a plate using an implicit method of finite difference.

In this study, we have shown that the energy balance depends not only on the temperature and the degree of cure, but also on several other intervening factors characterizing both the matrix and the reinforcement, namely: the fiber

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284094

volume fraction and the the temperature cycle of the mold wall. In our numerical approach, we have opted for a one-dimensional finite difference method to study the temperature transfer in composite and the rate of the curing reaction in the RTM process. To validate our results, we have compared the numerical solutions obtained with the data available in the literature, this gives good agreement.

2. Mathematical model

2.1 Heat transfer equation

The one-dimensional equation that governs heat transfer in the composite laminate is given as follows [4, 5 and 6]:

$$\rho C_p \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial z^2} + \rho_r . (1 - v_f) . \Delta H_r . \frac{d\alpha}{dt} \quad (1)$$

k, ρ, v_f , ΔH_r , α and C_p are respectively the thermal conductivity, the density, the volume fraction, the heat reaction, the degree of cure of resin and the heat capacity. According to Choi et al. [2], the thermal properties of the resin, the fiber and the cured part are defined as follows:

The mass fractions of the resin (w_r), the cured product (w_p) and the fibers (w_f) are respectively given by [2, 7, 8]:

$$w_r = w_{r0} . (1 - \alpha) \quad (2a)$$

$$w_p = w_{r0} . \alpha \quad (2b)$$

$$w_f = w_{f0} \quad (2c)$$

The volume fraction of the resin (v_r), the cured product (v_p) and the fibers (v_f) are given respectively by [2, 7]:

$$v_r = w_r \frac{\rho}{\rho_r} \quad (3a)$$

$$v_p = w_p \frac{\rho}{\rho_p} \quad (3b)$$

$$v_f = \frac{\rho}{\rho_f} \quad (3c)$$

The expressions of density, specific heat capacity and thermal conductivity are given, respectively, by [7, 9]:

$$\frac{1}{\rho} = \frac{w_r}{\rho_r} + \frac{w_p}{\rho_r} + \frac{w_f}{\rho_f} \quad (4)$$

$$C_p = w_r . C_{p,r} + w_p . C_{p,p} + w_f . C_{p,f} \quad (5)$$

$$lnk = V_{rp} . lnk_{rp} + V_f . lnk_f \quad (6)$$

2.2 Kinetic model

The equation of the curing rate reactions used in our case is defined as follows [10, 11]:

$$\frac{d\alpha}{dt} = (K_1 + K_2 \alpha^m)(1 - \alpha)^n \quad (7)$$

With:

$K_i = A_i . \exp\left(-\frac{E_i}{RT}\right)$

Or: i=1, 2

From where T is the temperature, α is the degree of cure and $\frac{d\alpha}{dt}$ indicates the reaction rate, K_i are reaction rate constants. E_i present the activation energy, R the universal gas constant, m and n are empirical exponents in the cure kinetic model..

3. Boundary conditions

Initial condition (t=0): $T = T_0$

The side wall: $T = T_w(t)$

In the center of the wall: $\frac{\partial T}{\partial n} = 0$

For the polymerization of Polyester / Glass composite laminate, the initial condition [11]: $\alpha = \alpha_0$ at t=t_0.

4. Numerical model

In our study we have opted with the explicit finite difference method to discretize the two equations (1) and (7). Temperature and degree of cure distributions in the entire domain are calculated at each time step increment as a function of the temperature history of the curing cycle [12].

The discretization of the curing equation by the finite difference method gives:

$$\alpha_j^{t+\Delta t} = (\frac{d\alpha}{dt}_j^{t+\Delta t} + \alpha_j^t).\Delta t \quad (8)$$

The stability condition used in this work is expressed as follows [13]:

$$\Delta t \leq \Delta z^2/(2.a) \quad (9)$$

The discretization of the heat equation by the finite difference method is given as follows:

$$\frac{T_j^{i+1} - T_j^i}{\Delta t} = \frac{k}{\Delta z^2}(T_{j+1}^i - 2T_j^i + T_{j-1}^i) + \rho_r.(1 - v_f).\Delta H_r(\frac{d\alpha}{dt})_j^{t+\Delta t} \quad (10)$$

This gives:

$$T_j^{i+1} = T_j^i + a.\Delta t/\Delta z^2(T_{j-1}^i - 2T_j^i + T_{j+1}^i) + \rho r.\left(1 - v_f\right).\Delta H_r.\Delta t/\Delta z^2.(\frac{d\alpha}{dt})_j^t \quad (11)$$

5. Results and Discussions

5.1 Validation of results

Considering a flat rectangular mold with the dimensions of 10 cm in length and 1 cm in thickness. The mold is filled with fiber glass preforms and saturated with the polyester resin which is heated through the mold wall to initiate curing. For the polyester / glass composite laminate, we adopted in our work the kinetic parameters adopted by Choi et al. [2] illustrated in Table 1, and the thermo-physical properties shown in Table 2.

Table 1. Cure kinetics for Glass / polyester [2].

$\boldsymbol{k_1(1/s)}$	5,68667. $\mathbf{10^{12}}$
$\boldsymbol{k_2(1/s)}$	8,61167. $\mathbf{10^8}$
$\boldsymbol{E_1}$**(cal/M)**	25570
$\boldsymbol{E_2}$**(cal/M)**	17930
n	1,42
m	0,58
ΔH (cal/M)	47,8

Table 2. Material properties of Glass / polyester [2,7].

Materials	$\rho(g/cm^3)$	C_p**(cal/g.K)**	**k(cal/cm.s.K)**
Uncured Polyester	1.10	0.45	$4.05.10^{-4}$
Cured Polyester	1.20	0.45	$1.54.10^{-4}+9.46.10^{-7}T$
Class fiber	2.54	0.199	$2.07.10^{-3}$

The numerical model that we developed with MATLAB calculation code for the polyester resin and fiberglass system is based on the coupled resolution of the heat equation and the rate of the curing reaction. This allows describing the evolution of the temperature and the chemical kinetics of the resin. This coupling makes possible to calculate both the temperature and the degree of curing α (T, α) at each point in the thickness of the laminate in the mold and at each moment [14].

From the results illustrated in figure 1 and 2 we can observe that the temperature and the degree of cure develops very quickly at the center of the polyester / glass composite laminate which causes an exothermic peak. We then validated our results with Choi et al. [2] and a good agreement is observed.

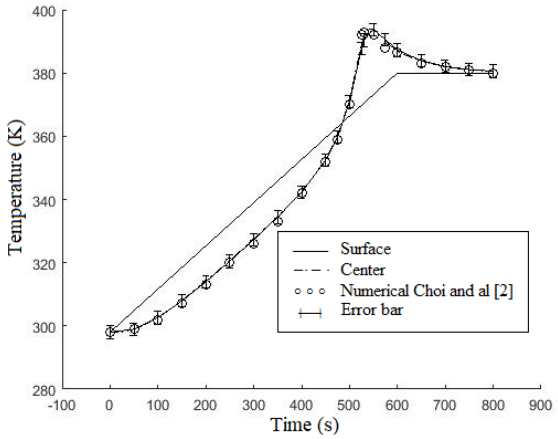


Fig.1. Temperature profile at the central node for the glass / polyester laminate.

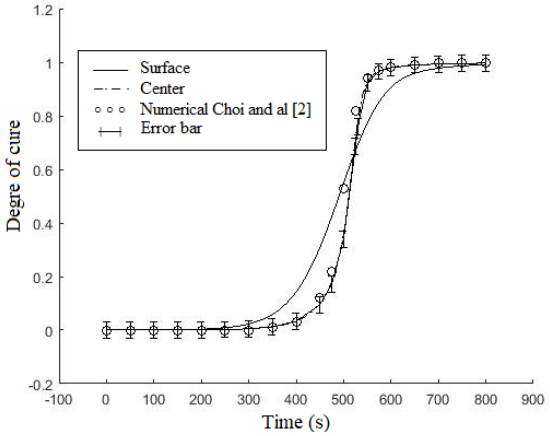


Fig. 2. Degree of cure at the central node for glass / polyester laminate.

Thermal gradients during the RTM process give rise to several concerns, among which: residual stresses that could lead to shrinkage and lack of dimensional tolerance, thermal stress that can deform the composite part....

The following part presents the one-dimensional numerical model used to simulate the curing stage and to predict the temperature profile in the RTM process and in particular for the minimization of the indesirable thermal gradient .

5.2 Strategies to reduce the thermal gradient in the composite laminate

- **Optimization of the temperature cycle**

In order to reduce and control thermal gradients in the RTM process for a polyester resin and fiber glass, we have opted for the optimization of the temperature cycle. The variables that have to be controlled are the degree of curing or the adjustment of the temperature boundary conditions of the mold wall. In our case, the methodology used is based on adjusting the temperature boundary conditions of the mold wall; this was done in two steps: In the first step, which is called the conduction zone, the mold wall temperature was raised to a plateau and held until the temperature across the thickness is close to uniformity. In the second step, called the reaction zone, the temperature of the mold wall was raised to a higher temperature to complete the cure reaction. For a clearer view, figures 3 and 4 show the temperature evolution in time for the surface and center of the composite laminate.

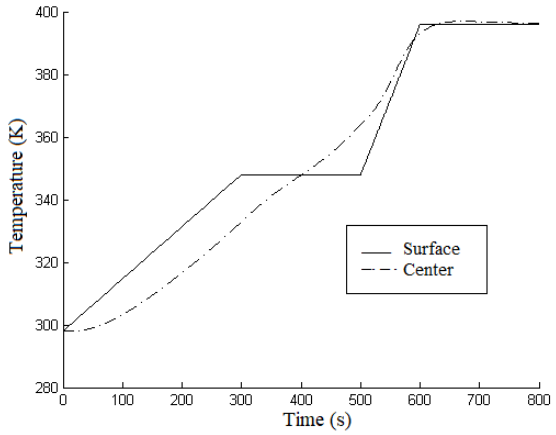


Fig. 3. Temperature profile at the central node for the glass / polyester laminate in the case of an optimized cure cycle.

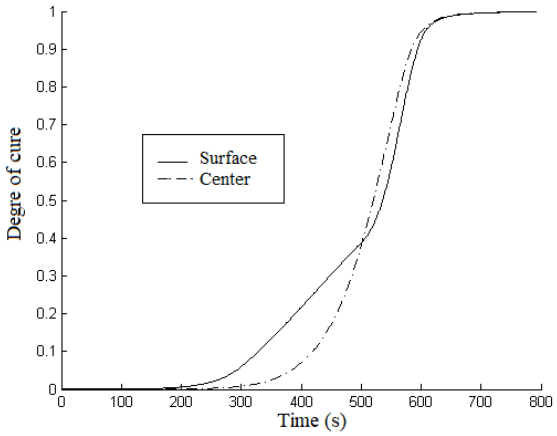


Fig. 4. Degree of cure at the central node for the glass / polyester laminate in the case of an optimized cure cycle.

The results of this procedure as showed in Figures 3 and 4 are very satisfactory with a disappearance of the thermal gradient. Thus the temperature will not exceed its value at the wall of the mold. In return, the curing process takes a constant time to complete. On the other hand, the curing reaction develops uniformly giving rise to a minimum of residual stresses in the final piece.

- **Choice of the fiber volume fraction**

One of the most important parameters affecting the properties of composites is the fiber volume fraction. In our work we have studied the effect of varying the fiber properties especially fiber content on the temperature and curing rate profile.

Figures 5 and 6 show the profile of the temperature and the degree of curing in the center of the mold for differents volume fraction of fibers: $v_f = 0.2$, $v_f = 0.4$ and $v_f = 0.5$.

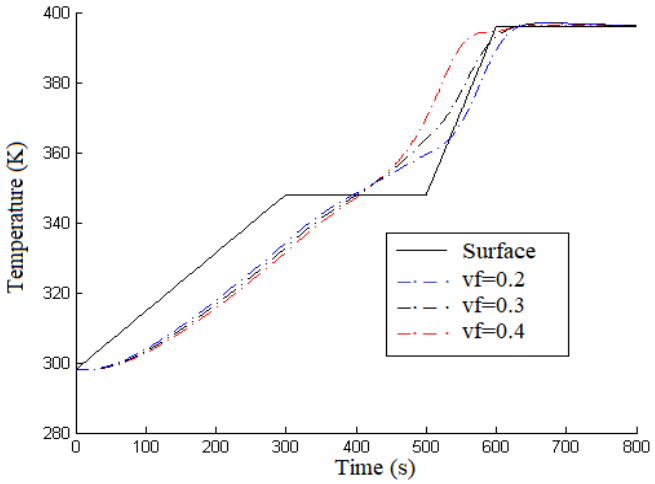


Fig.5. Profile of the Temperature at the central node of the Polyester /Glass composite laminate in the case of optimized polymerization cycle for different cases of volume fraction of fibers.

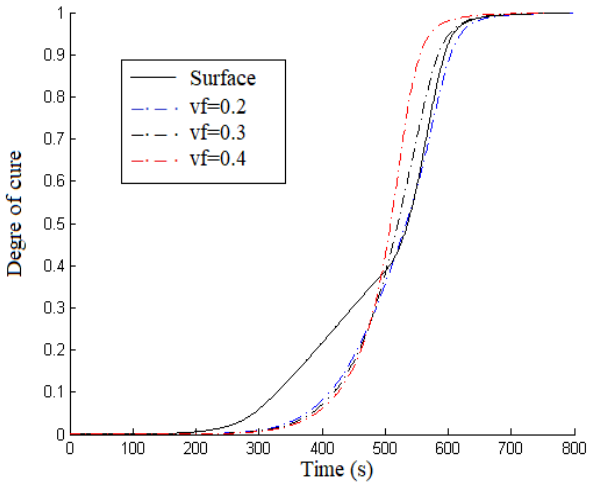


Fig.6. Degree of curing at the central node of the Polyester /Glass plate in the case of an optimized polymerization cycle in different cases of fiber volume fraction.

It's clear from the figures 5 and 6 that by increasing the amount of fiber present in the system, the peak of exothermic reaction was reduced and the thermal conduction was improved. This was explained by the fact that the fiber glass had a higher thermal conductivity than the polyester resin. It is thus of great interest to inrease the fiber volume fraction of the reinforcement used to first minimize the thermal gradient and then to increase the mechanical performance of the composite part.

6. Conclusion

In this work, we have developed a calculation program under MATLAB, we opted for the explicit finite difference methods in the one-dimensional case applied to the case of a polyester / glass composite laminate at the macroscopic scale. This code has the ability to predict the temperature profile and degree of cure in the mold. Thanks to our numerical simulations, we have found that the thermal gradient, which is detrimental to the quality of the final product, can be considerably reduced through a series of preventive measures, among others: The optimization of the temperature cycle in the mold and the choice of the convenient volume fraction of the fibers.

With the results we obtained by our calculation code, the methodology developed proved to be able to effectively solve the problem of thermal gradients. We are considering in the following, the research of other methods that can reduce the thermal gradient in order to improve the mechanical properties of the composite part.

References

[1] A. Saad, A. Echchelh, M. hattabi, M. El Ganaoui, 20111, An Improved Computational Method For Non Isothermal Resin Transfer Moulding Simulation, Thermal sciences, 15, pp. 275-289.

[2] Choi, M.A., Lee, M.H., Lee, S.J.: Numerical studies during cure in the resin transfer molding. Vol 6, pp. 211-218.Korea polymer journal, (1998).

[3] Vincenza, A., Michele, G., Kuang-Ting, H., Suresh, G.A.: A methodology to reduce thermal gradients due to the exothermic reactions in composites processing. Vol 45, pp.1675–1684. International Journal of Heat and Mass Transfer, (2002).

[4] Loos, A.C. , Springer,G.S.: Curing of epoxy matrix composites. Vol 17, pp. 135-169. Journal of Composite Materials, (1983).

[5] Yi, S., Hilton, H.H., Ahmad, M.F.: A finite element approach for cure simulation of thermosetting matrix composites. Vol. 64, pp. 383–388. Civil-Comp and Elsevier Science, (1997).

[6] Kim, J.S., Lee, D.G.: Development of an Autoclave Cure Cycle with Cooling and Reheating Steps for Thick Thermoset Composite Laminates. Vol. 31, pp. 2264-2282. Journal of Composite Materials, (1997).

[7] Han, C., Lee, D., Chin, H.: Development of a Mathematical Model for the Pultrusion Process. Vol. 26, pp. 393-404. Polymer Engineering and Science, (1986).

[8] Aouatif Saad, Adil Echchelh , Mohamed Hattabi and Mohammed El Ganaoui : The identification of effective thermal conductivity for fibrous reinforcement composite by inverse method. Vol. 33(23) 2183–2191. Journal of Reinforced Plastics and Composites, 2014.

[9] Saad, A., Echchelh, A., Hattabi, M. and El Ganaoui, M.: Optimization of the cycle time in resin transfer molding process by numerical Simulation. Vol.31, pp. 1388-1399. , Journal of Reinforced Plastics and Composites, (2012).

[10] KAMAL, M.R., SOUROUR, M.R.: Kinetic modelling of epoxy-amine cure process. Vol. 13, pp. 59-64. Polymer Engineering and Science, (1973).

[11] Cheung, A., Yu, Y., Pochiraju, K.: Three-dimensional finite element simulation of curing of polymer composites. Vol. 40, pp. 895–912. Finite Elements in Analysis and Design, (2004).

[12] Bogetti, T.A., Gillespie, J.W.: Two-dimensional cure simulation of thick thermosetting composites. Vol. 25, pp. 239–273. Journal of Composite Materials, (1991).

[13] Goncalvès, E. : Résolution Numérique, Discrétisation des EDP et EDO. Institut National Polytechnique de Grenoble, (2005).

[14] Bogetti, T.A., Gillespie, J.W.: Process-Induced Stress and Deformation in Thick-Section Thermoset Composite Laminates. Vol. 26, pp. 627-660. Journal of Composite Materials, (1992).

Optimisation of asset allocation and reinsurance strategy by the mean of a genetic algorithm : NSGA-II

Mehdi Echchelh
{mehdi.echchelh@ensae.org[1]}

ENSAE Paristech

Abstract. The management of insurance companies is a complex problem in which several indicators intervene and for which the interactions between these indicators and the management levers (reinsurance strategy, asset allocation, etc.) are not simple. We apply one of the most popular multi-objective optimization methods, namely NSGA-II. The implemented algorithm allows to find almost 98% of the hypervolume of the Pareto front in 20 iterations. In addition, the quality of representation of the Pareto front seems to be independent of the number of business of lines to be optimized, which may suggest that the work in this study may be scaled to larger companies for business use.

Keywords: Multi-objective Optimization, Genetic Algorithm, NSGA-II, Insurance, Reinsurance, Asset allocation.

1 Introduction

The optimization of reinsurance and asset allocation strategy is not an obvious problem as these instruments have an impact on both the profitability and the risk profile of the company. In this study, we adress the problem with one of the most popular multiobjective evolutionary algorithms, namely NSGA-II.

We note that the optimization of reinsurance treaties has already been addressed in previous work (see in particular [14], [7], [11] and [5]). The previous approaches were mainly concerned with the optimization of layers of reinsurance treaty for a given line of business. This optimization problem can not be solved analytically, thus Schlottmann, Oesterreicher and Miteschele [11] have highlighted the interest of evolutionary algorithms for this purpose.

However, to our knowledge, there is no previous work dealing with the simultaneous optimization of asset allocation and reinsurance strategy, moreover using the NSGA-II algorithm. Our work examines this problem while taking into account the usual key indicators of insurers (P&L, ROE, Solvency Ratio, SCR).

The results obtained show that the implemented algorithm reach 98% of the hypervolume of the Pareto front in 20 iterations which is achieved in 17 min of computation with the hardware used in our experiments. In addition, the algorithm implemented seems to find the Pareto front independently of the number of the lines of business studied.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284089

2 Reinsurance and asset allocation

2.1 Reinsurance

The two types of contracts studied in this article are:

- *Per risk XL (Excess of Loss):* the reinsurer agrees to settle $min(max(S - \min_{XL}, 0), \max_{XL})$. $\min_{XL}$ et $\max_{XL}$ are respectively named as the priority and the limit of the reinsurance contract. In return, the insurer must pay P^R to the reinsurer. We set this premium as being equal to the *pure premium*:

$$\mathbb{E}(\, min(max(S - \min_{XL}, 0), \max_{XL})) \tag{1}$$

  The parameters that one seeks to optimize in the case of an XL type contract are therefore $\min_{XL}$ and $\max_{XL}$.
- *Quota-Share:* The Quota-Share reinsurance treaty is a pro rata reinsurance contract in which the reinsurer agrees to pay α_{qs}% of the claims amount in return for the insurer's payment of $\alpha_{qs}(1 - \beta_{qs})$% of the written premiums. β_{qs} is called the ceding commision and is subject to negotiations between insurer and reinsurer. After studying the market data, we propose the following modeling:

$$\beta_{qs} = 80\% \max(1 - S/P\ ,\ 0) \tag{2}$$

  With P and S respectively the historical premium and loss amounts of the insurer. The premium perceived by the reinsurer is therefore written:

$$P^R = \alpha_{qs}(1 - \beta_{qs}) \cdot P \tag{3}$$

  The optimization of a quota-share reinsurance contract therefore lies in the optimization of the parameter α_{qs}, also called the cession rate.

2.1 Asset allocation

The amount of funds to invest is divided here into three different asset classes: stocks, bonds and cash. These assets have different risk and return profile, so the choice of the amounts invested in each of these assets will impact both the risk and the profitability of the whole portfolio. Generally French insurers invest from 80% to 90% of their assets in bonds (see [1], [2], and [3]), which is considered a low-risk investment but produces a lower return than equities.

3 Methods

3.1 Objectives mapping

We define in the following two parts the objective function of our optimization problem, namely:

$$F : \omega \in \Omega \rightarrow \ (f_1(\omega), \dots f_n(\omega)) \tag{4}$$

With Ω the decision space.

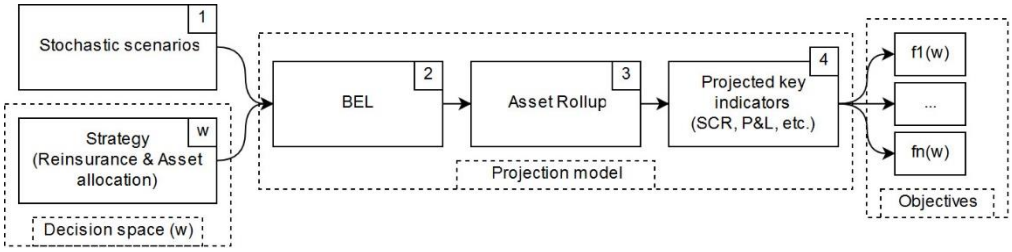


Fig. 1. Objectives mapping.

The calculation of objective values relies on stochastic simulations and on a projection model giving the distribution of different key indicators (P&L, ROE, etc.). Projection models are commonly used in insurance, they enable the assessment of business plan impact on the projection of key indicators. A projection model is defined over a fixed time horizon and with a specific time step, in this study we use a 5 years projection with an annual time step.

The projection model implemented here can be broken down into four stages wich are represented on the figure **Figure 1. [1]**:

1. *Simulation of stochastic scenarios:* we simulate claims' amounts for each line of business as well as the main economic variables (zero-coupon rate, stock return). These scenarios are then used to compute the projection of liabilities and assets (see steps 2 and 3 below). In this study we simulate 1000 stochastic scenarios.
2. *Projection of liabilities:* under the *Solvency II* regulatory framework the insurer's liabilities are composed of three elements: the *Best Estimate Liabilities* (BEL) which is the present value of the insurer's undertaking, the *Risk Margin* (RM) – which we take proportional to the BEL in order to reduce the calculation time – and the Net Asset Value, which is calculated as the difference between the market value of the assets and the sum of the BEL and the RM.
3. *Projection of assets:* once the first two steps have been completed, we project the market value of the company's investments. This projection is carried out by iteratively projecting the market value of the investments, taking into account the evolution of the economic variables, the claims payments as well as the allocation strategy of the company.

4. *Calculation of the key indicators:* Once steps 1, 2 and 3 have been completed, we calculate the value of the SCR and the different key indicators for each projection year and simulation.

3.2 Objectives and constraints

The defined projection model allows us to obtain the distributions of specific key indicators. The objectives are then defined by applying certain statistics (mean, standard deviation, VaR or quantile, etc.) on these indicators.

In the following we will consider two objectives, namely:

- $\mathbb{E}(P\&L)$
- $\mathbb{E}(SCR)$

The constraints taken into account - so called risk appetite constraints - are defined by:

- $VaR_{10\%}^{5\ years}(ROE) > 1\%$
- $VaR_{10\%}^{5\ years}(Solvency\ Ratio) > 200\%$
- $VaR_{90\%}^{5\ years}(S/P) < 100\%$

Let us note that the expectation and the VaR are estimated by using the empirical average and the empirical quantile on the outputs from the projection model.

3.3 Optimisation algorithm

Evolutionary algorithms are naturally suitable algorithms for the optimisation problem studied here since they do not require any assumptions of derivability or convexity of the objective functions (see [9]). Hence we use one of the most popular algorithm, namely NSGA-II.

NSGA-II provides a so-called "truly" multi-objective approach, as it makes it possible to simultaneously optimize several objectives and thus to find the entire Pareto front. As for all genetic algorithms, NSGA-II relies on genetic operators. The crossover and mutation operators used in this study are respectively *SBX (Simulated binary Crossover)* and *polynomial mutation*. The interested reader may refer to Kalyanmoy Deb et al. for more details (see [12]).

Appication to asset allocation and reinsurance optimisation. The crossover and mutation operators initially proposed by K. Deb for the NSGA-II algorithm are defined for real-coded individuals represented by vectors of $[0,1]^n$, with n the number of parameters or "genes". This representation suits perfectly for the optimization of the quantitative parameters (asset allocation, priority and limit of XL treaties). However when taking into account qualitative parameters such as the choice between an XL or a Quota-Share coverage, the real-coded representation of individuals doesn't naturally fit. The solution implemented here is to enumerate all the possible choice for each qualitative parameter and then to optimize the quantitative parameters for each of these choices. The algorithm is as follows:

- Enumerate all possible choices for the types of reinsurance coverage. If we have N_{lob} lines of business and n different types of reinsurance contracts, the number of possibilities is N_{lob}^n.

- For each choice previously enumerated, we optimize the quantitative parameters with the algorithm NSGA-II launched on a parallel worker.

The calculation in this study were done on a computer equiped with an intel i7 8th generation processor with 6 cores. So by using the package *future* of R, we parallelized calculations on the 6 cores, which improved the computation time by a factor of 4.

Dealing with constraints. To take into account the constraints of the optimization problem, we have chosen one of the most popular approach, namely the "multi-objective transformation" approach. It consists in integrating the constraints into the objectives set, thus it preserves the multi-objective aspect of the problem.

The optimization problem is thus as follows:

$$\begin{aligned} \min(&+\mathbb{E}(P\&L), \\ &-\mathbb{E}(SCR), \\ &+VaR_{10\%}^{5\ years}(ROE), \\ &+VaR_{10\%}^{5\ years}(Solvency\ Ratio), \\ &-VaR_{90\%}^{5\ years}(S/P)) \end{aligned} \tag{5}$$

4 Application and results

4.1 Impact of hyperparameters on normalized hypervolume

In this section we study the impact of the number of generations and of the population size on the quality of the Pareto front approximation, which is measured by the *normalised relative hypervolume* ($NRHV$).

Number of generations. The **Figure 2. [2]** gives the evolution of the average $NRHV$ over 30 runs of the optimization algorithm. We observe after the first 20 iterations that the difference between the $NRHV$ of the population and that of the reference front is less than 2.3%. We notice that the iteration 21 is reached on average after 17min24s, which can be considered as a reasonable time for a business use.

NB. - *The reference frontier corresponds to the set of non-dominated stratgies obtained on all the runs. The number of non-dominated strategies is 3,777.*

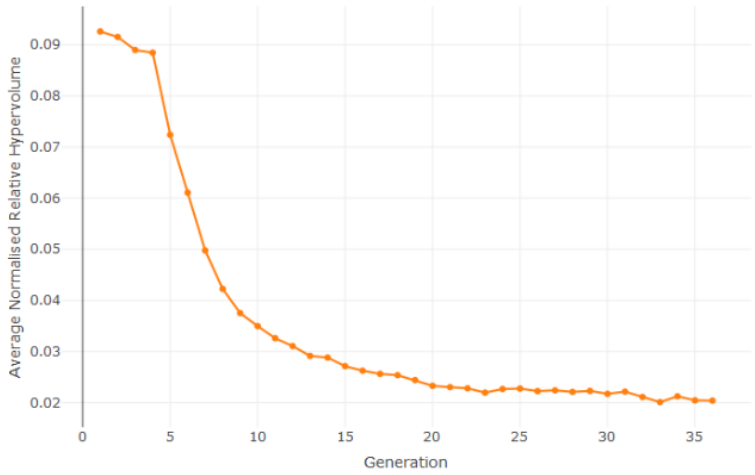


Fig. 2. Evolution of normalized relative hypervolume with the number of iterations.

Population size. We observe in the figure 3 that the normalized relative hypervolume does not seem to be able to decree until 0. This is because the number of individuals is a limit to the quality of the Pareto front approximation. The **Figure 3. [3]** gives an illustration of the impact of the population size on the quality of representation of the Pareto front.

The increase in population size shifts the whole curve downwards. However, in the proposed implementation, the calculation time increases linearly with the number of individuals. Thus the choice of the population size should result in practice from an trade-off between calculation cost and quality of representation of the Pareto front.

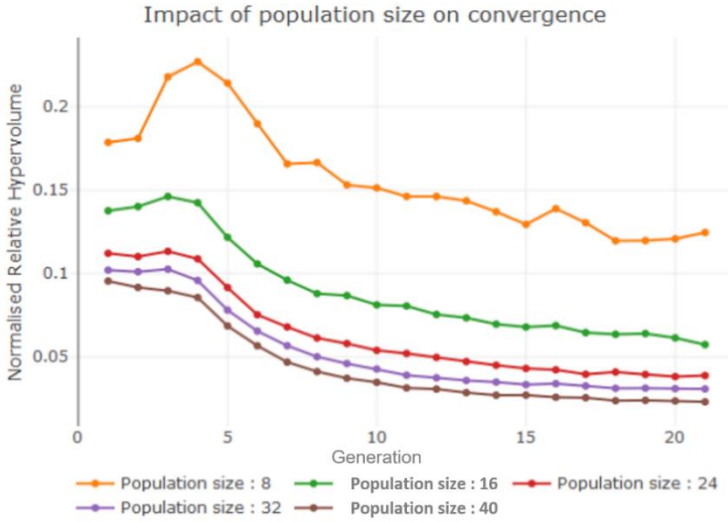


Fig. 3. Impact of population size on convergence.

4.2 Impact of the number of LOB on normalized hypervolume

We study the impact of the number of lines of business on the Pareto front approximation quality. So we optimize the company's strategy in the case where it reinsures one, two or three different business lines. The **Figure 4. [4]** does not show obvious differences between the three curves, besides a one-way ANOVA test allows us to conclude that one can not reject the hypothesis of equality of the Pareto front approximation quality for one, two or three lines of business (p-value : $0.27 > 0.05$).

So the quality of the the Pareto front approximation seems independent of the number of lines of business considered in the problem.

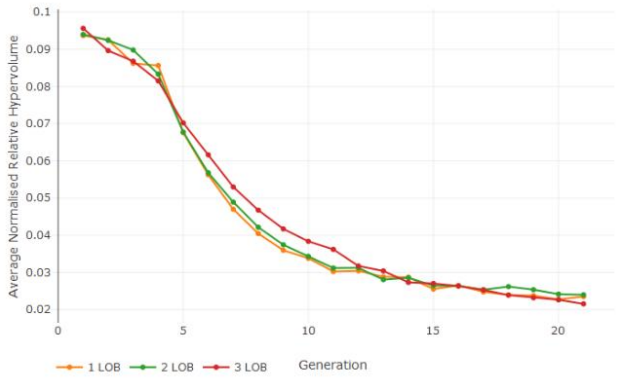


Fig. 4. Impact of the number of business lines on convergence.

4.3 Analysis of the relevance of the Pareto Front

In this section we analyse the Pareto front obtained with the optimization algorithm. The company considered consists of three lines of business and has three assets to its portfolio and its optimization problem is given by the equation (4).

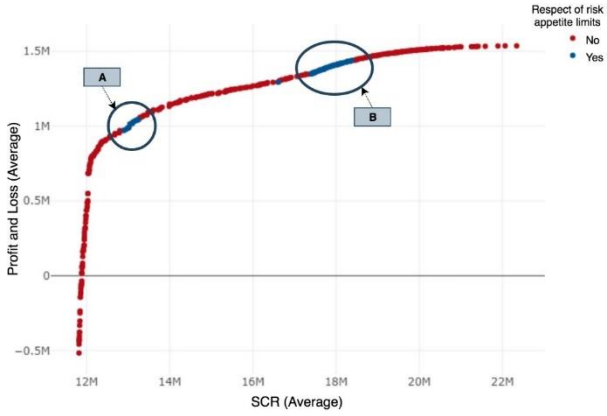


Fig. 5. Illustration of the Pareto front obtained.

On **Figure 5. [5]** we represent the Pareto front. The points for which the constraints are respected are shown in blue, while the points for which the constraints are not respected are shown in red. We observe that the Pareto front has two zones for which the risk appetite constraints are respected, the table 1 gives an overview of these strategies. By comparing the asset allocation strategies obtained with the algorithm with the allocations observed on the French insurance market, we see that the amounts of equity, bonds and cash are in similar proportions for type B strategies, which reinforces the relevance of the projection model and the results obtained with the optimization algorithm.

Table 1. Overview of optimal strategies. Strategies 1 and 2 belong to set A, strategy 3 and 4 belong to set B.

N	$\mathbb{E}(P\&L)$	$\mathbb{E}(SCR)$	Asset Allocation			Reinsurance		
			Equity	Bond	Cash	LOB 1	LOB 2	LOB 3
1	976k	12.9M	1%	95%	0%	460k XL 5k	QS: 75%	QS: 17%
2	1.01M	13M	3%	96%	1%	840k XL 5k	QS: 73%	QS: 15%
3	1.35M	17M	8%	90%	1%	711k XL 5k	328k XL 9k	430k XL 71k
4	1.43M	18.3M	16%	83%	1%	941k XL 5k	856k XL 9k	466k XL 57k

Table 2. Asset allocation of French insurers. Shares invested in equities, bonds and cash by French insurers. Source : ACPR (see [ ]).
NB. – M_i refers to the amount invested in asset i.

Semester	$\frac{M_{stock}}{M_{stock}+M_{bond}+M_{cash}}$	$\frac{M_{bond}}{M_{stock}+M_{bond}+M_{cash}}$	$\frac{M_{cash}}{M_{stock}+M_{bond}+M_{cash}}$
December – 2017	13%	84%	3%
December – 2016	13%	82%	5%

5 Conclusion

Finally, the proposed method highlighted the benefit of evolutionary algorithms in insurance and more particularly in the decision-making process. This type of study is all the more important now that the regulatory context is becoming more and more restrictive for the management of insurance companies.

The implemented algorithm reached 98% of the hypervolume of the reference front after 20 iterations, which was achieved in 17min with a consumer laptop, equipped with an Intel processor i7 8700k. The method implemented here seems therefore relevant in business-use applications, however we note that the proposed parallelization reach its limits when the number of lines of busines becomes too large in comparison to the number of cores of the processor.

In addition, we observed that the quality of representation of the Pareto front is the same when we consider 1, 2 or 3 lines of business. It therefore seems that the proposed approach can be generalized and still offer a good representation of the Pareto front when the number of business lines becomes larger.

However it is important to note that the projection model developed is based on certain assumptions and may require adjustments in a real application. In addition, we have restricted ourselves to the optimization of two objectives - apart from risk appetite constraints - in order to facilitate the visualization of the Pareto front. It might be interesting to explore multicriteria analysis methods to guide decisions when the Pareto front can not be viewed in its entirety.

References

[1] ACPR: La situation des assureurs soumis à *Solvabilité II* en France au premier semestre 2017. (2017)

[2] ACPR: La situation des assureurs soumis à *Solvabilité II* en France au premier semestre 2018. (2018)

[3] ACPR: La situation des assureurs soumis à *Solvabilité II* en France à fin 2017. (2018)

[4] Omar Carmona Cortes and Andrew Rau-Chaplin.: Enhanced multiobjective population-based incremental learning with applications in risk treaty optimization. Evolutionnary Intelligence (2016)

[5] Omar Carmona Cortes and Andrew Rau-Chaplin.: Reinsurance analytics using serial and parallel computation on the multiobjective evolutionary algorithm SPEA2. (2018)

[6] Omar Carmona Cortes and Andrew Rau-Chaplin.: Tackling Reinsurance Contract Optimization by Means of Evolutionary Algorithms and HPC. (2018)

[7] Omar Carmona Cortes and Andrew Rau-Chaplin, Pedro Felipe Do Prado.: A new vector evaluated PBIL algorithm for reinsurance analytics. (2015)

[8] Omar Carmona Cortes and Andrew Rau-Chaplin, Rafael Ferandes Lopes.: A survey on reinsurrance contract optimization using evolutionary and swarm computation. IEEE Latin America Transactions. (2016)

[9] Van Veldhuizen Coello, Lamont.: The book of MOEA: Evolutionary Algorithms for Solving Multi-Objective Problems. (2015)

[10] Detlef Seese Frank Schlottmann, Andreas Mitschele.: A multi-objective approach to integrated risk management.

[11] F. Schlottmann. I. Oesterreicher, A. Mitschele.: Comparison of multiobjective evolutionary algorithms in optimizing combinations of reinsurance contracts.

[12] Sameer Agarwal-T. Meyarivan Kalyanmoy Deb, Amrit Pratap.: fast and elitist multiobjective genetic algorithm: NSGA-II. (2002)

[13] Journal officiel de l'Union européenne. Règlement délégué (UE) 2015/ 35 de la commission - du 10 octobre 2014.

[14] Duane Wilson, Jürgen Gaiser-Porter, Omar Andres Carmona Cortes, Andrew Rau-Chaplin.: On PBIL, DE and PSO for optimization of reinsurance contracts. (2014)

Extensive Analysis of Clustered Routing Protocols For Heteregeneous Sensor Networks

Y. Yazid[1], I. Ez-zazi[1], M. Salhaoui[1,2], M. Arioua[1], A. El Oualkadi[1], A. Guerrero González[2]
{yazidyassine8, mounir.arioua, eloualkadi }@gmail.com

[1]Laboratory of Information and Communication Technologies, National School of Applied Sciences,

Abdelmalek Essaadi University, Tangier, Morocco
[2] Polytechnic University of Cartagena, Spain

Abstract. Energy efficiency has been widely considered as a critical issue in wireless sensor networks (WSN) due to the limited battery powered nodes. Therein, communication process is the most energy demanding in sensor nodes. Accordingly, using energy aware routing protocols to minimize the communication cost and prolong the network lifetime is of paramount importance. In this paper, an extensive analysis of routing protocols based on clustering approach is carried out. We have mainly focused on the performant clustered routing protocols dedicated for heterogeneous WSNs namely Stable Election Protocol (SEP), and its most recent based routing algorithms. The performances of these protocols based on various parameters such as energy efficiency, network lifetime, throughput and stability have been studied. The simulation results show that Threshold SEP clustered routing algorithm provides the best performance in terms of energy, network lifetime and throughput compared to other SEP routing algorithms.

Keywords: WSN, Energy Consumption, Routing, Clustering, SEP.

1 Introduction

In the last few years, wireless sensor networks (WSN) have experienced a tremendous upsurge. Notably, this technology has found application in a large number of fields, e.g. industry, monitoring, smart building and health care **[1] [2].** WSN consists of a large number of sensor nodes which are supposed to constantly sense, supervise and control the surrounding environment. In sensor networks, wearable nodes are likely to be battery powered and thus rely on a limited energy supply to sense, process and communicate the data to the base station (BS) **[3].** Consequently, it is of paramount importance to design energy aware protocols and algorithms, as replacing depleted batteries may be a costly and difficult operation.

Wireless data communication has been deemed the most critical energy demanding among other energy sources in sensor nodes **[4].** For this reason, routing approaches have been recently considered as one of the crucial energy efficient techniques used in WSN to lower the communication energy burden. Cluster-based routing architectures have been widely used in WSN due to their energy efficiency and load balancing in the network **[5].** Sensor nodes in cluster architecture are grouped into clusters in which a cluster head (CH) is elected, and a group of source nodes are directly attached to the CH. In recent works, various clustered routing protocols have been developed for the sake to prolong the longevity of the network **[6][7].** However, most of these protocols are designed based on energy homogeneous nodes, which are supposed to have equal initial energy such as LEACH, PEGASIS and HEED **[8][9][10].**

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284208

Typically, sensor nodes are energy heterogeneous in many applications. For this reason, providing energy-centric heterogeneous routing protocols is highly required.

This paper provides a performance analysis of the most energy-aware heterogeneous clustered routing protocols based on Stable Election Protocol (SEP). To this end, we have thoroughly investigated the performances of routing protocols according to their energy efficiency, network lifetime, throughput and stability.

The rest of the paper is organized as follows. Section 2 describes the WSN system model and the radio energy model used. Section 3 discusses the main clustered routing protocols based on SEP algorithm and their performances. Section 4 shows the simulation results and discusses the performance of various SEP based protocols for heterogeneous WSN. Section 5 concludes the paper.

2 System model

We consider a heterogeneous sensor network composed by N sensor nodes that constantly monitor a phenomenon of interest and communicate data to a distant BS in a clustered fashion. Each sensor node in the network is assumed to be battery-powered with a limited energy supply. Sensor nodes access the channel in a Time Division Multiple Access (TDMA) method, in which time is partitioned into frames. Each frame corresponds to the time interval [tk; tk+1]. In each frame, nodes convey the sensed data to the attached CH, and then to the BS which is assumed to be in the center of the network in order to reduce the communication distances. Moreover, nodes are supposed to be fix in the network without any mobility system.

2.1 Radio energy model

The wireless communication process is the most energy consuming in sensor networks compared to sensing and processing. The first order radio model is used in this work to evaluate the energy consumption of sensor nodes **[8]**. This energy model highlights the energy consumed in transmission, reception and circuitry. It depends mainly on the distance between transmitter and receiver, the packet size and the channel pathloss. Consequently, the energy required to transmit a *k*-bits information over a distance *d* is given by

$$E_{Tx}(k,d) = \begin{cases} k \times E_{elec} + \varepsilon_{fs} \times d^2 , & if\ d < d_0 \\ k \times E_{elec} + \varepsilon_{mp} \times d^4 , & if\ d \geq d_0 \end{cases} \quad (1)$$

Where E_{elec} is the energy required to run the transmitter and receiver circuitry, ε_{fs} and ε_{mp} are the energies per bit used for the transmission amplifier in free space ($\approx d^2$) and Two ray multipath ($\approx d^4$), respectively, and d is the distance between the transmitter and receiver. Note that when the inter- node distance $d < d_0$, the transmission energy is consumed according to the Free space model, while when $d \geq d_0$ the transmission energy follows the Two-ray ground model. The transition from Free space to Two-ray ground propagation is carried out according to a threshold distance d_0, which is expressed as function of both free space and two-ray amplifiers as follows:

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{2}$$

Where ε_{fs} and ε_{mp} are respectively the power amplifier when using free space(fs) model and the multipath (mp) model.

The energy consumption required for reception depends on the circuitry energy E_{elec} and received packet size k , which is given by

$$E_{Rx}(k) = E_{elec} \times k \tag{3}$$

3 Clustered based routing

Clustering has been recently considered a pivotal approach to save transmission energy in WSN **[11][1].** Indeed, the sensor network is divided into several groups namely clusters. In each cluster, a cluster head is elected among nodes, and several nodes are attached to the elected CH. Accordingly, each cluster contains a CH and several member nodes, and the network is formed by a group of clusters. The member nodes sense the environment and forward the collected data to the CH, which is supposed to process and transmit the aggregated data to the BS. The clustering process enables to balance the energy among all nodes in the network and extends the network lifetime. **Figure 1** shows the typical mechanism of clustering that consists of clusters, cluster heads, cluster members and a distant BS.

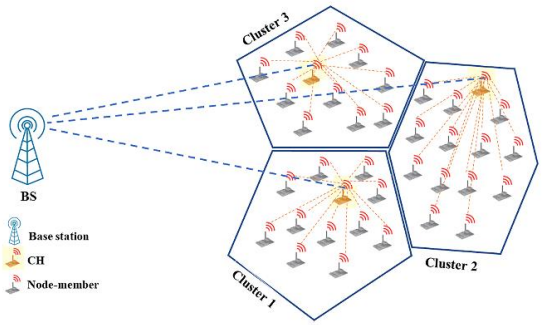


Fig. 1. Clustering technique in WSN

Among prior works, several approaches have been investigated to select efficiently the CH. For instance, in LEACH protocol, a node is elected to be a CH based on a determined probability **[11].** Other protocols are based on the energy level of nodes to decide

the elected one. To this end, many clustered protocols have been developed according to various conditions of CH selection.

3.1 Stable Election Protocol (SEP)

SEP algorithm **[12]**, improves the stable region of the hierarchy process by using the characteristic parameters of the heterogeneity, namely the fraction of the advanced nodes m and the additional energy factor between the advanced and normal nodes (α). In order to extend the stable region, SEP tries to maintain the constraint of a well-balanced energy consumption. Intuitively, advanced nodes must become cluster leaders more often than normal nodes, which equate to a fairness constraint on energy consumption. The total energy of the new heterogenization is equal to:

$$E_{Total} = n \times (1-m) \times E_0 + n \times m \times E_0 \times (1+\alpha) = n \times E_0 \times (1+\alpha \times m) \tag{4}$$

Where n is total number of nodes, m is the fraction of advanced nodes, E_0 is the supposed initial energy of nodes and α is the energy of advanced nodes. And the probabilities of normal and advanced nodes are respectively given by

$$P_{nrm} = \frac{P_{opt}}{1+\alpha \times m} \tag{5}$$

and

$$P_{adv} = \frac{P_{opt} \times (1+\alpha)}{1+\alpha \times m} \tag{6}$$

Where P_{opt} is the optimal probability of each node to become a CH. Furthermore, those must satisfy the conditions below to be CHs in the network:

$$T(s_{nrm}) = \begin{cases} \dfrac{P_{nrm}}{1 - P_{nrm} \times (r \bmod \dfrac{1}{P_{nrm}})}, & if\ s_{nrm} \in G' \\ 0, & Otherwise \end{cases} \tag{7}$$

$$T(s_{adv}) = \begin{cases} \dfrac{P_{adv}}{1 - P_{adv} \times (r \bmod \dfrac{1}{P_{adv}})}, & if\ s_{adv} \in G'' \\ 0, & Otherwise \end{cases} \tag{1}$$

Where G' and G'' are the sets of nodes that have not been CH in last $\frac{1}{P_{nrm}}$ and $\frac{1}{P_{adv}}$ rounds respectively for normal and advanced nodes , and r is the associated round.

3.2 Enhanced Stable Election Protocol (E-SEP)

E-SEP **[13]**, or enhanced type of the protocol SEP. Unlike SEP, this protocol deals with multi-level of sensor nodes in the network: normal, intermediate and advanced nodes. The total energy in the network is given by

$$E_{Total} = n \times (1-m-\lambda) \times E_0 + n \times m \times E_0 \times (1+\alpha) + n \times \lambda \times E_0 \times (1+\mu)$$
$$= n \times E_0 \times (1+\alpha \times m + \lambda \times \mu) \quad (2)$$

Where, α is the energy of advanced nodes, λ is the proportion of intermediate nodes that have μ times more energy than normal nodes where $\mu = \frac{\alpha}{2}$, n is total number of nodes, m is the fraction of advanced nodes, E_0 is the supposed initial energy of nodes .

As in SEP, in E-SEP CHs are selected depending on probability of each type of node. Probabilities of becoming CH for normal, intermediate and advanced nodes, respectively are given by

$$P_{nrm} = \frac{P_{opt}}{(1+\alpha \times m + \lambda \times \mu)} \quad (3)$$

$$P_{\text{int}} = \frac{P_{opt} \times (1+\alpha)}{(1+\mu)(1+\alpha \times m + \lambda \times \mu)} \quad (4)$$

$$P_{adv} = \frac{P_{opt} \times (1+\alpha)}{(1+\lambda)(1+\alpha \times m + \lambda \times \mu)} \quad (5)$$

3.3 Zonal Stable Election Protocol (Z-SEP)

Z-SEP shows two levels heterogeneity of nodes in the network as in SEP **[14]**. But the difference is that in Z-SEP normal nodes that are distributed near the BS use the direct transmission technique whereas advanced nodes use transmission via cluster heads. Generally, the network when using this technique is divided into three parties, zone 0, zone 1 and zone 2 (See **Figure 2**). In Zone 0 located near the BS when normal nodes are distributed monitor their environment in transmit data directly to the BS. In both zone 1 and zone 2, nodes use the clustering technique to send the aggregated data to the BS.

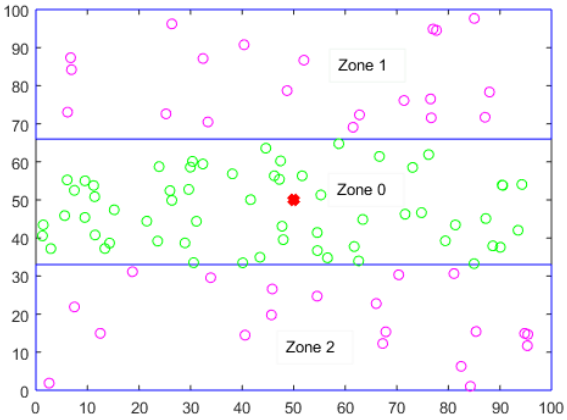


Fig. 2. The distribution of nodes in the network for the Z-SEP protocol

3.4 Threshold Stable Election Protocol (T-SEP)

T-SEP **[15]** is also one of proactive protocols that uses three different levels of heterogeneity where nodes are classed into normal, intermediate and advanced. Moreover, the probabilities of being CH is the same as E-SEP. Whereas to ensure better selection of CHs, T-SEP takes into consideration a threshold for such type of nodes to become a cluster head when each node in the network generates randomly a number between 0 and 1, if the generated value is less than threshold then this node becomes CHs. For each type of nodes, we got these different formulas for the calculation of threshold in function of their probabilities:

$$T(s_{nrm}) = \begin{cases} \dfrac{P_{nrm}}{1 - P_{nrm} \times (r \bmod \dfrac{1}{P_{nrm}})} & , \quad if\ s_{nrm} \in G' \\ 0 & , \quad Otherwise \end{cases} \tag{6}$$

$$T(s_{\text{int}}) = \begin{cases} \dfrac{P_{adv}}{1 - P_{\text{int}} \times (r \bmod \dfrac{1}{P_{\text{int}}})} & , \quad if\ s_{\text{int}} \in G'' \\ 0 & , \quad Otherwise \end{cases} \tag{7}$$

$$T(s_{adv}) = \begin{cases} \dfrac{P_{adv}}{1 - P_{adv} \times (r \bmod \dfrac{1}{P_{adv}})} & , \quad if\ s_{adv} \in G''' \\ 0 & , \quad Otherwise \end{cases} \tag{8}$$

Where G', G'' and G''' are the sets of nodes that have not been CH in last $\frac{1}{P_{nrm}}$, $\frac{1}{P_{\text{int}}}$ and $\frac{1}{P_{adv}}$ rounds respectively for normal, intermediate and advanced nodes.

4 Results and discussions

In this section, we evaluate the performances of SEP based routing protocols according to various parameters such as energy efficiency, network lifetime, throughput and stability. A clustered WSN is implemented in Matlab simulator to evaluate the routing schemes. We have considered 100 heterogeneous sensor nodes randomly deployed in a square field of dimension 100 *m* x 100 *m* and communicate with a BS which is located in the middle of the network. Nodes are assumed to have an initial energy level of E_o= 0.5 *J*. The simulation parameters are listed in Table 1. For our simulations, to simplify, we did not take into consideration the energy loss caused by signal collision and interference in wireless channel. Furthermore, the energy value of normal nodes is obtained by adjusting the initial energy E_o in function of the percentage of advanced and intermediate nodes in the network for each protocol. For the protocol Z-SEP, we choose m=0.4 (40 % of advanced nodes) with a=0.3, and 60% of normal nodes. To get the energy value of normal nodes we did this calculation: $\left(50 - Eo\left(1+a\right)\right)/60$. Although, for E-SEP, where three levels of heterogeneity are required in the network, we choose 30 % of nodes

are advanced with a=0.3, 20 % of nodes are intermediate with b=a/2 and 50% are normal, then we got the initial energy value of normal nodes by using this formula: $(50-Eo(30\times(1+a)\ +20\times(1+b)))/50.$

In addition, we assumed the following distributions of type of nodes: if the network contains two level of heterogeneity, we choose 60% of normal nodes and 40 % of advanced ones, however, in three levels of heterogeneity we choose 50% of normal nodes, 30% of advanced nodes and 20% of super or intermediate nodes according to the case.

Table 1. Simulation parameters.

Parameter	Value
Network's size	100 m x 100 m
Total number of nodes	100
Initial energy E_o	0.5 J
Message size	4000 bits
P_{opt}	0.1
E_{elec}	50 nJ/bit
E_{fs}	10 nJ/bit/m²
E_{amp}	0.0013 pJ/bit/m²
E_{da}	5 nJ/bit/signal
Soft Threshold	100
Hard Threshold	2

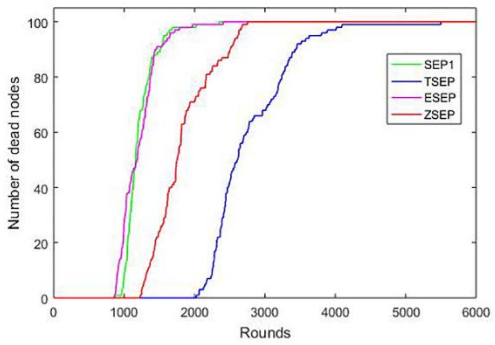


Fig. 3. Number of dead nodes versus rounds for SEP, T-SEP, E-SEP and Z-SEP

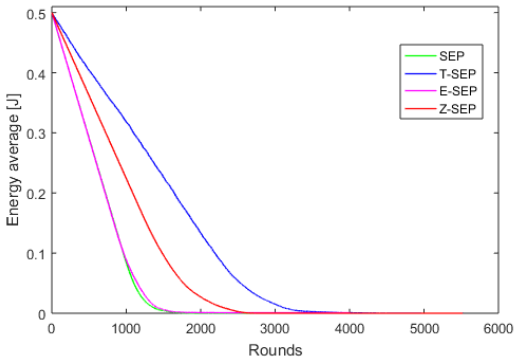


Fig. 4. Residual energy average of the network versus rounds for SEP, T-SEP, E-SEP and Z-SEP

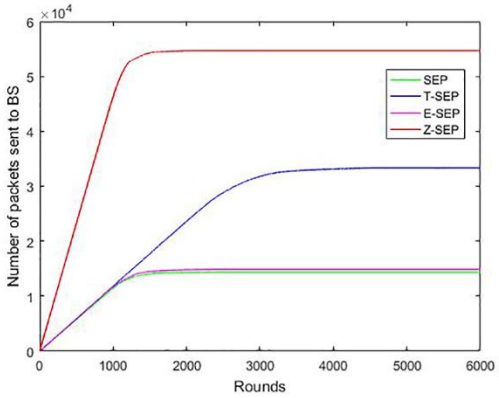


Fig. 5. Number of packets received by the BS using SEP, T-SEP, E-SEP and Z-SEP

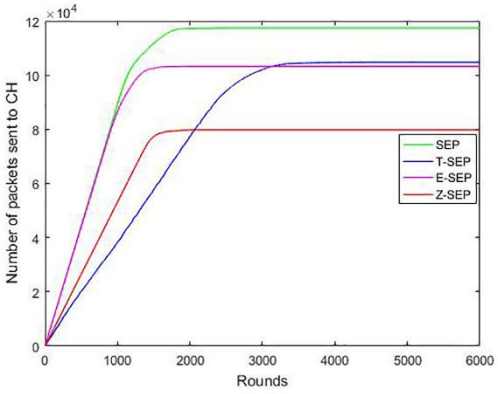


Fig. 6. Number of packets received by CH using SEP, T-SEP, E-SEP and Z-SEP

The average energy and number of dead nodes versus transmission rounds are shown in **Figures 3** and **4**, respectively. These **Figures** compare the performance of the studied routing algorithms i.e., SEP, E-SEP, Z-SEP and T-SEP. Table 2 summarizes the rounds when 1%, 50% and 100% of nodes are dead as well as the stability period of each protocol. The obtained results show that T-SEP performs better than the other protocols in terms of energy and network lifetime. When using T-SEP protocol, the percentages of 1%, 50 % and 100 % of dead nodes occur at 2026, 2604 and 5504 rounds, respectively. This corresponds to lifetime increase of 56 %, 55% and 57%, respectively, when compared to SEP protocol. Besides, T-SEP protocol outclasses SEP, Z-SEP and E-SEP in terms of stability as it lasts 2026 rounds before nodes start dying compared to 1228, 879 and 859 rounds when using Z-SEP, SEP and E-SEP, respectively. **Figures 5** and **6** show the number of packets transmitted successfully to the CH and BS, respectively. This is to assess the throughput performance of the studied SEP based protocols in heterogeneous WSN. Table 3 provides the maximum number of packets sent to corresponding CH and BS for each protocol. The results depict that SEP provides the best throughput at intra clusters (i.e. 117408 packets), flowed by T-SEP (i.e. 104767 packets), E-SEP (i.e. 103275 packets) and then Z-SEP (i.e. 5452 packets). For the throughput to the BS, Z-SEP is proved to have the best performance (i.e. 54673 packets) compared to T-SEP (i.e. 33328), E-SEP (i.e. 14838) and then SEP (i.e. 14360).

As we have noticed the SEP protocol has a large number of packets transmitted to the cluster heads, this is explained by the fact that this one uses a large number of clusters, which also means a large number of CH nodes and finally a reduction in the number of packets transmitted by the member nodes to their CHs in the associated cluster. More precisely, if there are many clusters in the network, it means that there are many cluster heads, which also means fewer member nodes (no CH nodes), therefore fewer packets transmitted to CH and more packets transmitted to the BS**.** Conversely, if there are few clusters, then few cluster heads, which means many non-cluster head nodes and consequently many packets are transmitted to the CHs and fewer packets are transmitted to the BS. However, the best performances of throughput to the BS provided by Z-SEP is explained by the fact that 60% of nodes deployed in zone 0 (**Figure 2**) transmit their aggregated data directly to the BS. Indeed, the other 40% of nodes are distributed equally into two groups zone 1 and zone 2 exactly 20% in each one. That

means a smaller number of CHs in those two zones, consequently less packets transmitted in a cluster by member nodes to their CH which explains likewise the poor performance of Z-SEP represented in red curve in **Figure 6**. Moreover, when the network losses all its sensor nodes, this means immediately no packets transmissions are occurred whether to their CH or to BS, the fact that explains the curves stabilization in **Figures 5** and **6**.

Generally, the nodes energy consumption is inversely proportional to the number of rounds ran down. Moreover, in first rounds, all nodes in the network are able to communicate but just after certain rounds, some ones become unable to carry out a radio communication operation due to insufficient energy residual in their batteries.

Then, round after round the network continues to lose node after node till the last one. This operation of node loss in the network after the course of the cycles affects also the number of packets transmitted whatsoever towards the CHs or to the BS. More precisely, in our simulation the number of packets transmitted to CH by T-SEP in the first round is 93 packets although only 23 packets are transmitted in round number 3000 which means a drop of 75.26% of performance carried out in the first round after 3000 rounds. On the other hand, 7 packets are transmitted to the BS in the first round as well just 5 packets are transmitted in the round number 3000 which also signifies a decrease of 28.57% of performance showed in the first round.

Table 2. Performance comparison of SEP based routing protocols

Protocol	**SEP**	**Z-SEP**	**E-SEP**	**T-SEP**
First node dead	879	1228	859	2026
50 % of dead nodes	1162	1747	1178	2604
100% of dead nodes/Network lifetime	2350	2726	2411	5504
Stability period	879	1228	859	2026

Table 3. Number of transmitted packets to cluster heads and to the BS

Protocol	**SEP**	**Z-SEP**	**E-SEP**	**T-SEP**
Max Nu. of packet sent to CH	117408	5452	103275	104767
Max Nu. of packet sent to BS	14360	54673	14838	33328

5 Conclusion

In this work, an extensive analysis of routing protocols based on SEP in heterogeneous WSN is presented. This paper has mainly focused on the imminent SEP based routing protocols dedicated for clustered architecture, namely, SEP, E-SEP, T-SEP and Z-SEP. We have provided a performance comparison of the studied routing protocols based on their energy efficiency, network lifetime, throughput and stability. The simulation results have shown that the T-SEP protocol outperforms SEP, Z-SEP and E-SEP in terms of energy efficiency, network lifetime and stability. However, SEP and T-SEP have proved to be efficient than Z-SEP and E-SEP in terms of throughput at CH nodes and BS, respectively. Consequently, T-SEP can be a suitable choice for energy constrained heterogeneous WSN in order to save energy and prolong the network lifespan.

References

[1] A. Berl et al., “Energy-efficient cloud computing,” Comput. J., vol. 53, no. 7, pp. 1045–1051, 2010.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” IEEE Commun. Mag., vol. 40, no. 8, pp. 102–114, 2002.

[3] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “In-network aggregation techniques for wireless sensor networks: a survey,” IEEE Wirel. Commun., vol. 14, no. 2, 2007.

[4] N. A. Pantazis and D. D. Vergados, “A survey on power control issues in wireless sensor networks,” IEEE Commun. Surv. Tutorials, vol. 9, no. 4, pp. 86–107, 2007.

[5] S. Bandyopadhyay and E. J. Coyle, “An energy efficient hierarchical clustering algorithm for wireless sensor networks,” in INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies, 2003, vol. 3, pp. 1713–1723.

[6] S. Tanwar, N. Kumar, and J. J. P. C. Rodrigues, “A systematic review on heterogeneous routing protocols for wireless sensor network,” J. Netw. Comput. Appl., vol. 53, pp. 39–56, 2015.

[7] A. Mahboub, M. Arioua, I. Ez-Zazi, and A. El Oualkadi, “Multi-zonal approach clustering based on stable election protocol in heterogeneous wireless sensor networks,” in Information Science and Technology (CiSt), 2016 4th IEEE International Colloquium on, 2016, pp. 912–917.

[8] M. Lauridsen, B. Vejlgaard, I. Z. Kovacs, H. Nguyen, and P. Mogensen, “Interference measurements in the European 868 MHz ISM band with focus on LoRa and SigFox,” in Wireless Communications and Networking Conference (WCNC), 2017 IEEE, 2017, pp. 1–6.

[9] S. Lindsey and C. S. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in Aerospace conference proceedings, 2002. IEEE, 2002, vol. 3, p. 3.

[10] O. Younis and S. Fahmy, “HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” IEEE Trans. Mob. Comput., vol. 3, no. 4, pp. 366–379, 2004.

[11] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on, 2000, p. 10–pp.

[12] G. Smaragdakis, I. Matta, and A. Bestavros, “SEP: A stable election protocol for clustered heterogeneous wireless sensor networks,” Boston University Computer Science Department, 2004.

[13] S. Faisal, N. Javaid, A. Javaid, M. A. Khan, S. H. Bouk, and Z. A. Khan, “Z-SEP: Zonal-stable election protocol for wireless sensor networks,” arXiv Prepr. arXiv1303.5364, 2013.

[14] S. Faisal, N. Javaid, A. Javaid, M. A. Khan, S. H. Bouk, and Z. A. Khan, “Z-SEP: Zonal-stable election protocol for wireless sensor networks,” *arXiv Prepr. arXiv1303.5364*, 2013.

[15] A. Kashaf, N. Javaid, Z. A. Khan, and I. A. Khan, “TSEP: Threshold-sensitive stable election protocol for WSNs,” in Frontiers of Information Technology (FIT), 2012 10th International Conference on, 2012, pp. 164–168.

Query expansion using Wikidata attributes' values

Sarah Dahir[1], Abderrahim El Qadi[2], Hamid Bennis[3]
{sarah.dahir2012@gmail.com[1], elqadi_a@yahoo.com[2], hamid.bennis@gmail.com[3]}

High School of Technology of Meknès, Moulay Ismail University of Meknès, Morocco[1], High School of Technology of Salé, Mohammed V University of Rabat, Morocco[2], High School of Technology of Meknès, Moulay Ismail University of Meknès, Morocco[3]

Abstract. One of the well known techniques that are used to increase the relevancy of user queries is query reformulation through the expansion of queries. Also, Linked data are being used in various fields and available for different domains. This paper aimed at improving Information Retrieval results using Linked data from Wikipedia in general and Wikidata in particular to expand the queries with attributes' values. The experiments that we have done, using Medline collection and Indri search engine to compare our suggested approach with baseline, lead to improvements in precision at different ranks.

Keywords: Information Retrieval, Query Expansion, Linked Data, Wikipedia, Wikidata.

1 Introduction

Information Retrieval (IR) is based on a matching between the user's query and a collection of documents that results in returning a subset documents judged as relevant and that contain the terms of the query [1].

A common problem in this domain is the vocabulary mismatch problem that refers to the use of different terms, that can be synonyms, polysemes, or inflections, to refer to the same concept which may lead to low recall (i.e. the non retrieval of relevant documents) in the case of synonymy as well as inflections, and low precision (i.e. the retrieval of non-relevant documents) in the case of polysemy [2].

One of the well known methods to improve the relevancy of the results is: query expansion that is done through adding new terms to the initial query based on association rules between the terms [1]. Query expansion is either interactive i.e. expansion terms are suggested to the user and it is up to him or her to use them or not, or automatic which means that the expansion terms are automatically added to the query without notifying the user [3]. However, adding too many terms to the query can affect negatively the results of the expansion method more than adding few terms [3].

In this paper, we suggest using linked data from Wikidata as expansion terms by searching for corresponding Wikipedia [4] articles to determine valuable top links to use in the Wikidata knowledge base; for determining expansion terms from the available attributes values. For instance, Wikidata is a free collaborative and secondary database with 53,615,165 data items that can be read and edited by both humans and machines since it contains linked data and uses linked data standards [5].

This paper is organized as follows: Section 2 discusses related work. Section 3 presents methodological details of our approach, and section 4 addresses its evaluation results, and gives an outlook on future work.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284070

2 Related work

In general, Automatic query Expansion (AQE) techniques can be classified into 4 categories:

- Linguistic analysis [6] - [7]: that can not solve ambiguity issues [2] since they deal with each term seperatly from the others using WordNet [4] - [8] for example. This lexical database which can be used in approaches that are ontology-based [9] - [10] has liabilites that include among others the limited coverage of concepts [11] as well as the very few number of the available relationships (e.g. synonymy) ;
- Query-log analysis: that exploits log files information e.g. the click activity of the users. The information in the logs of ancient queries that may be used to expand the user's query is the relationship between queries and selected documents [12]. In [13] the authors extract probabilistic correlations between query terms and document terms, by analyzing query logs, in order to determine expansion terms. Yet, this technique is not efficient for systems that do not have large logs [12]. In the last decades, the IR field has also known an integration of the context aspect to query expansion. In the work presented in [14], a new model for measuring similarity between web queries was proposed. The Language Model (LM) is used to build the query context, which is composed of the most similar queries to the query to expand and their top-ranked documents. Then, they apply a query expansion approach based on the query context and the Latent Semantic Analyses (LSA) method;
- Linked data techniques [15]: in the domain of health for example; linked data help in corresponding terms that patients use with those used by specialists of the medical domain. In [16], authors use the "Unified Medical Language System" (UMLS) database to dermine synonyms of users' queries phrases. In [17] authors show the importance of Linked Open Data Cloud properties in finding semantically similar alternatives, to user query keywords, to be used for expansion, but in there work they explore only a few number of Dbpedia [18] properties which means that valuable properties may not have been explored. In our work we do not select properties to use depending on their names, we select them depending on their attributes' values. In [3], authors use Wikipedia to solve ambiguity problems of queries buy finding correspondies articles to each query n-grams

3 Proposed Method

To expand queries using Linked data from Wikipedia we:

1. Looked for the query in the English Wikipedia. When the query contained many sentences (table 1); we opted for the advanced search to write the first sentence in the search text field and write the terms in the other sentences in the "One of these words" text field (as shown in **Figure 1**) to avoid not getting any results;

Table 1. Example of used queries, from Medline collection, that contain more than one sentence.

Query number	Query content
14	renal amyloidosis as a complication of tuberculosis and the effects of steroids on this condition. only the terms kidney diseases and nephrotic syndrome were selected by the requester. prednisone and prednisolone are the only steroids of interest.

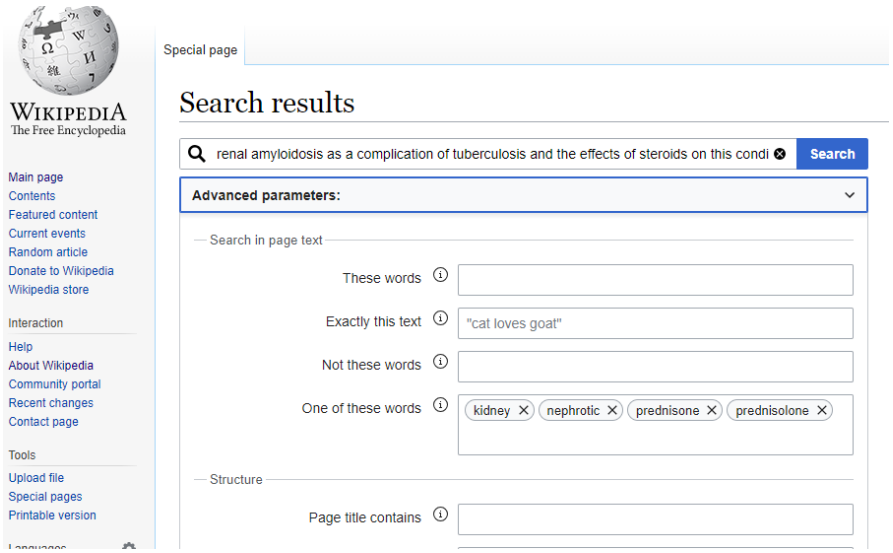


Fig. 1. Use of "One of these words" text field in the case of query number 14 that contains more than one sentence.

2. Selected from the top 10 links (because the most relevant results are believed to be the first ones): the titles that were equal to the query, the ones that contained a sub-part of it (if it was not a stop word), and the ones that contained a sub-part of the query written between parentheses but not preceded by expressions like: "redirects from", "category", and "section" ;
3. Used the "Wikidata item" link in the Wiki page (which is available in the left as one of the "tools" section links) to get its corresponding Wikidata page from which we determined: the attributes' values, from the section "Statements", that contained a query term (or a sub part of it), e.g. "renal carcinoma" which is the value of the property "subclass of", unless it was a stop word and deleted from them the expressions "science/", "subject/", "-pro" (**Figure 2**).

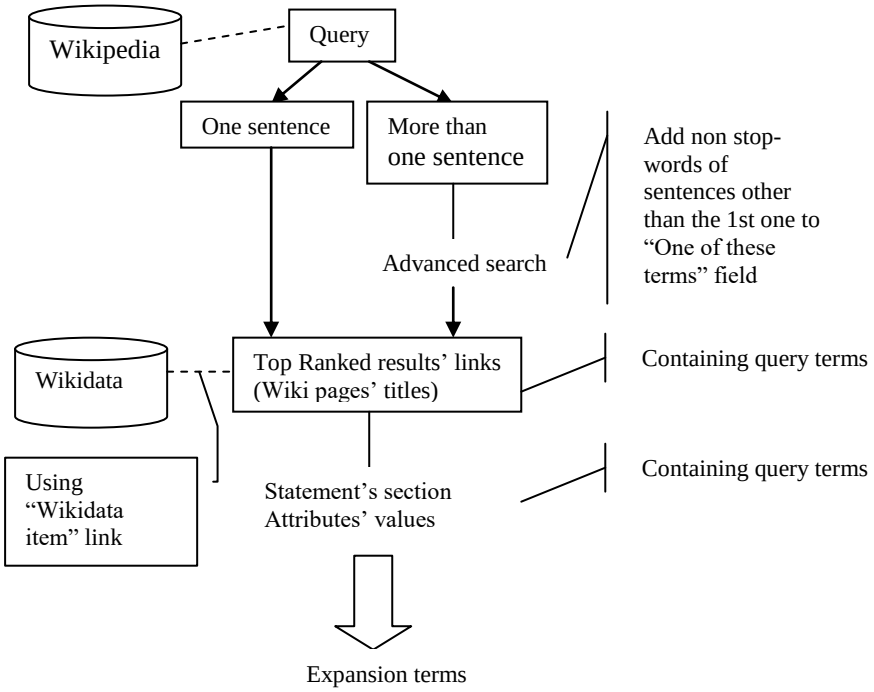


Fig. 2. The flowchart of our suggested approach

Our approach solved the problem of the determination of expansion terms' number to be used; since we used only the terms from Wikidata that contained a query term and not all of the attributes.

4 Results and discussion

To evaluate our approach, we used the collection Medline which is a collection of articles from a medical journal (containing 1033 documents, 30 queries and relevance assessments), that we indexed with a stop words list using Indri search engine.

4.1 Used retrieval model

In this work we used Kullback-Leibler (KL) IR language model for the implementation of our approach. And the general idea behind language models in IR, is about considering that a document D represents a sub-language for which we try to construct a language model M_D[19]. The score of D given a certain query Q is determined by the probability that the documents model generates the query. Or, we try to creat a language model for the query and give a score to a document depending on weather this document can be generated by the query's model [19]. Also, we may use smoothing e.g. Dirichlet to avoid getting a null result when a term is abscent in the constructed laguage model;

4.2 Evaluation metrics

Precision. is a measure (1) that shows how capable a system is of returning only relevant documents [20]:

$$Precision = \frac{\text{Number of relevant retrieved documents}}{\text{Number of retrieved documents}} \quad (1)$$

And Precision at rank N is evaluated by considering only top results returned by the system.

Recall. is a measure (2) that shows how capable a system is of returning all relevant documents [20]:

$$Recall = \frac{\text{Number of relevant retrieved documents}}{\text{Number of relevant documents}} \quad (2)$$

4.3 Results and discussion

To evaluate our method we compared it with a baseline for which we used the pre-defined KL model without any expansion of the queries (see **Figure 3**).

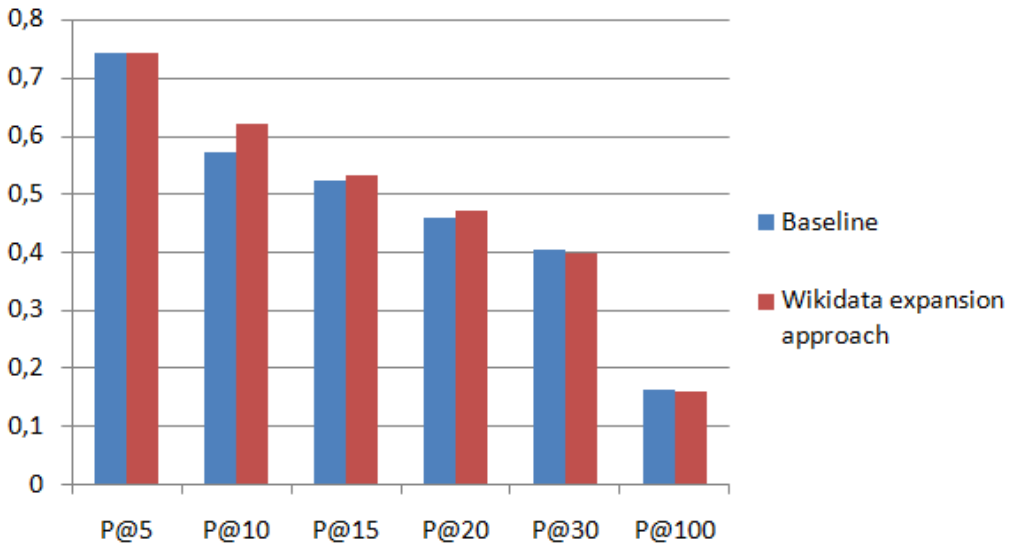


Fig. 3. P@N for our baseline and our suggested Wikidata query expansion approach that both use KL language model to retrieve results for Medline collection using Indri.

Fom the results, we noticed that the use of our suggested query expansion approach improved the baseline in terms of precision at different ranks. For instance, even though the P@5 of the baseline and the expansion method were the same; the P@10 of the suggested approach (which was 0,6214) was much better than that of the baseline (that was 0,5714). In addition to that, the P@15 of the expansion method (0,5333) was higher than that of the baseline (0,5238). And, the P@20 of the expansion approach that attained 0,4714 was better than that of the baseline (0,4607). Moreover, our approach increased slightly the recall at 0,00 since it was 0,8769 for the baseline and became 0,8867 for the Wikidata expansion approach.

Consequently, we believe that our approach is beneficial for Information Retrieval and query expansion; because it is simple and it benefits from linked data advantages through the use of Wikipedia in general and Wikidata in particular.

A possible reason for the unchanged results of P@5, might be the used collection of documents which is Medline that is domain specific and we know that the most appropriate database for this collection of documents is UMLS but the reason why we did not use it is the fact that we want our approach to be applied on any collection and not only on medicine collections. As a result, we think that we could have obtained higher results if we used another collection.

But in order to obtain even higher improvement; we think of using in future work, attributes' names instead of attributes' values and apply the method on other collections of documents. Also, the results vary depending on the retrieved number of documents, the more documents we retrieve; the higher the recall may get. And, the fewer documents we retrieve; the higher the precision may get.

5 Conclusion

To conclude, we may say that the use of Linked data in Automatic Query Expansion helps in the improvement of retrieval results.

Also, our suggested expansion approach that is based on the use of expansion terms from the Wikidata attributes' values increases the precision at different cut-off ranks and improves the recall at 0,00. Moreover, this proposed method is simple and can be used for any collection of documents and not only for collections of a specific domain like the medical domain because it uses Wikipedia and not a medical database like UMLS.

In the future, we will try to improve the relevancy of the results using attributes' names instead of attributes' values.

References

[1] Bouziri, A., Latiri, C., Gaussier, É.: Expansion de requêtes par apprentissage. Conférence en Recherche d'Informations et Applications (2016)

[2] Carpineto, C., and Romano, G.: A Survey of Automatic Query Expansion in Information Retrieval. ACM Comput. Surv., vol. 44, no. 1, pp. 1–50 (2012)

[3] Keikha, A., Ensan, F., and Bagheri, E.: Query expansion using pseudo relevance feedback on wikipedia (2017)

[4] Azad, H.K., and Deepak, A., A New Approach for Query Expansion using Wikipedia and WordNet. arXiv preprint arXiv:1901.10197 (2019)

[5] https://www.wikidata.org/wiki/Wikidata:Main_Page

[6] Moreau, F., Claveau, V., and Sébillot P.: Automatic morphological query expansion using analogy-based machine learning. ECIR'07 - 29th Eur. Conf. Inf. Retr., pp. 222–233 (2007)

[7] Bhogal, J., Macfarlane, A., and Smith, P.: A review of ontology based query expansion. Inf. Process. Manag., vol. 43, no. 4, pp. 866–886 (2007)

[8] Jain, A., Mittal, K., and Tayal, D. K.: Automatically incorporating context meaning for query expansion using graph connectivity measures. Progress in Artificial Intelligence, Volume 2, Issue 2–3, pp. 129–139 (2014)

[9] Hajmoosaei, A., and Skoric, P.: Museum Ontology-Based metatdata. In Tenth IEEE International Conference on Semantic Computing, ICSC 2016, Laguna Hills, CA, USA, February 4-6,2016, pp. 100-103 (2016)

[10] Grieser, K., Baldwin, T., Bohnert, F., and Sonenberg, L.: Using ontological and document similarity to estimate museum exhibit relatedness. Journal on Computing and Cultural Heritage (JOCCH). Vol. 3, Issue 3 (2011)
[11] Sinha, R., Mihalcea, R.: Unsupervised graph-based word sense disambiguation using measures of word semantic similarity. In: Proceedings of ICSC (2007)
[12] Guisado-Gámez, J., Dominguez-Sal, D., and Larriba-Pey, J.-L.: Massive Query Expansion by Exploiting Graph Knowledge Bases for Image Retrieval. Proc. Int. Conf. Multimed. Retr., no. i, p. 33:33--33:40 (2014)
[13] Cui, H., Wen, J.-R., Nie, J.-Y., and Ma, W.-Y.: Probabilistic Query Expansion Using Query Logs. Proceedings of the 11th International Conference on World Wide Web (2002)
[14] Egozi, O., Markovitch, and Gabrilovich, E.: Concept-Based IR using Explicit Semantic Analysis. ACM Trans. Inf. Syst., vol. 0, no. 0 (2000)
[15] Abbes, R. et al.: Apport du Web et du Web de Données pour la recherche d'attributs. Conférence en Recherche d'Information et Applications - CORIA (2013)
[16] Le Maguer, S., Hamon, T., Grabar, N., and Claveau, V.: Recherche d'information médicale pour le patient Impact de ressources terminologiques. COnférence en Recherche d'Information et Applications, CORIA 2015, Mar 2015, Paris, France. Actes de la conférence CORIA (2015)
[17] Augenstein, I., Gentile, A.L., Norton, B., Zhang, Z., and Ciravegna, F.:.Mapping Keywords to Linked Data Resources for Automatic Query Expansion. The Semantic Web: ESWC 2013 Satellite Events. Lecture Notes in Computer Science, vol 7955. Springer, Berlin, Heidelberg (2013)
[18] Dahir, S., El Qadi, A., and Bennis, H.: Enriching User Queries Using DBpedia Features and Relevance Feedback. Procedia Computer Science. Vol.127 Issue C, pp. 499-504 (2018)
[19] Boughanem, M., Kraaij, W., and Nie, J.Y.: Modèles de langue pour la recherche d'information. In : Les systèmes de recherche d'informations. majid Ihadjadene (Eds.), Hermes-Lavoisier, Lavoisier, 11, rue Lavoisier 75008, pp. 163-182 (2004)
[20] https://trec.nist.gov/pubs/trec10/appendices/measures.pdf

Numerical characterization of Symmetric Delamination in Orthotropic plate using Lamb waves

Salah Nissabouri[1], Mhammed El Allami[2], El hassan Boutyour[3]
{s.nissabouri@uhp.ac.ma[1],m.elallami@gmail.com[2],Boutyour.elhassan@gmail.com[3]}

Labo LMEET, Department of physics applied, FST-Settat, Morocco [1,3], Labo LMEET, Department of applied physics, FST-Settat, CRMEF-Settat, Morocco[2]

Abstract. In this paper we study numerical modeling of Lamb waves propagation and its interaction with the edges of symmetric delamination in unidirectional orthotropic plate $[0]_4$. The simulations were carried out using ABAQUS CAE by exciting the fundamental A_0 Lamb mode in the frequency 300 kHz. The delamination was then characterized by analyzing the signal picked up at two sensors using two technics: two Dimensional Fast Fourier Transform (2D-FFT) to identify the propagating and converted modes, and Wavelet Transform (WT) to measure the arrival times. The results showed that the mode A_0 is sensible to symmetric delamination. Besides, based on signal changes with the delamination edges, a localization method is proposed to estimate the position and the length of the delamination.

Keywords: Delamination, orthotropic, Lamb waves, converted modes.

1 Introduction

Composites have been widely used to improve mechanical proprieties of structures. They are used in many fields especially in aeronautic industry [1]. Their characteristics are influenced by the proportions of the matrix and the reinforcements. There are other parameters that also affect the properties of a composite like: size, orientation and distribution of the fibers. However, the drawback of composites is their heterogeneity which leads to their weakness and facilitates the appearance of internal and external damages such as: fiber breakage, matrix cracking, through-thickness hole, local delamination. Among these damages, delamination is especially easy to appear, it causes wave scattering, mode conversion, and multiple reflections. To understand these mechanisms, theoretical, numerical and experimental studies are conducted.

Bo Feng, Artur Lopes Ribeiro, Helena Geirinhas Ramos [2] analyzed the interaction of fundamental symmetric (S_0) and anti-symmetric (A_0) mode Lamb waves with the delamination using finite element simulations. Guo, Cawley [3] studied by finite element analysis and by experiment the interaction of the S_0 Lamb mode with delamination. Ng [4] presented a theoretical and Finite Element (FE) investigation of the scattering characteristics of the fundamental anti-symmetric (A_0) Lamb wave at delamination in a quasi-isotropic composite laminate. Ng, Veidt [5] studied the influence of stacking sequence on scattering characteristics of the fundamental anti-symmetric Lamb wave at through holes in composite laminates. Yang [6] investigated some aspects of numerical simulation of excitation and detection of Lamb waves using piezoelectric disks in plate like composite laminates.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284129

In a previous paper [7] we established the theoretical model of lamb waves' propagation in orthotropic plate. To get deeper insight of these phenomena, this work proposes a numerical study to simulate the lamb waves' interaction with symmetrical delamination

2 Numerical model using Finite Element Method

Many researchers have used the numerical simulation to study lamb waves. The finite element method has been extensively and successfully used to model the Lamb waves and study their interactions with defects in structures [1] [2] [8].

Chiua , Roseb, Nadarajaha [9] built a 3D continuum model in ABAQUS of the 8 ply laminate with a symmetric stacking sequence $[45/-45/0/90]_S$. Ng, Veidt [10] used a three-dimensional Finite Element (FE) method to simulate an eight-ply $[45/-45/0/90]_S$ quasi-isotropic composite laminate with a delamination. The simulations were carried out by ANSYS. Panda [11] presented 3D FE simulations that were carried out to visualize the wave propagation and their interaction with the defect at various depths of an 8-layered Glass Fiber Reinforced Polymer (GFRP) composite laminate using the commercial finite element software package ABAQUS/Explicit. Ramadas [12] investigated the interaction of the primary anti-symmetric mode, A_0 with symmetric delamination type defects in a quasi-isotropic laminated composite using 2D model ANSYS.

In this paper we try to predict the interaction of A_0 mode in orthotropic plate. The 2D numerical simulations were carried out using ABAQUS CAE. The plate has a length L= 400 mm (in the direction of propagation X_1) and a thickness h = 1.6 mm. The mechanical properties of each lamina are shown in Table 1. Once the geometry of the plate has been achieved, it remains to mesh and define a sampling sufficient time.

Table 1. Elastic coefficients C_{ij} (Gpa) and density (kg/m^3) of the materiel

ρ	1530
C_{11}	56.9
$C_{22}= C_{33}$	14.7
$C_{12}= C_{13} = C_{23}$	9.76
$C_{55}= C_{66}$	4.16

2.1 Model without defect

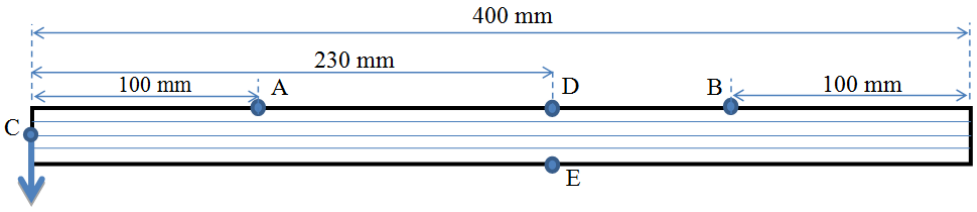


Fig. 1. Numerical Model without defect, C: actuator, A, B, D, E: sensors.

The out of plan displacements measured at the sensors are sinusoidal with the same amplitude. To verify the excitability of pure mode A_0, 2D-FFT was performed at the sensor A. A matrix was calculated at uniformly spaced points.

2D- FFT analysis

Propagating Lamb waves are sinusoidal in both the frequency and spatial domains. Therefore, a temporal Fourier transform may be carried out to go from the time to frequency domain, and then a spatial Fourier transform may be carried out to go to the frequency wave-number domain, where the amplitudes and wave Numbers of individual modes may be measured. Applying spatial Fourier methods in practice to data gained experimentally or numerically requires us to carry out a two-dimensional Fourier transform giving:

$$H(k,f) = \iint_{-\infty}^{+\infty} u(x,t)e^{-i(kx+\omega t)}\,dxdt \tag{1}$$

Where k is the wave number, f is the frequency, $\omega = 2\pi f$. The result of this transformation will be a two-dimensional array of amplitudes at discrete frequencies and wavenumber. [13]

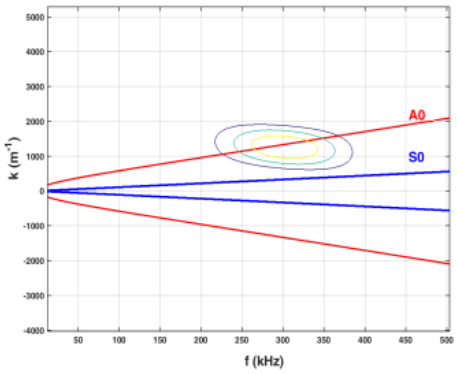


Fig .2. Superposition of FFT2D over analytical dispersion curves

The Fig. 2 shows that there is only one mode propagating which is A_0.

2.2 Model with symmetric delamination

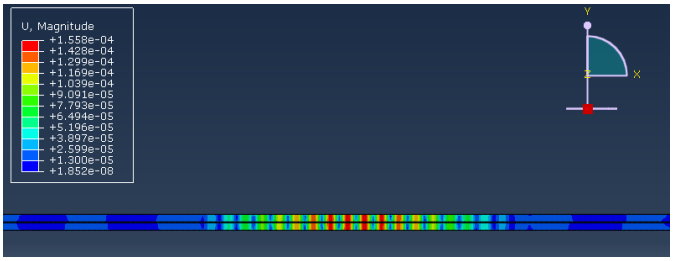


Fig .3. Snapshot of FE simulated out-of plane displacement in the upper and the lower sub-laminates separated by the delamination.

The Fig.3 is a snapshot of FE out-of plane displacement. The Lamb waves split into two parts and travel independently with different velocities in the upper and lower sub-laminates. Mode conversions happen at the two edges of a delamination due to the change of boundary conditions. To characterize the delamination, we analyzed the signal picked up at the sensors A and B.

3 Results

3.1 Signal picked up at Sensor A

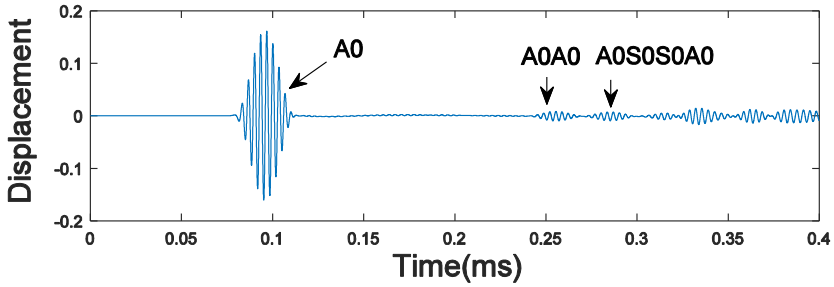


Fig. 4. Out-of-plane displacement, taken on damaged composite $[0]_4$ at x = 100 mm, case delamination length 60 between lamina 1 and lamina 2.

The Fig. 4 is A-scan image from the numerical simulations, taken at 100 mm for a delamination length of 60 mm. It showed also the incident wave A_0 and the two reflected waves: A_0A_0, $A_0S_0S_0A_0$.

3.2 Delamination Characterization from data acquired at sensor A

Wavelet transform is an important tool in the time frequency domain of transient signals. This information is independent of where in time this component appears. The continuous wavelet transform of signal u(t) is defined as follows:

$$cwt(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} u(t)\psi^*\left(\frac{t-b}{a}\right)dt \qquad (2)$$

The transformed signal is a function of two variables, a and b, the translation and scale parameters, respectively. Psi(t) is the transforming function, and it is called the mother wavelet. In this paper we use wavelet transform by 'Gauss1' to locate the pic which permits to determine the arrival time of the wave at specific frequency.

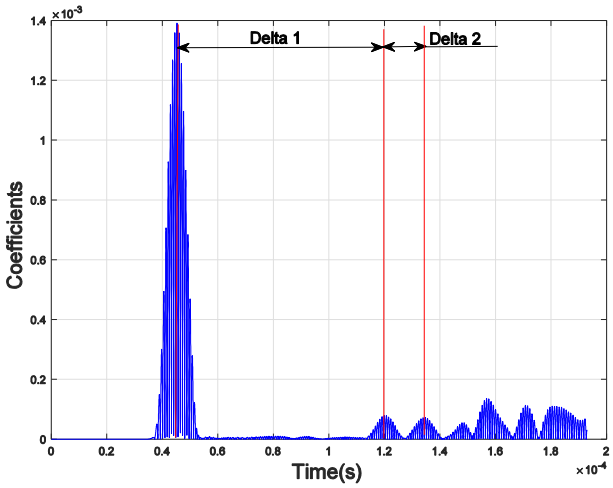


Fig. 5. Coefficients of wavelet gauss1, delamination length 60

The Fig. 5 presents wavelet coefficients versus time at sensor A. The time delay between the peaks permits to estimate the position and the length of delamination. Knowing the group velocity of the modes A_0 and S_0 : Vg(A_0,300khz)=1268 m/s, Vg(S_0,300khz)=5649m/s, we can calculate the delamination position from the sensor A noted P_D by Δt_1, and delamination length noted L_D by Δt_2 as presented in table below :

Table 2. Delamination charactrization : length and position (mm)

Length	P_D	P_{DTH}	Error %	L_D	Error %
60	103,2	100	3,2	62,2	3,7
90	106,6	100	6,6	94,1	4,5
120	103,1	100	3,1	120,9	0,8

The sensor A permits to evaluate the position and the length of symmetric delamination. The Errors associated with the predicted symmetric delamination position range from 3,1 % to 6,6 %. As far as the length prediction is concerned, the errors range from 0,8 % to 3,7 %. So we conclude that more the delamination is long more the length estimation is high.

3.3 Signal picked up at Sensor B

The Fig. 6 represents different wave modes in the case of A_0 mode is incident on the transmission side of the main laminate at sensor B. A new mode S_0 was generated when the A_0 interacted with the entrance of delamination as showed in the figure 6. This new mode traveled within the upper sub-laminates along the length of delamination. Then it reached the exit of delamination, interacted with the exit and generated a new mode A_0 in the forward direction.

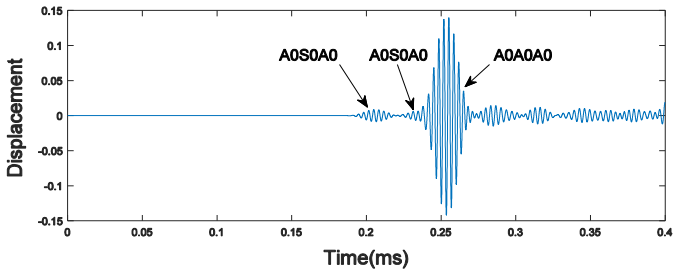


Fig. 6. Out-of-plane displacement, taken on damaged composite $[0]_4$ at x = 300 mm. case delamination length 60

Table 3. Delamination charactrization : length (mm)

Length	L_D	Error %
60	62,9	4,9
90	94,0	4,5
120	122,5	2,1

The sensor B permits to evaluate only the length of symmetric delamination. The Errors associated with the predicted symmetric delamination length range from 2,1 % to 4,9 %.

4 Conclusion

In this work, we have studied the numerical model describing the propagation of anti-symmetric mode A_0 within an orthotropic plate. Simulations based on Finite Element Method were carried out using ABAQUS Software in order to investigate the interaction of Lamb wave mode by symmetric through thickness delamination. The results showed the conversion phenomenon. To quantify the conversion and to characterize the delamination, we have computed the 2D FFT in order to identify the propagating and converted modes; after that, the WT was calculated at different sensors to measure the arrival times. Analyzing the signal obtained at sensors A and B, it was found that the mode conversions happen at the edges of a delamination for A_0 wave. In fact, this mode is sensitive to the position and to the length of delamination; this makes it become an effective tool to evaluate the delamination in orthotropic materiel. Besides, based on signal changes with the delamination edges, a localization method is proposed to estimate the position and the length of the delamination. These are the main results:

- The estimation is accurate when the delamination length increases.
- The optimal position to evaluate the beginning and the length of delamination is between the actuator and the defect: the pulse echo method.

References

[1] Bo Feng, Artur Lopes Ribeiro, Helena Geirinhas Ramos, Interaction of Lamb waves with the edges of a delamination in CFRP composites and a reference-free localization method for delamination, Measurement ,doi: https://doi.org/10.1016/j.measurement.2017.10.016. (2017)

[2] N. Guo a and P. Cawley,The interaction of Lamb waves with delaminations in composite laminates. The Journal of the Acoustical Society of America 94, 2240,(1993)

[3] Ching-Tai Ng, Analytical and finite element prediction of Lamb wave scattering at delaminations in quasi-isotropic composite laminates. Journal of Sound and Vibration, 4870–4883, (2012)

[4] Martin Veidt and Ching-Tai Ng. Influence of stacking sequence on scattering characteristics of the fundamental anti-symmetric Lamb wave at through holes in composite laminates. J. Acoust, 129, (2011)

[5] A De Luca, Numerical simulation of the Lamb wave propagation in impacted CFRP laminate, Procedia Engineering, 109 – 115, (2016)

[6] Chunhui Yang , Some aspects of numerical simulation for Lamb wave propagation in composite laminates, Composite Structures, 267–275,(2006)

[7] S. Nissabouri, M. El Allami, E. Boutyour, A. Errkik, Lamb wave interaction with delamination in orthotropic plate, International Conference on Computing and Wireless Communication Systems, , Laarche, Maroc, ISBN: 978-1-4503-5306-9 , doi>10.1145/3167486.3167545, (2017)

[8] R. Basri and W.K. Chiu, " Numerical analysis on the interaction of guided Lamb waves with a local elastic stiffness reduction in quasi-isotropic composite plate structures", Composite Structures 66 (2004) 87–99.

[9] W.K. Chiua , L.R.F. Roseb and N. Nadarajaha, "Scattering of the fundamental anti -symmetric Lamb wave by a midplane edge delamination in a fiber-composite laminate". Procedia Engineering 188, 317 – 324,2017.

[10] Ching-Tai Ng and Martin Veidta, "Scattering of the fundamental antisymmetric Lamb wave at delaminations in composite laminates". Acoustical Society of America 132, 115, 2012.

[11] Rabi Sankar Panda, "Lamb wave interactions with delaminations in composite laminates using air-coupled ultrasonic visualization". 14th Asia-Pacific Conference on Non-Destructive Testing, 2013.

[12] C Ramadas. "Interaction of the primary anti-symmetric Lamb mode (A0) with symmetric delaminations: numerical and experimental studies". Smart Materials and Structures, 2009.

[13] D. Alleyne and P. Cawley , "A two-dimensional Fourier transform method for the measurementof propagating multimode signals", J. Acoust.Soc. Am. 89 (3), March 1991.

Data mining for road accident analysis in a big data context

Fatima Zahra El Mazouri[1], Mohammed Chaouki Abounaima[2], Said Najah[3], Khalid Zenkouar[4]
{Fatimazahra.elmazouri@usmba.ac.ma[1], medchaouki.abounaima@usmba.ac.ma[2],
Said.najah@usmba.ac.ma[3], khalid.zenkouar@usmba.ac.ma[4]}

Laboratory of Intelligent Systems and Applications, Faculty of Sciences and Technologies, Sidi Mohammed Ben Abdellah University, Fez, Morocco [1, 2, 3, 4]

Abstract. Data Mining techniques and extracting association rules from a big dataset play an interesting role in knowledge discovery. Therefore, the decision makers encounter a huge number of resulting association rules that can make them unable to choose and decide rationally between these different extracted rules, also the time of the generation of these association rules brings a new challenge, we propose to overcome these challenges a learning model based on FP-growth algorithm using Apache Spark framework, in order to analyze data and extract interesting association rules by taking into account some quality measures. Experimental results on road accident data in France show that the proposed approach can provide useful information that could help the decision makers to choose the appropriate strategies in the perspective of improving road safety.

Keywords: Data Mining, Association Rules, FP-growth, Big Data, Apache Spark, Road Accident.

1 Introduction

This work is considered as a continuity of work [1] in which we have combined data mining and multicriteria decision analysis approach, in this work we propose an approach based only on data mining with a refinement step in order to extract the most relevant association rules.

In the last years, a big interest is addressed to analyzing data, in order to keep significant information and extract useful knowledge for further use by the decision maker. Data mining [2] is considered as a very important step of Knowledge Discovery in Database process [3], data mining is a set of algorithms and methods for exploration and analysis large databases in order to extract rules, associations, unknown trends, etc. [4]. Searching for relations and correlations between objects in a large database is a very powerful technique of data mining for discovering relevant information [5].

Extracting association rules from transactional databases have received intensive research since its introduction by Rakesh Agrawal [6], it has been successfully used in various domains and applications such as basket analysis [7] bioinformatics [8] chemo-informatics [9] biomedical [10] image [11].

Since the datasets are extremely large, parallel algorithms are recommended, for these reasons we chose apache spark [12] as a powerful framework to process and analyze big datasets by exploiting machine learning algorithms like FP-growth (Frequent Pattern growth), this algorithm proposed by Han et al. [13] is a very efficient technique thus provides a solution to the problem of searching for frequent patterns and association rules in a large transactional database. In addition, it generates a big number of association rules when the transactional database is large. Therefore, it is necessary to help the decision maker to select the most interesting rules among a large number of extracted rules. We can exceed this difficulty by implementing a post-processing stage of extracted rules, which consists to use a refinement process. This process that we propose in this work is based on 2 steps of refinement, in the first step, we retain that the rules of associations which have the best support and confidence thresholds, in the second step, we retain this time the best association rules, result of step 1, which have the best lift threshold.

Many statistics and works [14,15] show the gravity of road traffic accidents and remain worrying for everyone, and we are seriously trying to find effective solutions to fight against these losses of human lives. In this context, our study provides help for the stakeholders and decision-makers and finds correlations between human damage and the various conditions that characterize all recorded accidents. The aim of this work is to treat the case of road accidents in France by using the BAAC (Accident Analysis Bulletin Corporal) database [16-17], this database made by the French ministry of interior and the ministry of transport, our purpose is to use BAAC database in order to inform decision-makers of the most common conditions at the origin of accidents. To do this, we propose to extract the association rules by using the FP-growth algorithm in a big data environment using Apache Spark. Given a large

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284124

number of the generated association rules, therefore, their exploitation becomes more and more difficult, which leads us to use a refinement process based on support, confidence and lift measures.

The rest of this paper is organized as follows: the two first sections "association rules" and "Apache spark" are designed to give an overview of the association rule techniques and apache spark framework. "Proposed approach" section describes the proposed methodology. The results and discussion are presented in the "Results and evaluation" section. In the last section, we concluded by summarizing the work and giving some perspective in the "conclusion" section.

2 Association rules

2.1 Preliminary

Consider a transaction database D, in which each row contains a set of items (itemset/pattern) I with its identifier (Tid), The Items: A, B, C … each item can design a product (in other context books, cars, projects.)

A pattern or itemset is a set of items: C, AB, BCD...

The generation of association rules from a transactional database consists to extract a rule in the form A → B (A, B ⊂). Every rule is composed of two different sets of items, also known as itemset, A and B, where A is called antecedent or left-hand-side (LHS) and B consequent or right-hand-side (RHS).

An example of a rule A → B mean IF A THEN B: which represent a relation between A and B.

- The support of rule: support(A→B) = support(A∪B), means the support of the union of the itemsets in A and B. with:

$$\text{Support(A)}=\frac{|\{t \in T; A \subseteq t\}|}{|T|}, \quad (1)$$

the support of A with respect to T (with T is the table of transactions) is defined as the proportion of transactions t in the dataset which contains the itemset A.

- The confidence determines how frequently items in B appear in the transactions that contain A, ranges from 0 to 1.

$$\text{Confidence (A}\rightarrow\text{B)} = \frac{\text{support (A}\cup\text{B)}}{\text{support(A)}}. \quad (2)$$

- The lift measures how far from independence are A and B. It ranges within [0, +∞]:

$$\text{Lift(A}\rightarrow\text{B)} = \frac{\text{support (A}\cup\text{B)}}{\text{support(A)}*\text{support(B)}}. \quad (3)$$

2.2 Association rules Algorithms

Many algorithms [**18**] have been designed for generating association rules, we can cite that APRIORI and FP-Growth, are the two most popular for finding association rules in data sets, Moreover, they prove their performance on a transactional database where each transaction design a pattern (an itemset). In [19] a study shows that FP-growth provides much more consistent and quicker performance then APRIORI, because that FP- growth uses divide and conquer strategy and it needs only two passes over the datasets so it is much more scalable.

2.3 FP-Growth algorithm

FP-growth is a data mining algorithm proposed by Han and al. [13], based on a data structure known as the FP-tree (It represents the frequent pattern tree), FP-growth has a major step for the extraction of frequent itemset from this structure: it compresses the frequent itemset represented in the database using an FP-Tree (frequent-pattern tree) whose branches contain the possible associations of the items. Each association can be divided into fragments (pattern fragment) which constitute the frequent itemset. The FP-Growth method transforms the problem of finding the longest frequent itemset by searching for the smaller one and its concatenation with the corresponding suffix (the last frequent item in the branch leading to the item in question) which leads to reduce the cost of research.

FP-growth algorithm:

- Use a compressed presentation of the database using an FP-tree.
- Once an FP-tree has been constructed, it uses a recursive divide-and-conquer approach to mine the frequent itemsets.

Building the FP-Tree

1. Scan data to determine the support count of each item.
   Infrequent items are discarded, while the frequent items are sorted in decreasing support counts.
2. Make a second pass over the data to construct the FP-Tree.
   As the transactions are read, before being processed, their items are sorted according to the above order.

3 Apache Spark

Apache Spark is an open-source cluster computing framework for real-time processing [12] this platform offers a variety of machine learning and data mining programs; these programs are very efficient and able to process and to analyze a large amount of data. Apache Spark has an RDD (Resilient Distributed Dataset) an immutable distributed collection of data, partitioned across nodes in a cluster that can be operated in parallel. Apache spark provides high-level APIs in Java, Scala, Python and R, and an optimized engine that supports general execution graphs.

Spark runs in master-slave mode, that is, a master and one or more workers. There is a driver that talks to a single coordinator named master that manages workers in which executors run. More clearly, when you start treatment on the Spark framework, you go through the Driver which is in a way the master, itself communicates with the cluster manager, the latter manages the resources of the workers (the workers execute the actual processing) (see Figure 1).

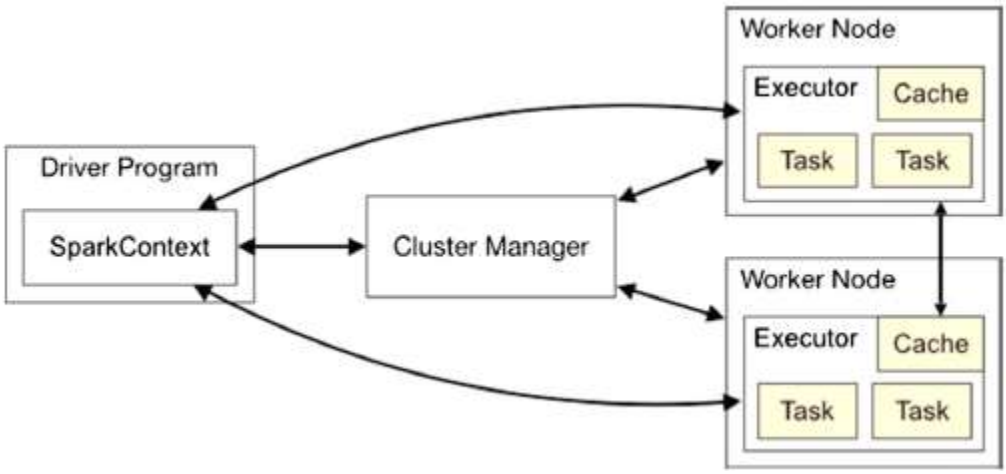


Fig. 1. Spark architecture.

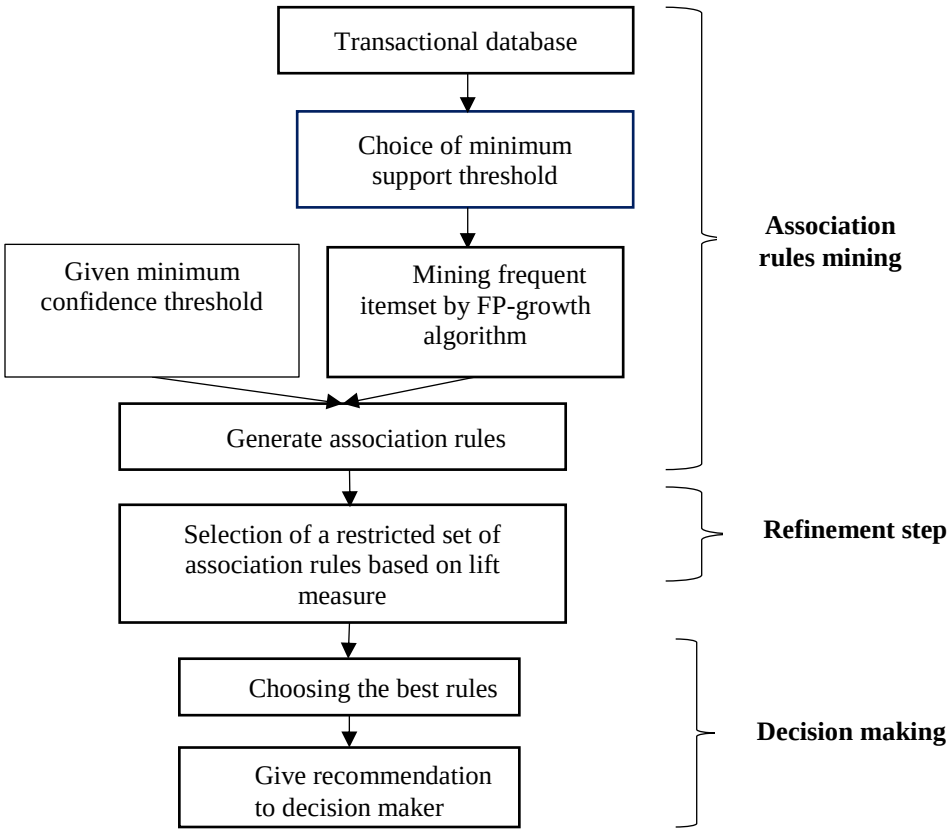
4 Proposed approach

In data mining, association rules algorithms produce a large number of rules that make the decision about the most interesting and useful association rules very difficult for the decision maker. To deal with this problem the integration of an evaluation of those association rules using some quality measures like support, confidence and lift can minimize the number of the generated association rules.

The purpose of our proposed approach is to extract knowledge about the most relevant conditions that cause the majority of road accident in France in the year 2017. According to the report published by ONISR (National Inter-Ministerial Observatory of Road Safety) on accidents occurred on the roads of France in 2017 [20], the cost of the corporeal accidents in metropolitan France would be 39.7 billion euros (B€) distributed as follows:

- 11.3 B€ in mortality;
- 23.1 B€ for hospitalizations;
- 4.0 B€ n for light victims;
- 1.2 B€ for the material damage of these bodily injuries.

The obtained association rule by the application of our approach is very useful to adopt for the right strategies to improve road safety in France.



The approach is summarized in 5 steps, the flowchart given in Figure 2 shows the full process.

Fig. 2. Flowchart of the proposed approach.

This proposed approach is described in the following steps:

Step1: collect data

We propose in this work, to use the databases collected on all the accidents that were recorded during the year 2017. The BAAC databases are available on the net [16-17], is composed of the 4 files in CSV format:

- Characteristics: This table describes the general circumstances of the accident. There are 16 attributes that are retained with 60 702 records.
- Places: This table describes the location where the accident occurred. There are also 18 attributes that are retained with 60 702 places described.
- Vehicles: This table describes the vehicles involved in the accident. There are also 9 attributes that are used for this table and we have 103 547 vehicles recorded.
- Users: This table gives all the information on the users involved in the accidents, a user can be a driver or a pedestrian. This table contains 12 attributes with 136 022 users registered.

Our first task was to create a database under the ORACLE DBMS for better exploitation of the data of 4 files (tables), we proceeded to the joining of the 4 tables to have a global vision on all the accidents. Finally, we obtained a transactions database with 135696 rows (accidents) and 31 attributes. Figure 3 shows an excerpt of the data.

	A	B	C	D	E	F
1	MONTH	HOUR	Light	Urbain	Intersection	Atmospheric_condition
2	January	Evening	Night-with-public-lighting-lights	In-urban	Out-of-intersection	Normal-ATM
3	January	Evening	Night-with-public-lighting-lights	In-urban	Out-of-intersection	Normal-ATM
4	January	Evening	Night-with-public-lighting-lights	In-urban	Out-of-intersection	Normal-ATM
5	February	Evening	Fullday	In-urban	Intersection-in-T	Normal-ATM
6	February	Evening	Fullday	In-urban	Intersection-in-T	Normal-ATM
7	March	Morning	Fullday	In-urban	Other_intersection	Normal-ATM

Fig. 3. An extract of the transactional database.

Step 2: Application of FP-growth algorithm on Apache Spark environment for association rules extraction

To have an appropriate model for extracting association rules from our database, we used the FP-growth algorithm on Apache spark environment. First, we get the frequent itemset by using minimum threshold support, then, the generation of the association rules based on the obtained frequent itemset and the minimum confidence threshold. In this study, we did the experiment with several thresholds (50%, 60%, 70%, 90%, and 100%) as shown in Table 1.

Step 3: Step of refinement

In this step, we try to choose the most relevant rule by crushing the thresholds of support and confidence until we obtain a set of minimal rules, and then, we get the most appropriate rule based on the rule with the highest lift.

Step 4: Recommendation of the final solution

At the last step, according to the classification of the association rules obtained by the application of the previous steps, we recommend to the decision-maker the most relevant associations in order to define, for example, the security measures to be considered in the short-term, median-term or in the long-term.

5 Result and discussion

At this stage, we present the analysis result of the treatment performed for association rules extraction, and we end by a refinement process to select the most interesting rule among the set of the obtained association rules. The experiments of our proposed approach are done by using road accident data in France in the year of 2017. The system was built on Spark by using single nodes.

5.1 Building the transactions table T

The obtained table of transactions T by merging the 4 tables (Characteristics - Places - Vehicles - Users) from the BAAC database, with an SQL join, as shown in Figure 3. This table contains 135 696 transactions, and each transaction is described in maximum by 220 items. Seen the size of the table of transactions, we present only an excerpt in Figure 3.

5.2 Application of FP-growth on apache spark environment

In this step, we chose the FP-growth algorithm in a big data environment using Apache Spark in order to generate association rules. Then, we look for the minimum support threshold and minimum confidence threshold giving the minimum number of association rules. The choice of thresholds minconf = 0.8 and minsup = 0.6 leads to the extraction of a minimum of association rules, a total of 15 association rules, which seems a more logical choice of the thresholds, see Table 1. But these 15 rules are extracted only at the base of the two measures support and confidence thresholds.

Table 1. Experimentation of several thresholds of confidence and support.

Support thresholds	Confidence Threshold	Number of extracted rules
0.5	0.5	116
0.6	0.5	19
0.7	0.5	2
0.8	0.5	0
0.9	0.5	0
0.5	0.6	116
0.6	0.6	19
0.7	0.6	2
0.8	0.6	0
0.9	0.6	0
0.5	0.7	94
0.6	0.7	19
0.7	0.7	2
0.8	0.7	0
0.9	0.7	0
0.5	0.8	60
0.6	**0.8**	**15**
0.7	0.8	2
0.8	0.8	0
0.9	0.8	0
0.5	0.9	16
0.6	0.9	4
0.7	0.9	2
0.8	0.9	0
0.9	0.9	0
0.5	1	0
0.6	1	0
0.7	1	0
0.8	1	0
0.9	1	0

Table 2 gives the 15 rules resulting from this step of association rules extraction with a minimum confidence threshold equal to 0.6 and a minimum support threshold equal to 0.8.

Table 2. The obtained association rules.

	LHS	RHS	Support	confidence
R1	{Drawing_in_plan=Straight-part}	{declivity_of_the_road=Dish}	0.606529	0.852739
R2	{Drawing_in_plan=Straight-part}	{Accident_Situation=On-the-pavement}	0.626383	0.880652
R3	{State_of_surface=Normal-surface}	{declivity_of_the_road=Dish}	0.636648	0.81309
R4	{declivity_of_the_road=Dish}	{State_of_surface=Normal-surface}	0.636648	0.81042
R5	{State_of_surface=Normal-surface}	{Accident_Situation=On-the-pavement}	0.636759	0.813231
R6	{State_of_surface=Normal-surface}	{Atmospheric_condition=Normal-ATM}	0.732776	0.935858
R7	{Atmospheric_condition=Normal-ATM}	{State_of_surface=Normal-surface}	0.732776	0.90892
R8	{declivity_of_the_road=Dish}	{Accident_Situation=On-the-pavement}	0.644666	0.820627

R9	{Accident_Situation=On-the-pavement}	{declivity_of_the_road=Dish}	0.644666	0.801176
R10	{declivity_of_the_road=Dish}	{Atmospheric_condition=Normal-ATM}	0.637931	0.812053
R11	{Accident_Situation=On-the-pavement}	{Atmospheric_condition=Normal-ATM}	0.649854	0.807624
R12	{Atmospheric_condition=Normal-ATM}	{Accident_Situation=On-the-pavement}	0.649854	0.806066
R13	{State_of_surface=Normal-surface,Accident_Situation=On-the-pavement}	{Atmospheric_condition=Normal-ATM}	0.600921	0.943719
R14	{Atmospheric_condition=Normal-ATM,State_of_surface=Normal-surface}	{Accident_Situation=On-the-pavement}	0.600921	0.820062
R15	{Atmospheric_condition=Normal-ATM,Accident_Situation=On-the-pavement}	{State_of_surface=Normal-surface}	0.600921	0.924701

5.3 Refinement Step

We rank the rules obtained in Table 2 using the lift measure to find the best association (Table 3).

Table 3. Association rules with lift measure.

Rules	Lift
R1	1.085493097
R2	1.094453198
R3	1.03502142
R4	1.03502142
R5	1.010664323
R6	**1.160819038**
R7	**1.160819038**
R8	1.019855463
R9	1.019855463
R10	1.007253184
R11	1.00175951
R12	1.00175951
R13	**1.170568809**
R14	1.019152938
R15	**1.180974197**

The first refinement step gave us 15 rules (Table 2) based on the 2 criteria: support threshold and confidence threshold. After we keep only 4 rules
R6:{State_of_surface=Normal-surface} → {Atmospheric_condition=Normal-ATM},
R7:{Atmospheric_condition=Normal-ATM} → {State_of_surface=Normal-surface},

R13:{State_of_surface=Normal-surface,Accident_Situation=On-the-pavement}→{Atmospheric-condition= Normal-ATM},
R15:{Atmospheric_condition=Normal-ATM,Accident_Situation=On-the-pavement} → {State_of_surface=Normal-surface},
because they got the best performance of the lift threshold (see Table 3).

In this work, we considered 3 criteria as our previous work [1]: support, confidence and lift, but this time, we proceed by a process of refinement, which starts by considering only the thresholds of support and confidence to extract set of association rules, then, our process ends with the consideration of the lift measure which allows us to keep just the rules having the best measure of lift.

5.4 Discussion and recommendation

The obtained association rules can be very useful for the decision-maker to do the appropriate strategies according to the conditions most related to road accidents, such as the R6 and R7 rules which show a relation between State of surface and Atmospheric conditions. These association rules indicate and show that the majority of French drivers cause more accidents in normal roads and normal weather conditions. This can be justified by several reasons, the first can be the excess speed in these best atmospheric and road conditions, the second is the increase in driver travel, especially for leisure travel. The third reason is related to the negligence of the driver by the consumption of drugs and the underestimation of the danger of the road.

For the rules R13 and R15 we notice a strong relation between the items: normal surface , on the pavement and normal atmospheric condition, these correlations show that the majority of road accident victims are unfortunately in normal roads with a normal atmospheric. Worse still, most of these pedestrians were hit by vehicles on the pavement. Moreover, hitting a pedestrian on the pavement can only be justified by the imprudence and the carelessness of the drivers. This imprudence returns to several reasons: the very young age of some drivers, the abuse of alcohol and drugs, using the phone while driving, etc.

We note that the results obtained in this present study on road traffic accidents incurred in 2017 are almost the same as those obtained in our work [1] concerning the accidents incurred in 2016.

These association rule needs to be further analyzed with the help of road safety specialist to choose the appropriate policies in the perspective of improving road safety. In conclusion, this knowledge is so relevant for road safety investigators and deserves to be studied further.

6 Conclusion

In this work, we tried to illustrate our contribution about reducing the number of the generated association rules from a large database and considering multiple measures of quality, which allow to the decision maker to select the most interesting rule. To do this we used Apache Spark and machine learning, especially the FP-growth algorithm to extract frequent itemsets and generate association rules. As expected, Apache Spark provided its fast execution engine

For distributed processing. We performed the experiment on road accident data in France in 2017 to extract the main causes of the accidents, by applying FP-growth algorithm in Apache Spark environment to extract the minimum number of association rules with a given minimum support threshold and minimum confidence threshold, and for more facilizing the task of selecting the most useful and relevant rule, we proceeded to a phase of refinement by selecting the rule with a high value of lift measure from the minimal set of rules obtained previously. The obtained rule can be easily exploited by the decision maker to choose the appropriate policies and strategies in the perspective of improving road safety.

For future work, a new methodology should be addressed based on deep learning for a good prediction and finer analysis of road accidents, that may occur in the future.

References

[1] El Mazouri, F.Z., Abounaima, M. & Zenkouar: Data mining combined to the multicriteria decision analysis for the improvement of road safety: case of France. J Big Data (2019) 6:5. https://doi.org/10.1186/s40537-018-0165-0.

[2] Lovell, M. C. (1983): Data Mining. The Review of Economics and Statistics. 65 (1): 112. doi:10.2307/1924403. JSTOR 1924403.

[3] Piatetsky-Shapiro, G., Frawley, W. J.: Knowledge Discovery in Databases. AAAI/MIT Press (1991).

[4] Fayyad, U.M., Piatesky-Shapiro, G., Smuth P., Uthurusamy, R.: Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press (1996).

[5] Tiwari, A., Gupta, R., Agrawal, D.: A survey on frequent pattern mining: current status and challenging issues. Inf. Technol. J. 9 (2010) 1278–1293.
[6] Agrawal, R., Imielinski, T., Swami A.: Mining association rules between sets of items in large database. In SIGMOD. pp. 207–216 (1993).
[7] Aguinis,H., Forcum, L. E., Joo, J.: Using Market Basket Analysis in Management Research. Journal of management. https://doi.org/10.1177/0149206312466147, (2012).
[8] Zaki, M. J., Sequeira, K.: Data mining in computational biology IN Handbook of Computational Molecular Biology. Chapman & Hall/CRC Press, ch. 38, pp. 1–26 (2006).
[9] Auer, J., Bajorath, J.: Emerging chemical patterns: A new methodology for molecular classification and compound selection, J. Chem. Inf. Mod, vol. 46, no. 6, pp. 2502–2514 (2006).
[10] Neesha, J., Nur'Aini A.R., Wahidah,H.: Data Mining in Healthcare – A Review. Procedia Computer Science. 72 (2015) 306 – 313.
[11] Sajedi,H.: Steganalysis based on steganography pattern Discovery, Journal of information security and applications. 30 (2 0 1 6) 3–14.
[12] http://spark.apache.org. Accessed in 2019.
[13] Han, J., Pei, J., Yin, Y., Mao, R.: Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Mining Knowledge Discovery. 2004;8(1):53–87.
[14] Wong, J., Chung, Y.: Comparison of methodology approach to identify causal factors of accident severity. Transp Res Rec 2083:190–198 (2008).
[15] Ait-Mlouk,A, Gharnati, F., Agouti, T.: An improved approach for association rule mining using a multi-criteria decision support system: a case study in road safety. Eur. Transp. Res. Rev. (2017) 9:40. DOI 10.1007/s12544-017-0257-5.
[16] https://www.data.gouv.fr/fr/datasets/base-de-donnees-accidents-corporels-de-la-circulation. Accessed in 2018.
[17] http://www.securite-routiere.gouv.fr/la-securite-routiere/l-observatoire-national-interministeriel-de-la-securite-routiere. Accessed in 2018.
[18] Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules. In: Proc. 20th Int. Conf. Very Large Data Bases, pp. 487-499 (1994).
[19] Shah, M., Shah, N., Shetty, A., Shah, D., Gotmare, P.: A comparative study of Pattern Recognition Algorithms on Sales Data. International Journal of Computer Applications (0975-8887), Vol 141-No.1 (2016).
[20] Direction de l'information légale et administrative, La sécurité routière en France : Bilan 2017, ONISR, Paris 2018, ISBN : 978-2-11-077443-9.

Comparative Study of ZigBee and 6LoWPAN Protocols: Review

I. Alaoui Ismaili[1], A. Azyat[2], N.Raissouni[3], N. Ben Achhab[4], A. Chahboun[5], M.Lahraoua[6]
{ iuoala.imane@gmail.com[1], azyat2037@gmail.com[2], naoufal.raissouni.ensa@gmail.com[3] }

Abdelmalek Essaadi University, Tetouan, Morocco[1]
National School of Applied Sciences of Tangier, Morocco[2]
National School of Applied Sciences of Tetouan, Morocco[3]

Abstract. Several types of research have been made on short-range and low-level wireless networks energy consumption since their appearances until today. To examine the specificities of each, we have done a comparative study to highlight the strengths and weaknesses based on energy consumption, scope and reliability of mesh architecture to conclude open international standard, and more adapted to industrial needs. This paper addresses the low power mechanisms provided by 6LoWPAN and the ZigBee Protocol with their two version (ZigBee-Pro and ZigBee IP), providing comparative assessments based on the results obtained by available in specialized literature and different researchers. The results show that ZigBee IP can be the most appropriate protocol in case we want to connect multiple nodes with fast communication and optimized power consumption

Keywords: 6LoWPAN, ZigBee, ZigBee-Pro and ZigBee IP

1 Introduction

The Internet of Things (IoT) is considered to be the next big challenge for the communication research community. It deployed two main categories of networks: short-range low power networks (called last 100 meters of connectivity) and long-range low power networks (called Low-Power Wide Area Networks (LPWAN)). Both short-range and long-range low power networks will play a crucial role in the IoT world where each network type has its unique applications and characteristics [23]. This paper has focused on a comparative study of two short range wireless protocols such as ZigBee and 6LoWPAN in terms of energy consumption, scope and reliability.

Several researches have been made on short-range and low-level wireless networks energy consumption since their appearances until today. For better understand the specifics of standard including ZigBee and 6LoWPAN. The comparison we made in this study does not value one of the technologies discussed, but mostly helps engineers in the design of new wireless network applications. We have presented a state of the art in research and development of architecture and topology of ZigBee and 6LoWPAN. We also proposed a simple study comparing between the ZigBee, ZigBee IP, ZigBee-Pro and 6LoWPAN protocols by considering the criteria of modulation type, network, topology, power, range, cryptography, spreading and coexistence

The paper is organized as follows. First part outlines related works while Second part describes the two (ZigBee and 6LoWPAN) protocols with a focus on their architecture and topology. The third part shows the experimental results and comparison between the ZigBee standard, ZigBee IP and 6LoWPAN while the last part provides our conclusions

2 Related works

Recently, many papers review, compare and evaluate ZigBee and 6LoWPAN with other wireless technologies. In [1], Lee, Su, & Shen in the study of the most used wireless networks (Bluetooth, UWB, ZigBee and Wi-Fi), they evaluated the main characteristics and behaviors in terms of various measures, including the efficiency of data coding, the delay of transmission, complexity of operation and energy consumption. They came up with results that claim that ZigBee is taking more time compared to others but still the better suited solution for mobile applications of battery-operated wireless sensor networks. In [2], Baker has done comparative study to highlight the strengths and weaknesses of ZigBee and Bluetooth in industrial applications by certifying that the ZigBee standard based on IEEE802.15.4 is more efficient in terms of power consumption, reliability and scope. They conclude that ZigBee is intended to be an open international standard, more adapted to industrial needs. In

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284215

[3], Vinay Kumar & Sudarshan Tiwari was done on the different routing algorithm and IP mobility in 6LoWPAN and comparing the basis of different metric like energy consumption, memory uses, mobility, scalability, and so forth. It was found that the routing protocols have their own advantages depending upon the application where they are used. Hi-Low routing protocol provides an advantage of memory saving by which the networks become more scalable. DYMO-Low provides more routing delay compared to other routing protocols. S-AODV provides benefits in terms of traffic reduction, power consumption, and network lifetime extension, for 6LoWPAN. In [4], Mazzer & Tourancheau compared the different implementations of 6LoWPAN on WSN and showcased the wireless technology that has become the primary means of communication for sensor networks. To overcome the problems of conforming to current standards and ensure the accessibility of nodes sensors the 6LoWPAN protocol was designed. In [4], Toscano & Lo discusses the adaptation of the 6LoWPAN protocol and mechanisms of the ZigBee standard to operate at low power. Generally, these mechanisms put nodes on standby to reduce the service cycle. Although both technologies are based on same lower layers, but the ZigBee approach is quite different from that adopted by 6LoWPAN. 6LoWPAN protocol based on 1Pv6 is asynchronous, requires channel listening mechanisms (for power saving purposes) because it does not support sleeper nodes and adopts a mesh topology. ZigBee defines the beacon mode that allows nodes to stay in low power mode for most of the time and periodically sends nodes a sync tag through the use of tree routing. In [4] they tried to make a comparative study between three wireless short-range wireless protocols, Bluetooth and ZigBee to deploy a wireless sensor network. The main intention of their study is to find the most appropriate protocol applicable to wireless sensor networks. In This study analyzed 3 protocols based on a transmission time, rate of bit error, signal strength received, and packet delivery ration and energy consumption. Thus, can we conclude that ZigBee is the best protocol among the three protocols studied, ZigBee has a stable transmission time, better spectral efficiency, a higher packet transmission ratio and minimal power consumption. In [5], Lucia Lo Bello in the study of the performance of 6LoWPAN and ZigBee protocols for high speed industrial networks. The results of the experimental evaluations showed that the ZigBee network is capable of handling shorter run cycles, generating maximum end-to-end delays and update times slightly below the theoretical values of 6LowPAN. On the other hand, the 6LoWPAN network has medium end-to-end delays and higher reliability (a lower percentage of packet loss). In [6], Franceschinis & al are tried to compare the two piles of ZigBee-Pro protocol and ZigBee IP developed in order to integrate ZigBee into the Internet of Things

3 Network Architecture

This section will give a thorough description for network architecture of ZigBee and 6LoWPAN. Commonly, the layers or architecture and topologies.

3.1 ZigBee

The ZigBee network (IEEE802.15.4) provides high-level wireless data transmission for the purpose of communicating small devices over radio links. It was created to be a standard for low-cost, high-level communication protocols and will be used in applications that require secure network devices, low data throughput, and longer battery life. ZigBee can support different types of topologies, such as mesh, star, and tree network topologies.

The architecture of the ZigBee standard is composed of 4 layers: physical, MAC, network and application as shown in Fig. 1:

- Physical layer translates sent and received frames into bits and manages the use of radio transmission and channel communication. The standard provides for two different physical layers (one for 868 / 915MHz and one for 2.4 GHz) using discrete sequence spread spectrum OQPSK (DSSS) modulation.
- MAC layer: synchronizes the network and ensures the reliability of communications using CRC and retransmissions, manages the transmission of frames and reserved time slots, and controls access to the wireless medium
- ZigBee layer: provides routing, sends / receives data to and from the application layer, addresses nodes, builds and maintains network topology, and provides security.
- Application Layer: Determines how all other layers behave. It is associated with several modules such as:
  - Security Service Provider which manages security functions
  - Application Sub-Layer support that supports the linking of devices
  - ZigBee Device Object messaging services that enable discovery of devices

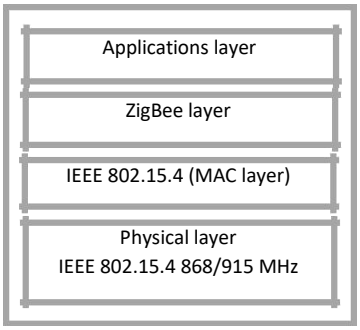


Fig. 1. ZigBee standard architecture

ZigBee includes the complete physical layer, the IEEE802.15.4 MAC specification for personal wireless networks, and other standard layers added based on the version of the ZigBee network used, such as ZigBee Pro and ZigBee IP.

3.1.1 ZigBee Pro

The ZigBee-Pro version is a significant enhancement to the ZigBee standard and advanced support for large networks of many nodes. ZigBee-Pro enables the transmission of large messages using fragmentation and reassembly and uses a gateway for data transfer from ZigBee nodes to the Internet. ZigBee-Pro is based on a mesh topology and not beacon mode. The main feature of this version is present in the table [7] below

Table 1. Features of ZigBee-PRO [7]

Features	Comments
Addressing	The new devices that add up to the network will automatically have an address and if two nodes have the same address, the network layer intervenes based on the MAC address of each node.
Links management	Each node of the mesh topology is able to communicate with its neighbor
Changed channel	The channel manager selects a new channel and notifies the rest of the network multiple ZigBee nodes detect overlap or noise interference Channel change
Fragmentation	ZigBee-Pro provides the fragmentation of larger packets into smaller ones to facilitate sending.
Power Management	Put devices in sleep mode to consume less power (When they wake up, they update themselves in the case of a change in the network)
Routing	ZigBee-Pro offers two types of routing multicast and many-ta-one.
Security	There are two security modes in ZigBee-Pro: Standard Mode and Superior Security Mode

The architecture of ZigBee-Pro as Franceschinis and al defined in 2013 is as follows

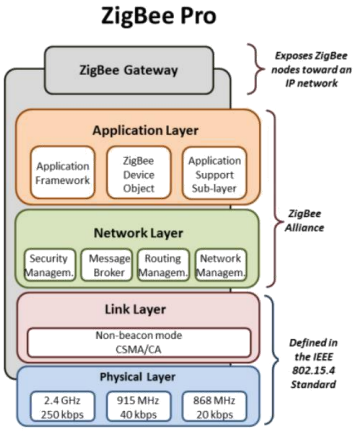


Fig. 2. ZigBee Pro network architecture [6]

3.1.2 ZigBee IP

ZigBee IP is for low-speed, short-range wireless mesh networks that support IPv6. It ensures the establishment of Internet connections seamlessly with low power and low-cost devices

ZigBee IP is based on the physical and MAC layers of the IEEE802.15.4 standard and improves network and application layers as recommended by the Internet Engineering Task Force. It provides end-to-end IPv6 networking without the use of intermediate gateways, uses standard protocols such as IPv6, 6LoWPAN, RPL, TCPIUDP and provides security through Transport Layer protocol. Figure 3 shows the protocols used at each layer of the ZigBee IP architecture. The physical and Link layers are based on IEEE 802.15.4 in the 2.4 GHz band. The NWK layer includes the RPL, IPv6, ICMP protocols. RPL is associated with 6LowPAN, the adaptation layer for the exchange of IPv6 packets on an IEEE 802.15.4-based network. TRN layer may adopt either TCP that is required to support HTTP or UDP that is optional to support CoAP

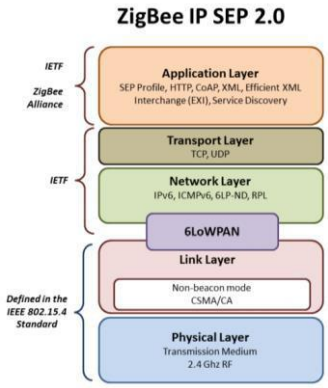


Fig.3. ZigBee IP network architecture [6]

3.2 6LoWPAN

The IPv6 over low-power wireless (6LoWPAN) is the short-range, low-power and personal area networks. It can be connected directly to another IP network without intermediate entities (gateways). The connection between 6LoWPAN elements is implemented via IEEE 802.15.4. It supports different lengths of addresses. It is also low bandwidth and low-cost power consumption. 6LoWPAN can support different types of topologies, such as mesh and star topologies.

The architecture of the 6LoWPAN is composed of 3 layers: host node, router node and edge router as shown in Fig. 2:

- Host node sense the physical environment and can actuated devices
- Router node Transfer data packets from hosts to edge routers or to a 6LoWPAN network destination. The connection between 6LoWPAN elements is implemented via IPv6 over IEEE 802.15.4.
- Edge router provide interconnection between the 6LoWPAN network and other IP networks. Sending and receiving packets between 6LoWPAN elements and IP nodes of other networks. Each 6LoWPAN element is identified by a unique IPv6 address.

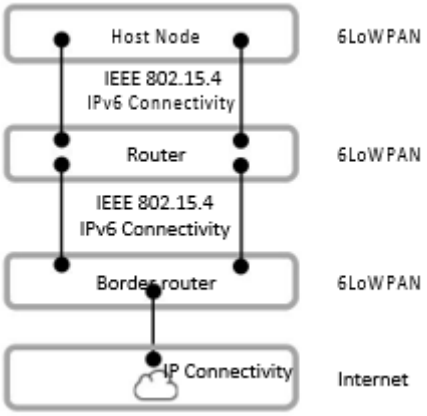


Fig.4. 6LLoWPAN architecture

Sending packets in a 6LoWPAN topology begins with the RFD endpoint that routes the same packet to an FFD router node that forwards, hopped the packet to the 6LoWPAN gateway. This gateway that is connected to the internet with the 1Pv6 domain will then forward the packet to its destination using the IP protocol.

4 Comparative study

This section, present different criteria are used to benchmark the differences between the ZigBee, ZigBee IP, ZigBee-Pro and 6LoWPAN protocols. Such criteria include standard, modulation type, network, topology, power, range, cryptography, spreading and coexistence

4.1 ZigBee-Pro and ZigBee IP

In [21], I. BERDAI compared the results of the third topology of the ZigBee-Pro network obtained with the XBee devices with the results of the first topology of the ZigBee IP developed on Contiki. For ZigBee-Pro, she emulated with real hardware and ZigBee IP was simulated virtually using Contiki. Table 1 shows the simulation results

Table 2. Comparison of results between ZigBee-Pro and ZigBee IP [22]

	ZigBee-Pro	ZigBee IP
Delay	213.073ms	127ms
Energy	0.00031J	2,28 x lo-s J

Test results shows that ZigBee-Pro, takes more time and more energy. On the other hand, data exchange for ZigBee IP is faster and energy consumption has been optimized. This is due to the use of advanced mechanisms like IPv6, 6LoWPAN

4.2 ZigBee and 6LoWPAN

The comparative study of ZigBee standard and 6LoWPAN reveals that 6LoWPAN adds IP functionality to WPANs and consumes little power, while ZigBee supports many more nodes operates at low power and requires a low cost. ZigBee can be used in HAN networks, as well as for smart metering, if it is used in a mesh structure. It can also provide remote intelligent meter monitoring and other devices. ZigBee has reliable security and uses powerful encryption techniques. It has a networking technique far superior to that of other technologies that avoids the collision of channels. On the other hand, 6LoWPAN is suitable for IP-based low power devices such as sensors and controllers. The main attributes of these technologies have been summarized in Table 2. In terms of power consumption, 6LoWPAN and ZigBee are designed for portable devices and limited battery power. Thus, it offers low power consumption. In the context of the comparison between ZigBee IP and 6LoWPAN, ZigBee IP is considered the first open standard of ZigBee compatible with IPv6 protocol. It provides seamless connection with low-power, low-cost devices. The 6LoWPAN adaptation layer built into ZigBee IP ensures interoperability between the link layer defined by the 802.15.4 standard and the network layer that supports IPv6, this is to bind any which device to the internet.

Table 3. Comparison between the 6LoWPAN and ZigBee

	6LoWPAN [6],[18], [20],[8]	ZigBee [18], [6],[21],[8]	ZigBee IP [9], [10], [12-20]
Radio Frequency	868MHz,915MHz, 2,4GHZ	868MHz,915MHz, 2,4GHZ	868MHz,915MHz, 2,4GHZ
RF Data Rate	250Kbps	250 Kbps	250 Kbps
Peak Current Consumption	Rx 20-35 mA Tx 12-25 mA	Rx 20-35 mA Tx 20-30 mA	
Network Topology	Star or Mesh	mesh, star	Cluster tree, Mesh
Suitable for WBANs	Suitable for application with low data rate and on-body sensors	portable devices	portable devices and for application with low data rate and on-body sensors
Max Outdoor Range	10-200 meters	500 meters	10-1000
Security	128b AES and application layer user defined	AES	AES block cipher (CTR, counter mode)
Max Nodes	100	65,536	> 65000
Standard	IEEE 802.15.4	IEEE802.15.4	IEEE802.15.4
Common Applications	Monitor and Control via internet	Home industry monitoring and controlling	,Monitor and Control via internet
Mesh Network Support	Many	Zigbee	Many
Internet Connectivity	Bridge/Router	Zigbee Gateway	Zigbee Gateway/ Bridge/Router
RF Radio Support	802.15.4	802.15.4	802.15.4
Network Size		65K	
Header Overhead	2-11 bytes	8-16 bytes	
Code Size with mesh	22K	32K to 64K+	

5 Conclusion

This paper presented state-of-the-art in research and development of architecture and topology of ZigBee and 6LoWPAN. The intension of this paper is not to draw any conclusion regarding which one is superior to others but to find the most appropriate protocol applicable for wireless sensor networks. In our work we have analyzed and compared ZigBee, ZigBee IP and 6LoWPAN protocols based on Network Topology, Max Outdoor Range, Security, and Max Nodes. According to our analysis we can conclude that ZigBee IP is the best protocol among those 3 as it has stable transmission time, lowest power consumption, highest packet delivery ratio, can bind any which device to the internet and have the Max Outdoor Range.

References

[1] Lee, J. S., Su, Y. W., & Shen "A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. IECON Proceedings" (Industrial Electronics Conference), 46-51. ttp://doi.org/10.1109/IECON.2007.4460126

[2] Baker "ZigBee and Bluetooth: strengths and weaknesses for industrial applications" Computing and Control Engineering, 16(2), 20-25. http:/ /doi.org/1 0.1 049/cce: 20050204

[3] Mazzer, Y., & Tourancheau" Comparisons of6LoWPAN implementations on wireless sensor networks". Proceedings -2009 3rd International Conference on Sensor Technologies and

[4] Toscano, E., & Lo Bello, L "Comparative assessments of IEEE 802.15.4/ZigBee and 6LoWPAN for law-power industrial WSNs in realistic scenarios" In 2012 9th IEEE International Workshop on Factory Communication Systems (pp. 115-124). IEEE. http://doi.org/10.1109/WFCS.2012.6242553

[5] Emanuele Toscano and Lucia Lo Bello "Comparative assessments of IEEE 802.15.4/ZigBee and 6LoWPAN for lowpower industrial WSNs in realistic scenarios" https://www.researchgate.net/publication/261203734, 22 October 2016

[6] Md. Taslim Arefin, Mohammad Hanif Ali, A.K.M. Fazlul Haque "A Comparative Analysis of Short Range Wireless Protocols For Wireless Sensor Network "International Journal of Scientific & Engineering Research, Volume 8, Issue 4, April-2017

[7] Mirko Franceschinis, Claudio Pastrone, Maurizio A. Spirito "On the Performance of ZigBee Pro and ZigBee IP in IEEE 802.15.4 Networks" 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)

[8] G. Mulligan. The 6LoWPAN architecture. In Proceedings of the 4th workshop on embedded networked sensors, pages 78–82. ACM New York, NY, USA, 2007

[9] Helen Fornazier et al., "Wireless Communication: Wi-Fi, Bluetooth, IEEE 802.15.4, DASH7", ROSE 2012 ELECINF344 / ELECINF381, Télécom ParisTech, web site: http://rose.eu.org/2012/category/admin

[10] Jin-Shyan Lee et al., "A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee", and Wi-Fi, The 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON), Taipei, Taiwan, November 5-8, 2007

[11] Klaus Gravogl et al., "Choosing the best wireless protocol for typical applications", 2nd Workshop on Ultra-low Power Wireless Sensor Networks (WUPS 2011) February 2011, Como, Italy, http://geodes.ict.tuwien.ac.at/PowerSavingHandbook/

[12] Z. Mammeri, " Réseaux sans fils Caractéristiques et principaux standards ", M1 Info Cours de Réseaux, IRIT, Université Paul Sabatier, Toulouse http://www.irit.fr/ Zoubir.Mammeri/Chap WLAN.pdf

[13] Ghobad Heidari, "WiMedia UWB: Technology of Choice for Wireless USB and Bluetooth", edition John Wiley & Sons Ltd 2008, ISBN 978- 11-470-51814-2 (HB)

[14] Ms. Dharmistha, D. Vishwakarma, "IEEE 802.15.4 and ZigBee: A Conceptual Study", International Journal of Advanced Research in Computer and Communication Engineering, ISSN: 2278 – 1021, Vol. 1, Issue 7, September 2012

[15] Vaddina Prakash Rao, "The simulative Investigation of Zigbee/IEEE 802.15.4", Master Thesis of Science, DRESDEN UNIVERSITY OF TECHNOLOGY, FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY, Department of Electrical Engineering and Information Technology, Chair of Telecommunications, September, 2005

[16] http://www.zigbee.org/Specifications/ZigBeeIP/Overview.aspx

[17] Reen-Cheng Wang et al., "Internetworking Between ZigBee/802.15.4 and IPv /802.3 Network", ACM 978-1-59593-790-2/07/0008, IPv '07, August 31, 2007, Kyoto, Japan.

[18] Rohan Tabish, Adel Ben Mnaouer, Farid Touati and Abdulaziz M. Ghaleb "A Comparative Analysis of BLE and IEEE802.15.4 (6LoWPAN) For U-HealthCare Applications" Conference: Conference: IEEE GCC, At Doha, Qatar

[19] Farid Touati & Rohan Tabish "U-Healthcare System: State-of-the-Art Review and Challenges" JMed Syst (2013) 37:9949 DOI 10.1007/s10916-013-9949-0

[20] d. Taslim Arefin, Mohammad Hanif Ali, A.K.M. Fazlul Haque "A Comparative Analysis of Short Range Wireless Protocols For Wireless Sensor Network" International Journal of Scientific & Engineering Research, Volume 8, Issue 4, April-2017 840 ISSN 2229-5518

[21] I. Berdai « Étude Comparative Des Protocoles Zigbee Pro ET Zigbee Ip » University Of Quebec in Montreal, Memory, October 2016

[22] H. A. A.Al-Kashoasha and Andrew H.Kemp "Comparison of 6lowpan and Lpwan for the Internet of Things" Australian Journal of Electrical and Electronics Engineering, 2017https://Doi.Org/10.1080/1448837x.2017.1409920

Artificial Intelligence based Composition for E-Government Services

Amina ADADI[1]. Mohammed BERRADA[1]. Nabil EL AKKAD[1]

{ amina.adadi@gmail.com }

[1]Computer and Interdisciplinary Physics Laboratory (LIPI), ENS Fez
Sidi Mohammed Ben Abdellah University
Fez, Morocco

Abstract. Given the complex nature of the public sector with several distributed governmental institutions and manifold semantic differences of interpretation, the achievement of interoperability and integration is a key challenge for a comprehensive and effective Electronic Government (e-Government). Promising technologies that could be used to tackle this issue are: the powerful concept of ontology and the advanced Artificial Intelligence (AI) systems. Ontologies contribute to a common understanding of heterogeneous resources, while AI techniques make process integration dynamic and automated. However, up until now the use of AI along with ontologies has been fairly limited in e-Government. There is still, then, untapped potential in this field which worth to be exploited. In this paper, we present a dynamic approach for semantically integrating e-Government Web services based on AI techniques. The overall objective of our approach is to improve the citizen centric e-Government vision by providing a conceptual framework for automatically discovering, composing and optimizing e-Government services. Within the proposed approach, special emphasis is put on personalization aspects and evaluation criteria for e-Government platform.

Keywords: Semantic Web Service Composition; e-Government; Ontology; Multi-agent systems; AI Planning; Machine Learning; Reinforcement Learning.

1 Introduction

Current trend in e-Government applications calls for joined-up services that shaped around and responding to the needs of the citizen, and not merely arranged for the public institutions' convenience. Setting up such a seamless e-Government services requires making the shift from isolated silos in public administration to one integrated government that encapsulates the size and the complexity of the administrative body and which provides interoperable and transparent services to citizen through a single point of access in the so-called one-stop government [1][2].

Technology is a key enabler of the envisioned one-stop government as it can bring real solutions to integration, interoperability, coordination and security issues. Indeed, several studies have been undertaken all over the world by different research communities in the field of e-Government services composition and integration [3][4][5]. Most of the recent e-Government projects support the one stop shop e-services vision and use semantic technologies to deal with interoperability issues. But, to the best of our knowledge, no one of such projects has explored the potential of using distributed planning algorithms along with ontologies based on a multi agent architecture, to address the issues of semantic interoperability and service integration of e-Government.

Toward dynamic semantic Web service composition is the subject of our researches, as it has been argued in [6] our goal is to automate the whole composition process including discovery, composition, execution and interoperation of Web services. Our proposal stems from the following basic underlying idea: The synergy between four basic technologies

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284071

namely; (i) Semantic Web [7], (ii) Multi-Agent System [8], (iii) Autonomic Computing paradigm [9] and (iv) Artificial Intelligent Planning (AIP) techniques [10], can lead to the development of a comprehensive platform that supports personalized, integrated and interoperable Web services. E-Government has some specific features as opposed to traditional e-business scenarios, the e-Government domain is a large, heterogeneous, dynamic and shared information space with various semantic differences of interpretation. Thus, e-Government seems to be an obvious and promising application field for our semantic Web service composition approach. And since e-Government services require a high degree of interoperability and distribution due to the numerous entities they imply, this is where our semantic Web service composition approach can produce the best results to solve a significant number of e-Government issues. Therefore, the combination of these two areas is very much natural and beneficial.

The present paper is an extension of our previous work [6][11][12] that propose a redesigned version of our semantic Web service composition approach for e-government use with a focus on the personalization aspects.

The remainder of the paper is organized as follows: in Section 2 presens the layered architecture of the proposed platform. Section 3 details our proposed machine-learning model for enabling personalized services. Section 4 proposes a qualitative evaluation of the proposed approach. Section 5 gives our concluding remarks.

2 The proposed architecture

Fig. 1 shows the high-level overview of the proposed approach architecture. It is organized in a layered architecture that traces the service lifecycle from its request until it delivery. As depicted in Fig. 1 the platform consists of four layers, namely i) Legacy system layer, ii) Ontology layer, iii) Composition layer and iv) Communication layer.

- **Legacy system layer:** contains the legacy applications available in each of the public administrations (PA) involved in the system.
- **Ontology layer:** Ontologies [13], as powerful knowledge representation formalism for semantic modelling and annotation, were chosen as a basic mechanism to capture semantic content of e-Government Web services in a manner that can be automatically processed by a machine. Our system integrates: (a) e-Government domain Ontology that contains the domain's types, concepts and relations among them (a full implementation of our e-Government domain Ontology can be found in [11]) . (b) e-Government Service Ontology plays the role of a UDDI registry with extended functionalities allowing the storing of a rich semantic description of Web services involved in PA services. (c) QoS Ontology specifies shared knowledge and vocabularies about the quality of service proprieties (QoS) and their relationship with respect to semantic services. (d) User profile Ontology incorporates concepts and properties used to model the user preferences and extra information, the latter plays a major role in the personalized process of the delivered services.
- **Composition layer:** Composition layer is the core component of the proposed architecture that provides the required functionalities such as discovery, composition and execution of services. It is composed of a set of intelligent agents capable of interacting and cooperating in order to co-build automatically a composition plan. The composition process is depicted under four phases namely: a) analysis, b) planning, c) evaluation and d) execution. For the sake of completeness the phases are just outlined here an extensive technical description of each phase can be found in [6]:
  - *Analysis*: starting from an abstract specification of the user request, the aim of the analysis phase is to define semantically the requested service.
  - *Planning*: during this phase a composition plan is generated in two steps namely, discovery and plan construction, the first step aims to find all candidate services that

can potentially lead to satisfying the desired functionalities. The search is based on a matchmaking algorithm [6] using the service ontology repository and the domain ontology repository. In the second step, the Planner agent coordinates a set of Manager Agents that act on behalf of discovered semantic Web services. They work cooperatively using AIP techniques [6] to co-build an abstract composition plan.

- *Evaluation*: given a composition plan, an important challenge for the system is to find the "best" composite service execution plan with the respect of user specifications. The evaluation phase deals with this issue. It uses QoS ontology to rank the alternative concrete composite services through the use of quality of service techniques.
- *Execution*: finally, in the execution phase, the system invokes the selected service instances according to the execution plan and delivers the result back to the user.

- **Communication layer:** is the user interface that manages the interaction between the user and the system. This layer ensures a front office integration by proposing a single and unique access point to the e-Government services that integrates and merges any other access channel to these services.

The role of the proposed system is to lead the citizen from his/her problem to the appropriate service(s) by encapsulating the complexity of public institutions. Indeed, with an architecture based essentially on intelligent agents and ontologies, the middleware element formed by ontology and composition layers provides mechanisms to coordinate existent services in order to provide a new personalized service. Starting from an informal, incomplete and ambiguous request, intelligent agents explore e-Government services repository, where PA services are published in a semantic manner, to find the right combination of services that satisfies the user requirements, the use of AIP techniques makes this process dynamic and automatic. Then agents take in charge the execution of the composition process in a changing heterogonous and distributed environment based on an autonomic computing architecture and QoS techniques. As a result, the system provides a value added services that are more convenient for users' needs and perceptions and increase efficiency in terms of costs and productivity by reusing autonomous PA services while keeping their internal processes and legacy systems intact.

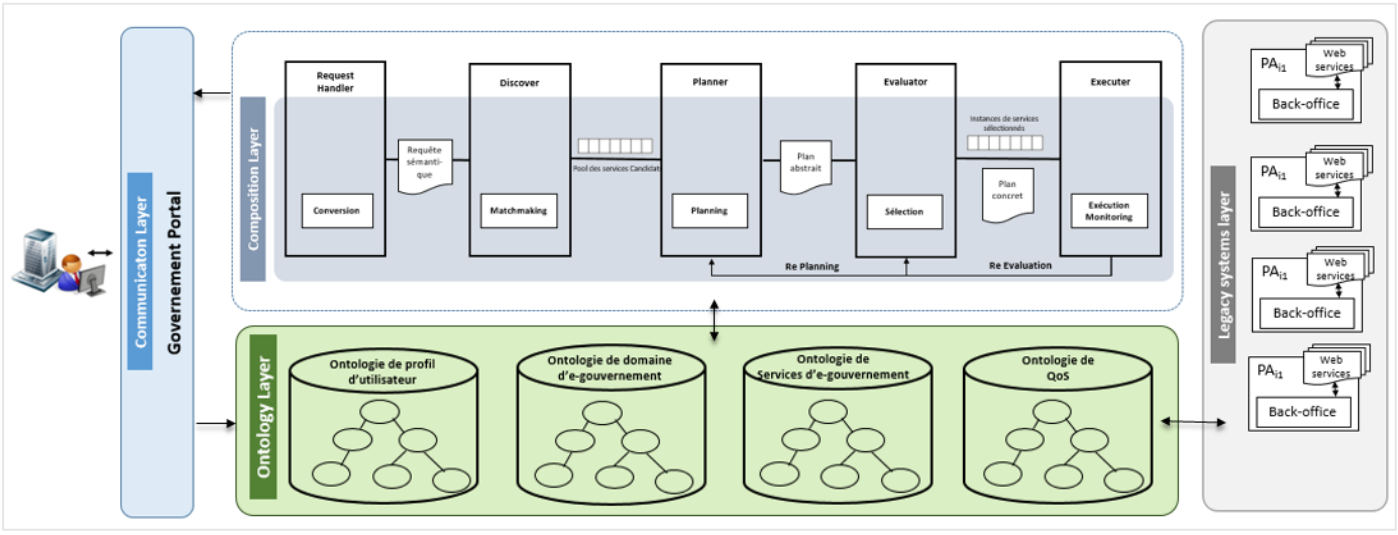


Fig. 1. The high level architecture of the proposed approach.

3 Recommendation layer: machine learning based personalization

The proposed approach, as presented, focuses on interoperability and integration that represented long-standing challenges of e-government. In the era of the fourth industrial revolution, citizen collaboration is also seen as a key enabler for an efficient electronic government, thus more attention should be given to the personalized aspects of government services delivery.

This section aims at studying the use of machine learning algorithms for supporting personalized e-government services. To this purpose, we propose to combine our composition approach with a machine learning technique called: Reinforcement Learning (RL) [14]. RL is a process where an agent learns through experience by exploring the environment based on reward or the feedback. Typically, agents are not told which action to take, as in most forms of machine learning, but instead they must discover themselves which actions yield the most reward by trial and error. Thus, in this type of machine learning, the focus is shifted from pattern recognition to experience-driven sequential decision-making, which is suited for our problem that do not define a particular goal to be reached but a reward (defined by citizens) that has to be maximized.

Therefore, inspired by the successful use of RL in a range of fields including education, finance, manufacturing, and e-commerce. We advocate herein the promising potential application of RL in e-government to solve personalization related issues. Conceptually, a recommender system in form of a layer called "recommendation layer" will expand the capabilities of the actual composition approach to facilitate adaptivity by making "recommendations" to the system based on citizen feedback. As result, similarly to e-business where recommendation systems are used to help client locate information or products that they would like to make offers. Recommendation layer will allow citizens to benefit from the same experience in the e-government settings.

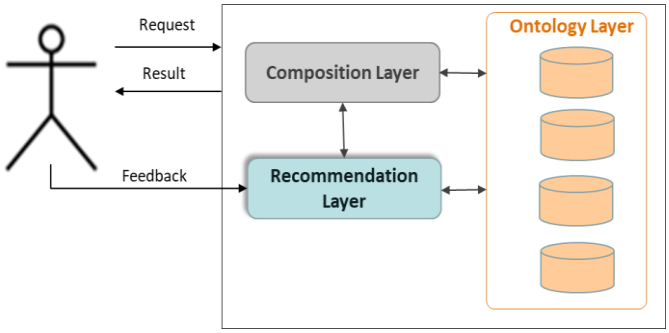


Fig. 2. The target combined approach with "Recommandation layer".

Fig.2 shows the main components and interaction processes of the target combined approach. The composition layer (that encapsulates the actual composition approach) and the recommendation layer are two distinct modules that interact with each other in order to generate personalized service that i) are formally correct with respect to user requests and ii) approximate implicit user requirements over time based on user ratings. This proposal is a starting point of a potential future work that needs to be deeply explored and developed in order to prove its relevance.

4 Evaluation and discussion

Recognizing that e-Government should be transformative and become more citizen – rather than government – focused in delivering public services, investing in the development of an e-government interoperability and integration framework is fundamental and promising, especially for developing countries. Indeed, in the Moroccan context, interoperability concepts and strategies are crucial agenda of the national e-Government development to achieve integrated, safe and efficient utilization of public services [15]. However, at a technical level, very few academic research works and studies have been focused on these issues. Thus, our work can be seen as an initiative that promotes an Artificial Intelligence based conceptual framework in response to interoperability and integration challenges of e-Government.

From our point of view, the seamless flow of data from one public administration to another can be achieved through ontologies. To support this idea, we developed, WebGov [11], an ontology for e-Government public services. WebGov fulfils the key requirements for a public services ontology to be useful [16]. Firstly, it covers multiple aspects of services, including administrative responsibility, involved documents, legislation and meta-data etc... Secondly, since it is designed and implemented based on METHONLOGY [17] method, the proposed ontology is fully extensible in the sense that it facilitates incremental ontology development, by incorporating basic concepts and relationships at an initial phase and then defining new ones, or creating instances as needed. The semantically rich network of interrelated concepts formulating by WebGovis mainly used in supporting the e-Government service composition tasks.

Furthermore, we proposed a conceptual framework populated by a toolkit of composition techniques that organize the e-Government integration process in a lifecycle. The use of Artificial Intelligence Planning algorithms and Multi-agent systems makes the proposed planning model dynamic and automated. While machine learning in form of RL has been identified as a key technology to meet the challenge of personalization aspects.

As a result, the overall approach represents an end-to-end conceptual framework that implements the following citizen/government relationship schema (Fig.3):

(i) C2G and G2C interactions that mark respectively the beginning and the end of the administrative process, they illustrate a front-office integration.

(ii) G2G interaction that encapsulates the mechanism of integrating (composing) the different public administration services in order to accomplish the administrative procedure requested by the citizen. This interaction illustrates a back-office integration.

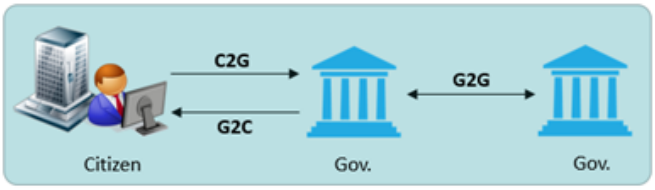


Fig. 3. The implemented Citizen/government relationship schema.

Evaluation plays a vital role in an approach validation process. At this stage of our work, we believe that a qualitative evaluation (QE) is the best tool that can lead us to reliable conclusions and can help us to identify the potential areas of improvement. Thus, in order to validate our approach, we conduct qualitative criteria based evaluation.

After having theorized extensively about evaluating Web service composition approaches [18][19][20][21] [21][23][24][25][26][27][28], We have noted that several evaluation criteria exist in the literature, however none of these evaluations take into account the e-Government specific requirements. The e-Government environment is highly heterogeneous in nature and manages a sensitive relationship (between government end citizens). Therefore, in order to perform a services composition in this domain, a composition approach must fulfils some

particular criteria such as service relevance and security, citizens and business satisfaction and preservation of public trust, in addition of other standard requirements such as flexibility, reliability, accessibility, availability, processes automation, lower cost, short delivery time etc…Despite their diversity, by collecting Web service composition technical requirements spread across the literature and by analyzing their relationship, we have identified five core aspects of requirements: i) Expressiveness ii) Interoperability, iii) Transparency and iv) Adaptivity. These criteria are used to check how far the proposed approach supports a successful e-Government service composition process.

- **Expressiveness**

A granular expressiveness in service request and management is one of the important aspects that has to be taking in account.
The presented approach relies on a rich expressiveness of semantic descriptions of Web services (OWL-S), hence it supports functional and non-functional parameters such as Quality of Service (QoS). The approach also considers preference parameters to help the user in describing their requests at a granular way.

- **Interoperability**

As interoperability is the primary goal of most e-Government frameworks, ensuring interoperability needs to be a key focus item.
As explained before, our approach is driven by ontologies as a key instrument to ensure interoperability. In the same vein, dealing with service composition in heterogeneous and distributed systems like public administrations, is also an important and challenging issue. To face this challenge our approach is based on a multi-agent system architecture that allows a decentralized e-Government services integration.

- **Transparency**

A composition approach for e-Government services must be user (citizen and business) centric and hide the complexity of the service production processes, hence it has to optimize the human intervention (automation).
In the proposed approach, the citizen starts with a simple user interface where information from domain ontology and user profile ontology can be automatically loaded. Based on the planning model technique, the composition process is performed in an autonomic way and return back the result to the user who is shielded from the complexities of the background composition process.The composition process is also performed transparently from the service provider (public administration) without requiring intervention from their part, hence their internal systems remain intact during the composition cycles.
The utilization of proposed system is then completely transparent from the citizen and the public administration.

- **Adaptivity**

Adaptivity aspect has recently been attracting interest in research communities and involves the ability of a service (atomic or composite) to adapt itself. Adaptation is the process of modifying service-based applications in order to satisfy new requirements and to fit new situations dictated by the environment on the basis of predefined adaptation strategies. Adaptation goes one step further from dynamic criteria, in the sense that the former also deals with changes in the requirements set by the requester (citizen in our case) which the latter cannot handle. Moreover, adaptive approaches deliver personalized services which increases user satisfaction.
In our approach self-healing is a key aspect that takes in charge handling faults and recovering from problems reported during execution which increases the adaptivity of the system.
Moreover, by adding the recommendation layer, the approach will support user's feedback which can potentially help in delivering more personalized services. It will also involve learning from previous composition experiences, which implies an optimization of the composition process.

As summarized in the table below, our approach is characterized by supporting QoS parameters, ontologies, self-healing and automated dynamic generation of composition schema based on the AI planning. However, the learning capacity is yet to be improved in order to improve the personalization aspect. Which reinforces, confirms, and motivates our proposal regarding the integration of the "Recommendation layer".

Table 1. Qualitative evaluation of the proposed approach.

Aspects	Level of fulfilment
Expressiveness	+
Interoperability	+
transparency	+
Adaptivity	+/-

+ high +/- medium - down

5 Conclusion

In this paper, we studied how our approach for semantic Web services composition based on Artificial Intelligence techniques can be effectively applied to deliver citizen centric e-Government services. Then we focused how to enable personalization based on machine learning.Finally, we evaluated qualitatively our proposal.

Our future work includes the completion of the proof-of-concept prototype but before, we plan to pursue this work by implementing applicative e-Government scenarios to extensively evaluate our approach's effectiveness. The use of machine learning is another extension way to be explored towards a more intelligent e-Government.

References

[1] Chatzidimitriou, M. Koumpis, A.: Marketing One-stop e-Government Solutions: the European OneStopGov Project, IAENG International Journal of Computer Science, Vol. 35, No. 1, (2008).

[2] United Nations.: Taking a whole of-government approach, E-Government Survey 2012, Chapter 3, (2012)

[3] Stojanovic, L. G. Kavadias, G.Apostolou, D. Probst, F. Hinkelmann, .: Ontology-enabled e-GovServiceConfiguration, Delivrable Ontology Project – IST Project 507237,http://www.ontogov.com

[4] eGOIA - Electronic Government Innovation and Access, http://www.egoia.sp.gov.br/

[5] E.C 6th Framework Programme.: Electronic and Secure Municipal Administration for EuropeanCitizens – eMayor, IST-2004-507217, (2004), www.emayor.org

[6] Adadi, A.Berrada, M. Chenouni, D.: A phased two stages conceptual framework for Web service composition, Journal of Engineering and Applied Sciences (JEAS), 13 (7), (2018).

[7] Yu, W.Wen,G. Chen,G. Cao , J.: Distributed Cooperative Control of Multi-agent Systems, John Wiley & Sons, (2016), 350 pp, DOI: 10.1002/9781119246213

[8] Laufer, C.: Semantic web guideline, government of the state of sãopaulo , Web Technologies Study Center, (2015)

[9] Abeywickrama, D.B. Ovaska, E.: A survey of autonomic computing methods in digital service ecosystems , Service Oriented Computing and Applications, Vol.11, Issue 1, pp: 1–31, (2017)

[10] M.Ghallab, M..Nau, D. Traverso, P. : Automated Planning and Acting, Cambridge University Press, (2016)

[11] Adadi, A.Berrada, M. Chenouni, D. Bounabat, B: Ontology based composition of e-Government services using AI Planning, the 10th International Conference on InelligentSystems : Theories and Applications (SITA'2015), (2015).

[12] Adadi, A.Berrada, M. Chenouni, D.:A Multi-Agent Planning Architecture for Semantic Web Service Composition, International Review on Computers and Software (IRECOS) , (2014)

[13] L. Zhu,L. Tao,W. Min, X. Gao ,W. : Theoretical Characteristics of Ontology Learning Algorithm in Multi-dividing Setting , IAENG International Journal of Computer Science , Vol. 43 Issue 2, p184-191, (2016).

[14] Arulkumaran, K. Deisenroth, M.Brundage, M.Bharath, A,:Deep Reinforcement Learning: A Brief Survey,IEEE Signal Processing Magazine , 34 (6), (2017)
[15] Vrandecic, D.: Ontology Evaluation: thesis, Karlsruhe Institute of Technology (KIT),Germany, 2010.
[16] CHAKRI, S.:Maroc digital 2020: Stratégiesnationales pour le développement de l'économienumérique, 2016, https://www.itu.int/en/ITU-D/Regional-Presence/ArabStates/Documents/events/2016/DigitalTransformation/Pres/Strat%C3%A9gies%20nationales%20pour%20le%20d%C3%A9veloppement%20de%20l%E2%80%99%C3%A9conomie%20num%C3%A9rique%2C%20samia%20chakri.pdf
[17] Sawsaa,A. F. Lu,J. : Building Information Science Ontology (OIS) with Methontology and Protégé, Journal of Internet Technology and Secured Transactions (JITST), Vol.1, Issue 4, (2012)
[18] Dustdar, S. Schreiner,W. .: "A survey on web services composition", Int. Journal of Web and Grid Services, vol. 1, no.1, pp. 1–30, (2005).
[19] Mantovaneli Pessoa, R. .Silva, E. Sinderen, M.V. Quartel,D. Pires, L.F. :"Enterprise Interoperability with SOA: a Survey of Service Composition Approaches", in Workshop on Enterprise Interoperability, pp. 32-45, (2008).
[20] Rao. J.. Su, X.: "A Survey of Automated Web Service Composition Methods", in Proceedings of the First International Workshop on Semantic Web Services Web Process Composition, (2004).
[21] Khadka, R. Sapkota,B. .: "An Evaluation of Dynamic Web Service Composition Approaches ", In: 4th International Workshop on Architectures, Concepts and Technologies for Service Oriented Computing, (2010).
[22] Tabatabaei,S.G.H. Wan Mohd,K. Suhaimi,I. : "A Comparative Evaluation of State-of-the-Art Approaches for Web Service Composition", in Proceedings of Third International Conference on Software Engineering Advances, pp. 488-493, (2008).
[23] Beek, M. Bucchiarone,A. Gnesi,S. : "Formal Methods for Service Composition", Annals of Mathematics, Computing &Teleinformatics, vol. 1, no. 5, pp. 1-10, (2007).
[24] Benatallah., B. :Service Composition: Concepts, Techniques, Tools and Trends in Service-Oriented Software System Engineering-Challenges and Practices, Idea Group Publishing, USA , pp. 48–66, (2005).
[25] Zuniga,J.C. Pérez-Alcazar, J. Digiampietri,L. : "Implementation Issues for Automatic Composition of Web Services", Workshops on Database and Expert Systems Applications, pp. 201-205, (2010).
[26] Sheng,Q.Z. Qiao, . X. Vasilakos, A.V. Szabo, C. Bourne,S. Xu,X.: Web services composition: A decade's overview, Information Sciences , Vol. 280, pp: 218–238, (2014)
[27] lagares, A. Florian, D. Boualem,B. : Web Service Composition: A Survey of Techniques and Tools, ACM Computing Surveys (CSUR), Vol. 48, Issue , (2016)
[28] AlSedrani,A. Touir,A. Web service composition in dynamic environment: a comparative study, The Fourth International Conference on Database and Data Mining, pp. 75–84, (2016), DOI: 10.5121/csit.2016.60508

Innovative Models Proposals for the Implementation of Adaptive Hypermedia Systems

Mehdi TMIMI[1], Mohamed BENSLIMANE[2],Mohammed BERRADA[3] and Kamar OUAZZANI[4]

{mehdi.tmimi@usmba.ac.ma[1], benslimane_mohamed@live.fr[2], mohammed.berrada@gmail.com[3] kamar_ouazzani@yahoo.fr[4] }

Laboratory of computing and interdisciplinary physics, ENS, USMBA, Fez, Morocco[1 3 4]
Transmission and Treatment of Information Laboratory, EST, USMBA, Fez, Morocco[2]

Abstract. This article is the continuation of our work concerning the development of an adaptive hypermedia system. In previous work, we proposed three new fundamental models for the implementation of an adaptive hypermedia system. such as: learner model, domain model and adaptation model. The purpose of this article can be summarized first by presenting our three proposed models while discussing the related work that we have studied and the new ideas that we have invented and implemented, and then the development of the global architecture of our future adaptive hypermedia that we intend to present it in details in future work.

Keywords: Adaptive Hypermedia Systems, Learner Model, Adaptation Model, Domain Model, E-learning, SCORM, CMI5, MUNICH, AHAM, ALEM.

1 Introduction

The brain is immensely complex to the point that until now it represents a mystery for researchers and scientists. This complexity involves also the learning process which is one of the main functions of the brain. Learning is defined as a relatively permanent change in knowledge and skills resulting from experiences. Indeed, we have found in literature several learning methods, namely: Learning by imitation [1], associative learning [2], learning by trial and error [3], observational learning, learning by repetition, etc. All of these methods are present in both traditional (classic) and distant learning (E-learning).

The first type of learning refers to the classical method where there is a teacher who brings together several learners in a class, or amphitheater. While the second type describes the new method of learning that unlike the classical one does not have constraints of time and place. Nowadays this new type of learning is known as E-learning which in collaboration with technological progress and internet has successfully challenged and remedied the major problems of classical learning.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284212

In literature, the distant learning has an old history of almost three centuries that can be summed up in three eras [4]:

1. First era: learning by correspondence.
2. Second era: learning using radio and television
3. Third era (E-learning): learning using the Internet and new multimedia technologies.

The European Commission has defined the E-learning as "the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services as well as and remote collaboration ". [5]

Given the success of this new type of learning, many systems and approaches have taken the step towards online learning and have dominated distance learning in recent years. However, these systems and approaches are today considered ‘traditional’ because on the one hand, they do not respect the technological progress of Internet, and on the other hand they are based on the tradition of "one design for all" [6] where the same course with the same structure and presentation is delivered to all types of learners.

As a result, a new generation of e-learning systems has emerged, known as adaptive learning systems, which gives learners an active role in acquiring knowledge and skills while also adapting the content of learning according to theirs needs and preferences.

These adaptive learning systems have a minimal architecture generally composed of three interdependent models. Which are:

1. Learner model: describes learner characteristics such as: skills, knowledge and learning styles, etc. [7]

2. Adaptation model: describes the set of construction and presentation rules that are responsible for constructing the content to be provided to the learner.

3. Domain model: describes the structure of the information content of the application [8].

Since the main objective of our research is the design and development of an adaptive hypermedia system, we will present in this article all the three fundamental models that we have developed while discussing the related work we have studied and the new ideas that we invented and implemented.

2 Learner model

The learner model is a very promising solution for representing and describing information about a learner, in order to provide a complete and faithful description of all aspects related to the learner's behavior in the learning phase.

In our previous work [9], we proposed the basic elements of our new learner model that we deduced from a set of existing approaches and ontologies. Our proposal for the learning model is based on six facets, which are:

- Personal information: represents the personal data of the learner, namely his identity, sex, age, etc.
- Knowledge and skills: contains all the knowledge, skills, and competencies of the learner.
- History: This facet is responsible for saving all the actions and interactions of the learner within the adaptive hypermedia.
- Psychological profile: Represents the learner’s preferences and psychological profile. it is so often based on one of the psychological types theories.

- Cognitive ability: refers to all specialized treatments that require the collaboration of brain spaces. Such as: attention, memory and strategies, vision, space, language, gestures and coordination.
- Emotional state: represents the emotions experienced by the learner during the learning phase. Such as: frustration, satisfaction, etc.

These facets were the subject of a detailed study published in a previous work [10] where we first presented our research made on each one of the six facets and then proposed our own learner model shown in Figure 1.

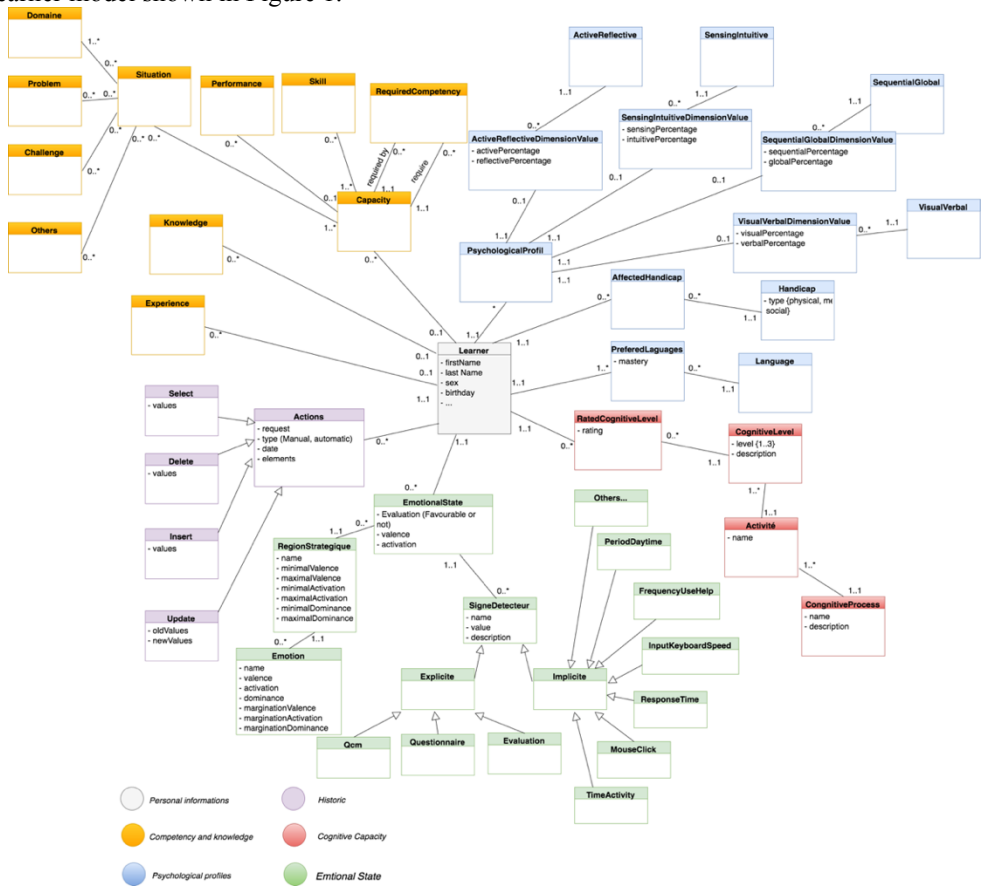


Fig. 1. Proposed learner model [10]

3 Domain model

The domain model describes the structure of the course elements such as: knowledge, concepts, etc. And also describes the resources needed to present its elements. Namely: videos, articles, homework, questionnaires [11], etc.

283

To develop this model, we first studied the reference models for adaptive hypermedia: AHAM [12] and MUNICH [13] and then the online learning technical standards such as SCORM [14] and CMI5 [15].

Regarding The domain model proposed by the two reference models AHAM and MUNICH. We've found that their domain model describes the structure of hypermedia as a collection of connected components, which form a network expressed by the composite design pattern [16]. Typically, a component is an abstract representation of concepts (knowledge, information, exercise, etc.), concepts relationships, pages and fragments.

However, we have deduced several remarks that can be resumed mainly first in the lack of semantics which refer to the objective or goal behind learning a concept and secondly in the type of architecture that they used, which we described it as a content-oriented architecture. This architecture is mainly built around the content (component).

Before discussing the online learning technical standards such as SCORM and CMI5. We need to shortly introduce them.

So, The Shareable Content Object Reference Model (SCORM) is a set of technical standards for e-learning software products developed by ADL (Advanced Distributed Learning [17]). It was mainly designed to present a reference model for defining and constructing learning course content that have the capacity to be shared with other SCORM compliant systems.

Unlike SCORM, CMI5 was not exclusively developed by ADL. Indeed, CMI5 was co-developed by both ADL and the Aviation Industry Computer-Based Training Committee (AICC). CMI5 intended to take advantage of both SCORM and Experience API (xApi) while responding for their limitations and adding some new capabilities.

These two online learning technical standards was the focus of our research where we first studied their specifications and then designed their UML class diagrams. these conceptual representations allowed us to notice that SCORM and CMI5 are based too on a content-oriented structure and were intentionally only designed to deliver content to a learner (traditional e-learning system).

Finally, after this analytical and critical study, we designed a new domain model - shown in Figure 2 - that responds to all the criticisms cited above.

284

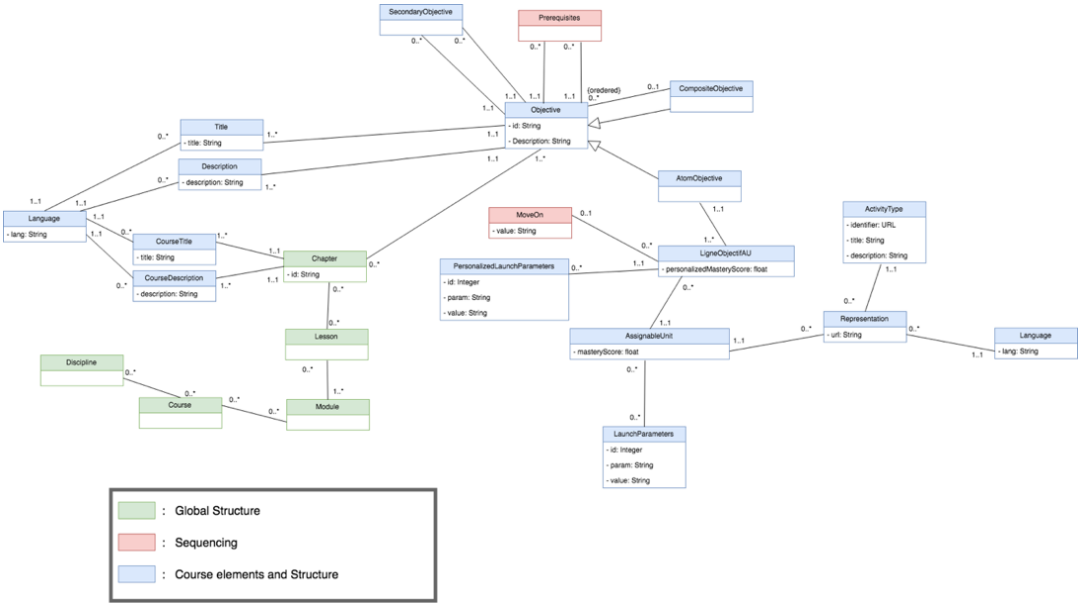


Fig. 2. Proposed domain model

4 Adaptation model

The adaptation model represents the mechanisms responsible for customizing and adapting learning content, links, representation and the structure of concepts and knowledge to be delivered to learners.

Indeed, there are several aspects that the adaptation and the personalization can be carried out. Which are:

Link adaptation: it concerns the management of the links by allowing either to limit the possibilities of navigation in the hypermedia, or to offer to the learner new links leading to other concepts.

Content adaptation: this type of adaptation concerns the selection of the most suitable and appropriate content to be delivered to the learner.

Structure adaptation: allows to display the same content under several structures and in different forms.

Adaptation of the presentation: this type of adaptation concerns the display of the same content of the page in different custom styles according to the learner's preferences. Such as: font weight, color, etc.

So, in a similar way to the learner and domain model, we studied the existing adaptation models proposed by MUNICH and ALEM [18].

4.1 MUNICH

285

In Figure 3, we present the UML class diagram of the adaptation meta-model proposed by Munich that mainly constructed around the 'Adaptation' and 'Rule' classes.

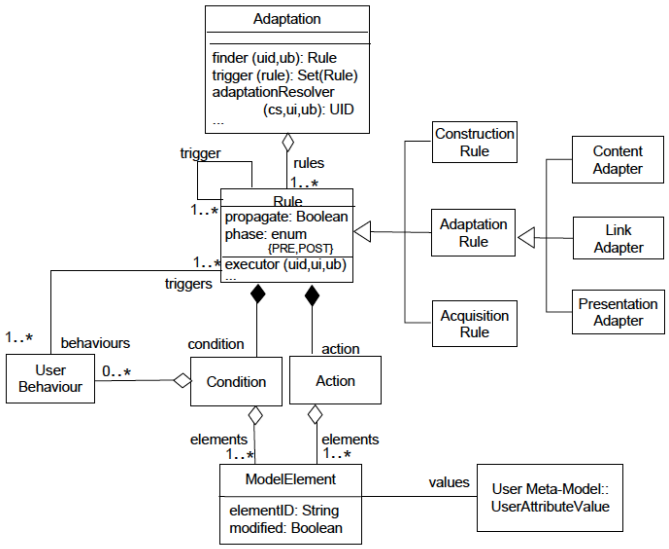


Fig. 3. Adaptation meta-model proposed by MUNICH [13]

The Adaptation class consists of three methods: an adaptation resolver, a finder, and a trigger. The first method resolves the specification of a component to a suitable component. The second and third methods implement a mechanism that finds all the rules that must be triggered by a given rule. The first rule to use is triggered by the user's behavior. The executor method of the 'Rule' class allows the system to select the appropriate components and to adapt the content, layout and link, as well as update the learning model.

Each object of class 'Rule' contains attributes and many conditions and actions related to it by the mean of a composition relation. The attributes refer to several pieces of information used for determining whether the rules are applied before or after updating the learning model.

Finally, the rules ‘class Rule’ are classified according to their mode of intervention in four types: rules of construction, rules of acquisition and rules of adaptation. This distribution is expressed by means of the inheritance mechanism provided by the UML class diagram.

4.2 ALEM

As shown in Figure 4, the adaptation model proposed ALEM is designed in a similar way to MUNICH. Indeed, ALEM is designed using the UML class diagram and built mainly around the Rule “Règle” class.

286

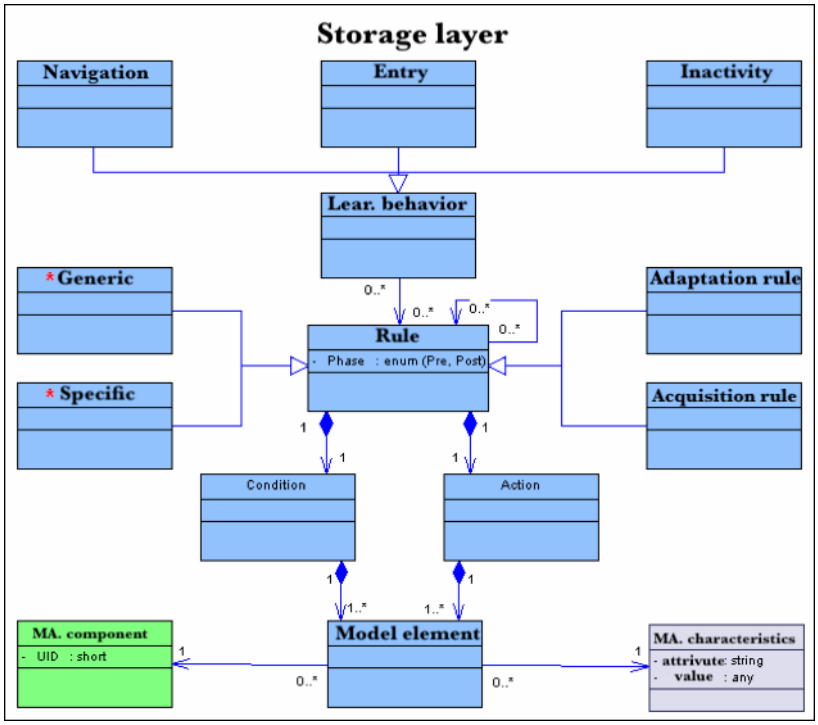


Fig. 4. adaptation meta-model proposed by ALEM [18]

However, in addition to most of the element supported by MUNICH, ALEM implemented a new concept for the adaptation rules expressed by the two classes:

- Generic rule "générique": refers to adaptive rules predefined by the designer of adaptive hypermedia.
- Specific rule "spécifique": this type of rule concerns the new rules created by the tutors and the authors of the learning content.

Finally, we noticed that learner behaviors are also introduced into the adaptation meta model and are classified into three classes: Navigation, Entry and Inactivity.

4.3 Remarks and reproaches

In the two proposed adaptation model by MUNICH and ALEM, we found several remarks and reproaches that make the proposed models incomplete or really difficult to implement. Indeed, although there are 4 types of adaptation that have different characteristics and levels of intervention, the two models (Munich and ALEM) have just included these 4 types of adaptation at the same conceptual level and in the same way without indicating how to implement and how to use them for adaptation. Also, ALEM assumes that each object of the class 'Rule' can refer to several types of rules namely: adaptation rule, acquisition rule, generic rule and specific rule. With the design shown in Figure 4, it’s forbidden to have a rule that combines two types at a time. Example: an adaptation rule can not be also a generic rule.

287

4.4 Proposition

Following the criticisms and reproaches cited above, we first developed a new adaptation model (presented in Figure 5) and then we implemented and tested it in a web platform.

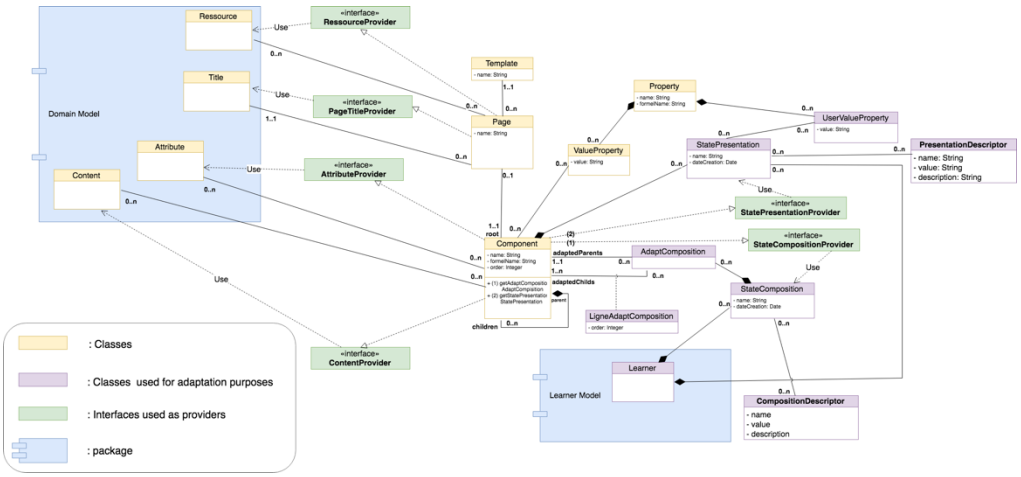


Fig. 5. Proposed adaptation model

4.5 Teste and implementations

In order to use our proposed model, we have developed a set of tools (see Figure 6 and 7) allowing the exploitation of the model and affirming its validity. The detailed description of these tools can be found in our published article [19].

- Tool for creating template pages:

Given the complexity of our proposed model and the need to respect its recommendations by the pages. we have developed a creation tool that takes on all the complexity and allows tutors to create pages that respect the structure proposed by our adaptation model

In figure 6, we show the web tool that we have developed and that allow the creation of pages in real time, the following is a short list of the functionality supported by our tool:

- Component management.
- Component Presentation management
- Preview of the built page

288

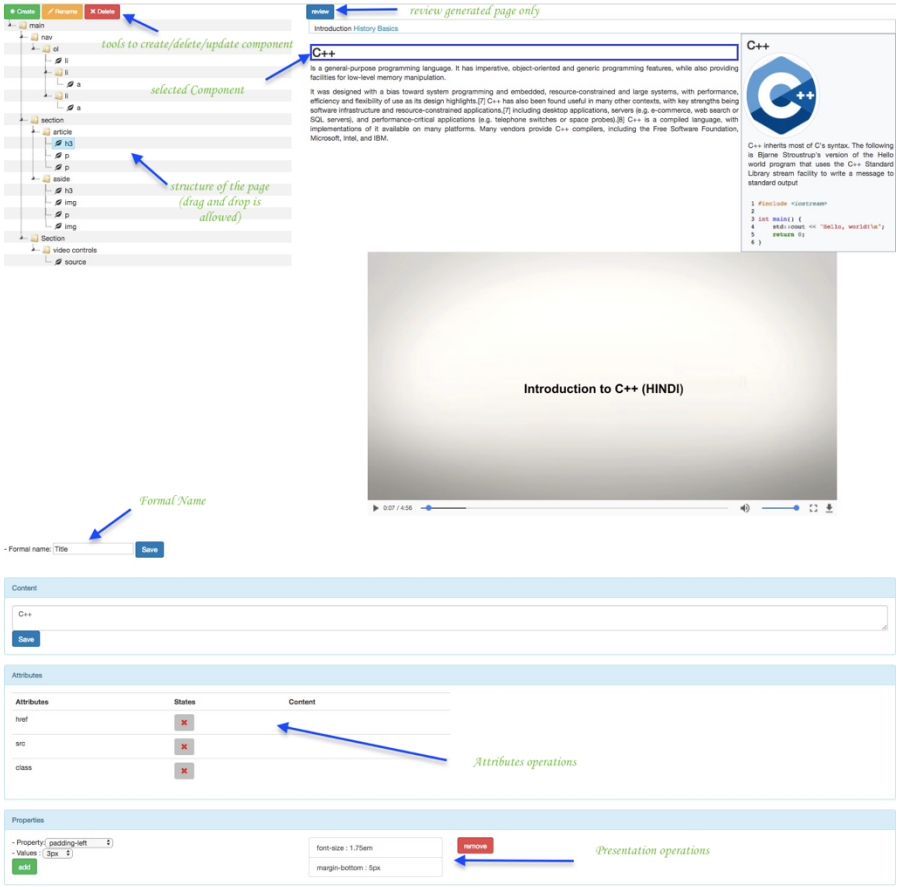


Fig. 6. Screenshot of our web app that represents the tools to create template pages

- Tool for Adaptation and customization:

In Figure 7, the system generates the built page to the learner while offering mechanisms for adaptation of structure and representation. In Figure 7, both the original structure and presentation have changed, this new structure (the video is on the top of the page) is only applicable to this learner.

289

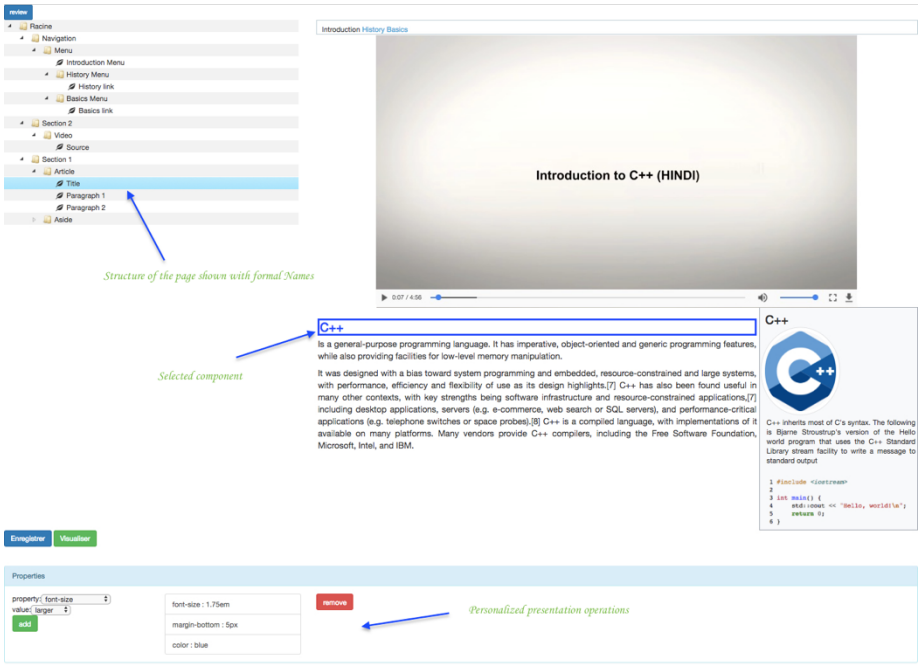


Fig. 7. Screenshot of our web app that represents a custom and adapted page of a learner

5 Global architecture

The three models we have developed form the core of our adaptive hypermedia system that we want to develop in the near future. Indeed, we have developed its global architecture (shown in figure 8) using the JEE design pattern and composed of our three models elaborated and other processes and engines required for its operation.

290

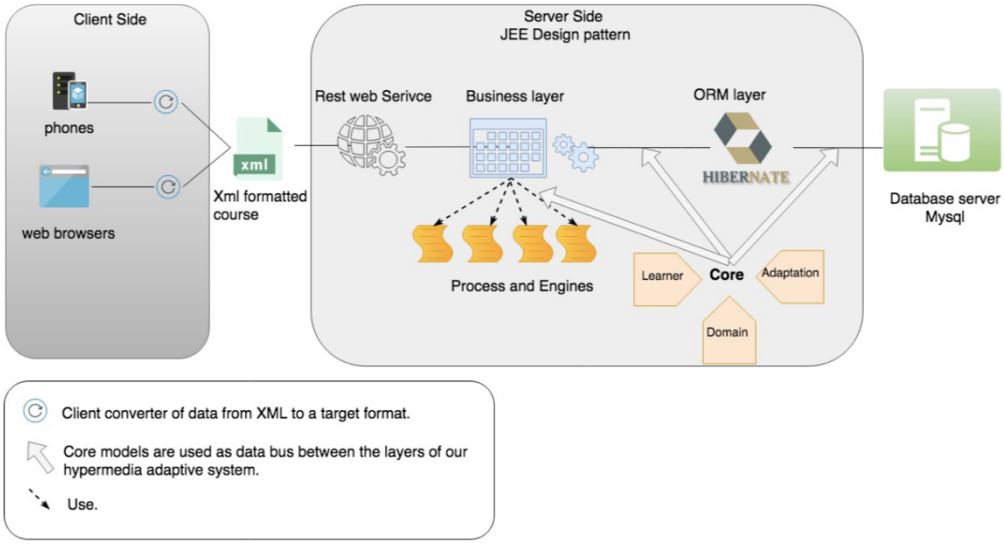


Fig. 8. Global architecture of our adaptive hypermedia system

6 Conclusion and perspectives

In this paper we presented our three designed models that represent the fundamental and indispensable elements for the development of an adaptive hypermedia. Indeed, in a future work we will first work on interconnecting these three models, then explain the conceptual and technological choices that we will use to implement the global architecture of our adaptive hypermedia system.

References

[1] F. Borsali: Apprentissage par observation chez le jeune enfant avec autisme. Psychologie. Université Paul Valéry - Montpellier III. Français. 〈NNT : 2015MON30022〉 . Doctoral thesis (2015).

[2] J. Favrel: Psychologie des apprentissages, approches cognitive et comportementale. Université de Bourgogne, L2 EFEC. (2018)

[3] H. E. Gendolla: Behaviorisme IIa, Le conditionnement opérant I, “Motivation et apprentissage VII”, Université de GENÈVE. [Online] available at: http://tecfaetu.unige.ch/etu-maltt/volt/venni6/bases/2/7_behaviorisme_IIa(1).pdf

[4] W.Wright, C. Jeffs, J. Wood: Distance Education : Must it be ‘ Out-of-Sight , Out-of- Mind ’ Education ?. available at: http://www.ascilite.org/conferences/melbourne95/smtu/papers/wright.pdf

[5] Commission of the European Communities: Communication from the Commission to the Council and the European Parliament - e-Learning Action Plan - Thinking about tomorrow's education. Bruxelles, Brussels,COM(2001)172 final. (2001) available at : https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0172:FIN:EN:PDF

291

[6] G. Kaya and A. Altun: A Learner Model for Learning Object Based Personalized Learning Environments. Commun. Comput. Inf. Sci., vol. 240, pp. 349–355. (2011).

[7] S. Drissi, T. Bensebaa, Y. Lafifi: Une nouvelle approche pour l'adaptation d'un hypermédia pédagogique au profil cognitif de l'apprenant en utilisant XML. in 2nd Conférence Internationale sur l'Informatique et ses Applications (CIIA'09). (2009).

[8] H. Wu, G.J.P.M. Houben, P. De Bra : Supporting user adaptation in adaptive hypermedia applications. Proceedings Conferentie Informatiewetenschap 2000,pp. 88-98.(2000)

[9] M. Tmimi, M. Benslimane, M. Berrada, K. Ouazzani: Elaboration d'une ontologie apprenant pour les hypermédias adaptatifs. Conférence Internationale sur les Approches Pédagogiques & E-Learning, Morocco, Appel'2016. (2016).

[10] M. Tmimi, M. Benslimane, M. Berrada, K. Ouazzani: International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 24, pp. 16008-16016. (2017)

[11] M. Madaio: Adaptive Learning Systems the State of the Field. in C21U Seminar, UnConference Room at C21U. (2015)

[12] P. De Bra, G.-J. Houben, H. Wu: AHAM, A Dexter-based Reference Model for Adaptive Hypermedia. in tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots: returning to our diverse roots, pp. 147–156. (1999).

[13] N. Koch, M. Wirsing: The Munich Reference Model for Adaptive Hypermedia Applications An Overview of the Reference Model. in 2nd International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, pp. 213–222. (2002).

[14] Rustici Software: What is SCORM? [Online]. Available at: http://www.scorm.com/scorm-explained/

[15] E. Cohen: cmi5 Specification Profile for xAPI. AICC. (2016) [Online]. available at: https://github.com/AICC/CMI-5_Spec_Current/blob/quartz/cmi5_spec.md

[16] T. Sellarès: Composite Pattern. Universitat de Girona. Available at: http://ima.udg.edu/~sellares/EINF-ES1/CompositeToni.pdf

[17] ADL: Advanced Distributed Learning, 2010. [Online] available at: www.adlnet.gov/help/Pages/CommonQuestions.aspx.

[18] M. Tadlaoui, A. Chikh, K. Bouamrane: Alem: A reference model for educational adaptive web applications. in Intelligent and Adaptive Educational-Learning Systems, Springer, pp. 25-48. (2013).

[19] M. Tmimi, M. Benslimane, M. Berrada, K. Ouazzani: Implemented and Tested Conception Proposal of Adaptation Model for Adaptive Hypermedia. iJET Vol. 14, No. 2 (2019)

292

A New Approach to Integrate The RFID in the Internet of Things Using The MQTT Protocol and 6lo-RFID Framwork

Y.Mouzouna[1] , M.Taouzari[2] ,A. Mouhsen[1] , S.Benalla[2] ,J.Elaoufi[1] , H. Nasraoui[1]

{ mouzouna.youssef@gmail.com , med.taouzari@gmail.com, mouhsen_ahmed@yahoo.fr }

1 Laboratory IMII, Faculty of Science and technical Settat, Hassan 1st University
2 Laboratory LISA, National School of Applied Sciences, ENSA Berrechid, Hassan 1st University

Abstract. The RFID Tag evolution from legacy passive Tag to smart Tag with sensing capacity plays a vital role in a smart environment especially in the internet of things, but their integration in this field still meet several difficulties. In this paper, we present a novel approach to integrate the RFID smart and legacy Tag in the internet of things using the publish/subscribe pattern with the MQTT (Message Queueing Telemetry Transport) protocol and a proposed 6lo-RFID framework. This approach allows integrating an RFID system continuously to the internet of things and reporting event in real time.

Keywords: RFID, Internet of Things, IoT, MQTT, Broker, 6lo-RFID Framework.

1 Introduction

Radio Frequency Identification (RFID) technology has experienced a continuous evolution in the last decade, by moving from simple transponders for tracking purposes to sensor-equipped smart Tags, which are able to implement sensing and processing functions [1] [2] [3]. Smart Tags have the potential to substitute or complement existing wireless sensor networks (WSNs) due to their tiny size, low power, and very cost-effective identification mechanisms. They may bring undisputed benefits to various IoT application scenarios, where networked Tags with sensing and computing capabilities create smart environments, such as in enhanced supply chain, factory automation, smart-home, smart-city and personal healthcare scenarios [3]. However, to fully exploit the potential offered by Smart RFID Tag, it becomes important to integrate them to the internet of things to enabling them to be accessible from and to communicate with any other networked devices in the Internet to create a fully smart environment.

The Internet Engineering Task Force (IETF) is driving the standardization of IPv6-based protocols for the IoT. The IPv6 over Networks of Resource-constrained Nodes (6lo) working Group (WG) [4] is currently enabling IPv6 connectivity for a variety of link-layer technologies in constrained node networks with limited power, memory and processing resources, such as Bluetooth Low Energy (BLE) [5] and Near Field Communication (NFC). The 6lo standardization effort has not yet addressed UHF RFID among the link-layer technologies of

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284209

interest, albeit the potential of modern smart Tags to be included as fully-fledged components in the IoT.
In this article, we design an IPv6-based internetworking solution, which natively enables smart RFID Tags and legacy tags to communicate directly with any other devices connected to the Internet. Kevin Ashton, the co-founder of the MIT´s Auto-ID Center, which developed the global standard for RFID, introduces the term IoT and other sensors, describing the Internet of Things as a system where the Internet is connected to the physical world through ubiquitous sensors [6]. The Internet of Things can be defined as a dynamically changeable and self-configurable, global network infrastructure having special characteristics based on interpretable communication protocols. In this infrastructure, physical and virtual "things" are intelligent and have unique identities, physical attributes, and are seamlessly integrated into the information network [7].
This paper describes an innovative solution based on the publish/subscribe pattern and 6lo-RFID framework to integrate the RFID Tag on the internet of things. this paper is organized as follow: a related works is presented in section 2, in section 3 the proposed approach is presented, in section 4 the 6lo-RFID framework for MQTT is described. In section 5 the communication pattern is presented and a conclusion and recommendations are developed in section 6.

2 Related Works

The integration of RFID technology in the IoT domain still requires further attention. There are many works in the literature focusing on the integration of RFID in the IoT Network by leveraging the IP stack, but their goals are only partially achieved and, typically, they lack flexibility. In [8] the concept of virtual endpoint was introduced, in [9] a proposed CoAP-compliant Solution for Efficient Inclusion of RFID in the Internet of Things is developed. In [10] an ultralow-power hybrid-sensing network composed of 6LoWPAN nodes that integrate UHF RFID functionalities is designed for e-Health applications. This solution requires that hybrid RFID Gen2 Tags are equipped with an additional IEEE 802.15.4 interface to exchange 6LoWPAN-compliant packets, thus increasing both device complexity and production cost. In [11] A Framework for Full Integration of Smart UHF RFID Tags into the Internet of Things was introduced by designing and testing a unified IPv6-based framework to access and manipulate RFID Tags' resources by means of the standard CoAP. [12] Introduces a semi passive, reconfigurable UHF RF identification (RFID) sensing Tag operating as the generic sensing platform (GSP) and the sensing nodes of the future Internet of Things (IoT). In [13] a new design and implement for integrating RFID to the Internet of Things where RFID Tags' resources are accessible via a reader acting as a CoAP proxy. The solution is backward compatible with legacy RFID systems, although it enforces the reader to interpret CoAP application requests and consequently manage Tags' resources. However, flexibility and scalability are limited by vendor lock-in issues and by the need for reader's firmware updates whenever Tags have new resources to offer. All this work uses the CoAP protocol, which is very useful when we need to access or update RFID Tag data. However, the CoAP protocol is not useful for RFID applications with requirements for continuously reporting events of RFID Tag reads (EPC codes) from many different origins. In [14] the RFIPv6 architecture makes a global access to the information stored in the EPCglobal platform, but with no real-time communications with RFID Tags.

294

3 Proposed approach

In our approach, we used the publish/subscribe pattern to integrate the RFID Tag in the internet of things. The protocol used is the MQTT (Message Queueing Telemetry Transport). The MQTT is a many-to-many communication protocol, which allows multiple clients to exchange messages using a central broker who is in charge of routing incoming messages based on published services "see Figure.1". MQTT is more matured and stable than CoAP. It is best used when a node sends data at irregular intervals of time i.e. event driven system which is good for RFID applications with requirements for continuously reporting events of RFID Tag reads (EPC codes) from many different origins [15]. MQTT is an asynchronous messaging protocol.

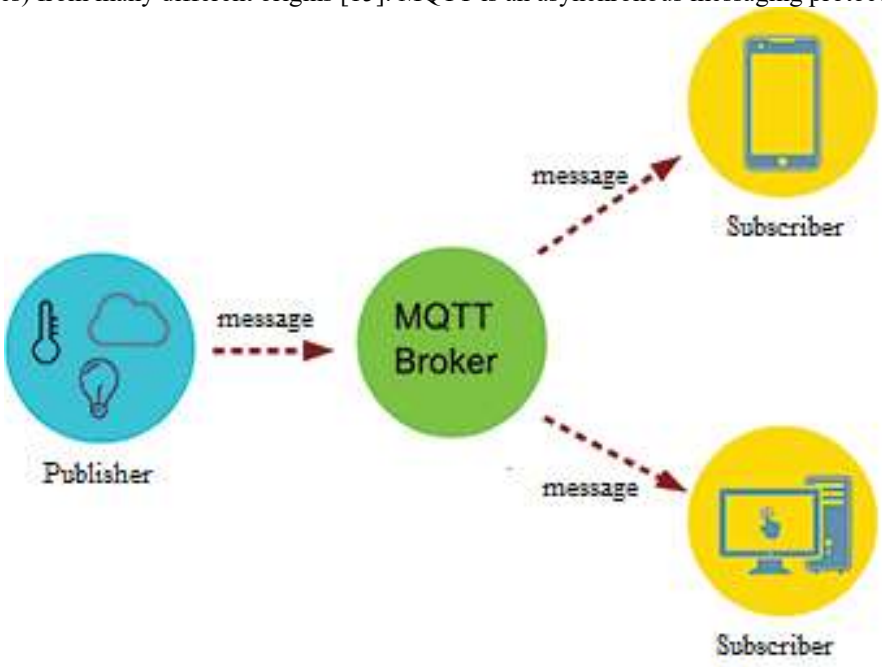


Fig. 1. The Principe of the fonctionement of the MQTT Protocol

The MQTT protocol is designed for light Machine-to-Machine communications in constrained networks [2] or high latency networks [3]. The MQTT Publisher publishes messages to an MQTT broker, which are subscribed by other clients (Subscriber) or may be retained for future subscriptions. Each message is published at an address, known as Topic [16]. Customers can subscribe to several Topics and receive each message posted to each Topic. MQTT is a binary protocol and normally requires a fixed 2-byte header with a small message payload of up to 256 MB [17]. It uses TCP as transport protocol and hence perfect for scenarios where connectivity is required all through, also it use TLS / SSL for security [17]. Thus, the communication between the client and the broker is a connection oriented. Another great feature of MQTT is its three levels of Quality of Service (QoS) for reliable delivery of messages [18].
The proposed system consists essentially of smart Tag to send their EPC and sensing data to the reader and legacy Tags their EPC to the reader. The reader will send this data to an MQTT broker or receives data from client to write them on the Tag memory; The MQTT broker can be

295

any hardware with a specific software for example a raspberry pi with a software like Mosquitto MQ. Many different brokers already support MQTT, such as ActiveMQ, Apollo, HiveMQ, IBM Message Sight, JoramMQ, RabbitMQ, MosquitoMQ and VerneMQ. The RFID reader will play the role of both a publisher by publishing continuously the data collected from the legacy Tags and smart Tags, and will also play the role of subscriber to receive data form another publisher. The clients or others devices in the network will play the role of the subscriber to read data from the Tags published in the broker in near real time and also the role of a publisher to send data to the Tags as depicted in " Figure.2". The messages will be published in Topics in the MQTT broker, the name of each Topic will be the EPC of the Tag of each Tag.

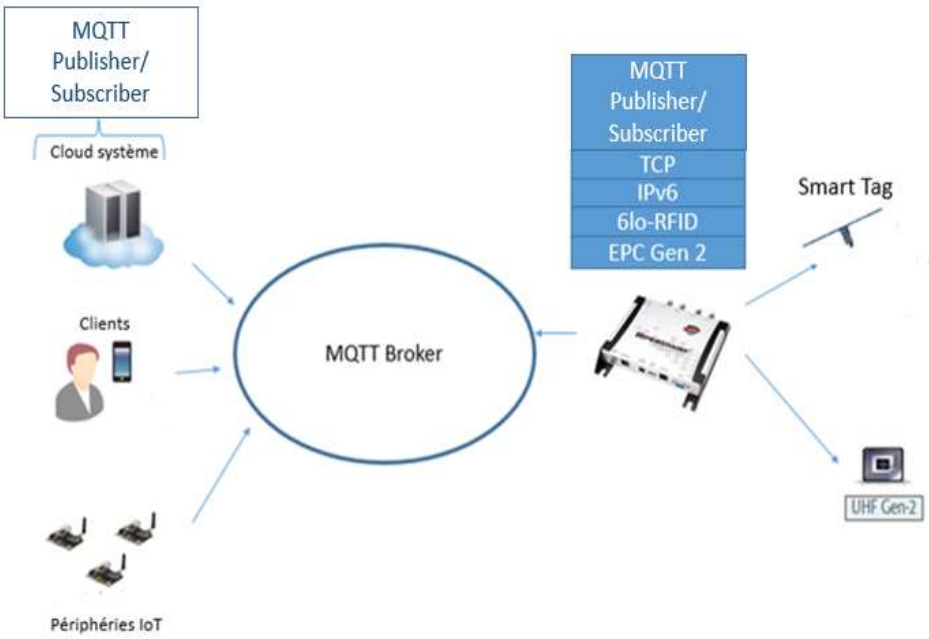


Fig. 2. The proposed system based on publish/subscribe pattern

4 6lo-RFID Framework

In the proposed system, RFID smart Tag will communicate by using IPv6 Address. For that purpose, we developed a framework for using the MQTT protocol and the Ipv6 for addressing each Tag. To allow the exchange of IPv6 packets generated by (or directed to) a given Tag, we propose a new organization of the user memory of the RFID Tag (UM), shown in Table. 1. According to the EPCglobal Tag Data Standard, the first eight bits of Unified Messaging contain the DSFID (Data Storage Format IDentifier) field, which specifies the format of the rest of the Unified Messaging store. This field includes the data format to indicate which data system predominates in the MU. Clearly, a new data format code must be associated with IPv6-based

296

RFID communications; this maintains compatibility with other RFID data standards. The following bits are reserved for future use (RFU) to manage the protocol version, priority management, or congestion management of the queue. Since the read and write commands operate on 16-bit word units, we propose to store the 6lo-RFID frame starting with the second 16-bit word of the MU to avoid unnecessary read / write of the field. The following fields are included in the 6lo-RFID header (2 to 4 bytes) 6lo-RFID header " see Figure. 3".

Table 1. The Proposed 6lo-RFID packet format

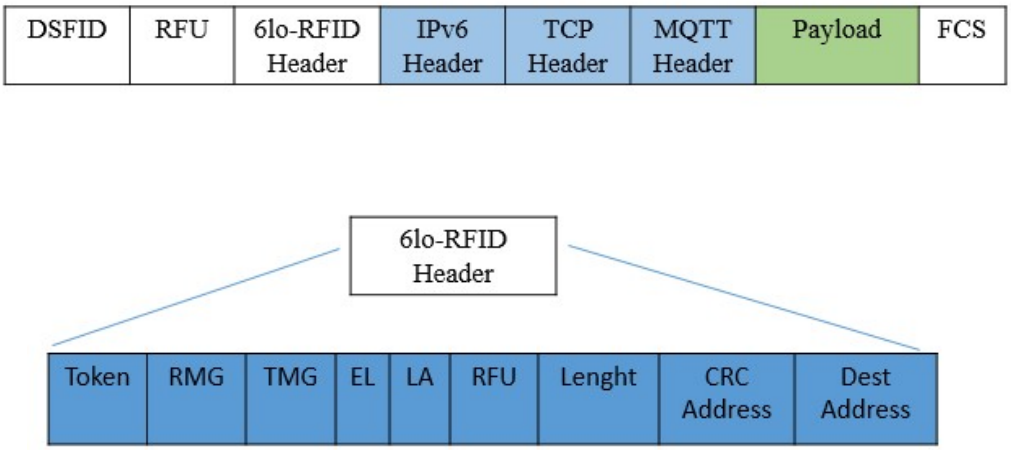


Fig. 3. 6lo-RFID Header Packet Format

Token (1 bit): Used to guarantee exclusivity in the write operation and, thus, manage simultaneous access to the UM of the Tag for the reader's RFID reader and sensor. Its value 0 allows access to the RFID reader to the MU, while the value 1 allows the access of the sensor to the UM. To avoid blocking, the reader resets the UM to start a new exchange of information if the sensor does not clear the token bit before an optimal time. **Reader Message (RMG**) (1 bit): Indicates if the reader has written a message into the Tag. The Tag subsequently clears this bit after reading the relevant reader message. Tag Message **(TMG) (1 bit)**: Indicates if the Tag has inserted a message in the UM. The reader clears this bit after reading the Tag message. **Extended Length (EL) (1 bit):** If the packet length is more than 255 bytes and the UM contains it, then this bit must be set to one to use an Extended Length field of 16 bits. **Local Addressing (LA) (1 bit)**: Its value 1 indicates that a local addressing strategy is performed. In this case, the header includes source and destination addresses (more details are given in the following). **(RFU)** Reserved for Future Use (3 bits). **Length (8–16 bits):** Indicates the number of bytes that compose the 6lo-RFID packet (the field allows to exchange up to 65535 byte-packets). A frame check sequence (FCS) of 32 bits is added at the tail of the 6lo-RFID frame to detect tampering or inconsistencies due to errors or partial writing of the UM. Considering the limited frame size, header compression is highly advisable to efficiently exchange data. To save valuable bits in the memory, The IPv6 And TCP headers shall be compressed according to the standard encoding

297

schemes defined in IETF RFC 6282 [4]. Since IPv6 defines a maximum transmission unit (MTU) of 1280 bytes, the 6lo-RFID layer must also implement fragmentation and reassembly functionality, as defined in RFC4944 [4], when the UM is not large enough to contain the whole IPv6 packet.

298

5 Communication Pattern

The proposed approach is a master-slave approach to accessing the UM and managing the reader-Tag interaction, where the reader acts as a master and controls the flow of data from / to all RFID Tags in its coverage area. The proposed solution is compatible with the Gen-2 protocol because it can easily exploit the already defined Gen-2 commands. The reader initiates the label singulation phase by specifying the EPC code of the desired label in a SELECT command, so that only the label with the specified EPC can participate in the inventory phase and the collisions, communication between two Tags is done after reading of each tag in its area. Then, the Writing and Reading commands respectively allow the transfer of data in the downlink and uplink directions. All these data collected by the reader are published in the broker. This operation is carried out continuously in order to evaluate any event in the RFID system. The communication pattern is described "in the Figure.4".

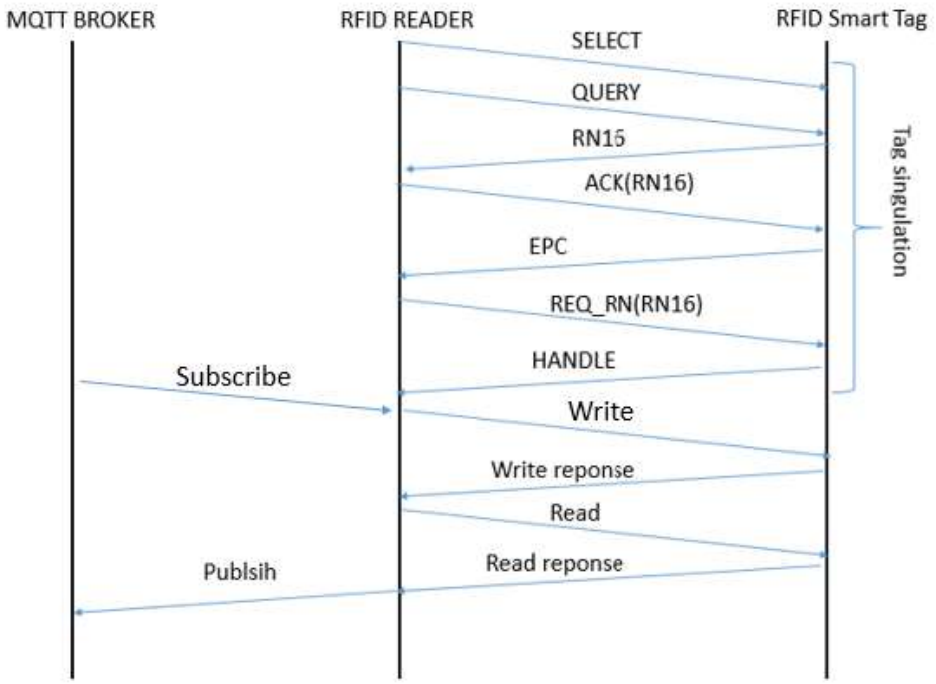


Fig. 4. The proposed system based on publish/subscribe pattern

299

6 Conclusion

In this article, a novel approach based on IPv6 framework and publish/subscribe pattern using MQTT protocol has been designed to transparently integrate RFID systems into the IoT domain and enable IPv6 communication and networking over UHF RFID smart Tags. This approach uses a new packet format of data stored in the UM of the RFID Tag called 6lo-RFID. This framework includes and IPv6 header, TCP header, MQTT header. This approach is compatible with the standard EPCglobal Gen-2 protocol which make it flexible to use.

300

References

[1] J. Smith, A. Sample, P. Powledge, S. Roy et A. Mamishev, “A wirelessly-powered platform for sensing and computation,” *Ubiquitous Computing, Lecture Notes in Computer Science; Springer: Orange County, CA, USA,,* Vol. 4206, pp. 495-506, 2006.

[2] D. De Donno, L. Catarinucci et L. Tarricone, “.A battery-assisted sensor-enhanced RFID Tag enabling heterogeneous wireless sensor networks,” *IEEE Sens.,* pp. 1048-1055, 2014.

[3] M. Khan, M. Islam et D. Hai, “Design of a reconfigurable RFID sensing Tag as a generic sensing platform toward the future Internet of Things.” *IEEE Internet Things,* pp. 300-310, 2014.

[4] IETF 6LoWPAN Working Group, https://tools.ietf.org/wg/6lowpan.

[5] J. Nieminen ; C. Gomez, “Networking Solutions for Connecting Bluetooth Low Energy Enabled Machines to the Internet of Things,” *IEEE Network,* Vol. 28, No. 16, pp. 83-90, 2014.

[6] Kevin Ashton, "That 'Internet of Things' Thing, RFID Jouranl, 2009.

[7] F. Hussein, Internet of things: Building blocks and business models, Canada,: springer, 2017.

[8] Jara AJ, Moreno-Sanchez P, Skarmeta AF, “IPv6 Addressing Proxy: Mapping Native Addressing from Legacy Technologies and Devices to the Internet of Things (IPv6)” *Sensors,* Vol. 13, No.15, pp. 6687–6712., 2013.

[9] Ivan Farris ; Antonio Lera ; Antonella Molinaro ; Sara Pizzi, “A CoAP-Compliant Solution for Efficient Inclusion of RFID in the Internet of Things” *IEEE Global Commun. Conf.,* pp. 2795–2800., 2014.

[10] L. Catarinucci ; D. de Donno ; L. Mainetti ; L. Palano “An IoT-Aware Architecture for Smart Healthcare Systems” *IEEE Internet of Things,* Vol. 2, No16, pp. 515-526, 2015.

[11] S. P. Ivan Farris, M. Merenda, A. Molinaro, R. Carotenuto et a. A. Iera, “6lo-RFID: A Framework for Full Integration of Smart UHF RFID Tags into the Internet of Things,” *IEEE Network,* 2017.

[12] S. Islam ; a. D. Ha, “Design of a Reconfigurable figurable RFID Sensing Tag as a Generic Sensing Platform Toward the Future Internet of Things,” *IEEE Internet of Things ,* Vol. 1, No.14, pp. 300-310, 2014.

[13] I.Varis; S.Pizzi; A.Molinaro ”Design and Implementation of a CoAP-Compliant Solution for RFID Inclusion in the Internet of Things” J.Sens.Actuator Netw,Vol.5, No.4, 2016.

[14] Yao-Chung Chang; Jiann-Liang Chen; Yu-Shan Lin ; Shi-Ming Wan “RFIPv6 - A Novel IPv6-EPC Bridge Mechanism” International Conference on Consumer Electronics IEEE, 2008

[15] F.Valente; A. Neto “Inteligent Steel Inventory Tracking with IoT/RFID” IEEE International Conference on RFID Technology & Application (RFID-TA) ,2017.

[16] Youngchoon Park, Building an effective IoT ecosystem for your business, Milwaukee, Wisconsin, USA: Springer, 2017.

[17] A. Gilchrist, Industry 4.0: The industrial internet of things, Apress, 2016.

[18] Bhattacharyya, «Lightweight internet protocols for web enablement of sensors using constrained gateway devices,» *IEEE International Conference on Computing, Networking and Communications (ICNC),* pp. 334-340, 2013.

301

Application of game theory to the optimization of wireless systems Algorithmic solutions

[1]Sara Riahi, [2]Azzeddine Riahi
{ [1]riahisaraphd@gmail.com, [2]riahikh@gmail.com }

Department of Computer Science, Chouaib Doukkali University , Faculty of Sciences , El Jadida, Morocco
IMC Laboratory, Department of Physics, Chouaib Doukkali, University, Faculty of Sciences, El Jadida, Morocco

Abstract: This paper is primarily based on a study of some research that addressed resource management in wireless systems. The study presents a classification of schedulers in the uplink and is interested in the class of schedulers based QoS because of the importance of delay parameters and rate in optimizing the management of resources by the formalism of game theory. Then, some scheduling algorithms in the ascending link are exposed to make a complete analysis of the different aspects adopted in the scheduling. Second, the courtesy algorithm resource optimization in the uplink in fixed wireless systems is presented. The algorithm defines a priority management policy to improve the low-priority traffic service without affecting the high priority traffic QoS. Finally, a critical assessment of existing solutions is carried out to a design of a robust scheduling mechanism.

Keywords: Game theory, schedulers, resource allocation, QoS , Fairness, utility.

1. Introduction:

Scheduling as well as resource allocation has attracted much research because of its importance in improving system performance. Wireless networks, and more particularly Satellite Telecommunication Systems, have not escaped this craze, even if they have distinct characteristics, and therefore constitute a field of application in its own right. As shown in our work **[1]**, the introduction of flow with very different characteristics and temporal versatility of the radio bearer is more complicated than scheduling homogeneous flow (MPEG-2 in the case of satellite) is easy. This problem is not inherent to the satellite, but the characteristics of the system are very demanding, especially on the joint evolution of capacity and demand, and encapsulation methods, and must be taken into account in the design of the algorithms. Our objective in this paper is to summarize the main aspects of the scheduling and resource allocation solutions that have been developed for satellite systems in particular **[2]**. An essential part of our study will, therefore, be devoted to determining how to adapt the presented solutions to our reference system. Note that, although very different, solutions designed for terrestrial systems can also be studied, if their adaptation seems relevant. Because of the eminently dynamic nature of the capacity and the demand of the studied system, we will pay particular attention to how the solutions proposed in the literature take into account these parameters. Most scheduling algorithms use one or more metrics to evaluate the state of the system and make decisions accordingly. We will first briefly focus on estimating these metrics. We present in a second time a general classification of the methods of scheduling where three mathematical tools, widely acclaimed in the literature, are exposed. Then, a state-of-the-art methods developed for the satellite systems are presented. This study makes it possible to highlight the difficulties related to our problem, as well as the insufficiencies of the solutions proposed until then.

2. Metrics:

The use of metrics to observe the state of a telecommunications system is necessary, therefore, that this system evolves in time, which is more difficult to predict manner. A metric is a numeric value that represents the state of the system. This definition also includes a temporal aspect, we find average metric, defined over a certain period of time, or a posteriori, and instant metrics, whose value has meaning at a given time **[3]**. For example, the average rate observed by a Terminal or delay of a frame in a queue of the Gateway. However, the estimation of these metrics is problematic: it must take the

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284113

least resources (memory, calculation) possible, while being relevant. We therefore propose a compromise between these two objectives.

2.1 Definitions:

First, we define the scheduling time, as when the last BBFrame is passed for the Go path. This moment comes back regularly, more precisely periodic for the return path, depending on the transmission time of the BBFrame for the Go path. Having an incremental numbering n of the BBFrames according to a given origin, we note $t(n)$ the instant at which the BBFrame n begins its transmission. We thus define a discrete measure of time, $t(0), t(1), ..., t(n)$, which will serve as a clock for our system **[4]**. For clarity, we will use interchangeably time $t(n)$ and n in the rest of this discussion. This modeling is presented in Fig 1 It seems important to note that this representation is not only arbitrary, but also strongly depends on the system considers (frequency of BBFrame, ModCod), although adopted in many works.

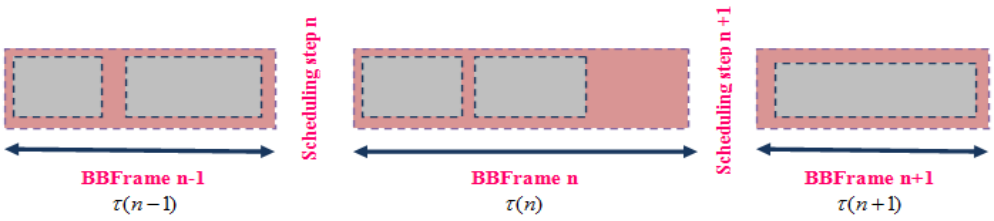


Fig 1: discrete time measurement

It is also said that the algorithm is self-timed because it does not use an external clock as opposed to algorithms such as Weighted Fair Queuing (WFQ). On the Return path, this time measurement corresponds to the period of the Super frames DVB-RCS2. Based on these values, more precisely on the system state observed at time n (the set of queues, and the variables associated to the access layer), several scheduling solutions are determined. They correspond to the different possibilities available to the access layer for scheduling or allocating resources **[5]**. We will call scheduling decision the selection of the scheduling solution (respectively resource allocation for the return channel). The way to make this decision depends on the algorithm, or an evaluation will be made of each scheduling solution, according to criteria specific to the algorithm. The tool for making this decision is called the scheduling criterion. Finally, we define by scheduling scenario all the values of the influence parameters for the scheduling in our system, namely **[6]**:

- The number of Terminals;
- The type of traffic, its characteristics, and the path considered (Forward and Return);
- The distribution of ModCod Terminals.

In summary, any change in the parameters of demand or capacity as defined in previous work constitutes for our system, a different scenario.

2.2 Metrics:

If the algorithm considered makes use of average values, the value of the average variable $\bar{x}$ used at time n will be $\bar{x}(n)$, and the instantaneous value measured on the system will be denoted by $x(n)$. We say that the algorithm enhances the metric x if the value of the latter is used in the scheduling decision, even in the expression of the scheduling criteria. The instants of observation of the system

corresponding to those of scheduling, the algorithm therefore has access only measured instantaneous values, $x(n)$. There are several ways to estimate the average value of x from these instantaneous values, we chose the exponential average, which filters the events with high amplitude relative to the average, thereby reading the measurement. This type of measurement, widely used, offers a compromise between accuracy and simplicity of calculation. When considering the metric, its exponential average is given by **[7]**:

$$\tilde{x}(n+1) = \alpha \tilde{x}(n) + (1-\alpha)x(n) \qquad (1)$$

Where α is a damping factor, close to 1. Note that this measurement does not depend on the time elapsed between two instants, which can be detrimental when this duration is variable, for example on the Go path, when two different BBFrames of ModCod follow each other. They will not have the same transmission time, so the time between two instants will be variable. The authors of **[8]** propose to adapt α to the duration between two instants, by posing:

$$\alpha(n) = \alpha^{\tau(n)} \qquad (2)$$

Where $\tau(n)$ is the duration between two scheduling instants, ie the transmission time of the chosen BBFrame. This formula will be used by default in the rest of our discussion.

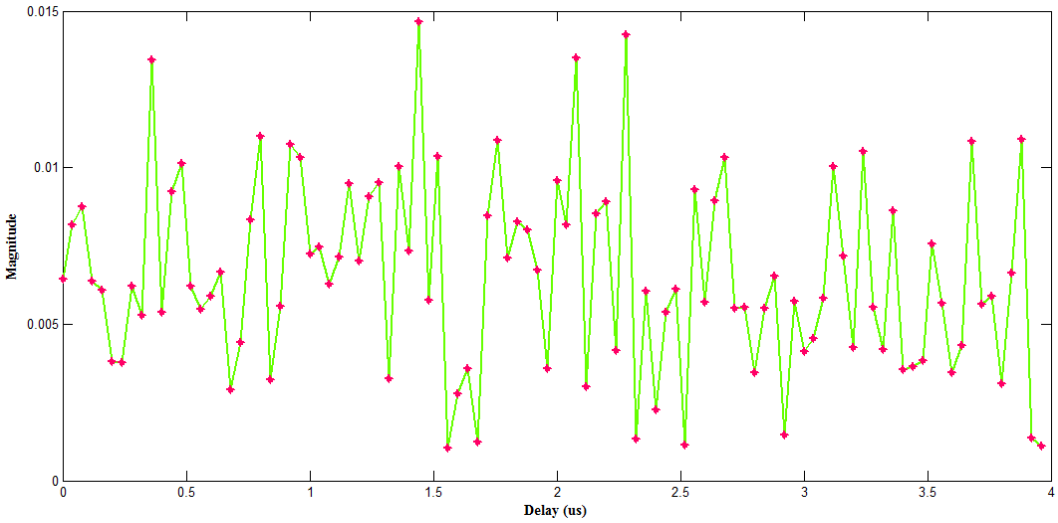


Fig 2: Power delay

3. Analytical scheduling tools:

In this part, we expose three methods for evaluating scheduling or resource allocation solutions, each based on a different mathematical modeling. Besides presenting a classification of schedulers, we have also investigated how methods that are a priori very different can lead to the same result, through a common example. Our reference system is constituted by the access layer of the Gateway, on the Go - path or on the Return-path. The scheduling is the reference algorithm of this study, but the explanations given here are, by default, also valid for the Go-path **[9]**. Both processes are distinguished when this is no longer the case. We consider here any number of users, indexed by the variable i . For the sake of generality, we do not further explain these users, which may be either queues located in the Gateway, Terminals or even flows.We consider a generic metric, denoted $x(n)$, average or instantaneous, representing the state of the system, and we assume further that can be clearly defined, and for each of

these users, a state $x_i(n)$. One example is the average rate, or the loss rate. This value is instantly available for the reference system, and updated at each new scheduling time.

3.1Explicit criterion:

This category of solutions includes a large number of developed algorithms for scheduling in wireless systems. Their principle can be summarized as follows **[10]**:

- Select one or more metrics, x_i representative of the status of each user in the system.
- Define an evaluation of this state at each moment, $f_i\left(x_i(n)\right)$, which will be associated with the scheduling criterion.
- Evaluate the different available scheduling or resource allocation solutions, and make the scheduling decision.

3.1.1 Empirical definition:

This definition remains general and could give rise to very different algorithms. It is said to be explicit, or empirical, because the evaluation uses a f function that must be explicitly given, and which is generally based on empirical considerations. Also, the way of using the criterion thus defined must also be explained. In the case where only one user is served at each scheduling instant, a very simple example of a scheduling algorithm built on this criterion is given by **[11]**:

$$\arg\max_i f_i\left(x_i(n)\right) \quad (3)$$

In other words, the user i^* is served at time n only if the value of his criterion is the highest. This formulation has the advantage of being simple and easy to implement, a simple calculation is sufficient to determine the next user to serve. In addition, its complexity increases linearly with the number of users, and its implementation can be easily parallelized. Another possible use of this criterion can be materialized through a conventional scheduling algorithm, such as the RR, where each user is also successively served, in a defined order. Here, we can suppose the users being sorted by descending criterion, this order being updated regularly, with the new values of the criterion **[12]**. In a system with only packets of constant size, the RR is fair and efficient, and has a constant algorithmic complexity, $o(1)$.It is also possible to use this criterion in the RR priority variants. The Weighted Round Robin (WRR) makes it possible to differentiate the treatment between the different users, by assigning a weight to them. This weight may correspond, for example, to a QoS constraint. Mainly developed for ATM, weight can be expressed as the relative time given to a user during a RR cycle. This algorithm has a certain interest when used with packets of constant size, as is the case of ATM cells, since its behavior can be predicted. DRR, presented in **[13]** is a variant of WRR adapted to variable size packets. The deficit is thus the volume allocated to each user in each round of the RR. This deficit can also be adapted according to QoS constraints. This deficit is consumed when the user packets to be transmitted and accumulated when the deficit accumulates is not sufficient to transmit the next packet. In this way, the DRR also allows smoothing of the traffic. Like RR and WRR, the algorithmic complexity of DRR is low $(o(1))$. We can finally mention the WFQ, although this algorithm is more complex. It is based on a fair sharing of capacity among users, according to an implementation time-sharing model with a processor between several spots. This algorithm creates many searches. The fact that this algorithm assumes an external clock to the system and its complexity, make it difficult to use for a system whose capacity varies with time. Moreover, it has a higher algorithmic complexity, $o(log(N))$. In summary, this formulation allows a very high flexibility and low complexity implementation, two very important assets for systems like satellite gateway, where the scheduling decisions must be made in a few microseconds. But this very simple formulation hides a great complexity of design, in the choice of f but also in the use that is made of the criterion **[14]**. Indeed, since there is no particular constraint on these two choices, the number of possible solutions is enormous. However, the evaluation work of each

criterion is complex: an analytical assessment is possible for a single criterion, but if becomes too complex criteria expression contains several different metrics, each having its own impact on f .

3.1.2 Scheduling rules:

The scheduling rules are a partial response to the problems of the explicit criteria mentioned above. A scheduling rule is a particular category of explicit criteria. These rules take the form of criteria whose expression, derived from reference works **[15]**, is defined, thus allowing a more homogeneous formulation of the scheduling algorithms. The parameters used by these rules are average values of conventional metrics, such as the outgoing throughput, the delay of the packet at the head of the queue (Head of Line (HoL)), or the instantaneous quality of the medium. A particularly popular rule is Proportional Fairness (PF), whose expression is:

$$i^* = arg\ \max_i \frac{r_i(n)}{\tilde{r}_i(n)} \qquad (4)$$

Where:

- r_i is the instantaneous rate achievable by the user i , depending on the transmission conditions;
- $\tilde{r}_i$ is the average rate obtained by the user i . PF achieves a compromise between spectral efficiency and fairness.

Through the incorporation of instant channel quality, PF privileged users with a good achievable throughput, which are called opportunistic behavior. A rule that only takes into account this parameter would maximize spectral efficiency, provided that the users served always have data to transmit or receive **[16]**. It would, however, degrade fairness since a user with poor transmission conditions for a long time may not be served all this time, creating a starvation phenomenon. That's why this influence is balanced by the inclusion of the obtained average rate. Thus, even if a user has very good conditions of transmission, the value of his criterion will fall as it is served, thus making it less of a priority. This rule easily adapts to the Go-path as well as to the Return-path: it is possible to choose one or several users, sorted by decreasing PF criterion. It is mainly adapted to so-called elastic traffic, that is to say whose bandwidth requirements can adapt to the available capacity. Inelastic traffic, usually at constant speed (VoIP, video), requires a more complex rule, such as Modified Latest Weighted Deadline First (M-LWDF):

$$i^* = arg\ \max_i \gamma i\, \bar{\omega}_i(n) r_i(n) \qquad (5)$$

Where:

- γ_i is a QoS priority;
- $\bar{\omega}_i(n)$ is the average waiting time in queue head of the user i .

This rule allows a compromise between throughput, through r_i , and the delay. We can also mention the exponential rule, or Exponential Rule - PF (exp-pf), or Earliest Deadline First (EDF), both adapted to real-time traffic. The scheduling rules seem to provide a good compromise between complexity, performance and ease of design. The abundance of research on these specific criteria is also an advantage. However, there are two disadvantages to the use of scheduling rules. First, it is difficult to obtain stability results from these rules, other than by extensive simulations, taking into account a large number of cases. This can be critical in high-load systems, where the risk of destabilization of the system is great. Secondly, the scheduling algorithms defined here are said to be myopic, since they only consider the past of the system to make their decision. In other words, it is not possible to easily predict the impact of scheduling decisions on the system, and thus to know the state to which scheduling leads the system **[17]**.

3.2 Game theory:
In order to give a rigorous mathematical framework to the scheduling, many works have used the Game Theory to model situations where several users share the same telecommunications system. Game Theory was introduced in the mid-twentieth century to formalize competitive situations between economic actors. These players, called Players, make decisions according to a strategy so as to maximize their earnings. Applied to wireless networks, we can easily see how it is possible to apply this theory, for example to model contention access protocols, where the gain would be the rate received according to the strategy, which would be the way in which the user accesses the support.

3.3 Utility functions:
One of the major results of Game Theory is the Nash equilibrium, introduced by John Nash **[18]**.It stipulates that, in the case of players having a rational estimate of their winnings according to their strategy, and having a complete knowledge of the strategies of the other players, each player will reach a balance where his gain will be maximum, knowing the strategies adopted by other players. In other words, in this equilibrium, each player takes into account the strategies of all other players, and decides the best strategy. This state is a balance, since no player can increase his gain by changing strategy. This solution is called Nash Bargaining Solution (NBS). A NBS can be formulated as follows, for users whose gain is the average received rate $\bar{r}_i$, the Nash equilibrium is the solution of the following optimization problem **[15]**:

$$max \prod_i \bar{r}_i \quad , \qquad \sum_i \bar{r}_i \leq C \qquad (6)$$

Where C is the total capacity of the system, expressed here in terms of throughput. This formulation is particularly useful for capacity sharing problems, as in wired networks, or in a static context, for example for the design in the networks Orthogonal Frequency Division Multiple Access (OFDMA). In addition, the Game Theory provides guarantees for the stability of the equilibrium achieved, which is a definite advantage over the explicit criteria. We will note later $\bar{r}^*$ the solution of this problem. However, there is still a big problem before we can use game theory effectively: adapt this formulation to the dynamic nature of our problem. Indeed, the previous formulation provides a balance on average values, independent of time. However, not only can scheduling be processed without a temporal dimension, but it is even necessary to adapt it to changes in the system. For example, we have here expressed the throughput capacity. As seen in **[11]**, this ability evolves over time, it would be more accurate to replace C by $C(n)$, which has the effect of changing the solution (4), since the capacity constraint changes. For the rest, even by explaining the values of $\bar{r}^*$, how to make the connection between this value and the scheduling decision taken at the instant n ? Or, how does the valorization of the flow rate appear as a gain in the scheduling? The Game Theory provides no answer here, at least by adopting the formulation of (4), and the relationship between the equilibrium arising from the NBS and the effective strategy that must be adopted at each scheduling stage is absolutely not explicit. The advantage provided by game theory is therefore mixed here: although it provides rigorous mathematical tools, its formulation is not adapted to the dynamics of the system, especially that of scheduling. Moreover, assuming that an iterative method has been found in order to reach the defined equilibrium, the dynamic and limited nature of the resources will here again be a constraint. Indeed, assuming the instant n a system state, ie average values of speed $\bar{r}(n)$, it probably will not be achieved in one step $\bar{r}^*$ scheduling, because the speed at which the average flow rate can evolve is limited by the capacity $C(n)$. In other words, the set of reachable states in one scheduling step is limited, and this set does not necessarily contain the optimal **[8]**. Therefore, it will be necessary to perform several scheduling steps before reaching the defined equilibrium. Now, this equilibrium, as seen above, depends on time too, through $C(n)$: then we risk never being able to reach the defined equilibrium, since it changes over

time. Formally, if the algorithm converges towards equilibrium T steps will require $C(n) = C(n+T)$ to that equilibrium will actually be achieved; otherwise the system will not reach equilibrium at $n+T$, since it will have changed.

Proportional Fairness:
An interesting property of this formulation is the possibility of finding the PF criterion as previously reported **[10]**. Indeed, by considering the following problem, after passing logarithm equivalent to (4):

$$max \sum_i ln(\bar{r}_i) \quad , \quad \sum_i \bar{r}_i \leq C(n) \qquad \textbf{(7)}$$

We then obtain a nonlinear optimization problem, strictly convex, which can be solved using Lagrangian methods. The solution of this problem **[17]**, which is none other than $\bar{r}^*$, expressed proportionally fair, due to the property obtained for $n \to +\infty$:

$$\sum_i \frac{r_i - \bar{r}_i^*}{\bar{r}_i} \leq 0 \qquad \textbf{(8)}$$

Where R is a possible allocation of average bit rates, which means that the overall gain brought by a change, relative to the NBS, in the bit rate of one of the users (for example its increase) is zero or negative. This property ensures fair behavior, preventing a single user from capturing the majority of resources: players will work together to achieve the best compromise. In its application to scheduling problems, a gradient algorithm has been adopted **[18]**, to solve the problem iteratively. The expression of the scheduling criterion arising from this problem is as follows:

$$i^* = arg \max_i \frac{r_i(n)}{\bar{r}_i(n)} \qquad \textbf{(9)}$$

Exactly the expression (2), which means that this criterion is not only fair, but also optimal for elastic traffic, enhancing throughput. It is important to note that here, the scheduling decision is not myopic: we choose the one that leads as quickly as possible to equilibrium, which is an important property because of the dynamics of the system. This method also makes it possible to evaluate according to a defined metric, namely the gain at each step, the different possible scheduling solutions **[12]**.

Utility functions:
The utility functions are related to game theory, but can be a tool in its own right, usable without having to deploy the formalization necessary for game theory. Formally, a utility function associates a variable, or a metric, with a utility. It describes how the value of the variable or metric is valued by the algorithm. We can then define, for our reference telecommunications system, a set of variables and their associated utility: for example, one could measure user satisfaction with a certain level of service received, and ensure that scheduling maximizes this value. However, this solution has been retained in a number of scheduling methods and algorithms **[19]**. This success is explained by the flexibility offered by utility functions compared to game theory, and by the fact that it is often possible to obtain comparable results, without resorting to a complex and sometimes constraining theory.

3.3.3 Proportional Fairness derived from utility functions:

Suppose again users with elastic traffic, whose average rate allocated by the system is $\bar{r}$, and the utility corresponding to this traffic $u(\bar{r})$. If utility represents the satisfaction of the user vis-à-vis the service received, it will have to be strictly increasing with the flow for an elastic traffic. However, in order to

differentiate low flow rates from high flows, we put $u(\bar{r}) = ln(\bar{r})$. In this way, the difference between two allocations tends to decrease gradually as these allowances are increasing, which means that it will be more useful to increase the throughput of users with low average throughput, rather than those already having a high average throughput. The objective of the system is relatively simple; it consists in maximizing the usefulness of the system, namely the sum of the utilities of the users:

$$max \sum_i ln(\bar{r}_i) \quad , \qquad \sum_i \bar{r}_i \leq C(n) \qquad \textbf{(10)}$$

This definition, directly in the form of an optimization problem, did not need other hypotheses than the convexity of the problem, easily demonstrable, and the concavity of u. We thus find a formulation strictly equivalent to (5), to which we will apply the same resolution, to find the criterion PF, whose expression is given by (2).

3.3.4 Necessary compromises:

As we have seen with the example of PF, the utility functions need to provide an objective, here the maximization of the utility of the system. Once again, this problem is defined using average values, and will therefore pose the same difficulties as the game theory with regard to the resolution and the algorithm resulting from this resolution. The advantage of the utility functions is the flexibility provided: the purpose of the scheduling is explicitly defined, and it is possible to design a more complex utility function, which already represents a compromise between several variables, for example delay and jitter, or the loss rate. This increased flexibility, however, has a cost, since it can make solving the optimization problem more complex, or even lead to an unstable equilibrium. It will then be necessary to make explicit these properties. We can thus see utility functions as a compromise between explicit criteria and game theory **[20]**. In the first, we explicitly define the instantaneous scheduling criterion, without studying its impact on a time scale longer than a scheduling step. Game theory, it presupposes a defined user behavior, which derives the optimum. The definition of this behavior is certainly empirical, but more easily justified: we say that elastic flow enhances throughput, while a real-time stream will be attached to jitter and delay. However, the mathematical framework thus defined is relatively heavy, for a final result sometimes equivalent to the scheduling criteria. The utility functions presuppose two assumptions **[21]**: The needs of users, represented by the utility function, and the purpose of the system, the problem of optimization. Empirical choices are therefore more important with utility functions, but they allow more flexibility in scheduling and resource allocation algorithms. While it is possible to use the utility functions in Game Theory, which represent the gain but there is no guarantee that this additional formulation is useful for the scheduling algorithm itself, as seen with PF. We successively proposed three analytical tools for scheduling and resource allocation: the explicit criteria, the Game Theory, and finally the utility functions. Although related, we sought to isolate, for each tool, its specificities, its advantages and disadvantages. This classification highlights the difficulty of reconciling simplicity (in the design as well as in the implementation) with the guarantees provided by a rigorous analytical model. In terms of computational complexity, we can see that although the Game Theory and the utility functions are disadvantaged, some explicit criteria can be used by algorithms such as WFQ whose complexity is comparable to the optimization problem solving algorithms. We also highlighted the difficulty of linking the definition of an optimal with the algorithm to reach, in the case of Game Theory as in the utility functions. We can therefore question the possibility of actually achieving this optimal, and its relevance in a system evolving over time **[22]**. If the myopic scheduling seems to be less interesting, what gain is actually brought by methods anticipating the impact of scheduling decisions. In addition, the boundaries between each family of tools are not strict, even less if they achieve the same result, as PF. However, we can, from this description, analyze the different scheduling proposals from a different angle of a simple assessment of performance, including the relationship between the scheduling criteria that determines the decision to each scheduling instant and the objective of scheduling, in the longer term, which describes an average behavior of the system. We now have the necessary tools to analyze the scheduling and resource allocation solutions for wireless systems, proposed in the literature on the subject.

4. The methods of scheduling and resource allocation:

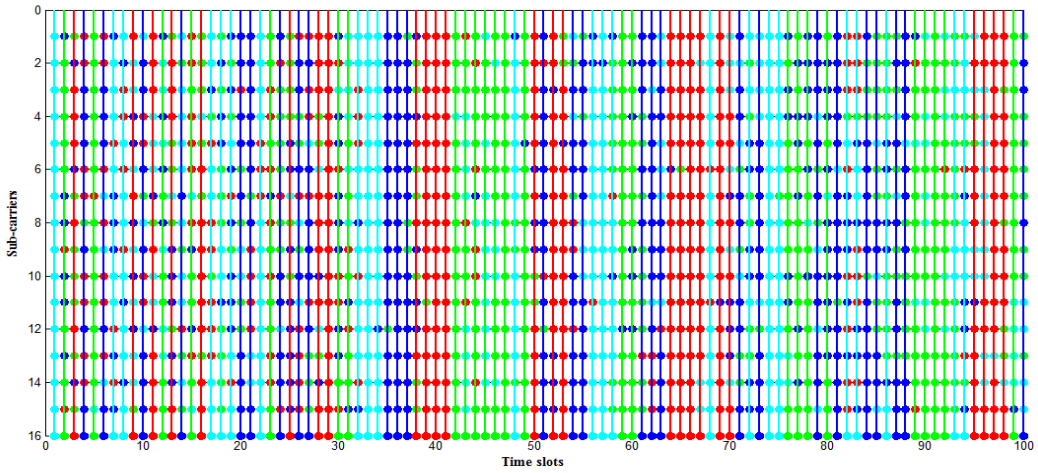


Fig 3: Subchannel allocation according to PFS

Scheduling in the uplink is complex for several reasons. First, the user equipment has a limited power source. Second, it is difficult to predict the number of radio resources that the user equipment needs to exchange data with the base station. Depending on the objective function considered and the classes of traffic carried in the radio channels, the authors of **[23]** define four families of schedulers. The QoS-based scheduler category considers the delay, maximum throughput, and number of users served to provide the required QoS to users. The QoS based scheduling has been the subject of several studies. The algorithms presented in the following section provide an overview of the different solutions proposed in the literature.

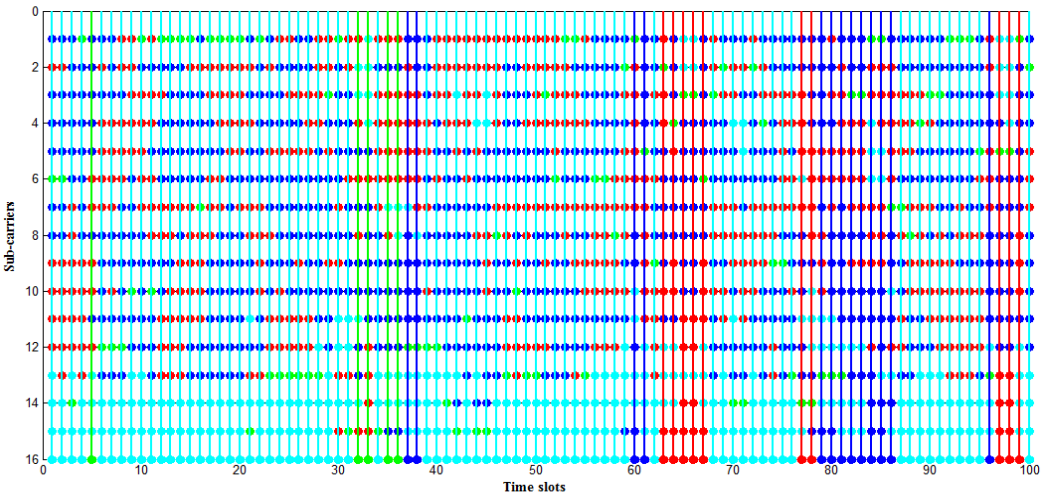


Fig 4: Subchannel allocation according to WFQ

An advanced version of the Proportional Fair scheduling algorithm is developed in **[24]**. Its purpose is to improve the throughput in the uplink of LTE-A users who are at the edges of the cell and with a low SINR. The algorithm proceeds in three steps.

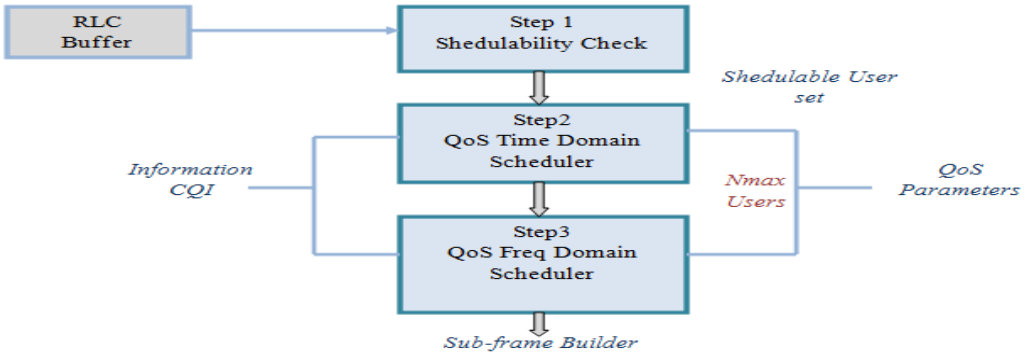


Fig 5: packet scheduler Diagram

Step 1: Users are selected based on the period and the value of the buffer to be scheduled at the current time t . Other users are selected at the next moment.

Step 2: the main purpose of this step is to provide appropriate QoS to users who have a low tolerance for delay. Users are divided according to the waiting time into two priority groups. If the value of the latency is less than a predefined time threshold, the users are grouped into group 1 and have the lowest priority. In the opposite case, users are grouped in group 2 and have the highest priority. Depending on the priority, $(N_{max}+1)$ users are selected for frequency domain scheduling because it is assumed that some users are unable to receive the data.

Step 3: The scheduler selects the frequency domain PRB and users according to their priorities. PRBs are allocated to maximize the throughput of each user. The simulation results show that the new Proportional Fair algorithm improves throughput for users who have low SINR by comparing with the classic Proportional Fair algorithm. In **[25]** the authors propose a solution to overcome the problem of QoS degradation at the edges of the cell caused by interference with neighboring cells. The solution is based in the first place on a new call delimitation policy called New Call Bounding whose principle is to reject a new call when the number of new calls admitted in the cell exceeds a certain threshold M . The handoff call is rejected only when all channels in the cell are used. The second phase of the solution is to allocate resources to users located at the edges of the cell according to the Software Frequency Reuse scheme which divides the cell into two parts, the edges of the cell and the center part of the cell.

5. Conclusion:

Game theory allows placing scheduling and resource allocation within a rigorous mathematical framework, bringing the optimality guarantees and stability around an equilibrium. Unlike explicit criteria outlined above, game theory clearly defines a goal, a direction toward which the sequencing will lead. We are then able to determine and study this equilibrium, and thus better know the behavior of the system. But this definition of an equilibrium, and therefore of average values, does not directly provide the solution to the problem of scheduling, nor even to that of resource allocation, because of the rapid dynamics of the system. The classic solution to this problem is to move the system, at each scheduling step, to solve the problem. Therefore, the rewards of game theory seem to yield to the complexity of implementing the solution, which is particularly evident with the example of PF.

Acknowledgment:

We would like to thank the CNRST of Morocco (I 012/004) for support.

References :

[1] Sara Riahi, Azzeddine Riahi,'' Energy Efficiency analysis in wireless systems by game theory '',5th international IEEE Congress on information Sciences and Technology IEEE CISt'18,Marrakech ,Morocco ;October 21-24,2018.

[2] Q. Fan, Q. Wu, F. M. es, N. Xiong, A. V. Vasilakos, A. Y. He, "Game and balance multicast architecture algorithms for sensor grid", *Sensors*, vol. 9, no. 9, 2009.

[3] J. Shneidman, C. Ng, D. C. Parkes, A. Auyoung, A. C. Snoeren, A. Vahdat, and B. Chun, “Why markets could solve resource allocation problems in systems,” in Proceedings of the 10th USENIX Workshop on Hot Topics in Operating Systems, 2005.
[4] S.Gayathri M.Tech, R.Sabitha,’’ A Survey on Resource Allocation in OFDMA Wireless Networks’’, International Journal of Computer Applications , 3rd National Conference on Future Computing ,February 2014.
[5] Xiuting Tao, Guoqiang Li, Daniel Sun, and Hongming Cai, “A game-theoretic model and analysis of data exchange protocols for Internet of Things in clouds,” Future Generation Computer Systems, 2016.
[6] Xin Xu, Huiqun Yu, Xin Pei, Xin Xu, Huiqun Yu, and Xin Pei, “A novel resource scheduling approach in container based clouds,” Proceedings - 17th IEEE International Conference on Computational Science and Engineering, CSE 2014, FCST 2014, 2015.
[7] Sara RIAHI, Azzeddine RIAHI , ’’ Applications of game theory to optimization of resource allocation of wireless systems: algorithmic solution ’’ ,International Journal of Online Engineering (iJOE), ISSN: 2626-8493, PP. 5-25,Vol 14, No 12 (2018).
[8] Lei Xu, Jun Wang, A. Nallanathan, Yaping Li, "Resource Allocation Based on Double Auction for Cloud Computing System", *High Performance Computing and Communications;14th International Conference on Smart City* , 2016.
[9] Yi Gao , Wei Dong , Chun Chen , Xiaoyu Zhang , Jiajun Bu , Xue Liu, Accurate Per-Packet Delay Tomography in Wireless Ad Hoc Networks, IEEE/ACM Transactions on Networking , v.25 n.1, February 2017
[10] Sara Riahi, Azzeddine Riahi,’’ Game theory for resource sharing in large distributed systems‘’, International Journal of Electrical and Computer Engineering (IJECE), Vol. 9, No. 2, April 2019, ISSN: 2088-8708.
[11] I. Kim, H. L. Lee, B. Kim, and Y. H. Lee, ”On the Use of Linear Programming for Dynamic Subchannel and Bit Allocation in Multiuser OFDM,” in Proc. IEEE Global Communications Conf., vol. 6, 2001.
[12] C. Ng and C. Sung, “Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems,” IEEE Trans. Wireless Comm., vol. 7, no. 5 Part 1, 2008.
[13] H. Lei, X. Zhang, and Y. Wang;, “Real-time traffic scheduling algorithm for MIMO-OFDMA systems,” IEEE ICC, Apr 2008.
[14] M. Sternad, T. Ottosson, A. Ahln, A. Svensson, "Attaining both coverage and high spectral efficiency with adaptive OFDM", *Proc. IEEE Veh. Technol. Conf.*, pp. 2486-2490, 2003-Oct.
[15] A.X. Han and I-Tai Lu. Optimizing beyond the carrier by carrier proportional fair scheduler. IEEE Sarnoff Symposium, pages 1–5, may 2011.
[16] Jianwei Huang, Vijay G. Subramanian, Rajeev Agrawal, and Randall A. Berry. Downlink scheduling and resource allocation for OFDM systems. IEEE Trans. on Wireless Communic., 8(1):288–296, 2009.
[17] G. Yu, Z. Zhang, Y. Chen, J. Shi, P. Qiu: 'A Novel Resource Allocation Algorithm for Real-time Services in Multiuser OFDM Systems', IEEE VTC Spring 2006
[18] Ameigeiras, P., Wang, Y., Navarro-Ortiz, J. et al. J Wireless Com Network (2012) 2012: 61.
[19] G. Barriac and J. Holtzman, “Introducing delay sensitivity into the proportional fair algorithm for CDMA downlink scheduling,” in Proc. Int. Symp. Spread-Spectrum Tech. Appl., 2002.
[20] G. Song and Y. Li, “Adaptive subcarrier and power allocation in OFDM based on maximizing utility,” in Proc. IEEE Veh. Technol. Conf., Apr. 2003.
[21] Sara RIAHI, Ali ELHORE, ’’ Estimation of QoS in 4th generation wireless systems’’, The International Conference on Learning and Optimization Algorithms: Theory and Applications (LOPAL'2018), 2-5 May 2018, Rabat -Morocco.
[22] L. C. Wang and W. J. Lin, “Throughput and fairness enhancement for OFDMA broadband wireless access systems using the maximum C/I scheduling,” IEEE 60th Vehicular Technology Conference, September 2004.
[23] Sara Riahi, Azzeddine Riahi,’’ Optimal performance of the opportunistic scheduling in new generation mobile systems ‘’, ICSDE'18 Proceedings of the 2nd International Conference on Smart Digital Environment, ISBN: 978-1-4503-6507-9 ,October 2018.
[24] Y. L. M. Tao and F. Zhang, “Resource allocation for delay differentiated traffic in multiuser ofdm systems,” in IEEE Transactions on Wireless Communications, 2008
[25] Holtzman, J. M. Asymptotic analysis of proportional fair algorithm. In *Proceedings of the 12th IEEE international symposium on personal, indoor and mobile radio communications*, San Diego, California, 2001 .

Comparative study of performing features applied in CNN architectures

Sara LAROUI[1], Hicham OMARA[2], Mohamed LAZAAR[3], Oussama MAHBOUB[4]

{ saralaroui0@gmail.com, Hichamomara@gmail.com, lazaarmd@gmail.com,
mahbouboussama@gmail.com }

[1,2,4]Abdelmalek Essaadi University, Tetuan – Morocco

[3]ENSIAS, Mohammed V University, Rabat, Morocco

Abstract. Convolutional neural network (CNN), a class of artificial neural networks that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including Medical image analysis. CNN is designed to automatically and adaptively learn spatial hierarchies of features through back-propagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. This paper presents an approach based on CNN for the classification of brain tumors, based on several characteristics that will be extracted automatically and some performing features that will be used in our CNN, This proposed approach provides efficient results at the level of automatic classification than the other usual methods.

Keywords: Deep networks, classification, convolutional neural network, brain tumor, Medical imaging.

1 Introduction

With advancements in computer vision (multimedia, computer-assisted editing, computer graphics) CNNs are finding their place in performing better, especially in medical imaging field. This progress is due to considerable work in this area and the availability of international image databases. CNN's applica tions in medical image analysis date back to the 1990s, when used in a variety of applications ranging from data analysis to decision systems [1]. Actually, by using CNNs the doctors are able to make more accurate and timely decision about patient's disease, its stage, and diagnosis.

We are interested in this article to problem of brain classification from MRI (Magnetic Resonance Imaging) [2]. There are many techniques available for diagnosis of brain tumor from the brain tissues, detection of brain tumor such as conventional radiology, ultrasonography, magnetic resonance imaging, computerized tomography and etc., but the process of diagnosing a number of CT-scan images manually becomes tiresome and also susceptible to error. Therefore, computer aided systems are used to assist the physicians as a second option to reduce the mistakes and errors, this raise the need of the automated com- puterized system, such as CNN [3]–[6]. In literature, there are many applications of CNN in the medical field and has played an important role for automated detection of cancerous cells from mammographic images, Computed topography (CT) and Magnetic resonance (MR) imaging are the most widely used radiographic techniques in diagnosis, clinical studies and treatment planning.

Many researchers has been applied CNN in medical image analysis to detect a brain tumor, such as: Verdhana et al. that were proposed in [3] a low-power architecture for edge detection to detect the bio- medical

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284238

images are presented. They employed two major edge detection algorithms that are the Sobel and Prewitt algorithms using CMOS logic at 180 nm technology. The design is applied for various CT scan images of lungs, MR images of brain, X-ray images, etc. Another example, we see the work of Shuchao Pang et al. in [7]; they proposed a computational model for image classification based on end-to-end classifier using domain transferred deep convolutional neural networks or CNN. The proposed domain transferred deep CNN (DT-DCNN) showed significant increase in the accuracy while classifying the medical images with high precision. The new thought of the authors for simple data augmentation based on classification has improved the performance of this CNN based hybridized model. Another work has been proposed for classification of echocardiography by Xiaohong W.Gao et al. in [8]. They have pro- posed one hybridized network architecture with two CNNs along space and time directions both in spatial and temporal information fused to obtain classification scores from both networks. This spatial CNN work upon the normalized images to learn spatial information and pre-processing of the all the images are done in temporal CNN before training/learning starts and finally, to get the velocity and acceleration of images, the optical flow has been applied two times. Finally, we cite the proposition of Wei Shen et al. in [4], that they considered Multi-crop CNN (MC-CNN) to automatically extract the suspicious information based on an n multi-crop pooling strategy which crops different regions from convolutional feature maps without using segmentation to produce multi-sale features and applied on computationally effective single net- work. The goal of this work was to discover a set of discriminative features from hierarchical neural net- work to capture the suspicious information.

The main challenge addressed by the authors is to generate an image space including both the healthy tissues and specific suspicious nodules at different scales by ex- ploring the capabilities of deep learning architecture in an integrated fashion named as MC-CNN.

In this paper, we propose a method based on automatic two steps processing:

- Extraction of characteristics for the brain tumor.
- Choice of the best parameter that makes the neuron network performs.
- Classification of medical images containing brain tumors according to the characteristics extracted previously.

The rest of the paper has been organized as follows; section 2 and 3 describes the architecture and functionality of deep CNN. Section 4 describes some evaluation metrics for medical image analysis system that will be used in our experimentations, and a discussion of result obtained when applied the proposed model to classify brain tumors from DICOM images. Finally, we conclude the paper.

2 Convolutional neural networks (CNN)

A convolutional neural network (CNN) is a type of artificial neural networks, proposed by Yann Le Cun in 1988, most commonly applied to analyzing visual imagery[1]. Their multistage architectures are inspired from the science of visual cortex. One of the most popular and efficient uses of this architecture is image classification. The algorithm started by identifying low level features and then learns to recognize and combine these features to learn more complicated patterns. These different levels of features come from different layers of the network. And each layer has specific number of neurons and presented in 3 dimensions: height, width, depth.

The first part of a CNN is the convoluted part. It functions as a feature extractor of images. In this part, an image is passed through a succession of filters, or convolution kernels, creating new images called convolution maps. Some intermediate filters reduce the resolution of the image by a local maximum oper- ation.

A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalized layers. The

description of each element is detailed in the rest of this part.

2.1 Convolution layer

A convolution layer is a fundamental component of the CNN architecture that used for feature ex- traction. Its purpose is to locate the presence of a set of features in the images received as input. It is com- posed of a stack of mathematical operations, such as convolution, a specialized type of linear operation.

In the case of digital images (whose pixel values are stored in a two-dimensional grid) a convolution- al filtering in the form of a small grid (called kernel) is applied at each image position to extract features. This procedure is repeated applying multiple kernels to form an arbitrary number of feature maps, which represent different characteristics of the input. The size of the kernels, number of kernels, padding, and stride are hyper-parameters that need to be set before the training process starts.

The individual grid outputs from the kernels are called feature maps (or activation maps). After each convolutional layer is applied, an activation function is applied to the 3-dimensional output. Typically, the activation function that is applied is called ReLU (or Rectified Linear Unit Layer) [8], [9]. ReLU gets rid of negative values, by setting any negative pixel values to 0. It is defined as follows:

$$ReLU(x) = max(x, 0)$$

2.2 Pooling layer

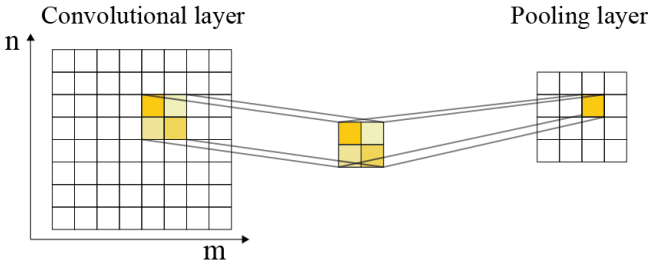


Fig.1. Convolution process

A pooling layer reduces the in-plane dimensionality of the feature maps in order to keep only the most relevant pixels. This process has a form of non-linear down-sampling. His main roles are to reduce progressively the spatial size of the representation, to reduce the number of parameters and amount of computation in the network, and hence to also control over fitting. It is of note that the filter size, stride, and padding are hyper parameters in pooling operations. The most common form is a pooling layer with filters of size 2x2 applied with a stride of 2 down samples at every depth slice in the input by 2 along both width and height, discarding 75% of the activations (as described in Fig. 1). In this case, every max op- eration is over 4 numbers. The depth dimension remains unchanged [10].

2.3 Max-pooling layer

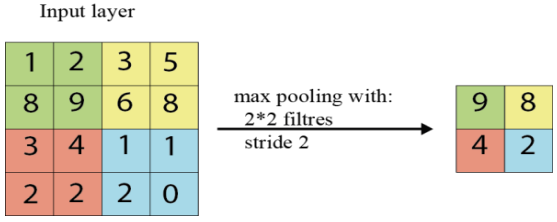


Fig. 2. A diagram showing a max pooling process.

The most popular form of pooling operation is max pooling (or down-sampling layer), which extracts patches from the input feature maps, outputs the maximum value in each patch, and discards all the other values. A max pooling with a filter of size 2 × 2 with a stride of 2 is commonly used in practice (Fig 2). This makes it possible to perform downscaling the dimension (height and width) of feature maps by a factor of 2, through pooling, without losing important information, however the depth dimension of fea- ture maps remains unchanged. Another example, in literature, consists to use the average pooling, which uses the average value from each of a cluster of neurons at the prior layer [5].

3 CNN training

The first step, in creating CNN, is to fix the architecture by fixing the number of chosen layers, their sizes and matrix operations that connect them. The training consists then, of optimizing the network's coefficients to minimize the output classification error. Similar to multilayer perceptron, CNNs are trained with the back-propagation algorithm by minimizing the following cost function with respect to the unknown weights W :

$$\zeta = -\frac{1}{|X|}\sum_{i}^{|X|}\ln\left(p\left(y_i \mid X_i\right)\right)$$

where $|X|$ denotes the number of training images, X_i denotes the ith training image with the corresponding label y_i, and $p\left(y_i \mid X_i\right)$ denotes the probability by which X_i is correctly classified. Stochastic gradient descent is commonly used for minimizing this cost function, where the cost over the entire training set is approximated with the cost over mini-batches of data. If $W_j(\mathrm{t})$ denotes the weights in j^{th} convolutional layer at iteration t, and L^ denotes the cost over a mini-batch of size N, then the updated weights in the next iteration is computed as follows:

$$\gamma^t = \gamma^{\left|\frac{tN}{|X|}\right|}$$

$$V_j(\mathrm{t+1}) = \mu V_j(\mathrm{t}) - \gamma^t \alpha_j \frac{d\zeta}{dW_j}$$

$$W_j(\mathrm{t+1}) = W_j(\mathrm{t}) + V_j(t+1)$$

where α_j is the learning rate of the j^{th} layer, μ is the momentum that indicates the contribution of the previous weight update in the current iteration, and γ is the scheduling rate that decreases learning rate α at the end of each epoch.

The iterative weight update begins with a set of randomly initialized weights. Specifically, before the commencement of the training phase, weights in each convolutional layer of a CNN are initialized by values randomly sampled from a normal distribution with a zero mean and small standard deviation. However, considering the large number of weights in a CNN and the limited availability of labelled data, the iterative weight update, starting with a random weight initialization, may lead to an undesirable local minimum for the cost function. Alternatively, the weights of the convolutional layers can be initialized with the weights of a pre-trained CNN with the same architecture. The pre-trained net is generated with a massive set of labelled data from a different application. Training a CNN from a set of pre-trained weights is called *fine tuning* and has been used successfully in several applications.

3.1 Proposed approach

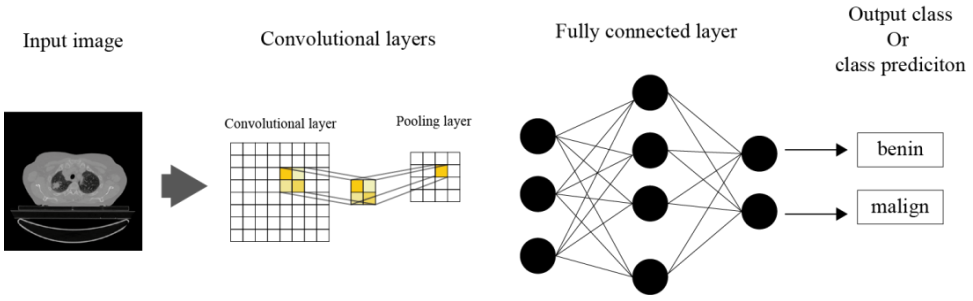


Fig.3. Global Architecture of the CNN used for medical image classification.

Fig.3 shows CNN architecture for classification of medical images. The network has convolutional, max pooling and fully connected layers. Each convolutional layer generates a feature map of different size and the pooling layers reduce the size of feature maps to be transferred to the following layers. The fully connected layers at the output produce the required class prediction. The number of parameters required to define a network depends upon the number of layers, neurons in each layer, the connection between neu- rons. The training phase of the network makes sure that the best possible weights are learned, that would give high performance for the problem at hand. The advancement in deep learning methods and computa- tional resources has inspired medical imaging researchers to incorporate deep learning in medical image analysis.

Some recent studies have shown that deep learning algorithms are successfully used for medical im- aging applications. A multi scale CNN based approach for automatic segmentation of MR Images is presented that classifies voxel into brain tissue classes. A total of five databases are used having T1 and T2 weighted images at a different postmenstrual age (PMA). A tri-planar CNN is used for segmentation of tibial cartilage in knee MRI images. Three 2D CNNs are integrated together for classification of a voxel in 3D images. A segmentation of brain tissue is presented through a CNN using multimodal MRI dataset by training the network on three patches that are extracted from the images. A deep learning based approach has been presented, in which the network uses a convolutional layer in place of fully connected layer to speed up the segmentation process. A cascaded architecture has been utilized, which concatenates the output of first network with the input of succeeding network. The network presented uses small kernels to classify pixels in MR image. The use of small kernels decreases network parameters, allowing to build deeper networks, without worrying about the dangers of over- fitting. Data augmentation and intensity normalization have been performed in pre-processing step to facilitate training process.

The approach proposed in this article avoids the design and the manual extraction of features. This approach is more typically faced with difficult problem such as:

- The design of features robust and easy to calculate.
- The evaluation of these characteristics and their relevance for the separation of classes,
- And the more general case of the selection of relevant characteristics.

Our approach based on convolutional neural networks avoids the delicate step related to the charac- teristics, and using multiple neural networks convolutional for the same problematic. The idea behind this is that each network will learn different characteristicsof the same input.

In order to have an efficient classification, we will create 3 CNN acting at different resolutions, choosing the most reliable activation function for this we use a medical images size to full resolution is from 512*512, the latter is successively divided by a factor $\sqrt{4}$ and $\sqrt{4^2}$ according to each axis so as to give images of sizes 256*256 and 64*64, respectively.

To learn features for our medical image classification task, we apply ConvNets combined with fine tuning technique. We used the pre-trained convolutional neural network AlexNet[11], [12] which is trained for 1000 possible categories on the large dataset ImageNet in the Large Scale Visual Recognition

Challenge (ILSVRC-2012), containing over then 1.2 million images. Results were remarkable and achieved a top- 5 error of 15.3%.

About our proposed algorithm, each image (or feature map) has dimensions W × H × D, where W is its width in pixels, H is its height in pixels and D is the number of channels (1 for a black and white im- age, 3 for an image in colors): The convolution layer has four hyper parameters:

- The number of filters K
- The size F of the filters: each filter has dimensions F × F × D pixels.
- The step S with which the window corresponding to the filter is dragged on the image. The zero-padding P is added to the input image of the layer a black outline of thickness P pix- els. Without this outline, the output dimensions are smaller. Thus, the more one stack of convolution layers with P = 0, the more the input image of the network shrinks. So we lose a lot of information quickly, which makes the task of extracting features difficult

And the pooling layer has only two hyper parameters:

- The size F of the cells: the image is divided into square cells of size F × F pixels
- Step S: The cells are separated from each other by S pixels

Our proposed algorithm is the gradient back propagation algorithm follows the Widrow-Hoff ap- proach. Define an error notion on an example and then calculate the contribution to this error of each of the synaptic weights. It is this second step that is not obvious, In order to be able to apply the gra- dient method, one needs to calculate derivatives.

4 Experiment

Medical imaging includes those processes that provide visual information of the human body. The purpose of medical imaging is to aid radiologists and clinicians to make the diagnostic and treatment process more efficient. Medical imaging is a predominant part of diagnosis and treatment of diseases and represents different imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasound, X-ray and hybrid modalities [7], [15]. These modali- ties play a vital role in the detection of anatomical and functional information about different body organs for diagnosis as well as research. Medical imaging is an essential aid in the modern healthcare and com- puter aided diagnosis (CADx) systems. Machine learning plays a vital role in CADx with its applications in medical image analysis, tissue or cancer detection and classification, image guided therapy, medical image annotation and medical image retrieval.

4.1 Dataset

We used in this article brain tumor image that working with DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format.

Our work is about some of the brain MR images containing tumor taken for testing our proposed al- gorithm are shown. These images are taken from the cancer archive imaging. Our brain tumor dataset contains 160 contrast-enhanced images of 512*512 with three kinds of brain tumor of the type .dcm.

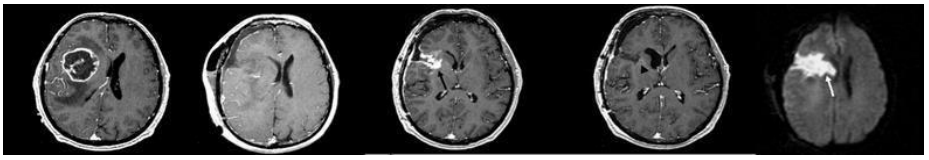


Fig. 5. Brain MR images containing tumor

4.2 Evaluation Metrics for Medical Image Analysis System

To evaluate proposed approach, we are used different key performance measures such as accuracy, precision, recall, sensitivity and specificity [13], [14]. Mathematically, these measures are calculated as,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + TN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Fscore = \frac{2 \times precision \times recall}{precision + recall}$$

where TP (True Positives) is correctly classified positive cases. TN (True Negative) is correctly classified negative cases. FP (False Positives) is incorrectly classified negative cases. FN (False Negative) is incor- rectly classified positive cases.

4.3 Result and discussion

In this section, we present experimental results about the effect of one of the hyper parameters of a deep neural network in our model that we will use it in our model proposed model.

The method we propose for our problematic is one of the best methods of deep learning; it is the convolutional neuron networks "CNN" have shown their effectiveness as a learning technique for the classification of data. They are indeed able to approximate complex non-linear functions in order to pro- cess large data. We are going to work on the classification of medical images, two approaches are merged within deep neural networks that realize both the automatic extraction of the characteristics and their classification, moreover, an approach is proposed using several networks depths at different resolutions, and whose outputs are merged, and to improve the performance of this model.

Our work is part of the analysis of medical images of MRI of the brain for the diagnosis of brain tu- mors (Glioma), the aim of our study is to provide physicians with a more precise automatic classification by

distinguishing the types of these brain tumors, which prove to be very important during an early diag- nosis of cancers.

We will implement most of the characteristics extracted from each medical image of the base, and then we build networks of neurons with different resolutions by reducing the image in order to obtain more details by using one of the performing parameter obtained in those below tables (Table 2 -3-4). The last layer, in the case of supervised learning, contains as many neurons as desired, and a softmax activa- tion function is used to obtain probabilities of belonging to each class.

The function is:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad pour\ tout\ j \in \{1,...,K\}$$

That is, the component j of the vector $\sigma(z)$ is equal to the exponential of the component j of the vector z divided by the sum of the exponentials of all the components of z.

The following terms are consistently employed throughout this article so as to avoid confusion. A parameter in this article stands for a variable that is automatically learned during the training process. A hyper parameter refers to a variable that needs to be set before the training process starts. A kernel refers to the sets of learnable parameters applied in convolution operations. A weight is generally used inter- changeably with parameter; however, we tried to employ this term when referring to a parameter outside of convolution layers

Table 1. A list of parameters and hyper parameters in a convolutional neural network (CNN)

	Parameters	Hyperparameters
Convolution layer	Kernels	Kernel size, number of kernels, stride, padding, activation function
Pooling layer	None	Pooling method, filter size, stride, padding
Fully connected layer	Weights	Number of weights, activation function
Others		Model architecture, optimizer, learning rate, loss function, mini-batch size, epochs, regularization, weight initializa- tion, dataset splitting

Note that a parameter is a variable that is automatically optimized during the training process and a hyper parameter is a variable that needs to be set beforehand.

Activation Function:

The outputs of a linear operation such as convolution are then passed through a nonlinear activa- tion function. Although smooth nonlinear functions, such as sigmoid or hyperbolic tangent (tanh) func- tion, were used previously because they are mathematical representations of a biological neuron behavior, the most common nonlinear activation function used presently is the rectified linear unit (ReLU), which simply computes the function: f(x)=max(0,x).

Our problem type is classification, this why we implemented this experiment about extracting the most efficient hyper parameters as to use from our approach as a activation function and learning rate which aims to avoid over-learning and minimize the error rate, In this table below, we used the learning rate to 0.1 and fixing the epoch at 400 to get this results per number of hidden layers:

Table 2. The loss test according to the learning rate of the one hidden layer

Activation Function	Learning Rate	Test Loss	Training loss
Sigmoid	0,1	0.007	0.008
Tangh	0,1	0.003	0.003
ReLU	0,1	0.001	0.001

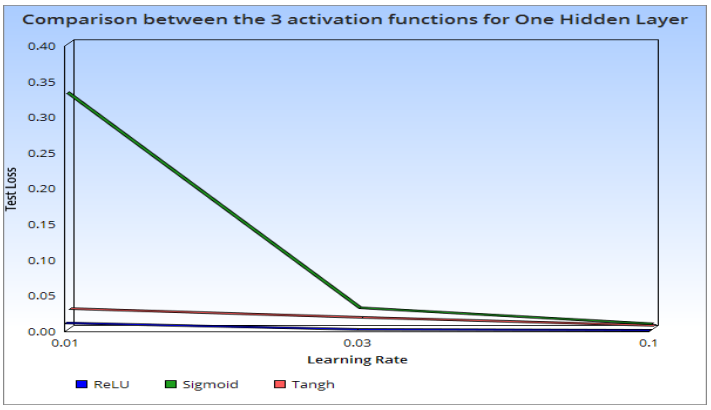


Fig. 6. Comparison between the 3 activation functions for one hidden layer

According to the table above, reproduces the results of the error test according to the learning rate in order to then know the most effective activation function when using a single masked layer, This **Fig.6** show more clearly the result of the table 2.

Now we will add the second hidden layer as you see in the table below:

Table 3. The loss test according to the learning rate of the two hidden layers

Activation Function	Learning Rate	Number of Neu- ron for The one hidden	Number of neu- rons for the two hidden	Test Loss	Training loss
Sigmoid	0,1	4	4	0.006	0.004
Tangh	0,1	4	4	0.000	0.000
ReLU	0,1	4	4	0.000	0.000

According to the table above, we added a second hidden layer and we obtain a more reduced error test than that of a single hidden layer on the table 2.

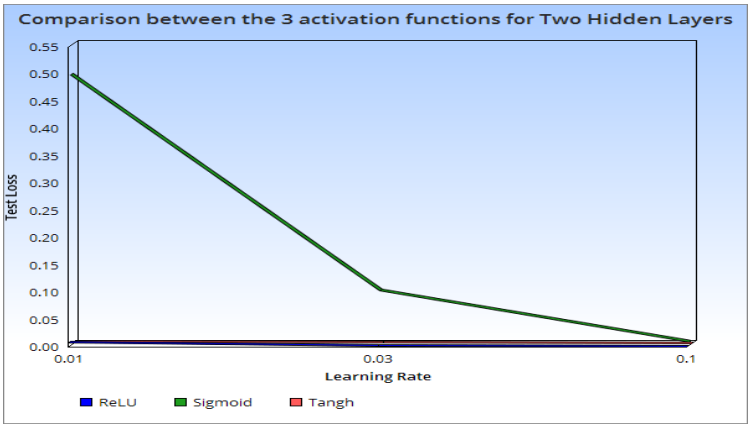


Fig. 7. Comparison between the 3 activation functions for two hidden layers

In order to obtain several precisions, we will add the third hidden layer as you see in the table below:

Table 4. The loss test according to the learning rate of the three hidden layers

Activation Function	Learning Rate	Number of Neu- ron for The one hidden	Number of neurons for the two hid- den	Number of neurons for the Three hidden	Test Loss	Training loss
Sigmoid	0,1	4	4	2	0.403	0.403
Tangh	0,1	4	4	2	0.000	0.000
ReLU	0,1	4	4	2	0.000	0.000

According to the table above, reproduces the results of the error test according to the learning rate in order to then know the most effective activation function when using a three masked layer is ReLU, This graph show more clearly the result of the table 4.

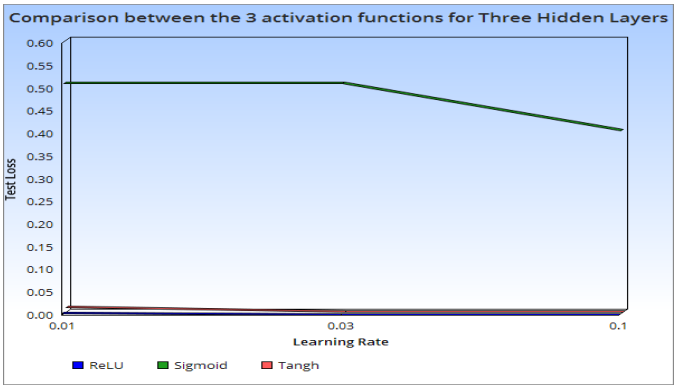


Fig. 8. Comparison between the 3 activation functions for three hidden layers

According to the results of Table 2,3 and 4, we find that the number of hidden layers play an im- portant

role in the reduction of the error rate in terms of activation function that allows us to see that ReLU function shows suitable results compared to the other Sigmoid and Tangh activation functions.

Number of parameters:

Now, when talking about neural networks, it is nearly always the case that the network has far more parameters than training samples.

Theoretically, a simple two-layer neural network with 2n+d2n+d parameters is capable of perfectly fitting any dataset of N samples of dimension D.

The obvious benefit of having many parameters is that you can represent much more complicated functions than with fewer parameters. The relationships that neural networks model are often very com- plicated ones and using a small network (adapting the size of the network to the size of the training set, i.e. making your data look big just by using a small model) can lead to the problem when your network is too simple and unable to represent the desired mapping (high bias).

On the other hand, if you have many parameters, the network is flexible enough to represent the desired mapping and you can always employ stronger regularization to prevent over fitting, so this below table represents the parameter count per layer used:

Table 5. Increasing in number of parameters depending on the hidden layer

Layer	Filter size, stride	Output W×H×N	Parameter Count per Layer
Input	-	512 × 512 × 1	400
Conv	20 × 20, 2	247 × 247 × 1	
ReLU	-	247 × 247 × 1	
Max-pool	2 × 2, 1	246 × 246 × 1	
Conv	10 × 10, 2	119 × 119 × 1	500
ReLU	-	119 × 119 × 1	
Max-pool	4 × 4, 1	116 × 116 × 1	
Conv	5 × 5, 5	24 × 24 × 1	525
ReLU	-	24 × 24 × 1	
Max-pool	2 × 2, 2	12 × 12× 1	
Fully Connected FC	-	12 × 12× 1	21 549

According to this table 5, we notice an increase in the number of parameters in terms of the convolutional layers which layer allows a deeper network of neurons and returns better results.

The receptive field:

Another parameter and the most important concepts in Convolutional Neural Networks (CNNs) is the receptive field is defined as the region in the input space that a particular CNN's feature is looking at (i.e. be

affected by). A receptive field of a feature can be described by its center location and its size. However, not all pixels in a receptive field is equally important to its corresponding CNN's feature. Within a receptive field, the closer a pixel to the center of the field, the more it contributes to the calculation of the output feature. Which means that a feature does not only look at a particular region (i.e. its receptive field) in the input image, but also focus exponentially more to the middle of that region?

We can compute the spatial size of the output volume as a function of the input volume size (W), the receptive field size of the Conv Layer neurons (F), the stride with which they are applied (S), and the amount of zero padding used (P) on the border. You can convince yourself that the correct formula for calculating how many neurons "fit" is given by:

$$(W-F+2P) / S+1$$

- Accepts a volume of size W1×H1×D1

- Requires four hyperparameters:

Number of filters K, Their spatial extent F, The stride S,

The amount of zero padding P.

For now, in this table below, we focus on calculating a particular receiver field based on the number of hidden layers:

Table 6. The performing activation function used

Layer #	Kernel size	Stride	Dilation	Padding	Input size	Output size	Receptive Field
1	20	2	1	0	512	247	20
2	10	2	1	0	247	119	38
3	5	5	1	0	119	23	54

As we can see from on - Table 6- The result is very clear with this graph :

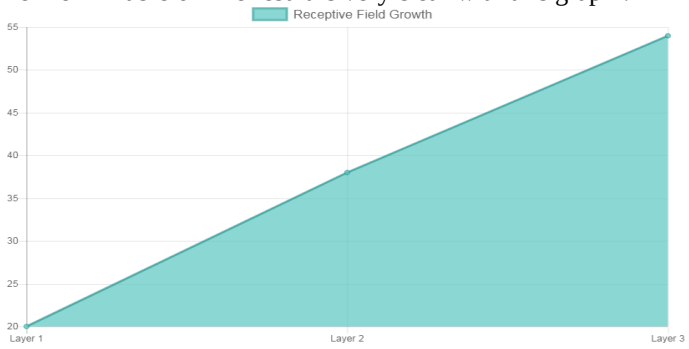


Fig. 9. Graph of Recetive Field Gtowth according to layers

Fig. 9, we can see the graph above this receptive field -in a convolutional neural network refers to the part of the image that is visible to one filter at a time- increases linearly as we stack more convolution- al layers or increases exponentially when we stack convolutions, When doing the forward pass, we can see that the central receptive field pixels can propagate their information to the output using many different paths, as they are part of multiple output unit's calculations, that help us to get an output more clearly to classify.

What is also very interesting is that the effective receptive field has a very close relationship with the foveal vision of the human eye, which produces the sharp central vision, effect of the high-density region of cone cells.

5 Conclusions

We presented in this article, the important parameters that we will need in our model of neural net- work, and we showed the results of each parameter so that we can then choose them and show the perfor- mance of our future model, and this results obtained show that all these parameters that we appreciated above, are important and robust parameters that allow us later to optimize our model with efficient results.

References

[1] A. Bhandare, M. Bhide, P. Gokhale, and R. Chandavarkar, "Applications of Convolutional Neural Networks,"
Int. J. Comput. Sci. Inf. Technol. IJCSIT, vol. 7, no. 5, pp. 2206–2215, 2016.

[2] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[3] M. Vardhana, N. Arunkumar, S. Lasrado, E. Abdulhay, and G. Ramirez-Gonzalez, "Convolutional neural net- work for bio-medical image segmentation with hardware acceleration," *Cogn. Syst. Res.*, vol. 50, pp. 10–14, Aug. 2018.

[4] W. Shen *et al.*, "Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classifi- cation," *Pattern Recognit.*, vol. 61, pp. 663–673, Jan.2017.

[5] S. Mittal, "A survey of FPGA-based accelerators for convolutional neural networks," *Neural Comput. Appl.*, Oct. 2018.

[6] X. W. Gao, R. Hui, and Z. Tian, "Classification of CT brain images based on deep learning networks," *Comput. Methods Programs Biomed.*, vol. 138, pp. 49–56, Jan. 2017.

[7] S.-C. Pang, Z. Yu, and M. A. Orgun, "A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images," *Comput. Methods Programs Biomed.*, vol. 140, pp. 283–293, 2017.

[8] X. Gao, W. Li, M. Loomes, and L. Wang, "A Fused Deep Learning Architecture for Viewpoint Classification of Echocardiography," *Inf Fusion*, vol. 36, no. C, pp. 103–113, Jul. 2017.

[9] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, USA, 2010, pp. 807–814.

[10] D. Scherer, A. Müller, and S. Behnke, "Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition," in *Artificial Neural Networks – ICANN 2010*, 2010, pp. 92–101.

[11] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Net- works," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, USA, 2012, pp. 1097–1105.

[13] O. Hicham, L. Mohamed, and T. Youness, "Improved self-organizing maps based on distance travelled by neu- rons," *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 5, 2018.

[14] O. Hicham, M. Lazaar, et T. Youness, « Self-Organizing Maps and Principal Component Analysis to Improve Classification Accuracy », Res. J. Appl. Sci. Eng. Technol., vol. 15, no 5, p. 190-196, mai 2018.

[15] H. Omara, M. Lazaar, et Y. Tabii, « Effect of Feature Selection on Gene Expression Datasets Classification Accuracy », Int. J. Electr. Comput. Eng. IJECE, vol. 8, no 5, p. 3194-3203, oct. 2018.

A comparative study of CHN-MNC, GA and PSO for solving constraints satisfaction problems

Bouhouch adil[1], Loqman Chakir[2], Bennis Hamid[3] and El Qadi Abderrahim[4]
{ bouhouch.adil@gmail.com[1]}

Department of Computer Science, Faculty Sciences, Chouaib Doukkali University, El Jadida, Morocco[1]
Department of informatics, Faculty Sciences, Dhar Mehraz, Sidi Mohammed Ben Abdellah Fez, Morocco [2]
Team TIM, High School of Technology - Moulay Ismail University, Meknes, Morocco [3]
High School of Technology - Mohammed V University of Rabat Morocco[4]

Abstract. Our approach CHN-MNC, based Continuous Hopfield neural network and Min-Conflict heuristic), have proved that is more efficient than using CHN alone to solve Constraints Satisfaction Problem (CSP). In This paper we study the performance of CHN-MNC by comparing it robustness with two evolutionary algorithms. We choose a Genetic Algorithm and Swarm optimisation to performers this study. Some numerical experiments are done over a variety of problems to verify the efficiency and fast convergence of our approach. abstract needs to summarize the content of the paper.

Keywords: CSP, Metaheuristics, GA, PSO, Min-Conflict Heuristic.

1 Introduction

A large number of real world problems can be represented as constraint satisfaction problems (CSP). For example, scheduling, qualitative and symbolic reasoning, diagnosis, temporal andspatial planning, hardware design and verification, real-time systems and robot planning. A CSP problem can be considered as of a finite set of variables, each one has a finite domain of values and a set of constraints. A solution to a CSP is a complete assignment of variables which satisfy all constraints. But, finding this solution on a finite domain is a NP-complete problem requiring a combination of heuristics and combinatory search methods in order to be solved in a reasonable time [1]. In general, approaches to solve CSPs can be classified in two main categories: exact approaches and heuristic ones. As for exact approaches, most of them have the backtracking algorithm (BT) as a main algorithm for solving constraint satisfaction problems. As far as heuristic approaches are concerned, we find a very different approach has taken investigating neural approach works to solve CSP[2]–[4]. In this neural network approach, the constraints are encoded in the network topology, biases strengths connection, and problem is formulated as quadratic cost function which is a Lyapunov function. Particularly, a very different approach has been taken investigating Hopfield network with continuous times for solving CSPs, as we can see in [5]–[8] authors propose mapping CSP to a quadratic model and giving appropriate parameters setting to reach an equilibrium point of CHN. In the practice, there are two important problem with approaches based on conventional neural network architectures, The first is that network is

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284084

partially mitigate the problem of getting stuck in local optimum, the second is due to dynamic Hopfield network which continuously explore the search space and will not stabilize at border 0 or 1, if the same case appear, we get low solution quality or an incomplete assignment of variables. In order to improve solution or to complete invalid solution, we propose to use Min-Conflict heuristic [8] after that CHN reached stabilisation. In this paper we extend our previous study by comparing CHN-MNC with Genetic algorithm [9] and Swarm optimisation [10].

2 CHN-MNC Algorithm

CHN-MNC is a collaborative hybrid algorithm, which benefit of the fast convergence of Hopfield neural network and the amelioration of the solution quality by the local search [11]

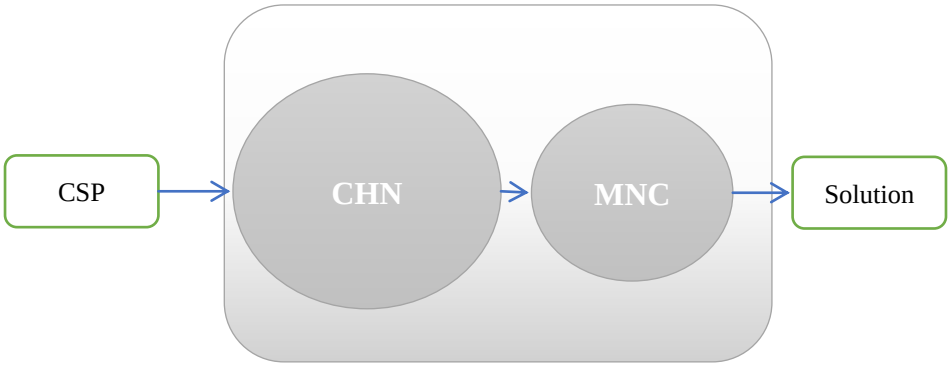


Fig. 1. Architecture CHN-MNC

2.1 CHN solver

A large number of real problems such as artificial intelligence, scheduling, assignment problem can be formulated as a Constraint Satisfaction Problem. Solving a CSP requires to finding an assignment of all variables problem under constraints restriction. The CSP can be formulated as three sets [12]:

- Set of N variables $X=\{X_i ; 1 \leq i \leq N \}$.
- Set of N variables domains: $D =\{ D_i ; 1 \leq i \leq di\}$ where each Di contains set of di range values for X_i.
- Set of M constraints: $C =\{ C_i ; 1 \leq i \leq M\}$.

Each constraint C_i associates an ordered variables subset which is called the scope of C_i. The arity of a constraint is the number of involved variables. We can easily reformulate CSP as a Quadratic Problem (QP), by introducing a binary variable x_{ik} for each CSP variable x_i, where k varies over the range of xi, given as follows:

$$x_{ik} = \begin{cases} 1, & if\ variable\ i\ takes\ value\ k \\ 0, & otherwise \end{cases} \quad (1)$$

For each binary constraint C_{ij} , between the variables y_i and y_j , we associate a state function defined as:

$$S_{ij}(x) = \sum_{r=1}^{d_i}\sum_{s=1}^{d_j} x_{ir}x_{js}Q_{irjs} \quad (2)$$

Where $x = \{x_{ik}, i \in [1 \ldots N], k \in d_i\}$ a vector of QP solution and the quadratic terms Qirjs defined as:

$$Q_{irjs} = \begin{cases} 1 & if\ (r,s) \notin C_{ij} \\ 0 & otherwise \end{cases} \quad (3)$$

From all the equations defined in (2), which correspond to problem constraints, we deduce the objective function of its equivalent QP:

$$f(x) = \sum_{i=1}^{N}\sum_{r=1}^{d_i}\sum_{j=1}^{N}\sum_{s=1}^{d_j} x_{ir}x_{js}Q_{irjs} \quad (4)$$

Furthermore, some strict linear constraints equations must be satisfied by the solution: $\sum_{r=1}^{d_i} x_{ir} = 1$, for $i = 1..N$ which can be written also as $Ax = b$ (A is a $N \times M$ matrix and b is a M dimension vector fully initialized to 1). So, the model is given as follows:

$$(QP) \quad \begin{cases} Min \quad f(x) = x^T Qx \\ with \\ \qquad Ax = b \\ \qquad x \in \{0,1\}^n \end{cases} \quad (5)$$

Systematically, to solve the last Quadratic Optimization Problem with Hopfield model, we need to build an energy function such as the feasible solutions of the problem corresponding to the minimal of CHN energy function:

$$E(x) = \frac{\alpha}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{r=1}^{d_i}\sum_{s=1}^{d_j} xirx_{js}Q_{ijrs} + \beta\sum_{i=1}^{N}\sum_{r=1}^{d_i} x_{ir} + \gamma\sum_{i=1}^{N}\sum_{r=1}^{d_i} x_{ir}(1 - x_{ir}) \quad (6)$$

2.2 MNC local search

There are many methods which combine two or more no exacts approaches to solve a given optimization problem [13]–[16]. In the same direction we introduce a hybrid approach based CHN and MNC. The MNC algorithm [11] is a very simple and fast local repairing method to resolve CSPs, which aims at assigning all the variables randomly. Next, it iteratively selects one variable from the set of the variables with conflicts which violates one or more constraints of the CSP. Then, it assigns a value to the selected variable, so that it can minimize the number of conflicts. MNC has demonstrated to be able to solve the queens problem in minutes [17]. MNC is widely used to construct hybrid algorithms with other optimizations [14][15][18][19]. In this way, the basic idea of our proposed approach is to use MNC to improve the solution reached by CHN. This will be done in tw o steps (see Figure 2). First, MNC visits all assigned variables; for each one, we apply Min-Conflict directly to the neural network structure, then, it returns the best assignment for the current variable (see Figure 3), the decision will be taken by the sum of all activated neurons weight. Second, we propagate this assignment to other set variables not yet assigned iteratively by applying the MNC heuristic to guarantee as much consistency as possible. The diagram of our proposed algorithm is described by (Figure 4).

```
Function CHNMNC ( CSP : Problem ) :
    V_a = CHN(CSPs)
    For (each cluster x_j ∈ V_a) do
        V_a = V_a \ (x_j, a) {a is the current affected value to x_j }
        a = Min-Conf(x_j, V_a) { new value assigned to x_j }
        V_a = V_a ⋃ (x_j, a)
    end For
    {propagate current sub assignment to other
    variables which not assigned }
    For (each cluster x_j ∉ V_a) do
        a' = Min − Conf(x_j, V_a)
        V_a = V_a ⋃ (x_j, a')
    end For
    return V_a
End
```

Fig. 2. Main function which improve solution by Min-conflict algorithm

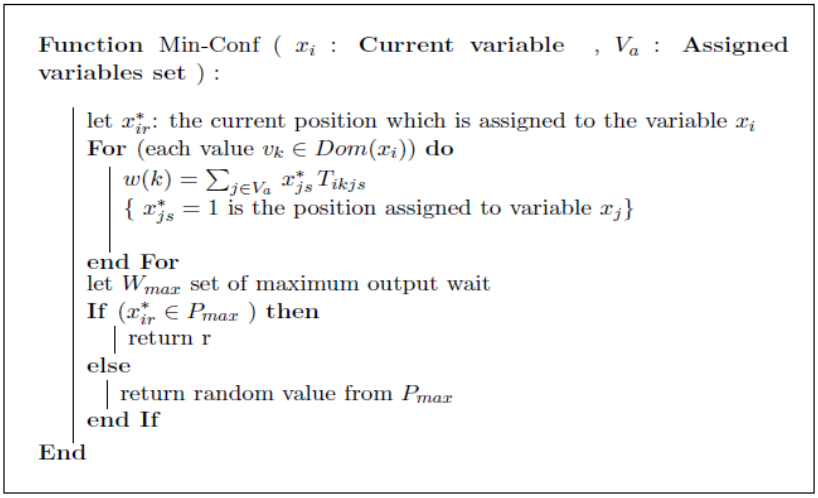


Fig. 3. Selected the most coherent neuron of current cluster with other variables clusters already affected

3 Evolutionary Algorithm

The main problem with investing metaheuristic to resolve optimization problem is the probability of becoming trapped at locally optimal points. So, the combination and the cooperation between these approaches may improve the solution quality. Many hybrids approaches have been developed, for our case we shoos two intended, approaches, for solving CSP. The first, one is a combination with genetic algorithm and heuristics reproduction operators. The second is an adapted Swarm optimization to combinatory probes.

2.1 GA

We focus on hybrids algorithms which hybrids genetic algorithm with local search. Most all approaches use an integrative hybrid [14]–[16]. So, the Genetic algorithm is executed to the exploration of solution space and local search is used to improve the quality of each intermediate generation of population individual. Others make change on reproduction to adapted GA to a specific problem. In this work we opted for comparison with approach in [20]. The last method use a hybrid search approach that combines the genetic algorithm with the min-conflicts hill-climbing (MCHC).

2.1 Swarm

Collective intelligence refers to the capabilities of a resulting community of interactions multiple members (or agents) of the community. Agents can thus accomplish complex tasks through a fundamental mechanism called synergy. Under special conditions, the synergy created by the collaboration between individuals brings out the possibilities of representation, creation and learning superior to those of isolated individuals. The PSO parameters are specific to the problem to be solved and must be determined in each problem. The first work

that employed the PSO for the resolution of the binary CSP problem was by Schoofs in [21]. The CSP problem is a combinatorial problem, where the concept of speed must be redefined. Thus, we present the redefinition of the usual operators of calculation (sum and product) between the positions and the speeds.

3 Numerical result

For showing the practical interest of our approach, we compares it its performance with others metaheuristics approaches over problems of different natures (random, academic and real-world problems). So we choose two algorithms elaborated specially to solve CSP [20], [21]. To perform a competitive comparison we do not settle for the authors original settings, thus we determined them empirically: for GA the population was 200 mutation rate equal to 5% and crossing rate equal to 72%, as for PSO [21] we choose $\varphi_1 = \varphi_2 = 1$ and population size fixed at 100. We run also 500 times each one.

We run some preliminary experiments on the randomly generated problems and we study the performances evolutions of the cited above approaches. We use a random generator based extended model B as it is described in [22]–[24]. This extended model which is called Model RB is able to generate a hard instance with forced mode which allow that instance have a solution. For each tightness value we generate 100 instances. From Figure 4 we can learn that our approach gives a good solution quality whatever the difficulty (Tightness) of the problem.

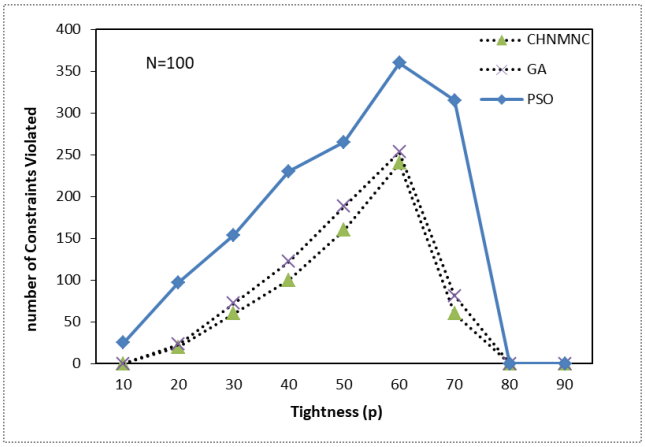


Fig. 4. Performance of GA, PSO and CHN-MNC random gendered instances.

Furthermore, Table 1 shows the results of CHN-MNC, PSO and Genetic Algorithm over 500 independent runs. The simulation is done on selected reals and academics instances of the benchmark [25]. According to the obtained means values our approach better and the standard deviation is more closed to the means values.

Table 1. Means values and standard deviation performed by CHN-MNC, GA, and PSO over selected instances.

Name of			CHN-MNC		GA		PSO	
Instance	V	C	Mean	SD	Mean	SD	Mean	SD
queens-10	10	45	1,01	0,16	2,00	1,02	13,50	5.5
queens-20	20	190,	2,40	0,32	4,00	1,06	39,20	3,11
queens-5-5-5	25	160	0,00	0,00	0,00	0,00	19,00	2,10
frb30-15-5-mgd	30	210	10,00	1,07	14,00	3.46	40,10	4,08
geom-30a-5	30	81	1,05	0,11	2,00	2,13	7,00	0,21
geom-30a-6	30	81	0,00	0,00	0,00	0,00	4,00	0,20
queens-30	30	435	4,00	0,94	6,00	1,05	83,00	7,39
frb40-19-3-mgd	40	308	14,00	0,93	21,00	2,14	53,00	3,56
geom-40-2	40	78	23,00	0,01	24,00	0,03	28,00	5,04
geom-40-6	40	78	0,00	0,00	0,00	0,21	4,01	1,06
myciel-5g-3	47	236	10,00	0,22	12,00	0,50	47,00	6,83
myciel-5g-4	47	236	5,00	0,41	8,00	0,95	29,00	3,58
myciel-5g-5	47	236	1,00	0,47	3,00	1,09	12,00	2,81
myciel-5g-6	47	236	0,00	0,13	1,00	0,31	12,50	1,51
driverlogw-01c	71	217	0,00	0,00	0,00	0,00	3,00	0,59
composed-25-10-20-5	105	620	13,10	1,93	14,00	6,75	56,00	7,20
dsjc-125-1-4	125	736	50,00	2,30	64,00	3,69	102,00	4,15
dsjc-125-1-5	125	736	19,00	1,80	29,00	1,83	85,00	4,04
Qw h-15-106-1	225	2324	20,00	1,60	23,00	3,68	66,60	5,11
qwh-15-106-4	225	2324	18,20	1,06	22,00	9,35	59,04	10,52
qwh-15-106-6	225	2324	22,00	2,31	23,00	5,31	68,00	8,82
driverlogw-04c	272	3876	3,00	2,32	5,00	2,33	11,00	3,75
driverlogw-02c	301	4055	3,00	1,17	5,00	1,89	9,00	2,01
qwh-20-166-0	400	5092	30,00	2,79	32,00	3,63	93,03	6,12
qwh-20-166-3	400	5092	29,00	1,88	31,00	2,63	89,00	4,28
qwh-20-166-6	400	5092	25,00	2,19	29,00	6,84	86,00	9,27
le-450-5a-3	450	5714	1173,00	5,69	1261,00	767,49	1566,00	17,03
le-450-5a-4	450	5714	712,00	1,89	745,00	8,95	1066,00	14,86
le-450-5a-5t	450	5714	441,00	2,29	470,00	5,26	874,00	13,72

3 Conclusion

In the last decade several areas have applied CSP model. So, many exact and heuristic methods were introduced. In this logic we focussed our contribution to develops a new hybrid approach based Hopfield neural network and local search. The main contribution in this work is the amelioration by the of the solution given by CHN Some numerical examples assess the effectiveness of the theoretical results are shown in this paper, and also the advantages of this new approach which improve considerably the solution quality and avoid network crush. Other studies are in progress to apply this approach to many problems such as timetabling and resource allocation.

References

[1] Carvalho, A. and Santos, C.: A Generator of Heavy-Tailed Search Trees, in Recent Developments in Modeling and Applications in Statistics, Springer. pp. 107–113 (2013)

[2] Tsang, E. P. K. and Wang, C. J.: A generic neural network approach for constraint satisfaction problems. in Neural network applications. Springer.pp. 12–22 (1992).

[3] Adorf , H. M.and Johnston, M. D. :A discrete stochastic neural network algorithm for constraint satisfaction problems. in Neural Networks. IJCNN. pp. 917–924 (1990)

[4] Nakano, T and Nagamatu, M. :Lagrange neural network for solving csp which includes linear inequality constraints, in Artificial Neural Networks: Formal Models and Their Applications--ICANN 2005. Springer. pp. 943–948 (2005)

[5] Bouhouch, A. Loqman, C and El Qadi, A. :Improving CHN by Min-Conflict algorithm to solve constraint satisfaction problems. Adv. Inf. Technol. Theory Appl, vol. 1, no. 1 (2016).

[6] Bouhouch, A. Loqman, C and El Qadi, A. :Scheduling meeting solved by neural network and Min-conflict heuristic, 4th IEEE International Colloquium on in Information Science and Technology (CiSt). pp. 773–778 (2016)

[7] Bouhouch, A. Loqman, C and El Qadi, A. :CHN and Min-Conflict Heuristic to Solve Scheduling Meeting Problems, in Bioinspired Heuristics for Optimization, Springer, pp. 171–184 (2019)

[8] Bouhouch, A. Loqman, C and El Qadi, A. :CHN and Min-Conflict Heuristic to Solve Scheduling Meeting Problems, Meta'2016 Marakech (2016)

[9] Loukakis, E. :A new backtracking algorithm for generating the family of maximal independent sets of a graph. Comput.Math.withAppl, vol. 9, no. 4, pp. 583–589 (1983).

[10] Kanoh, H. Matsumoto, M and Nishihara, S. :Genetic algorithms for constraint satisfaction problems. IEEE International Conference on in Systems, Man and Cybernetics, Intelligent Systems for the 21st Century, vol. 1, pp. 626–631 (1995)

[11] Minton, S. Philips, A. Johnston, M. D. Laird, and P. :Minimizing conflicts: A heuristic repair method for constraint-satisfaction and scheduling problems. J. Artif. Intell, Res., vol. 1, pp. 1–15, (1990)

[12] Montanari,U. :Networks of constraints: Fundamental properties and applications to picture processing. Inf. Sci. (Ny)., vol. 7, pp. 95–132 (1974)

[13] Bouhouch, A. Loqman, C and El Qadi, A. ;CHN and Swap Heuristic to Solve the Maximum Independent Set Problem. International Journal of Electrical and Computer Engineering , vol. 7, no. 6, pp. 3583–3592 (2017)

[14] Handa, H. : Hybridization of estimation of distribution algorithms with a repair method for solving constraint satisfaction problems, in Genetic and Evolutionary Computation Conference, pp. 991–1002 (2003)

[15] Liu, T. Liu, M. Zhang, Y.B and Zhang, L. :Hybrid genetic algorithm based on synthetical level of resource conflict for complex construction project scheduling problem. Proceedings of the International Conference on Machine Learning and Cybernetics, vol. 9, pp. 5699–5703. (2005)

[16] Duan, Q. T. Liao, W. and Yi, H. Z. A comparative study of different local search application strategies in hybrid metaheuristics Appl. Soft Computing, vol. 13, no. 3, pp. 1464–1477 (2013)

[17] Minton,S. Johnston, M. D. Philips, A. B and Laird, P ;Minimizing conflicts: a heuristic repair method for constraint satisfaction and scheduling problems. Artif.Intell., vol. 58, no. 1, pp. 161–205 (1992)

[18] Hayakawa, D. Mizuno, K. H. Sasaki, and S. Nishihara, Improving Search Efficiency Adopting Hill-Climbing to Ant Colony Optimization for Constraint Satisfaction Problems, in Knowledge and Systems Engineering (KSE), 2011 Third International Conference on, 2011, pp. 200–204. (2011)

[19] H. Zou and B. Y. Choueiry, : Characterizing the Behavior of a Multi-Agent Search by Using it to Solve a Tight, Real-World Resource Allocation Problem in Workshop on Applications of Constraint Programming, Kinsale, County Cork, Ireland, 2003, pp. 81–101 (2003)

[20] A. E. Eiben, P.-E.Raué, and Z. Ruttkay, : Solving constraint satisfaction problems using genetic algorithms. in Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on, 1994, pp. 542–547 (1994)

[21] L. Schoofs and B. Naudts : Swarm intelligence on the binary constraint satisfaction problem. in Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on, 2002, vol. 2, pp. 1444–1449 (2002)
[22] K. Xu and W. Li, "Exact phase transitions in random constraint satisfaction problems," J. Artif.Intell. Res. (2000)
[23] K. Xu, F. Boussemart, F. Hemery, and C. Lecoutre, A simple model to generate hard satisfiable instances, arXivPrepr.cs/0509032 (2005).
[24] K. Xu, F. Boussemart, F. Hemery, and C. Lecoutre, :Random constraint satisfaction: Easy generation of hard (satisfiable) instances. Artif.Intell., vol. 171, no. 8, pp. 514–534 (2007).

The Hybrid Machine Learning Model Based on Random Forest Optimized by PSO and ACO for Predicting Heart Disease

Youness Khourdifi[1], Mohamed Bahaj[2]

[1,2] Faculty of Sciences and Techniques, Hassan 1st University, Settat, Morocco
[1,2]{ykhourdifi,mohamedbahaj}@gmail.com

Abstract. In this paper, we used the hybrid Machine Learning model, for proposed PA-RF, a classification based on Random Forest model, optimized by Particle Swarm Optimization (PSO) associated with Ant Colony Optimization (ACO), and we use Fast Correlation-Based Feature Selection (FCBF) method to filter redundant and irrelevant characteristics, in order to improve the quality of heart disease classification. The proposed mixed approach is applied to the heart disease dataset. The results demonstrate the effectiveness and robustness of the proposed hybrid method in processing various types of data for the classification of heart disease. Therefore, this study examines the different automatic learning algorithms and compares the results using different performance measures, i.e. Accuracy, Precision, Recall, F1-Score, etc. The data set used in this study comes from the UCI's automatic learning repository, entitled "Heart Disease" Data set. We can be concluded that PA-RF has demonstrated efficiency and robustness compared to other classification methods.

Keywords: Machine Learning; Heart Disease; Random Forest; Ant Colony Optimization; Particle Swarm Optimization.

1 Introduction

Intelligent optimization algorithms are developed by simulating or revealing certain natural phenomena and are widely used in many research fields because of their versatility [1], [2]. The Particle Swarm Optimization (PSO) algorithm has been successfully applied to heart disease because of its simplicity and generality[3]. However, PSO easily fell into the optimal local solution. In addition, the ACO algorithm was originally introduced for combinatorial optimization. Recently, ACO algorithms have been developed to solve continuous optimization problems. These problems are characterized by the fact that decision variables have continuous domains, unlike discrete problems [4]. Using a single optimization algorithm has the disadvantages of low accuracy and generalizability in solving complex problems. To further explore the application of intelligent optimization in bioinformatics, PSO and ACO are combined in this article, meaning that exploitation and exploration capacity are combined for binary

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284088

and multi-class heart disease. In this article, the Fast Correlation-Based Feature selection (FCBF) method[5] used to remove redundant and irrelevant features, the results of the PSO optimization are considered the initial values of the ACO, and then the classification model for heart disease is constructed after the parameters are adjusted. In this study, algorithms such as K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF) and Artificial Neural Network (MLP) are used. The results demonstrate that the hybrid optimized by FCBF, PSO and ACO method presented in this work is robust and provides more accurate classification results. This work aims to provide heart disease classification results for reference and to contribute to the clinical diagnosis and treatment of different types of heart disease.

The main objective of this article is the prediction of heart disease using the weka data-mining tool and its use for classification in the field of medical bioinformatics. It first classifies the data set and then determines the best algorithm for the diagnosis and prediction of heart disease. Prediction begins by identifying symptoms in patients, then identifying sick patients from a large number of sick and healthy patients. Thus, the primary objective of this paper is to analyze data from a heart disease dataset using a classification technique to accurately predict the class in each case. The main contributions of this paper are:

- **Extraction of classified accuracy useful for heart disease prediction**
- **Remove redundant and irrelevant features with Fast Correlation-Based Feature selection (FCBF) method.**
- **Optimizations with Particle Swarm Optimization PSO then we consider the result of PSO the initial values of Ant Colony Optimization ACO approaches.**
- **Comparison of different data mining algorithms on the heart disease dataset.**
- **Identification of the best performance-based algorithm for heart disease prediction.**

The rest of the paper is arranged as follows: Recent work in this area is discussed in Section 2. Section 3 describes the detailed description of the proposed methodology. Section 4 explains in detail the experiments using the proposed machine learning models. Section 5 presents conclusions and future research directions.

2 Related Works

Several experiments are conducted on medical data sets using multiple classifiers and features selection techniques. There is little research on the classification of the heart disease dataset. Many of them show good classification accuracy[6].

Tan et al. [7] Proposed a hybrid method in which two machine learning algorithms, Support Vector Machine (SVM) and Genetic Algorithm (G.A), are effectively

336

combined by the wrapper approach. The LIBSVM and the WEKA data mining tool are used to analyze the results of this method. Five data sets (Iris, diabetes disease, breast cancer disease, heart disease and hepatitis) are collected from the Irvine UC machine learning repository for this experiment. After applying the hybrid GA and SVM approach, an accuracy of 84.07% is obtained for heart disease. For all diabetes data, 78.26% accuracy is achieved. The accuracy for breast cancer is 76.20%. The 86.12% accuracy is the result of hepatitis disease.

R. RanjaniRani et al. [8] Proposed a combination of Spider Monkey Optimization Algorithm along with the Support Vector Machine classification algorithm and employed for the microarray cancer gene expression data. It has two phases: first is to eliminate irrelevant and redundant genes and select the subset of genes from the large volume of genes using the Spider Monkey Optimization algorithm. The next phase is to classify cancer types using selected genes from the initial phase. The results conclude that the proposed algorithm outperforms the other existing methods in classification accuracy, and it selects a less number of genes.

Vembandasamy et al. [9] diagnosed heart disease using the Naive Bayes algorithm. Bayes' theorem is used in Naive Bayes. Therefore, Naïve Bayes has a powerful principle of independence. The data used are from one of the leading diabetes research institutes in Chennai. The data set consists of 500 patients. Weka is used as a tool and performs classification using 70% of the Percentage Split. Naive Bayes offers 86.419% accuracy.

3 Methodology.

3.1 Data Set and Attributes

The data is collected from the UCI machine learning repository. The data set is named Heart Disease DataSet and can be found in the UCI machine learning repository. The UCI machine learning repository contains a vast and varied number of datasets which include datasets from various domains. These data are widely used by machine learning community from novices to experts to understand data empirically. Various academic papers and researches have been conducted using this repository. This repository was created in 1987 by David Aha and fellow students at UCI Irvine. Heart disease dataset contains data from four institutions[10].

- Cleveland Clinic Foundation.
- Hungarian Institute of Cardiology, Budapest.
- V.A. Medical Centre, Long Beach, CA.
- University Hospital, Zurich, Switzerland.

For the purpose of this study, the data set provided by the Cleveland Clinic Foundation is used. This dataset was provided by Robert Detrano, M.D. Ph.D. Reason to choose this dataset is, it has less missing values and is also widely used by the research community [11].

337

Table 1. Attributes of the Heart disease dataset

Attribute	Representation	Information Attribute	Description
Age	Age	Integer	Age in years (29 to 77)
Sex	Sex	Integer	Gender instance (0 = Female, 1 = Male)
ChestPainType	Cp	Integer	Chest pain type (1: typical angina, 2: atypical angina, 3: non- anginal pain, 4: asymptomatic)
RestBloodPressure	Trestbps	Integer	Resting blood pressure in mm Hg[94, 200]
SerumCholestoral	Chol	Integer	Serum cholesterol in mg/dl[126, 564]
FastingBloodSugar	Fbs	Integer	Fasting blood sugar > 120 mg/dl (0 = False, 1= True)
ResElectrocardiographic	Restecg	Integer	Resting ECG results (0: normal, 1: ST-T wave abnormality, 2: LV hypertrophy)
MaxHeartRate	Thalach	Integer	Maximum heart rate achieved[71, 202]
ExerciseInduced	Exang	Integer	Exercise induced angina (0: No, 1: Yes)
Oldpeak	Oldpeak	Real	ST depression induced by exercise relative to rest[0.0, 62.0]
Slope	Slope	Integer	Slope of the peak exercise ST segment (1: up-sloping, 2: flat, 3: down-sloping)
MajorVessels	Ca	Integer	Number of major vessels coloured by fluoroscopy (values 0 - 3)
Thal	Thal	Integer	Defect types: value 3: normal, 6: fixed defect, 7: irreversible defect
Class	Class	Integer	Diagnosis of heart disease (1: Unhealthy, 2: Healthy)

3.2 Classification Task

From the perspective of automatic learning, heart disease detection can be seen as a classification or clustering problem. On the other hand, we formed a model on the vast set of presence and absence file data, we can reduce this problem to classification. For known families, this problem can be reduced to one classification only - having a limited set of classes, certainly including the heart disease sample, it is easier to identify the right class, and the result would be more accurate than with clustering algorithms. In this section, the theoretical context is given on all the methods used in this research. For the purpose of comparative analysis, five Machine Learning algorithms are discussed. The different Machine Learning (ML) algorithms are K-Nearest Neighbour (KNN), Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes and Artificial Neural Network (ANN). The reason to choose these algorithms is based on their popularity [12].

3.3Particle swarm optimization (PSO)

338

Swarm intelligence is a distributed solution to complex problems which intend to solve complicated problems by interactions between simple agents and their environment[13]–[15]. In 1995, Russel Eberhart, electrical engineer and James Kennedy, socio-psychologist were inspired by the living world to set up a metaheuristic: optimization by particle swarm. This method is based on the collaboration of individuals between them: each particle moves and at each iteration, the one closest to the optimum communicates its position to the others so that they can modify their trajectory. This idea is that a group of unintelligent individuals may have a complex global organization.

Due to its recent nature, a lot of research is being done on P.S.O., but the most effective so far is the extension to the framework of combinatorial optimization.

Figure 1 show the flowchart of the PSO algorithm.

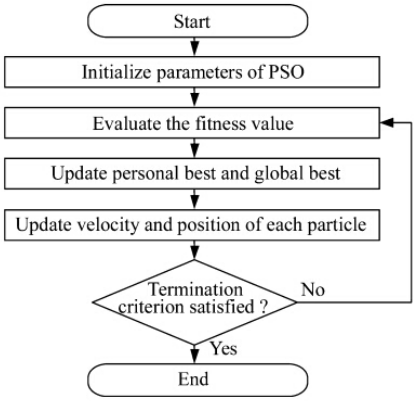


Fig. 1. The flowchart of the PSO algorithm

In particle swarm optimization, each individual of the population called particle. In standard PSO, after the initialization of the population, each particle update its velocity and its position in each iteration based on their own experience (pbest) and the best experience of all particles (gbest) as shown in Eql.(1 & 2). At the end of each iteration the performance of all particles will be evaluated by predefined cost functions.

$$v^i[\mathrm{t}+1] = \mathrm{w}.v^i[t] + c_1 r_1(p^{i,best}[t] - p^i[t] + c_2 r_2(p^{g,best}[t] - p^i[t] \quad (1)$$

$$p^i[\mathrm{t}+1] = p^i[t] + v^i[t+1] \quad (2)$$

Where, $i = 1,2,\dots,N$, N is the a number of swarm population. $v^i[t]$ is the velocity vector in $[t]th$ iteration. $p^i[t]$ represent the current position of the *i*th particle. $p^{i,best}[t]$ is the previous best position of *i*th particle and $p^{g,best}[t]$ is the previous best position of a whole particle. To control the pressure of local and global search, *w* has been used. c_1 and c_2 are positive acceleration coefficients which respectively called cognitive parameter and social parameter. r_1 and r_2 are random number between 0 and 1.

339

3.4Ant Colony Optimization (ACO)

Ant Colony Optimization method explores to find the optimal feature subset using some iterations [16]. The main objective of the Ant Colony Optimization method is to minimize redundancy between them by selecting a subset of feature. In this method, each ant in relation to the previously selected features selects the lowest similarity features. Therefore, if a feature is selected by most ants, it indicates that the features has the lowest similarity with the other features. The feature receive the largest number of pheromones, and the chances of its selection by other ants will be increased in subsequent iterations. Finally, by considering the similarity between the features, the selected main features will have high pheromone values. Thus, the ACO method selects the best features with a minimum of redundancy [17].
Figure 2 shows the illustration of the feature selection problem.

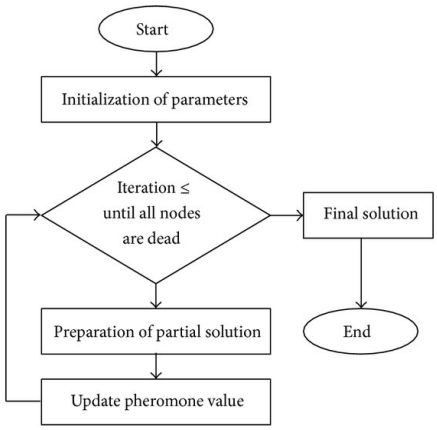


Fig. 2. The flowchart of the ACO algorithm

4 Experiments and Results

The aim of the entire project was to test which algorithm classifies heart disease the best with the proposed optimization methods.
The classification experiment in this paper was carried out under a Weka environment. In addition, due to the small number of selected features, 10-fold cross validation was used. For the purpose of avoiding instable operation results, each experiment was run 10 times, and the optimal classification accuracy was selected for comparison. We evaluate the effectiveness of all classifiers in terms of time to build the model, correctly classified instances, incorrectly classified instances and accuracy according to 3 steps:

1. Optimizations with PSO and ACO approaches for Random Forest, and represented by PA-RF.

340

2. Comparison of different data mining algorithms on the heart disease dataset.
3. Identification of the best performance-based algorithm for heart disease prediction.

4.1 Effectiveness

In this section, we evaluate the effectiveness of all classifiers in terms of time to build the model, correctly classified instances, incorrectly classified instances and accuracy. The results are shown in Table 2.

Table 2. Classifiers Performance

Evaluation criteria	K-NN	SVM	RF	NB	MLP	PA-RF
Time to build model (s)	0,01	0,07	0,16	0,01	0,89	0,03
Correctly classified instances	202	226	220	226	222	269
Incorrectly classified instance	68	44	50	44	48	1

In order to improve the measurement of classifier performance, the simulation error is also taken into account in this study. To do this, we evaluate the effectiveness of our classifier in terms of: Kappa as a randomly corrected measure of agreement between classifications and actual classes, Mean Absolute Error as the way in which predictions or predictions approximate possible results, Root Mean Squared Error, Relative Absolute Error, Root Relative Absolute Error, Root Relative Squared Error. The results are presented in Figures 3.

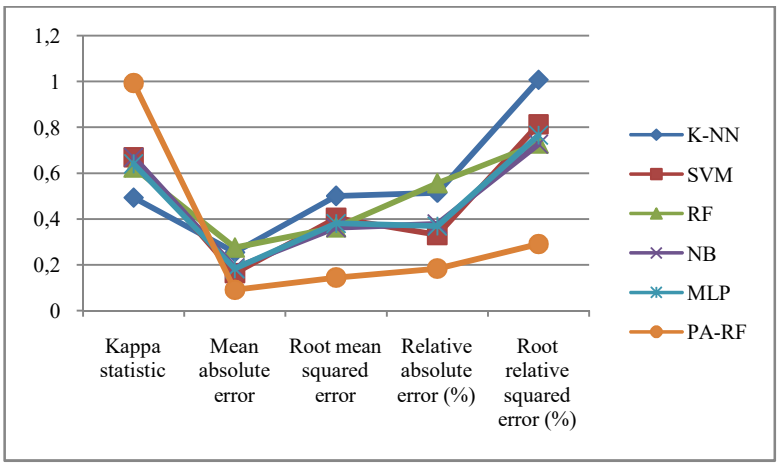


Fig. 3. Simulation Error

4.2 Accuracy Results

341

Once the predictive model is built, we can check how efficient it is. For that, we compare the accuracy measures based on precision, recall, TP rate and FP rate values for K-NN, SVM, RF, NB, MLP and PA-RF. The results are shown in Figure 4.

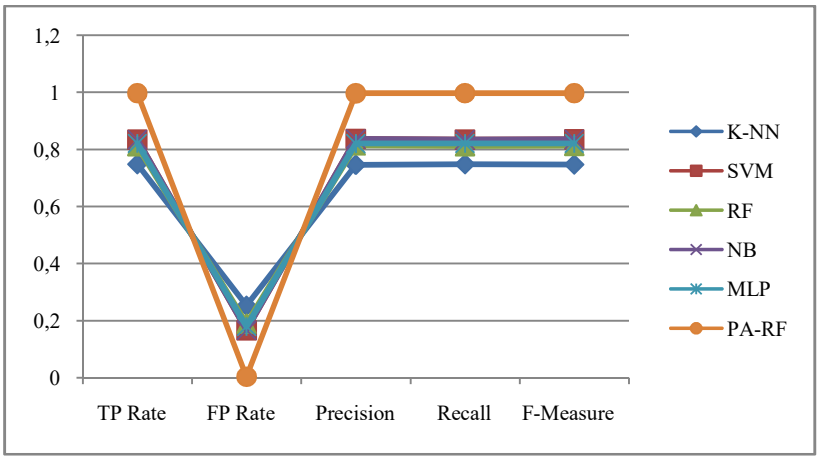


Fig. 4. Accuracy measured

From the different classifiers results presented in Figure 4. We can see that the best results are those generated by Classifiers optimized by PSO and ACO. The PA-RF model shows the best results in comparison with other classifiers algorithms.

Confusion Matrix

Confusion matrices represent a useful way of evaluating classifier; each row of Table 3 represents rates in an actual class while each column shows predictions.

Table 3. Confusion Matrix

	Absence	Presence	
K-NN	113	37	Absence
	31	89	Presence
SVM	130	20	Absence
	24	96	Presence
RF	127	23	Absence
	27	93	Presence
NB	130	20	Absence
	24	96	Presence
MLP	125	25	Absence
	23	97	Presence
PA-RF	149	1	Absence
	0	120	Presence

342

Results Discussion

In this paper, we applied machine-learning algorithms on heart disease dataset to predict heart disease, based on the data of each attribute for each patient. Our goal was to compare different classification models and define the most efficient one. From all the results above, different algorithms performed better depending upon the situation whether cross-validation, grid search, calibration and feature selection is used or not.
For the comparison of the dataset, performance metrics after feature selection, parameter tuning and calibration areused because this is a standard process of evaluating algorithms. The precision average value of the best performance without optimization it's for SVM and NB with 83,6% than RF with 81,35%. These shows SVM and NB are performing on average, after optimized by FCBF, PSO and ACO, we find the best one is PA-RF with 99,6 %.

Conclusion and Future work

The hybrid Machine Learning model proposed in this work was to optimize the Random Forest algorithm using Particle Swarm Optimization (PSO) associated with Ant Colony Optimization (ACO). The results of the Random Forest optimize (PA-RF) were compared to other algorithms with different performance measures. The results showed that the optimized Random Forest algorithm, PA-RF showed the better performance and effectiveness compared to K-Nearest Neighbour K-NN, Random Forest RF, Naïve Bayes NB, Support Vector Machine SVM and Artificial Neural Network MLP in all the data set used in this study.
This work can be the first step in learning the diagnosis of heart disease through automatic learning and can be extended for future research. There are several limitations to this study, mainly the tools used in this study such as the processing power of the computer and second the time limit available for the study. This type of study requires state-of-the-art resources and expertise in the respective fields.

For future work, we intend to conduct an in-depth study of these datasets by combining Machine Learning techniques with deep learning models on the application of more complex deep learning architectures to achieve better performance. In addition, we test our in-depth learning approach on larger data sets with more disease classes to achieve higher accuracy.

References

1. Kamkar, M. Akbarzadeh-T, and M. Yaghoobi, "Intelligent water drops a new optimization algorithm for solving the Vehicle Routing Problem," in 2010 IEEE International

343

Conference on Systems, Man and Cybernetics, 2010, pp. 4142–4146.
2. Z. Gazzaz, N. M., Yusoff, M. K., Ramli, M. F., Juahir, H., & Aris, “Artificial neural network modeling of the water quality index using land use areas as predictors,” Water Environ. Res., vol. 87, no. 2, pp. 99–112, 2015.
3. G. Zhang, Y., Wang, S., & Ji, “A comprehensive survey on particle swarm optimization algorithm and its applications,” Math. Probl. Eng., 2015.
4. H. M. Alshamlan, G. H. Badr, and Y. A. Alohali, “Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification,” Comput. Biol. Chem., vol. 56, pp. 49–60, 2015.
5. H. Yu, L., & Liu, “Feature selection for high-dimensional data: A fast correlation-based filter solution,” Proc. 20th Int. Conf. Mach. Learn., pp. 856–863, 2003.
6. M. Fatima, M., & Pasha, “Survey of machine learning algorithms for disease diagnostic,” J. Intell. Learn. Syst. Appl., vol. 9, no. 1, p. 1, 2017.
7. K. C. Tan, E. J. Teoh, Q. Yu, and K. C. Goh, “A hybrid evolutionary algorithm for attribute selection in data mining,” Expert Syst. Appl., vol. 36, no. 4, pp. 8616–8630, 2009.
8. Rani, R. Ranjani, and D. Ramyachitra. "Microarray Cancer Gene Feature Selection Using Spider Monkey Optimization Algorithm and Cancer Classification using SVM." Procedia computer science 143 (2018): 108-116.
9. E. Vembandasamy, K., Sasipriya, R., & Deepa, “Heart Diseases Detection Using Naive Bayes Algorithm,” IJISET-International J. Innov. Sci. Eng. Technol., vol. 2, pp. 441–444, 2015.
10. M. Lichman, “UCI Machine Learning Repositry [Online],” Available: https://archive.ics.uci.edu/, 2013.
11. U. H. Dataset, “UCI Machine Learning Repository,” [online]. URLhttps//archive.ics.uci.edu/ml/machine-learning-databases/heartdisease/ Hear.
12. L. Van Cauwenberge., “Top 10 Machine Learning Algorithms,” Data Sci. Cent., 2015.
13. R. Poli, J. Kennedy, and T. Blackwell, “Particle swarm optimization,” Swarm Intell., vol. 1, no. 1, pp. 33–57, 2007.
14. J. Kennedy, “Particle Swarm Optimization,” in Encyclopedia of Machine Learning (Springer), C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2010, pp. 760–766.
15. Y. Shi, “Particle swarm optimization,” IEEE Connect., vol. 2, no. 1, pp. 8–13, 2004.
16. M. Dorigo and M. Birattari, “Ant Colony Optimization,” in In Encyclopedia of machine learning (Springer), C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2010, pp. 36–39.
17. S. Tabakhi, P. Moradi, and F. Akhlaghian, “An unsupervised feature selection algorithm based on ant colony optimization,” Eng. Appl. Artif. Intell., vol. 32, pp. 112–123, 2014.

344

An Energy Balanced Algorithm for data communication in WSN

HAJAR LAGRAINI[1], HICHAM OULDZIRA[2] ,MOSTAFA CHHIBA[1], ABDELMOUMEN TABYAOUI[1], AHMED MOUHSEN[2]
{ h.lagraini@uhp.ac.ma , odzira@yahoo.fr, moschhiba@yahoo.fr, atabyaoui@gmail.com , ahmed.mouhsen@uhp.ac.ma }

[1]Radiations materials & instrumentations Laboratory, Faculty of Sciences & Technology FSTS, University Hassan 1st. BP 577, 2600, SETTAT, MOROCCO
[2]Management, industrial & innovation Laboratory, Faculty of Sciences & Technology FSTS, University Hassan 1st. BP 577, 2600, SETTAT, MOROCCO

Abstract. Wireless sensor networks are seeming as an emergent requisite for mankind. Such networks are still in research stage though, they have high prospective to be useful in practically every arena of life. A lot of study is done and lots of other is a waiting to be standardized. In this paper, clustering in WSNs is deliberate precisely. Furthermore, we improve one of the utmost popular WSN's routing algorithms "LEACH" as enhanced LEACH. Our improved algorithms, in comparison with basic algorithm overtake it by means of lifetime of cluster head and network lifespan. Subsequently, new thresholds are employed on improved algorithms that claim the performance even more. To finish a brief performance test of basic protocol, with ES-LEACH and ESD-LEACH is assumed bearing in mind energy leftover and distance metrics.

Keywords: sensor nodes, cluster routing, static sink, wireless sensor network.

1 Introduction

Data- data- data; we need instantaneous data in every feature of our lives. An automatic engineer desires to get data about his instrument running, a doctor needs to know about his patient's ECG interpretation, measured variations in earth and weather are required by geologists; list continue. To suffice this must, numerous networks are considered to transmit data. At first human to human, and then telephonic systems giving birth to WSs. Multi hop networks were intended to give additional autonomy of movement. And then arises the eon where machine-machine communication was announced. A treating device purchases data, process it and transmit it to another device. That data is auxiliary precised / gathered/ attached perceptively so that it can be accessible to us "humans". In WSNs, that gadget generally is named a sensor, or smart sensor node and has its own limits i.e. It is necessary proficient of sensing, treating and communicating. Each SN therefore need an electrical power source to ensure all these tasks. Taking into account applications of WSNs, fixing a battery on each SN is a the best option. Though, regulating consumption of power is one of the key tasks in these networks. These batteries need to be smart enough to give a SN extreme life regardless of being very small sized. Any technology that is in progress and expansion, provide lot of contests. Similarly WSNs do. Detecting, processing and communicating by very

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284093

small sized nodes with power restraint is a pretentious thing. Consequently this is the main unease for scientist researchers. To enhance SN's lifespan, we must focalize on such techniques that can make maximum out of restricted battery. In any type of networks specifically WSN, for effective performance; these algorithms should be effectual. Many algorithms are presented that discourse limited battery issue in WSNs. In an effective WSN, we need efficient routing algorithm that has low-slung routing overhead and well-structured information collection tools to protect battery life of SN. Therefore clustering of the network is an alternative technique to promote power energy effectiveness. In hierarchical architecture sensing zone is divided to a numeral clusters and a set of nodes are haphazardly chosen as (CH).CH collects the data from non-CH members, united it and then transmit it to the BS for further treating [2]. Clustering thus distribute the energy load, reduce the energy ingestion and sustain the WSN life [3], [4], [5], [6]. In this paper an enhanced energy balanced algorithms are introduced for WSNs. our method proposes a new cluster head (CH) selection technique based on tree elements: clustering, residual of energy and distance. This algorithm is an improvement of basic protocol presented in [7] and simulation result demonstrate that the enhanced protocol outperform LEACH in terms of network life time extension. This work is structured as follows. In Section 2 we discuss the literature review of existing algorithms. In section 3 we present clustering approach, CH selection and hierarchical routing algorithms. We further evaluated the proposal ES LEACH / ESD LEACH in section 4. To close, the conclusions are discussed in section 5.

2 Literature review:

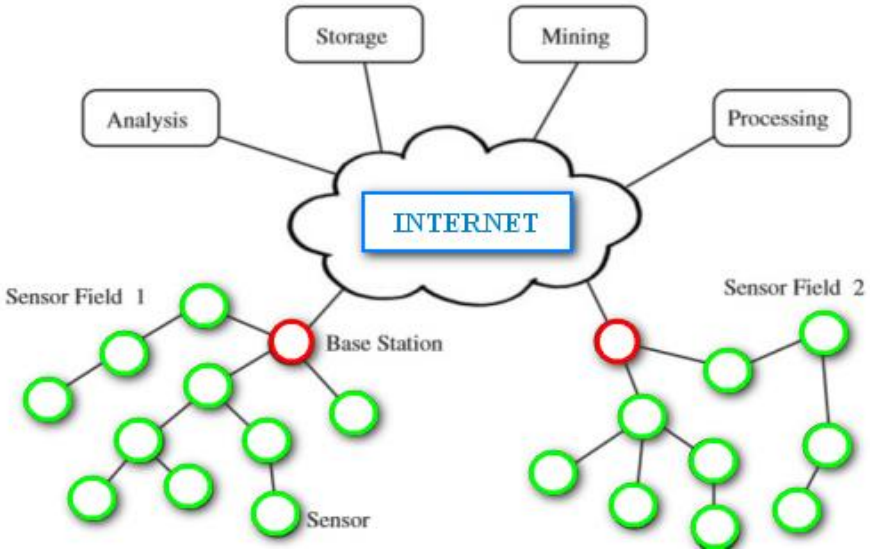


Fig 1: Overview of wireless sensor network with static sink

2.1. CLUSTERING AND CLUSTER HEAD SELECTION

In environmental sensing, clustering algorithms are easy and also effective. SNs are arbitrarily positioned in the monitoring area. Networks are clustered into distinct groups named clusters. Probabilities for SNs that are adjacent to each other are really elevated. Collection of information in clustering algorithms affords a power energy proficient technique for information aggregation. Correspondingly every cluster member node communicates a one-hop packet to the CH; merely then the combined data-packet will be diffused to the sink by the selected CH. whenever the collected information will go over information combination process just afterward that one data-packet will be created. The chosen CH will communicate the sensed information to the sink as an alternative of every SN distinctly. This process of arbitrarily turning the role of CH will regularly allocate the assigned work has been presented. To enhance the WSN lifetime the transmission capacity is necessary which is conceivable through using clustering technique. The entirely member-nodes transmit their information to the cluster heads owing to this battery is saved. Clustering approach allows proficient source distribution which primes to enhance conspiring of power source regulation. Any adjustment in the SN touches just that specific cluster and not the complete network.

2.2. HIERARCHICAL CLUSTERING

Similar to the alternative communication systems, one of the utmost imperative intentions characteristic of a WSN is extendibility. The gateway could become overburdened as the sensor substantiality rises adopting a one-hop transmission. Like time interval between the stimulation and response in communication is promising consequently overload and tracing of measures eventually becomes tough. One-hop network architecture is not ascendable if the WSN becomes expanded; it is not able to manage transmission. Clustering technique has been followed in a few routing methods to permit the network to deal with superfluous load and to envelope the large SN. The purpose for advancing hierarchical routing is to efficiently preserve the energy of the SNs by appealing them in multi-hop transceiving [25]. Packet gathering/combination is implemented in a WSN to reduce the amount of packets communicated to the BS. The CH choosing and cluster arrangement are founded on the residual of energy and the adjacency to the CH correspondingly. The principal aim of hierarchical routing technique is to conserve the energy ingesting of SNs. Establishing a cluster basically depend on level of residual or leftover of energy of SNs and adjacent to the cluster chief. This approach saves energy, based on clustering technique lifespan of the network is extended; scalability and energy ingesting are upgraded. Just one cluster head / cluster completes the routing mission and the rest of sensor members just transmit their information to the CH. In vast WSNs, clustering approach shows a significant gain for the network it is smooth to administrate many clusters. Adopting clustering methods just preservation of the CH is sufficient. The SNs in wireless sensor networks are resource self-conscious which means they have restricted energy. Low energy adaptive clustering hierarchy is the most popular hierarchical routing algorithm in WSNs.

2.3. PURE LEACH

LEACH "Low-Energy Adaptive Clustering Hierarchy" is one of the most popular hierarchical routing algorithms. The function process of this protocol is subdivided into rounds. every round begins with a setup phase, pursued by steady state. In the first phase, every SN attempts to elite itself as a CH conforming to probability model. For choosing a CH, every SN set up a arbitrary number in between zero and one. In case that the number is inferior than the threshold T (n), the SN choose itself as a CH for prevailing round. The threshold T (n) is calculated thusly:

$$T(n) = \begin{cases} \frac{P}{1-P*(r \bmod \frac{1}{P})} & ; if\ n \in G \\ 0 & ; otherwise \end{cases} \quad (1)$$

Where: p is the desired percentage of cluster heads in the sensor area and r is the prevailing round number, where G is the set of SNs that have not been chosen as cluster heads in the last 1/p rounds [2][8].

3 PROPOSED ALGORITHM

On examining the properties of LEACH, the SNs chosen as CHs are not an ultimate choice for communicating the information to the base station. In the interest to unravel this issue, this fragment gives an effective protocol to provide a considerable saving in battery ingestion. The suggested protocols are suitable for WSN; which has following qualifications: SNs are static, i.e., once they are randomly diffused, they remain in the same position. Every SN has a special ID and recognizes its location and it's residual of energy. Moreover, every SN has enough energy to communicate with the sink. The entire sensor nodes have the similar incipient power (Einit) and their power sources are not rechargeable or exchangeable. Nodes expire with the expiration of their battery. In our proposition, the possibility or chance for a SN to be cluster head is decided by its distance to the base station and by energy leftover of sensor node. though diminishing of broadcasting energy reduces exponentially with the augmenting diffusion range in wireless transmission. The level of remaining energy of SNs and the range between CHs and the sink are two indispensable metrics that possess an influence on total energy ingestion in every round. accordingly, we will have to consider the selection of the cluster head in order to distance of cluster heads to the BS and range of CMs to the CH are minimal, which will considerably enhance energy depletion for wireless sensor networks lifetime optimization. In our proposed protocol a SN could be chosen as CH basically on two metrics energy of the sensor and its distance from sink, the threshold value is calculated as follow:

$$T(n) = \begin{cases} C * \frac{E_{curr}}{E_{init}} + b * p * \left[\frac{D_{max}-D}{D_{max}-D}\right]^2 ; \text{if } n \in G \\ 0 \qquad\qquad ; \text{otherwise} \end{cases} \quad (7)$$

$$\text{While; } C = \frac{P}{1-P*(r \bmod \frac{1}{P})}$$

Where, E_{curr} is the energy leftover of the SN for the present round and E_{init} is the primary energy , on suggested protocol, cluster head is chosed based on range of SN and sink. The SN which is close to the BS will be chosen as the CH. This enhancement takes remaining energy and distance into account, further it considers the distance from node to cluster head BS and compares the distance from node CH and BS. D_{max}: Represents the maximum distance from SN to BS. D represents distance from sensor node to the BS. Based on left over of energy and distance from BS, threshold criteria T(n)is calculated.

4 RESULTS AND SIMULATION

The simulation results are provided of the proposed protocoles performed in which the Sink is fixed in MATLAB2017a simulator and compared with the results of the basic protocol. Comparison is done in the number of alive nodes, energy of the whole network and the number of received messages to the sink. Table 1 shows simulation parameters used in this work.

Table 1 : Simulation Environment Specifications

Parameter	Value
Simulator	Matlab2017a
WSN area / m2	200*200
Number of nodes	160
Number of rounds	5000
Initial energy of node : Ei	0.5 Joules

Desired percentage of CHs : P		0.05
Transmitter Amplifier Energy	Emp	0.0013 pJ/bit/m4
	Efs	10 pJ/bit/m2
Data Aggregation Energy Eagg		5 nJ/bit
Electric energy		50 nJ/bit

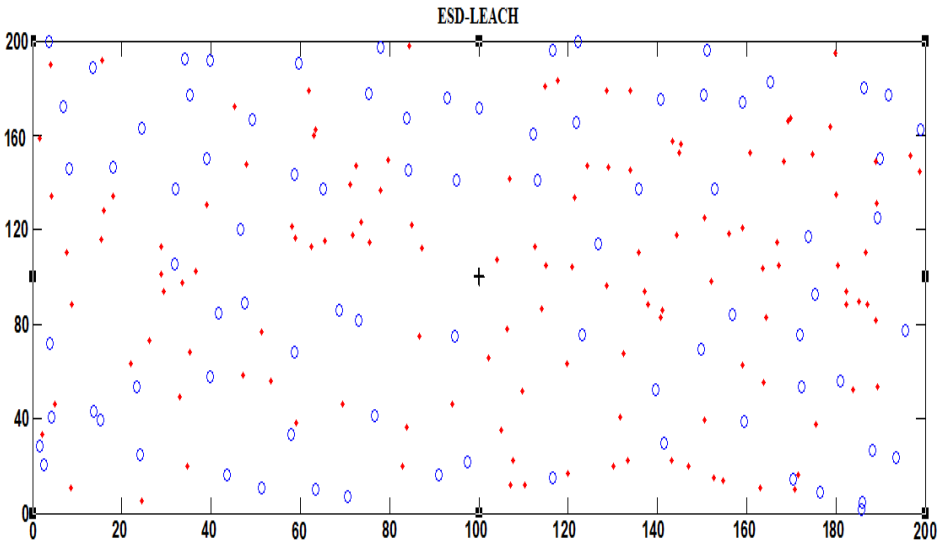


Fig. 2. alive node vs dead nodes in 5000 rounds using the enhanced algorithm

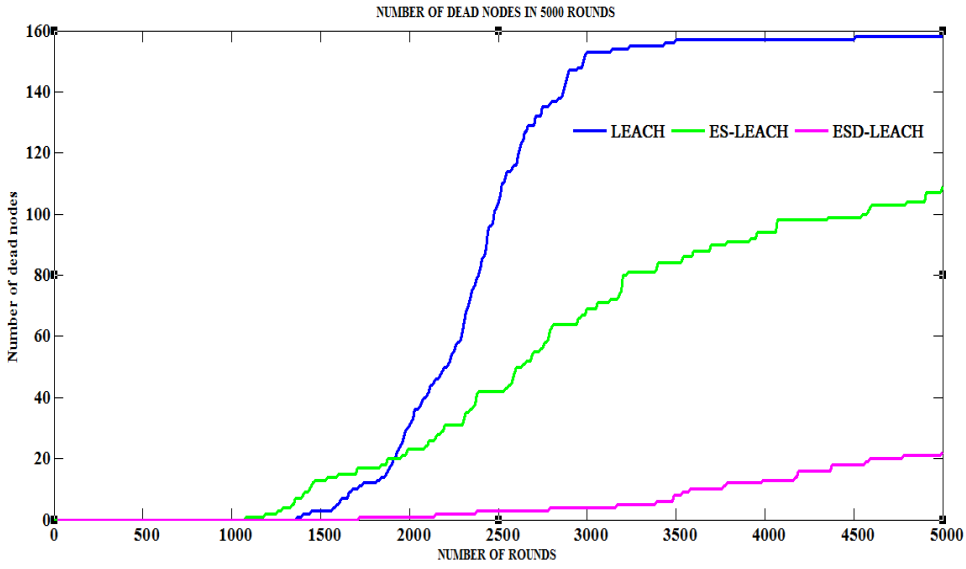


Fig. 3: Dead Nodes over rounds

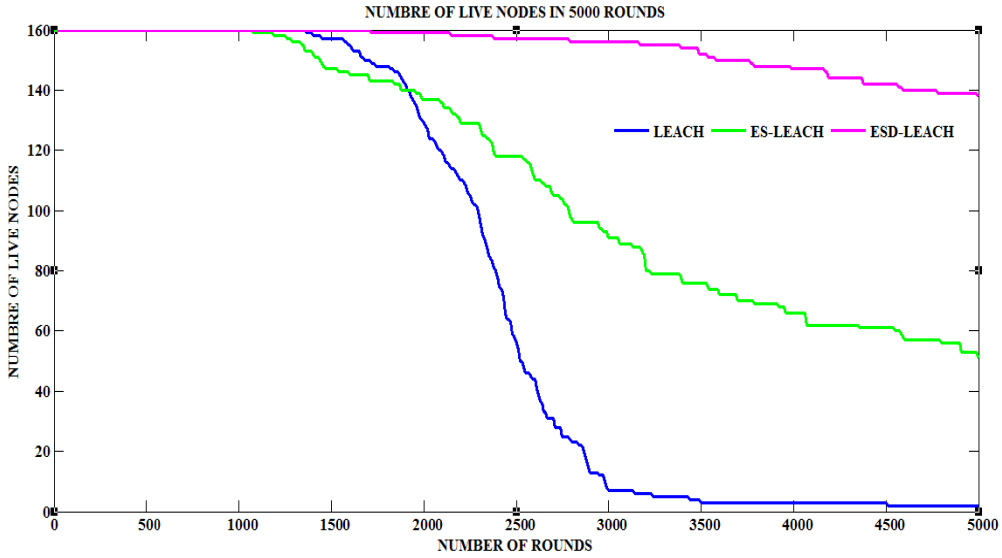


Fig. 4 Alive nodes over time in 5000 rounds

Figures 3, 4 show the comparison of pure LEACH, ES-LEACH and ESD LEACH algorithms in terms of amount of alive sensors in each round through all network life time of the 3 algorithms for 160 nodes diffusion in 200x200 areas.

The results are shown in Figure 3; SNs in ESD LEACH have a longer network battery life than the pure LEACH. The reduction in power ingesting of ESD LEACH is done by comparing it with that of the pure algorithm. We observe from the figure that the nodes of basic algorithm start to die approximately at 1600 rounds. However, ESD LEACH SNs start to die at 3500rounds. We can also clearly notice that the network lifetime of proposed algorithm prolonged until 5000 rounds, whereas that of pure LEACH ended at 3500 rounds.

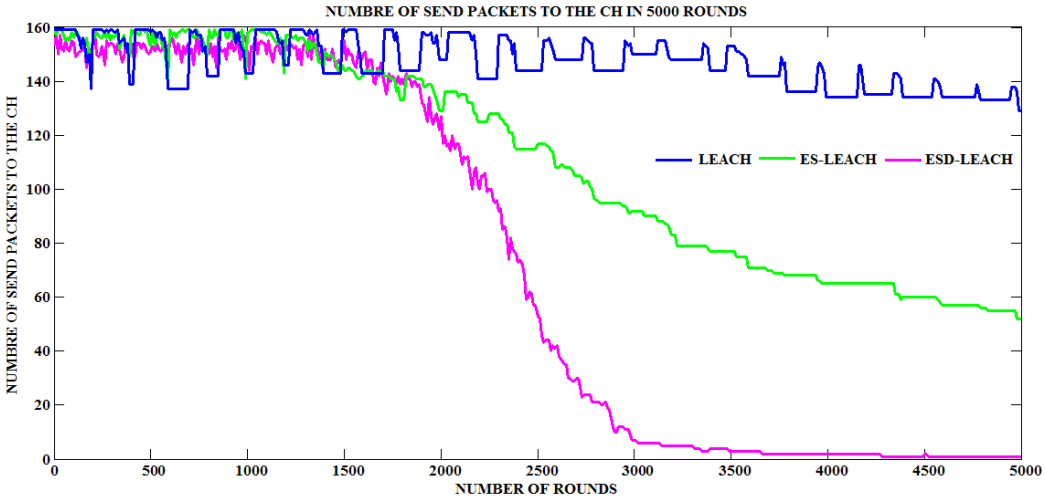


Fig. 5 : number of send packets to the CH in 200x200 m^2

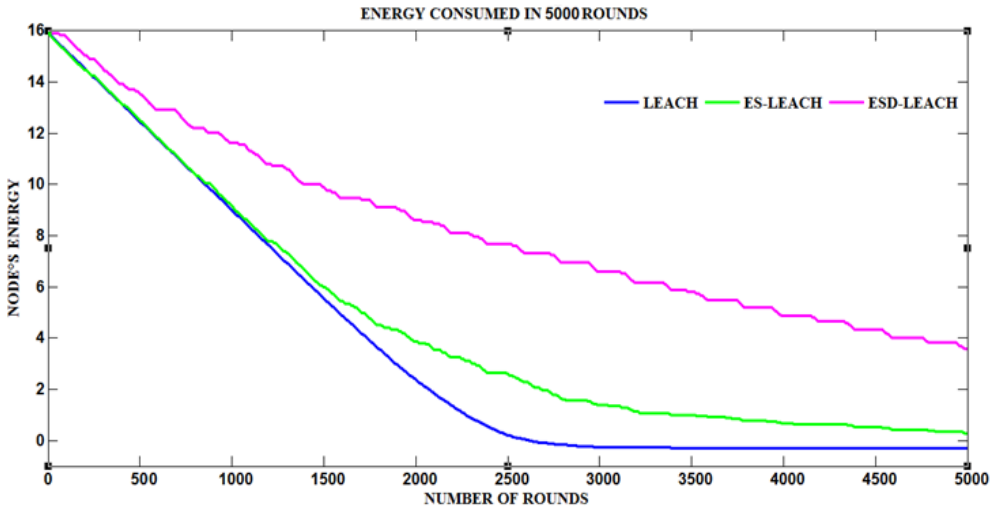


Fig.6: the comparison of energy ingestion rate in 5000 rounds.

We focus on the complete left over energy of SNs. Figure 6 shows the diminishing of this energy/ round. When comparing our algorithms to the basic algorithm, it's crystal clear that in our modified protocols, remaining energy diminishes leisurelier than basic one. In other words, our algorithms are proficient to balance energy ingesting much greater than pure algorithm. This verifies that we have picked an effectual way for choosing cluster heads based on outstanding energy level and the range between the SN and the SINK.

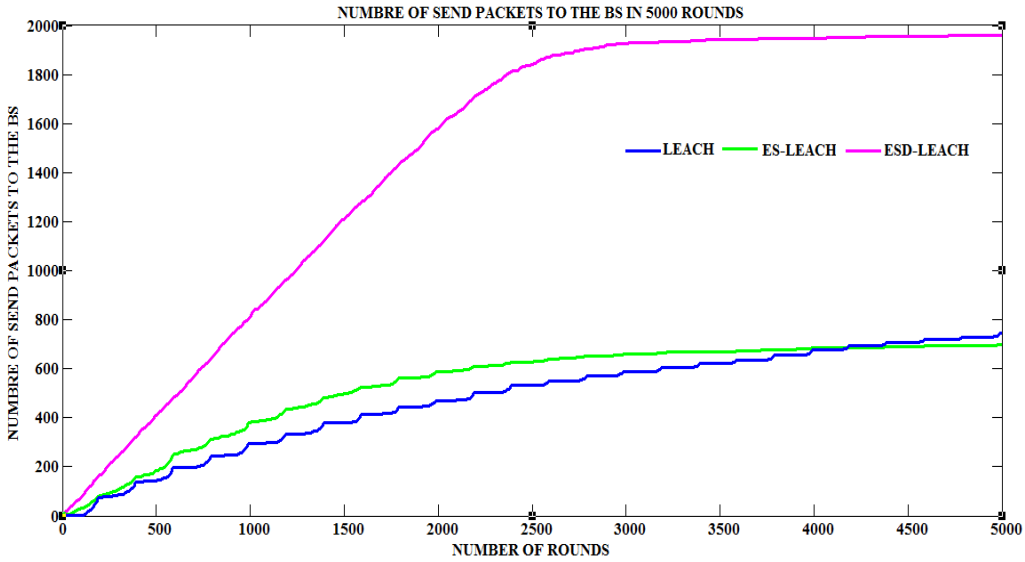


Fig.7 : Average number packets that received to the sink.

This Figure 7 shows delivery of data packets to the BS. ES LEACH , ESD LEACH packet transmission give better rate then pure LEACH.

5 CONCLUSIONS AND FUTUR WORK

In this paper, we proposed an energy balanced algorithm for wsn which is an enhancement to pure LEACH protocol, we presented a new technique to choose the CHs in each round which basically depends on two metrics distance and energy level od nodes. In ES LEACH / ESD LEACH, we think closest CH to the sink will consumes fewer energy than other nodes since the information transmission dissipates the ultimate energy in WSNs. Simulation graphs confirmed that the enhanced ESD LEACH surpass basic LEACH in stretching the networks life. present results in this work can achieve performance improvements, there are other areas for improvement to make our ESD LEACH best suited everywhere. As the smart sensor nodes have limited power energy, therefore the nodes die after a certain period limit. Future work directions are to take the nodes themselves as solar aware nodes which recapture energy it selves, so that our protocol will be more energy sufficient. And another provision is to make our algorithm in to hierarchal protocol by forming ”TOP clusters” out of the cluster head nodes and TOP clusters will process all the information from the CHs. Therefore the energy dissipation of every CH will be condensed to transfer data to the sink. This enhancement will make the ESD LEACH effective for an extensive range of smart networks.

6 References

[1] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. Computer networks, 38(4), 393-422.

[2] Akkaya, K., & Younis, M. (2005). A survey on routing protocols for wireless sensor networks. Ad hoc networks, 3(3), 325-349.

[3] Al-Karaki, J. N., & Kamal, A. E. (2004). Routing techniques in wireless sensor networks: a survey. IEEE wireless communications, 11(6), 6-28.

[4] Liang, C. K., Huang, Y. J., & Lin, J. D. (2008, March). An energy efficient routing scheme in wireless sensor networks. In Advanced Information Networking and Applications-Workshops, 2008. AINAW 2008. 22nd International Conference on (pp. 916-921). IEEE.

[5] Zheng, J., & Jamalipour, A. (2009). Wireless sensor networks: a networking perspective. John Wiley & Sons.

[6] Lindsey, S., & Raghavendra, C. S. (2002). PEGASIS: Power-efficient gathering in sensor information systems. In Aerospace conference proceedings, 2002. IEEE (Vol. 3, pp. 3-3). IEEE.

[7] Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. IEEE Transactions on wireless communications, 1(4), 660-670.

[8] Bian, Q., Zhang, Y., & Zhao, Y. (2010, May). Research on clustering routing algorithms in wireless sensor networks. In Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on (Vol. 2, pp. 1110-1113). IEEE.

[9] Long-long, X., & Jian-jun, Z. (2010, August). Improved LEACH cluster head multi-hops algorithm in wireless sensor networks. In Distributed Computing and Applications to Business Engineering and Science (DCABES), 2010 Ninth International Symposium on(pp. 263-267). IEEE.

[10] Hou, R., Ren, W., & Zhang, Y. (2009, October). A wireless sensor network clustering algorithm based on energy and distance. In Computer Science and

Engineering, 2009. WCSE'09. Second International Workshop on (Vol. 1, pp. 439-442). IEEE.
[11] Ali, M. S., Dey, T., & Biswas, R. (2008, December). ALEACH: Advanced LEACH routing protocol for wireless microsensor networks. In Electrical and Computer Engineering, 2008. ICECE 2008. International Conference on (pp. 909-914). IEEE.
[12] Chen, G., Zhang, X., Yu, J., & Wang, M. (2012, July). An improved LEACH algorithm based on heterogeneous energy of nodes in wireless sensor networks. In Computing, Measurement, Control and Sensor Network (CMCSN), 2012 International Conference on (pp. 101-104). IEEE.
[13] Ahmad, A., Latif, K., Javaidl, N., Khan, Z. A., & Qasim, U. (2013, May). Density controlled divide-and-rule scheme for energy efficient routing in Wireless Sensor Networks. In Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on (pp. 1-4). IEEE.
[14] Bakaraniya, P., & Mehta, S. (2013). K-leach: An improved leach protocol for lifetime improvement in wsn. International Journal of Engineering Trends and Technology, 4(5), 1521-1526. 2013.

Investigating fading impact on AODV, OLSR and DSDV protocols in NS3

[1]Hanene Brahmia, [2]Cherif Tolba
{ [1]hanenebrahmia@gmail.com, [2]ctolba@yahoo.fr }

[1,2]Laboratoire Réseaux et Systèmes LRS, Dept. computer science BADJI Mokhtar Annaba University, Annaba,Algeria

Abstract. Vehicular communication is an emerging research topic that has not been satisfied yet.Vehicular Ad hoc NETwork, VANET is a system that provides warning messages,road traffic information and advertisement data to drivers.Although It is a domain were the realisation of real experimentations is very expensive and may be dangerous so simulations are the most adapted investigation method to these situations.In our study we used Network Simulator 3 to investigate the influence of the NAKAGAMI fading on the network performances of the three routing protocols AODV,DSDV, and OLSR. Packet Loss ratio and mean delay augmentations were analysed for every protocol in different node densities..

Keywords: VANET, routing protocols, AODV, ITS, DSDV, OLSR, NAKAGAMI..

1 Introduction

Vehicular ad hoc networks have been an attractive research field to thousands of researchers and manufacturers due to the demand of road users that are now more exigent in terms of comfort, entertainment and safety. In recent decades, ensuring road safety has become a major challenge, as road traffic crashes cause a large number of deaths around the world. Intelligent transportation systems are enriched due to the integration of on board units like the several equipments deployed to ensure connectivity: Global Positioning System (GPS), sensors, antennas and the emerging technology using Light Emitting Diodes (LED) [1]. That equipments allow the vehicles to communicate with each other directly (vehicle to vehicle V2V) and provides traffic information to drivers or keeping him aware of the surrounding environment like other lanes, road side signs or traffic lights. The vehicles can also communicate via or with the infrastructure (V2I communication). Objects in the road can be either an obstacle or a relaying node in the vanet architecture.The trucks and the buildings are one of the most distracting obstacles that causes non line of sight (NLOS) situations[2].Since vehicles are moving with some behaviours: speed, acceleration, deceleration or sudden brake that may cause serious problems. The communication links might be broken and re-established several times per second and so the topology is changing frequently; a well-healing strategy is then a great challenge. In urban

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284207

areas the density of vehicles is usually higher and causes signal interferences. On the other hand, in rural areas low density can cause travel delays due to weak mesh between nodes. When the speed of vehicles increases in highways links establishment and maintenance will be a harder task. The behaviour of the transmitted radio waves also depends on the environment.It can be affected by several phenomenas like : Scattering, polarization, reflection, absorption, diffraction and refraction. Signal fading affects also the routing protocol's mechanism it may causes delay and drives the packets to be lost[3].There is a set of distributions that models the fading in wireless networks[4]. NAKAGAMI which is a statistical one, gives more accurate results.in this paper we investigate the impact of NAKAGAMI fading on the behaviour of the three routing protocols that are: AODV (Ad hoc On demand Distance Vector), OLSR (Optimized Link State Routing) and DSDV (Dynamic Destination-Sequenced Distance Vector) .The paper is organised as follows. Section II is devoted to VANET architecture, its principal features and major challenges then routing protocols will be presented in section III. The propagation models will be discussed in section IV. While the simulation setup of our study and the results and discussions will be devoted in section V and VI respectively . Finally, section VII concludes this paper.

2 Vanet overview

An overview of vehicular ad hoc networks is presented here, describing it's architecture, principal features and challenges.

2.1 VANET Architecture

VANET or Vehicular Ad-Hoc Networks are a subclass of MANET Mobile Ad-Hoc Networks that are a mesh of interconnected nodes that can communicate without the necessity of a fixed infrastructure. Figure. 1. Presents the most common wireless networks.

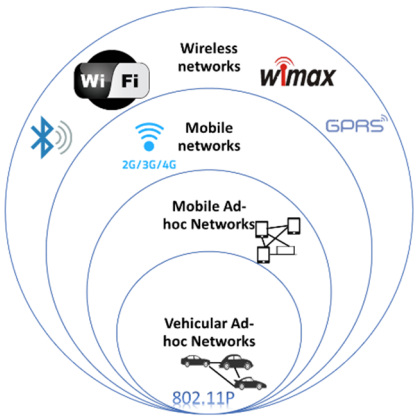


Fig 1. Vanet and wireless networks.

The node of a mobile ad-hoc network might be a smart phone, a laptop or a wireless printer,while in VANET it can be a truck, a car, or even buses and so on.They can communicate with each other V2V to exchange information on the road, broadcast warnings or advertisements. While this network may have fixed entities as Road Side Units (RSU) that manages traffic information and gather technical data about vehicles V2I like speed, direction or destination. The whole architecture is presented in Figure.2.

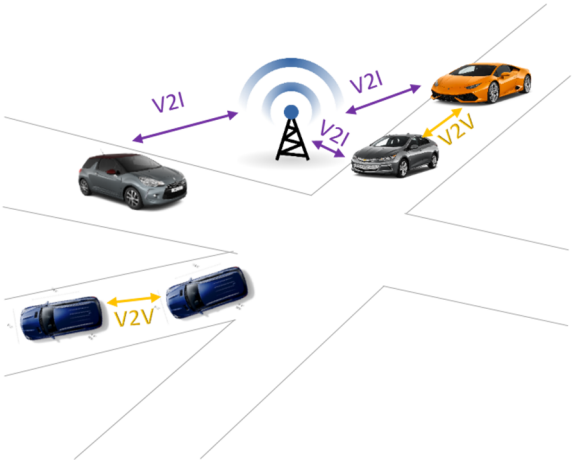


Fig 2. VANET architecture

The network then is divided into two modules:

- The study of the vehicle behaviour without its environment that is the microscopic level.
- The macroscopic level, when the entire surround environment is studied.

2.2 Principal features of vehicular networks

In spite of the affiliation of VANET to MANET, it has some specificities and differences from mobile networks that can be resumed in Figure.3 and discussed below:

- **High dynamicity:** The vehicles move with certain speed and dynamicity especially in highways where the speed is usually over 100 km/h.
- **Large and quickly scalable network:** The nodes like we have said before might be cars, trucks, buses and two wheels that travel together in the road, at pick hours the number of nodes can be huge and a bigger network might be formed quickly.
- **Predictable and directional mobility:** Vehicular ad-hoc networks may be distinguished from their subclass mobile networks by their predictable mobility where the movements of the vehicles are oriented and non-random.
- **No energy limitation:** One of the most challenging problems in mobile and sensor networks is the energy while it is an advantage in VANETs, where the nodes generate energy when moving.

Some specifications might be the cause of a lot of problems; some of them are the research challenges presented below.

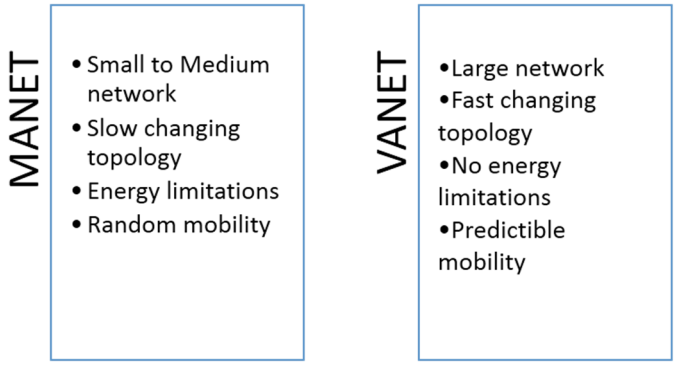


Fig 3.VANET Vs MANET

2.3 Vehicular networks' Challenges

There are a lot of challenges in VANET that researchers have to solve.This paper will resume some of them, but the research fields are not limited to the following.

- **Routing issues and optimizations:** The frequently changing topology affects directly communication links and break them, what makes the establishment of a solid connection between vehicles a great challenge.
- **Collision avoidance and congestion control:** The network is quickly scalable and can be very large what causes traffic jam,non line of sight situations and road accidents.
- **Simulating the networks:** Since the realisation of real experiments is an expensive and critical experience, where the implementation of the solutions in the real world can cause accidents, the use of an adequate simulator is necessary. To simulate a vehicular network we need two simulators: traffic and network simulator. There are many tools to do the task where the selection of the most adequate one depending on the scenario and the requirements is somehow difficult.

3 Routing Protocols

The existence of many MANET routing protocols does not involve as many vehicular ad hoc routing protocols. Some of them does not take in consideration vehicular network specificities or suppose a random movement while it is directed and oriented. Even for those that are suitable for VANET the obtained results and performances cannot reflect the reality.
For this purpose those protocols were studied, compared and ameliorated to be adapted to the vehicular networks [5],[6],[7],[8],[9].

3.1 Routing issues for vehicular ad-hoc networks

Due to the rapidity of link establishment or destruction between nodes, low or high density and road obstacles. Routing protocols will face many problems, from where we can mention:

- **Local maximum problem:** When the nodes cannot find the next hop to forward the message.

- **Limited buffers:** Where the messages are ignored due to the lack of space in the buffer, caused by the huge number of packets to be transmitted under time condition.
- **Delay and overheads:** Some protocols such reactive ones take much time to establish new routes while other ones uses many control packets what produces delay or submerge the network.
- **Common solution:** The different densities in different areas influence the algorithms and strategies of protocols.

This is the cause of the variation of simulation results depending on scenarios. Protocols that fail in some areas can perform better in other.

3.2 Routing for VANET

The routing strategy must take in consideration the bandwidth capacity, the number of nodes and network resources, and have to ensure a reliable path to route the data and maintain those routes. Table 1. Resumes the preferable areas, advantages and limitations of several routing protocols. The protocols were classified according to [10] in five classes: topology (reactive, proactive or hybrid) localisation, cluster, broadcast and geocast while in [11] they have been categorised depending on their message dissemination strategy: unicast, broadcast, multicast or geocast and cluster. In unicasting technique the messages are transmitted individually based on topology. The routes or next hops are maintained in a routing table of each vehicle, in two methods:

- Reactive (routes created on demand) like Ad hoc On-demand Distance Vector (AODV)[12]
- Proactive (routes updated periodically) such Destination Sequence Distance Vector (DSDV)[13] and Optimised Link State Routing protocol (OLSR) [14] that uses shortest path algorithm to provide a single route without loop to the destination which is stored in the routing table.

Unicasting can also be based on the position of the vehicle that is collected from GPS or other positioning systems like GPSR, GSR, GPCR. Broadcasting routing protocols are the most used for safety applications, they are divided into five classes :traffic (DVCAST)[15],area, location, cluster or probability. Cluster based routing protocols divides the network into smaller regions and delegate a cluster head that serves as a collector node that collects all data messages from his zone and as a bridge that transmits to other cluster heads in the other zones. Multicast is like broadcast technique but here the concept of groups is introduced, so instead of broadcasting to all nodes a zone of relevance is chosen, with a forwarding zone where packets are forwarded, not flooded. Multicasting is also sub categorized into topology, or location-based. In this document, our interest is on DSDV, AODV and OLSR protocols. The mechanism of each one is detailed below:

3.2.1 AODV Ad-hoc On-demand Distance Vector

Ad-Hoc On-demand Distance Vector is a reactive protocol, which means that it, establishes route when needed, rather than the routing tables in each node that must be updated in every topology change. Researchers give this protocol lot of interest, interpreted by the huge number of papers that presents its ameliorations in different phases [16].

3.2.2 DSDV Destination Sequence Distance Vector

DSDV implements a distance vector strategy that uses to shortest path and guarantee free loop routes. A sequence number labels each entry in the routing table with total number of hops from source to destination.

3.2.3 OLSR Optimised Link State Routing

Despite of DSDV, OLSR maintains a table of all possible routes to each node in the network what may cause network congestion. The routing tables are broadcasted in every topology change.

Table 1. Routing protocols advantages and limitations

categories	Sub categories	Most popular protocols	Preferred scenarios	Advantages	Limitations
Topology based	Proactive	OLSR	Urban	Desirable to use in a dense network	Bandwidth consumption by unused paths
	Reactive	AODV	Urban	Reduces memory requirements	Takes time to establish a route
	Hybrid	TORA	Highway	Reduce network overhead	Not scalable
Location based	Delay Tolerant	VADD	Urban	High PDR	If the topology is changing quickly, it causes delay
	Non-Delay Tolerant	GPCR	Urban	Good scalability	Depends on junction nodes
Cluster based		LORA-CBF	Highway	Good scalability	Causes delay
Broadcast based	Traffic	DV-CAST	Highway	Reliable transmission	Consume large amount of network bandwidth
	Area	UMB	Urban	Can overcome interference and higher packets load	Waste of bandwidth.
	Cluster	MDDV	Highway	Message delivery reliability	High control packet overhead
	Probability	REAR	Highway	Good scalability	Causes more delay and high control packet overhead.
Multicast or Geocast based		ROVER	Highway	reliable geographical multicast routing	High delay and network congestion

4 Propagation models

In a wireless environment, the propagation models characterize the way the signals are propagated over the network. Most simulators provides different propagation models to give a better study of the environment. Those models are classified into three types: abstract, deterministic and stochastic. The abstract models are to avoid in transportation system simulations because they cannot reflect the reality. Stochastic ones needs hard mathematical calculations but gives realistic mean. The computational efforts that does the stochastic models are heavy for the machines, for this purpose the middle class is preferred to give near realistic results. A clear comparison between propagation models is given in [14]. Over the time, the attenuation of signals vary depending on several variables such geographical position and radio frequency. It is often modelled as a random process. Fading causes communication failures, here we use NAKAGAMI to study the robustness of routing protocols. Since it is able to implement different levels of fading.
It is a stochastic model, where the signal power obeys a gamma Γ distribution with two parameters omega ω = $[x^2]$ the average fading power and mu μ > 0 the severity of fading, where the channel amplitude $x \geq 0$.
The NAKAGAMI Probability Density Function PDF is represented according to equation (1):

$$f(x; \mu, \omega) = \frac{2\,\mu^{\mu}}{\Gamma(\mu)\omega^{\mu}} x^{2\,\mu-1} e^{x\,\frac{-\mu}{\omega}x^2} \qquad (1)$$

5 Simulation setup

Most researches on routing protocols assumes a clear environment.So, the performances are somehow idealised.In our simulation we will expose those signals to fading effects using NAKAGAMI. We compare two performance metrics that are packet loss and mean delay. We will distinguish which of the AODV, OLSR, and DSDV is the most resistant protocol under fading circumstances.The simulation parameters are presented in table 2.

Table 2. Simulation setup

Platform	Ubuntu 15.10
NS Version	Ns-allinone-3.26
Area	300*1500 m
Routing protocols	Aodv, olsr, dsdv
Number of nodes	10,20,30,40,50
Mac layer	802.11p
Fading	Nakagami, none
Simulation time	50 seconds

6 Results and discussion

Here we work in a low-density area containing ten to fifty nodes, such in highway or rural environment; we focus on two important metrics: packet loss and mean delay to study the fading impact on three protocols from two classes, reactive and proactive. The augmentation of the mean delay varied from 0,01 to 0,21 seconds when applying fading on OLSR. It raises significantly at 20, 30, and 40 nodes and decrease again at 50. The routing strategy of OLSR produces more delay by submerging the network with control packets. It establishes routes that may not be used. In a vehicular network the safety application doesn't need to maintain lot of paths, with all nodes as it need to deliver warnings with less delay and better reliability. With OLSR the augmentation of packet loss under realistic fading circumstances is higher than the other protocols. DSDV protocol performances were less affected by Nakagami fading than OLSR.As it uses shortest path algorithm and the fading is modelled, as a random process, it could not resist its impact. The shortest path is not necessarily the best as we can deduce from the packet loss increase and decrease when exposing DSDV to fading. In figure .4,the invocation of Nakagami fading didn't affected AODV protocol, while it does increase the mean delay of OLSR and DSDV by 0,21 and 0,02 seconds respectively, when the node number is 40.With 10 nodes the mean delay of the three protocols was dramatically increased than in higher densities.

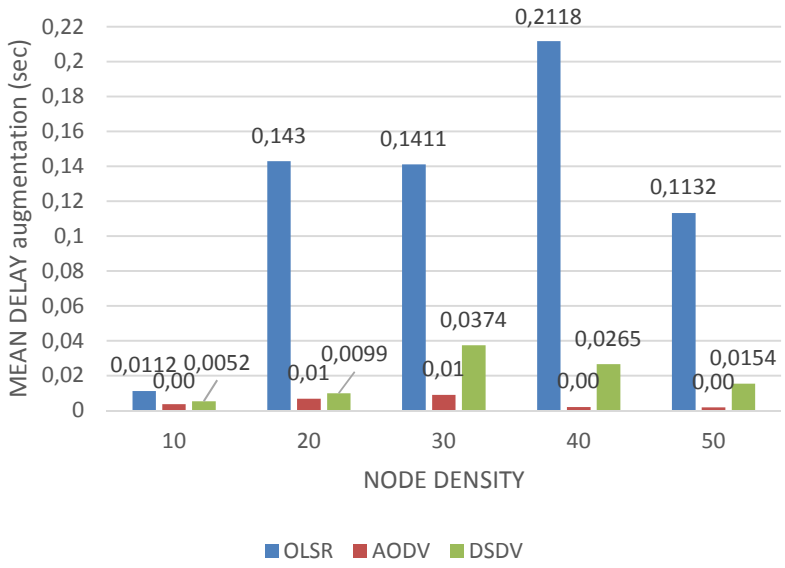


Fig 4.Mean delay increase vs node density

From figure.5,we can perceive that the reactive protocol AODV outperforms the proactive one (DSDV and OLSR) in the range of 20 to 30 nodes and at 50. At 10 node the DSDV protocol gives its best result what means that is more suited for low density areas like rural environment due to the low neighbourhood messages that have to be sent between nodes. This study should be useful to enhance those three protocols to be adapted to a realistic vehicular scenarios.

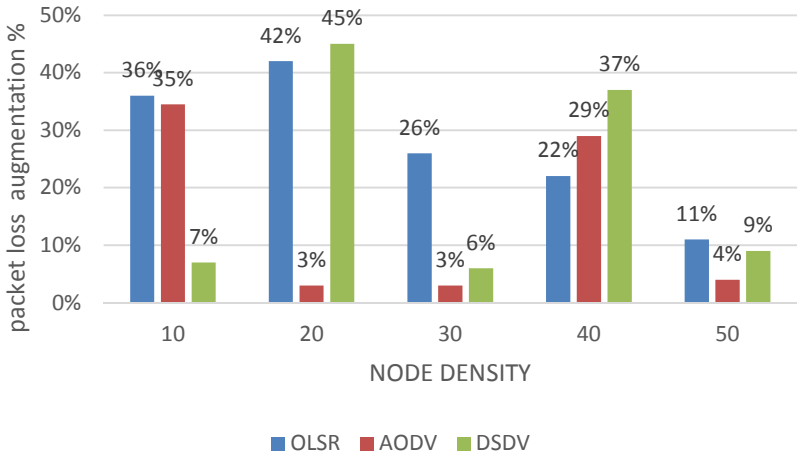


Fig 5. Packet loss increase vs node density

7 conclusion

In this study we evaluated the performances of AODV,DSDV and OLSR protocols using Nakagami fading in low density areas.We cited some VANET specificities and challenges then we over viewed the routing protocols and classified them according to their routing strategies. The Nakagami fading model was then clarified and applied on the routing protocols. The influence of fading model on the simulation results of VANET may vary depending on road types.The large and multi-lane roads in urban environment or wide ones in highways make a very diverse propagation environment, so many variables must be taken in consideration when simulating each scenario. as futur work we will focus on a specific zone on real map and study the behaviour of some relevant routing protocols beside realistic propagation model.

References

[1] Prabhu T. N, Adharsh.M, Ashok Kumar M, Gokul Krishna M., and Dhayanithi G : Vehicle to Vehicle Communication using Light Fidelity. International Journal of Computer Applications (0975 – 8887) Vol. 164 , No 2, Apr- 2017.

[2] Abumansoor O, Boukerche. A : A Secure Cooperative Approach for Nonline-of-Sight Location Verification in VANET. IEEE Transactions on Vehicular Technology, Vol. 61, No. 1, JAN- 2012

[3] Gozalvez J, Sepulcre M,bauza R: Impact of the radio channel modelling on the performance of VANET communication protocols. Telecommun Syst Vol.50, Issue 3, pp 149 – 167, 2012.
[4] Abbas T, Tufvesson F, Karedal J: Shadow Fading Model for Vehicle-to-Vehicle Network Simulations. COST IC1004 2012.
[5] Ali A, and Latha C.A : Performance Analysis of Ad-Hoc Routing Protocols for VANET Using NS2 Simulation. IJCST Vol. 6, Issue 2, April - June 2015.
[6] Maratha B.P, Sheltami T.R and Salah K: Performance Study of MANET Routing Protocols in VANET. King Fahd University of Petroleum & Minerals Vol. 42, Issue 8, pp 31153126 (2016).
[7] Khattak M, Ahmad G :Implementation of Routing Protocols of MANETs in VANETs Using Route Analysis. International Journal of Engineering Works Kambohwell Publisher Enterprises,4 (5), pp. 78-82 (2017).
[8] Dutta R, Thalore R. : A Review of Various Routing Protocols in VANET. International Journal of Advanced Engineering Research and Science (IJAERS), Vol. 4, Issue-4, Apr- 2017.
[9] Tamilselvi R,Kathiresan V: A review on routing protocols in VANET.International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Vol. 6, Issue 9, Sep- 2017.
[10] ur Rehman S, Khan M. A,. Zia T.A,Zheng L : Vehicular Ad-Hoc Networks (VANETs) - An Overview and Challenges. Journal of Wireless Networking and Communications, 29-38 (2013).
[11] Venkatesh, Indra A, Murali R : Routing Protocols for Vehicular Adhoc Networks (VANETs): A Review. Journal of Emerging Trends in Computing and Information Sciences Vol.5, No.1 Jan- 2014.
[12] Perkins C, Belding-Royer E, Das S, quet : Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561, July 2003.
[13] Altayeb M, Mahgoub I. : A Survey of Vehicular Ad hoc Networks Routing Protocols .International Journal of Innovation and Applied Studies ISSN 2028-9324 Vol. 3 No. 3, pp. 829-846 Jul- 2013.
[14] Clausen T.H and Jacquet P : Optimized Link State Routing (OLSR), RFC 3626, October 2003.
[15] Tonguz O. K, Wisitpongphan N and Bai F : DV-CAST: A distributed vehicular broadcast protocol for vehicular ad hoc networks. IEEE Wireless Communications, vol. 17, no. 2, pp. 47-57, Apr- 2010.
[16] Brahmia. H, Tolba .C : Vanet routing protocols: discussion of various Ad-hoc On-demand Distance Vector (AODV) improvements. Proceeding of the The 3rd International Conference on Pattern Analysis and Intelligent Systems, pp. 276-281, OCT- 24-25, 2018.
[17] Stoffers M, Riley G. "Comparing the ns–3 Propagation Models" Modeling,Analysis & Telecommunication Systems (MASCOTS), IEEE 20th International Symposium,61-67 (2012).

Numerical approach of determining the Stress Intensity Factor (KI) for a curved crack problem using Extended Isogeometric Analysis

SAID ELFAKKOUSS[1], HASSANE MOUSTABCHIR[2], AHMED ELKHALFI[3]
{said.elfakkoussi@gmail.com[1] , hmoustabchir@hotmail.com[2], aelkhalfi@gmail.com[3]}

Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University Fez. Morocco[1,3],
Laboratory of Systems Engineering and Applications (LISA), National School of Applied Sciences of Fez. Sidi Mohamed Ben Abdellah University Fez. Morocco[2].

Abstract. The fracture of cracked metal structures remains the main concern of engineers and researchers in various fields (automotive, oil, naval and aeronautics, etc.). They are always looking for ways and tools to locate and determine where and when the break occurs. In the industrial field there are many instruments and devices for measuring and locating defects in metal structures. Prevention tools, however, remain the only way to understand or limit the source of the problem. The most efficient of these tools is the computer which helps to model the fracture phenomenon of cracked metal structures. In this work, we will study the problem of cracking in pipes using new interpolations functions called NURBS (rational non-uniform B-Splines). These functions are added to the formulation of the X-FEM method to study a 2D problem in the linear elastic domain. The geometry studied is a half-tube with a semi-elliptic crack. It should be kept in mind that the main purpose of this study is to calculate the KI factor with good precision, which will be used in determining the rate of rupture.

Keywords: Finite element method (FEM), Extended finite element method (X-FEM), Extended Isogeometric analysis (X-IGA), Semi-elliptical crack, B-Spline, NURBS Function, Interaction integral, Subroutine UEL, the Stress Intensity Factor (KI).

1 Introduction

The stress intensity factor KI is a parameter of linear elastic mechanics of rupture. It is used in particular to determine the degree of damage of cracked structures. Several methods can be used to determine this factor, namely, analytical, semi-analytical and numerical (FEM and X-FEM). Numerical methods remain the most used tools in this field because of thier efficiency in terms of convergence, computation time and the actual modeling of the problem. The convergence of numerical methods is based on computer tools such as the power of machine servers, solvers, and the reliability of the calculation software used as well as the choice of programming language. To apply these tools in the industrial environment, especially to study the fracture of cracked pipelines, we will compute the KI factor numerically for a semi-elliptic crack in a half-tube (2D) by a new method named X-IGA implemented in the Abaqus / Standard software via the Subroutine UEL, based on the FORTRAN language [1]. El Fakkoussi et al. [1] already applied this implantation for a longitudinal crack plane in a half-tube (2D). It should be noted that the IGA and X-IGA approaches are only implemented in the Matlab software. So far, to our knowledge, this method has been partially implemented in the Abaqus / Standard software

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284078

only on structures without cracks. This is why we propose to apply the new approach extended isogeometric analysis to the evaluation of the structure containing a crack and to present a special case to calculate the stress intensity factor (KI) using the approach extended isogeometric analysis for a pipeline containing an external semi-elliptic longitudinal plane crack. Finally, to prove the efficiency of the new strategy, we carried out a comparative study with analytical results witch exists in the literature and numerical extended finite element method (X-FEM).

2 Theoretical Background.

In this section, we will introduce the context of the IGA isogeometric analysis as well as the analytical solution for 2D half-pipe.

2.1 Isogeometric analysis (IGA)

B-Spline function. The B-Spline functions are written in the form of Cox-de Boor recursion formula [2]:

For $p = 0$

$$N_{i,0}(\xi) = \begin{cases} 1 & \xi_i \le \xi \le \xi_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

For $p \ge 1$

$$N_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi) \tag{2}$$

The derivative of the B-Spline function is written as follows:

$$\frac{dN_{i,p}(\xi)}{d\xi} = \frac{p}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) - \frac{p}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi) \tag{3}$$

Non-Uniform Function Rational B-Spline (NURBS). These functions are a generalization of B-Spline functions based on rational functions: the NURBS functions are written as a fraction of two polynomials with real coefficients, in particular the B-Spline functions. These are identical to those used in CAD (computer-aided design). Indeed, these functions check the interpolation conditions [2]. It is important to point out that these functions were used in the isogeometric approach proposed in 2005 by Hughes [2].

$$R_{i,p}(\xi) = \frac{\omega_A . N_{i,p}(\xi)}{\sum_{B=1}^{n} \omega_A . N_{B,p}(\xi)} \tag{4}$$

$$R_A^{p,q}(\xi,\eta) = \frac{\omega_A . N_{i,p}(\xi) . M_{j,q}(\eta)}{\sum_{i=1}^{n}\sum_{j=1}^{m} \omega_A . N_{i,p}(\xi) . M_{j,q}(\eta)} \tag{5}$$

With:

$N_{i,p}(\xi)$ and $M_{j,q}(\eta)$ are the basic functions B-Splines with p and q are the order of the interpolations function, n and m are the number of control points and are the weights corresponding to the control points $T = \{\zeta_1, \zeta_2, ..., \zeta_{n+p+1}\}$ and $H = \{\eta_1, \eta_2, ..., \eta_{n+q+1}\}$ are knots vectors (nodes).

Surface NURBS. To create B-Spline surfaces, it is necessary to build two-dimensional (2D) B-Spline functions. To achieve this, it is sufficeint to use the tensor product which simply consists of multiplying the B-Spline functions into a single dimension (1D) in the two directions of space. The surface NURBS functions are represented by the following equation:

$$S(\zeta,\eta) = \sum_{A=1}^{n.m} P_A . R_A^{p,q}(\xi,\eta) \tag{6}$$

With:

P_A Is the control points, $R_A^{p,q}(\xi,\eta)$ is the bivariate NURBS basic functions.

3 IGA / X-FEM implementation in Abaqus/Standard software

The present work will apply this strategy to the implementation of the isogeometric analysis approach associated with the X-FEM method in the Abaqus / Standard software to study curved cracks in pipelines. Subsequently, the NURBS functions will be implemented in the Abaqus / Standard software via the User defined elements (UEL) subroutine which allows the finite element formulation in this software to be modified. Some other parameters of the X-FEM method have been added in this code, such as the enrichment functions at the front and at the tip of the crack and the Heaviside function. the code UEL operation principle is detailed in figure 2. To make it easier to add additional information to the UEL (User defined elements), we have used the subroutine UEXTERNALDB (User-Defined External Databases). This will subsequently allow the addition of the necessary information on the crack and the NURBS function (figure1). These parameters are then used in the main program UEL (User defined elements). We find that El Fakkoussi et al. [1] have continued the work of Duval et al. [3] , whose goal is to couple the isogeometric analysis approach with the Extended Finite Element Method (X-FEM) to study plane cracking in pipelines.

```
*begin
2          ! problem dimension
3, 3       ! degree of each dimension
4, 4       ! control points in each direction
-----------------------------------------------------------
*knots     ! knot spans
0.0, 0.0, 0.0, 0.5, 1.0, 1.0, 1.0
0.0, 0.0, 0.0, 0.5, 1.0, 1.0, 1.0
0.0, 0.0, 0.0, 1.0, 1.0, 1.0
-----------------------------------------------------------
*bnet     !nurbs control points
0.0000000000E+00 0.0000000000E+00 2.4875000000E+01 1.0000000000E+00
0.0000000000E+00 0.0000000000E+00 2.5000000000E+01 1.0000000000E+00
0.0000000000E+00 0.0000000000E+00 2.5125000000E+01 1.0000000000E+00
0.0000000000E+00 6.2500000000E+00 2.4875000000E+01 1.0000000000E+00
...
*end
```

Fig. 1. Input file NURBS

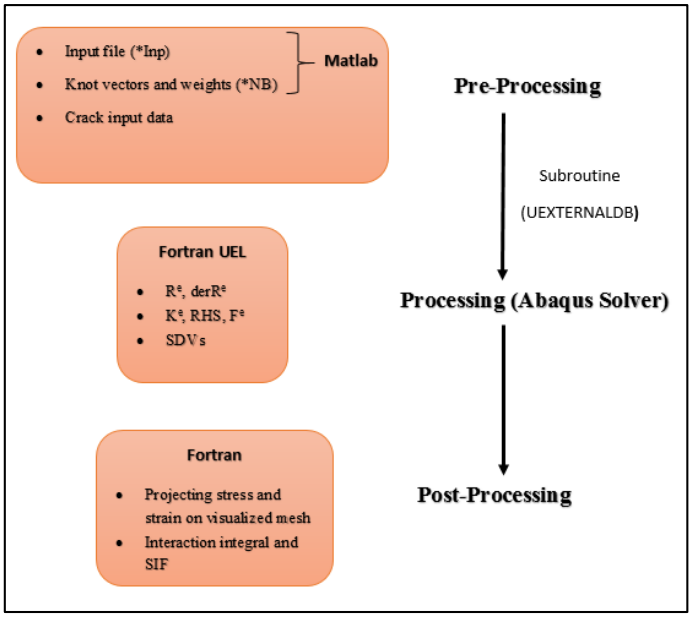


Fig .2. Implementation methodology of X-IGA approach in Abaqus/Standard software [1].

The calculation of results such as stresses and deformations determined by the extended isogeometric analysis (X-IGA) approach is based on the Bézier composition technique, which has been detailed in the references [4]. The extraction of the results and the main variables of the fracture mechanics are determined using the Fortran algorithm. This will help us program equation (7) to calculate, in a later step, the KI Constraint Intensity Factor as shown in Equation (10).

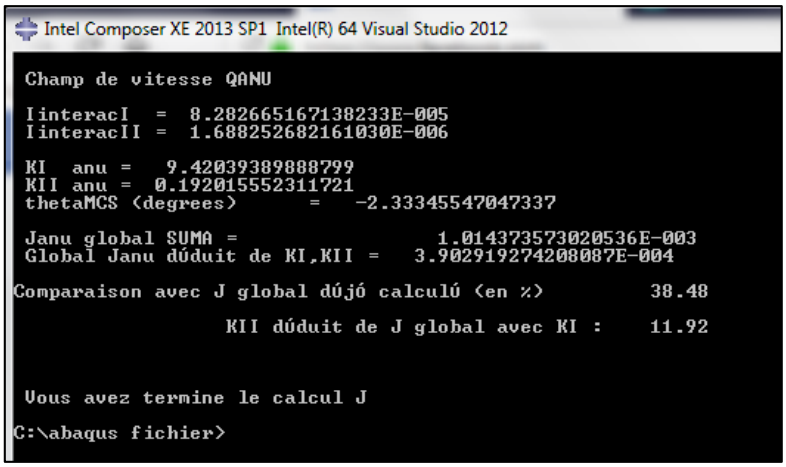


Fig. 3. FORTRAN algorithm for calculation the factor KI.

4 Calculation of the Stress Intensity Factor (KI) for the curved crack

The computation of the Stress Intensity Factor (KI) is based on the integral interaction method [5]. The mathematical formulation of this method for a two-dimensional problem is determined in the following way:

$$M^{(1,2)} = \int_A \left[\sigma^{(1)}_{ij} \frac{\delta u^{(2)}_i}{\delta x_1} + \sigma^{(2)}_{ij} \frac{\delta u^{(1)}_i}{\delta x_1} - W^{(1,2)} \delta_{ij} \right] \frac{\delta q}{\delta x_j} dA \tag{7}$$

With:

q : Weight function

$q = 1$: In the crack tip and $q = 0$ in the contour around the crack

$$W^{(1,2)} = \frac{1}{2}\left(\sigma^{(1)}_{ij} \varepsilon^{(2)}_{ij} + \sigma^{(2)}_{ij} \varepsilon^{(1)}_{ij} \right) = \sigma^{(1)}_{ij} \varepsilon^{(2)}_{ij} = \sigma^{(2)}_{ij} \varepsilon^{(1)}_{ij} \tag{8}$$

$$M^{(1,2)} = \frac{2}{\left(\frac{E}{1-\mu^2} \right)} \left(K^{(1)}_{I} K^{(2)}_{I} + K^{(1)}_{II} K^{(2)}_{II} \right) \tag{9}$$

For mode I, we have $K^{(2)}_{I} = 1$ and $K^{(2)}_{II} = 0$

Finally, we get:

$$K^{(1)}_{I} = \frac{M^{(1,I)} \cdot \left(\frac{E}{1-\mu^2} \right)}{2} \tag{10}$$

The theoretical Stress Intensity Factor (KI) for mode I, in the case of a two-dimensional plate with a curved crack is written as follows [6]:

$$K_I = F_I \sigma \sqrt{\pi c} \tag{11}$$

α / ρ/2c	15.0		30.0		45.0		60.0	
	F_I	F_{II}	F_I	F_{II}	F_I	F_{II}	F_I	F_{II}
0.1	0.9656	0.1616	0.8547	0.3257	0.6746	0.4571	0.4480	0.5327
0.2	0.9623	0.1741	0.8389	0.3557	0.6387	0.5006	0.3890	0.5820
0.4	0.9581	0.1907	0.8207	0.3934	0.6009	0.5532	0.3338	0.6386
0.6	0.9554	0.2024	0.8100	0.4183	0.5826	0.5847	0.3141	0.6675
0.8	0.9533	0.2116	0.8034	0.4361	0.5742	0.6042	0.3110	0.6797
1.0	0.9517	0.2190	0.7992	0.4492	0.5718	0.6152	0.3165	0.6805
Straight	0.9493	0.2544	0.8059	0.4653	0.5946	0.5946	0.3536	0.6124

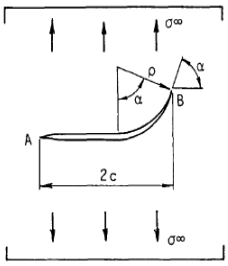


Fig. 4. Dimensionless KI of a crack that consists of a circular-arc part [6].

5 Numerical simulation

Numerical simulation is an important phase in the mechanical study of cracked pipelines, notably with the new X-IGA approach implemented in abaqus / standard software. To achieve this, we have used a half-tube (2D) with curved crack characterized by an inner radius and outer radius Ri = 10 mm, Re = 20 mm and a crack length a = 5 mm was used. The major details of a tube containing a curved crack and its boundary conditions were introduced in figure 5. The model studied is a static elastic linear, where the objective was to extract the values of the Stress Intensity Factor (KI). The Non-Uniform Functions Rational B-Spline used in this study are of the cubic order (p=q=3). The number of elements used to mesh the two-dimensional half-tube with curved crack are respectively 121 and 441 (figure 6). The material used in this study is generally used in gas and oil transport pipelines. the mechanical characteristics are shown in table 2 [7]:

Table 1 : Mechanical properties of the material [7].

Module d'Young	E= 207000 [MPa]
Poisson coefficient	ν =0.3
Yield strength	Re= 340 [MPa]
Ultimate tensile strength	Rm= 440 [MPa]
Elongation at break	A=35 %

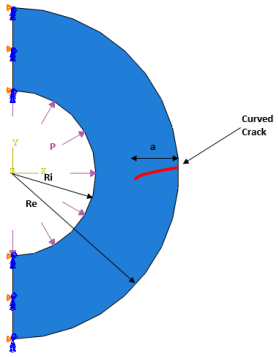


Fig.5. Annular tube with an axial crack and its boundary conditions.

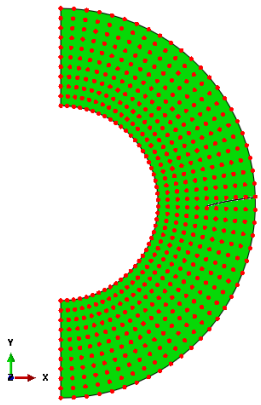


Fig. 6. Control point distribution for the X-IGA half-tube mesh with elliptic type crack.

6 Results and discussion

Numerical simulation was done in the Abaqus / Standard software to implement the new Extended Isogeometric Analysis (X-IGA) approach to study curved cracks. We did the post-processing of the results in the Abaqus / Standard software using a program written in Fortran language. This program calculates the stress intensity factor KI Equation (10) based on integral interaction method the Equation (7). We used three half-tubes, two dimensional with three different cases of the mesh, a mesh of 121, 441 and 691 quadrilateral elements. Then, we compared the results arrived at with the analytical methods, the finite elements FEM and the method extended finite element method (X-FEM). The obtained results are shown in Table 2. These results show a very good agreement with the other numerical results of the FEM method and the X-FEM method which do not exceed an error margin of 1.47%. This means that the implementation of the new approach was successfull.

Table 2 : Comparison of KI factors between the proposed method X-IGA and the analytical, FEM and X-FEM methods for the case of a crack with a = 5 mm.

KI [MPa.mm$^{0.5}$]				
Analytique [6]	*FEM Method*	*X-FEM Method*	*X-IGA 121 éléments*	*X-IGA 441 éléments*
4.75	*4.68*	*4.71*	*4.7209*	*4.7269*
Error (%)	*1.47*	*0.84*	*0.61*	*0.48*

Figure 7 presents the comparison of the Von Mises constraints for the finite element method (FEM) and the extended finite element method (X-FEM) with the new Extended isogeometric analysis approach implemented in the Abaqus / Standard software. We used in this comparison a two-dimensional half-tube with an elliptical crack located in the outer surface of the tube under internal pressure, the study was carried out in the linear elastic domain. The results obtained after the simulation show a good agreement between the X-FEM method and the new X-IGA approach. The difference is due to the singularity problem near the tip of the crack whereas the FEM method is unable to capture the stresses near the crack front.

Simultaneously, we found that the constraints of Von Mises for the X-IGA method are close to the X-FEM method because of the enrichment functions already added to avoid the problem of the singularity of the close stresses at the crack tip.

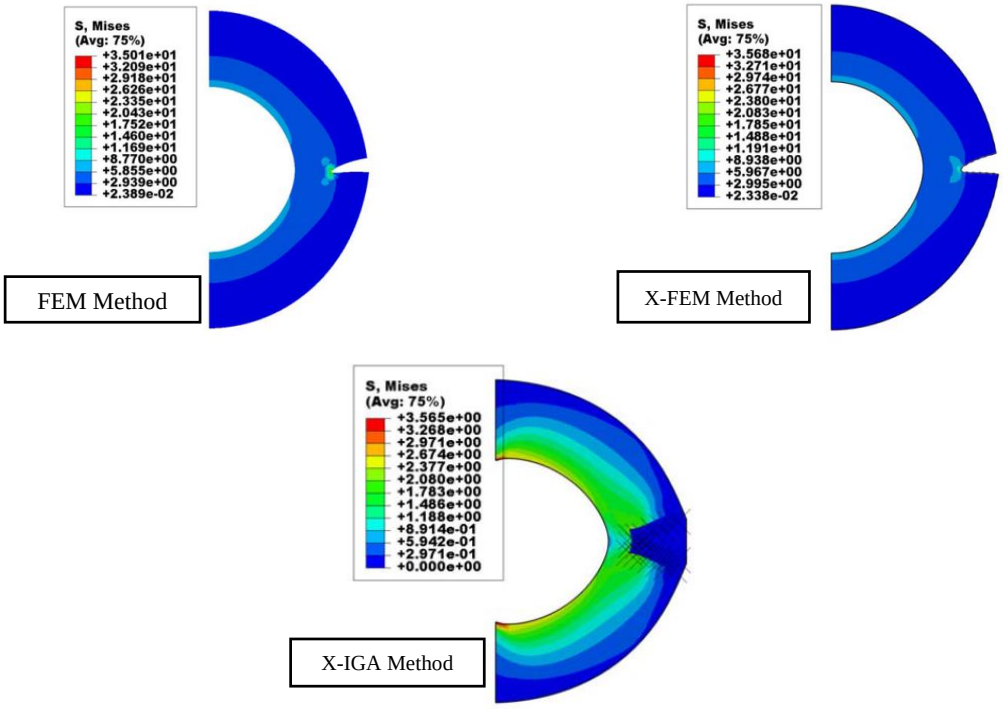


Fig. 7. The von Mises stress for the case of X-IGA, XFEM and FEM.

7 Conclusion

In this paper, we presented the theoretical study extended isogeometric analysis X-IGA. We also treated the formulation of the KI Constraint Intensity Factor for this new approach computed in the linear elastic domain in mode I, for the case of the half-tube with an elliptical crack. Then, we presented the steps of the implementation of the new X-IGA approach in the abaqus / standard software via Subroutine UEL. Finally, we carried out a comparative study between the new approach implemented and the other numerical methods (FEM, X-FEM) to study elliptic cracks.

The new X-IGA approach implemented under the Abaqus / Standard software has been successfully validated by corroborating its results against analytical and numerical computational models. The comparison of errors, shows that the calculation error of the KI factor using the new X-IGA approach is less compared to the other numerical methods. Therefore, We conclude that the use of the new X-IGA approach to study elliptic fissures has proved to be reliable in the study of cracked structures including in pipelines. The reliability generated by the proposed method can help solve the main problems of fracture mechanics and provide better accuracy at an economical cost.

References

[1] S. El Fakkoussi, H. Moustabchir, A. Elkhalfi, et C. I. Pruncu, « Application of the Extended Isogeometric Analysis (X-IGA) to Evaluate a Pipeline Structure Containing an External Crack », Journal of Engineering, vol. 2018, p. 1‑10, oct. 2018.

[2] J. A. Cottrell, T. J. R. Hughes, et Y. Bazilevs, Isogeometric Analysis: Toward Integration of CAD and FEA. Wiley, 2009.

[3] A. Duval, T. Elguedj, H. Al-Akhras, et F. Maurin, « abqNURBS: implémentation d'éléments isogéométriques dans Abaqus et outils de pré-et post-traitement dédiés », in 12e Colloque national en calcul des structures, 2015.

[4] A. Stahl, T. Kvamsdal, et C. Schellewald, « Post-processing and visualization techniques for isogeometric analysis results », Computer Methods in Applied Mechanics and Engineering, vol. 316, p. 880‑943, avr. 2017.

[5] F. Z. Li, C. F. Shih, et A. Needleman, « A comparison of methods for calculating energy release rates », Engineering Fracture Mechanics, vol. 21, nº 2, p. 405–421, 1985.

[6] N.-A. Noda, K. Oda, et K. Ishii, « Analysis of Stress Intensity Factors for Curved Cracks », JSME international journal. Ser. A, Mechanics and material engineering, vol. 37, nº 4, p. 360‑365, 1994.

[7] H. Moustabchir, J. Arbaoui, Z. Azari, S. Hariri, et C. I. Pruncu, « Experimental/numerical investigation of mechanical behaviour of internally pressurized cylindrical shells with external longitudinal and circumferential semi-elliptical defects », Alexandria Engineering Journal, juin 2017

Reference Model of Open Distributed Processing Basic Modelling Concepts in Event-B

Abdessamad Jarrar, Nawfal Ferfra, Taoufiq Gadi, Youssef Balouki
{ abdessamad.jarrar@gmail.com, nawfal.ferfra@gmail.com, gtaoufiq@yahoo.fr, balouki.youssef@gmail.com }
[1]University Hassan First, Faculty of Sciences and Technologies of Settat, Computing, Imaging and Modeling of Complex Systems Laboratory, Morocco
[2]CMR The Moroccan Pension Fund, Rabat, Morocco

Abstract. We present a set of recommendations to help engineers using the Event-B formal methods to specify the basic modelling concepts of the Reference Model of Open Distributed Processing (RM-ODP). This model is developed by the IUT and ISO in order to standarise the development of Open Distributed Processing (ODP). RM-ODP is criticized for inssuficient definition of the basic modeling concepts which limits the applicability of the model. Therefore, the IUT and ISO provide for users a sematic architecture formalization in several formal methods (LOTOS, ACT ONE, SDL-92, Z, and ESTELLE). However, these formal methods are very basic and suffer from the lack of predefined mathematical operators and also some of them are very poor in term of specification techniques. These issues encourage us to develop our recommendations for specifying and developing the ODP using a more sophisticated method called Event-B. This formal methods is very rich in term of predifined mathematical operator and provides sophisticated techniques that can be used during the specification process. Additionaly, Event-B provide a set of verification proofs that highly guarentee the absence of bugs.

Keywords: Open Distributed Processing, RM-ODP, Formal methods, Event-B, architecture semantic formalization.

1 Introduction

Verifying the absence of bugs in Open Distributed Processing systems is a challenging task in the context of software engineering [2]. Therefore, it is highly recommended to use powerful tools and techniques to study these systems. One of the most successful techniques in this domain is formal methods approach. The use of formal methods highly guarantees strong assurance of bugs' absence by means of mathematical modelization and proofs. However, using formal methods is, in general, not an easy task because it requires a very well understanding of the system processing. Therefore, we propose in this paper a set of recommendations in term of a pattern –similar to design patterns- to help engineering to develop systems using the Event-B formal method [1].

Event-B is a formal method that verifies systems' correctness based on theorem proving which is a method-by-method verification technique [3]. This method have been applied to several ODP systems such as Meteor line 14 driverless metro in Paris in which no bug has been detected [4]. Therefore, we chose the Event-B as a formalization and verification method.

This paper contributes to a stepwise facilitating the modelization and verification of ODP systems. We propose a pattern that illuminates how it is possible to model RM-ODP concepts

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284096

using Event-B. The approach proposed is this paper covers the most important basic modelling concepts and specification concepts.

The rest of the paper is organized as follows. Section 2 gives some background on the fundamental concepts of ODP and formal methods. Section 3 presents the contribution of the paper. Section 4 illuminate the approach proposed in term of a set of recommendations and guides on how to formally model RM-ODP basic concepts in Event-B. Finally, we conclude in section 5.

2 background and literature review

2.1 RM-ODP parts

The rapid growth of distributed systems encourages scientists to develop a coordinating framework for the standardization of ODP system. Therefore, the International Organization for Standardization (ISO), the International Electro-technical Commission (IEC) and the Telecommunication Standardization Sector (ITU-T) joint effort to develop the Reference Model of Open Distributed Processing [12]. This reference Model provides an architecture that supports distribution, interworking and portability. It is made up of four part, these are:

1. Overview: this first part contains motivation, scoping, justifications and explanation about the use of ODP architecture. It contains also guide lines of how the RM-ODP can be interpreted and a categorization of required areas of standardization [11].
2. Foundations: this brief part with only 18 pages contains the foundations of RM-ODP. It contains concepts of categories, basic interpretation concepts, linguistic concepts, modelization concepts, specification concepts, organization concepts, system and object properties, designation concepts, behaviour and management concepts, and principles of conformance to ODP standards [10].
3. Architecture: this part includes the specification of the characteristics required for a distributed processing system to be open. These are the constraints that the open distributed processing (ODP) must conform [9].

Architectural semantics: the recommendation presented in this part is an integral part of the RM-ODP. It contains a formalization of ODP modelling concepts obtained by the interpretation of each concept according to the constructions of the different standardized formal description techniques [6].

2.2 Formal methods

Modeling can be defined as the representation of a real-world system either graphically representation or mathematically. It is a stage before construction any system that illuminate its expected structure. Modeling helps also to study and test some system properties to reduce the risk of failures. For example, by modeling correctly an auto-driving cars system, we are able to prove theoretically that it is impossible for two cars to collide. Modeling can be done by means of graphical representation or mathematical equations. The first one is easy to read and can visualize the system structure better, which means it can be shown to clients. However, it is not very accurate in term of representation and its verification techniques are not powerful. On the other hand, mathematical representations need mathematical knowledge to be read and cannot be shown to most clients. It is also not easy to use due to the required accuracy and the huge number of proofs. Despite all that, it is very powerful for detecting

errors and ensures a strong assurance of bugs' absence. Therefore, we think that the use of formal methods based on mathematical representation is highly required.

In event-B, a system is developed as sequence of models. These models are refined in successive steps by making each model richer in term of details. In other words, we start with a very abstract model called initial model, and then we refine it to get a more concrete models. These models are made up of contexts and machines [7]. Contexts are the static parts of models; they are presented in term of sets, constants and axioms, whereas, machines are the dynamic part of models. In a machine, the dynamic status of the system is described by means of variables; the statuses of the system are constrained by invariants. Invariants are mathematical properties describing the necessary condition that must be preserved during the system life-time. Statuses transitions are described by events, which are a set of actions. Each action changes the value of certain variable. Events may have some necessary conditions to be triggered; these conditions are called guards. The figure below illustrates the development process when using event-B:

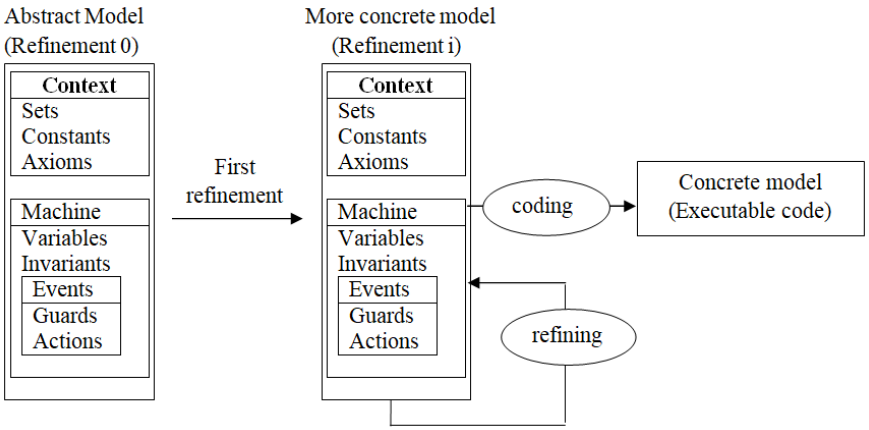


Fig. 2. Process of development in event-B [16]

To ensure that the system will not contain bugs, some mathematical proofs called proof obligations must be performed. One of the most important proofs among proof obligation is invariants preservation that ensures the preservation of all invariant before and after an event trigger. They include also dead-lock freedom to ensure that the system will never be in a status where no guard is verified, which means no event can be trigger. Almost all these proofs are established automatically by means of an eclipse platform called Rodin [8]. It provides the fundamental functionality for syntactic analysis and proof-based verification of Event-B models [8]. In some cases, we may be forced to interfere manually by guiding Rodin in order to perform some proofs. To do so, it may be necessary to specify the hypotheses that Rodin should consider in order to perform proofs. The reason why this should be done is that Rodin ignore some properties if he that they are not necessary for the proof. In other cases, we propose some additional hypothesis that will be proved later separately and will help proving the current desired proof.

3 Contribution

The Reference Model of Open Distributed Processing is mainly based on distributed processing development and the use of formal specification techniques. Whereas, RM-ODP documents are criticized of being very abstract and ambiguous, formal methods provide a very specific models by mean of mathematical representation of distributed systems. In the ITU-T X.904 – Computational formalization [5], the semantic architecture presented in several formal languages (LOTOS, ACT ONE, SDL-92, Z, and ESTELLE).The main contribution of this paper is presenting the architecture semantic using Event-B alongside recommendation on how to develop ODP systems using Event-B. this method provide several advantages compared to other methods such as the independence of temporal ordering, the richness in sense of mathematical predefined operators in addition to the use of techniques to model such as refinement.

Furthermore, Event-B is a formal method based on proof obligations that provide a strong assurance of bugs' absence. Therefore, Event-B is a very suitable formal method to develop ODP systems.

4 Computational formalization

Object modelling is a formalization approach based on abstraction and encapsulation. By means of abstraction we describe the system main functionality separately without considering details, while encapsulation hides all kinds of heterogeneity, security, mechanism, localization, etc [17].

The object modelling concepts cover basic modelling concepts and specification concepts. Basic modelling concepts define the most fundamental concepts of ODP system such as action, objects, interface, and activity [19]. Whereas, specification concepts stands for the reasoning approach used for ODP systems design such as refinement, invariant and compatibility.

In this section, we describe how to formalize the main basic modelling in event-B. We provide also an example of self driving cars along the description to facilitate the understanding of approach.

4.1 Object

An object may be any system component, for example an object may be router or a computer in a network and may be a car in an auto driving cars system. To describe an object in Event-B, a set OBJECTS is defined in the context. After that, we can introduce an object as variable and presented it as an object by mean of a typing invariant (inv 1). In this paper, an abstract recommended model for development of any ODP system is presented and it is up to ready to integrate the concept in the system [20].

```
CONTEXT
        RM-ODP
SETS
        OBJECTS
END
```

MACHINE
RM-ODP
VARIABLES
object
INVARIANTS
inv1 : object ∈ OBJECTS

The rest of concepts are presented along an example of self driving car modelling, therefore the OBJECTS are substituted by a set CARS to facilitate models understanding.

4.2 Environment of an object

The environment of an object is described in term of its relation with it. Any object input or output is considered as a part of the environment. In general, the environment is interacting with object within an event as an event parameter. For example, in an auto driving cars system, the environment which is weather is affecting the car (which is the object) in term of deciding either it is allowed to move or not [21]. The event parameter in this case is wind speed which should be less than about 35 miles per hour to allow car moving. This may be formalized below by an event that prevents car from moving and associate to it the state stopping.

moving_preventing
ANY
wind_speed
car
WHERE
grd1 : wind_speed ∈ ℕ
grd2 : wind_speed ≥ 35
grd3 : car ∈ CARS
THEN
act1 : state(car)=stopping
END

4.3 State of an object

The state of an object is described as a total function from the set OBJECTS to a set of STATES. The set of states is presented in term of enumeration [15]. For example, states moving/stopping can be an enumeration of car states in the auto driving cars system.

SETS
OBJECTS
STATES
CONSTANTS
moving
stopping
AXIOMS
axm1 : partition(STATES,{moving},{stopping})
END

The partition predicate is an easy way to enumerate sets. Mathematically, the partition predicate is defined as follows:

$$partition(S, x, y) \Leftrightarrow x \cup y = S \ \wedge \ x \cap y = \emptyset$$

Where x and y are two subsets of a set S.

The state function that associates to a car a certain state is defined by means of an invariant in the machine as follows:

$$\text{inv2} \ : \ \text{state} \in \text{CARS} \rightarrow \text{STATES}$$

4.4 Action

We must be careful here to avoid confusion between RM-ODP action and Event-B action. The first is an operation that changes the object states, whereas, the event-B action changes the value of a variable. Therefore, we will refer to them in this section as ODP_action and B_action.

An ODP_action is modelled in event-B by the performance of an event. The effect is the instantaneous change in state (or the null change) of the objects with which that ODP_action is associated. An object is defined by a set of variables that describe the current state of the object; the values of these variables may change by means of B_actions in an event. In the auto driving car example, we model the ODP_action that change the state of the car from moving to stopping using an event Stop_car as below:

Stop_car
ANY
car
WHERE
grd1 : car ∈ CARS
THEN
act1 : state(car)≔stopping
act2 : speed(car)≔0
END

4.5 Interface

An abstraction of the behaviour of an object obtained by identifying the operations associated with that object that is to form the substance of the interface. In the auto driving cars example, we may define an interface vehicle in the initial model that express all the possible operation that a vehicle perform such as moving, stopping, accelerating, etc. this can be expressed as an iUML-B representation [14].

4.6 Activity

While Event-B is based on graphical representation, the presentation of an activity as a single headed directed acyclic graph of action does not exist directly in Event-B. However, we can express an activity by forcing a certain order of ODP_actions. This can be done by defining an enumeration denoting the various steps of activity, then we use guard to ensure that the correct event is the only one that may occurs. The steps' transition is done by means of an action the associate to the current step variable the next wanted action.

For example, let A, B, C, and D four ODP_actions that should occur in this order. We define an enumeration of set that include all possible steps beside two additional ones that are: the start and the end. The steps enumeration set is defined as follows:

Partition (STEPS,{start},{step1},{step2},{step3},{step4},{end})

After that, a variable called current_step is defined indicating the current step. The current_step is initialized by start, then a guard in the A event ensure that the current step is the start. A B_action is also added that indicate the next step which is step1. The B event will have a guard that current_step is step1, so on and so forth.

4.7 Behaviour of an object

The behaviour of an object is the set of all possible activities that may occur. The actual activity that will occur depends on the environment of the object and the current state [13]. A very practical way to present the behaviour o an object is by using iUML-B which is a graphical representation of the different states of the object.

4.8 Communication

Communication may be modelled in event-B through the interaction between object. This communication is presented in several forms in the event-B model. In some cases, it is presented as an invariant (for example: two cars should keep a minimum distance between them) or as an event (for example: a car may follow another), etc. all interactions between two objects are guided by a communication between them. In general, the communication between objects is performed by exchanging outputs and inputs of associated variables.

5 Conclusion

Modelling an ODP system that follows the RM-ODP with a significant number of components is, in general, a challenging task that should be performed carefully. Moreover, the use of a sophisticated formal method such as Event-B rise also the complexity of the task, therefore, we propose a number of recommendations and guides that helps modelling such systems. These recommendations and guides illuminate how to model the main concepts of object modelling.

The recommendations proposed in this paper were used in several works [16, 18, 22-25] and all proof obligations were performed correctly using Rodin platform. Therefore, the use of these recommendations highly guarantees the correctness of the model thus the discharging of proofs.

References

[1] Qi Zhang, Zhiqiu HuangJian Xie: Distributed System Model Using SysML and Event-B, International Conference on Machine Learning and Intelligent Communications, MLICOM 2017: Machine Learning and Intelligent Communications pp 326-336(2018).

[2] Ciancarini, P., Poggi, F., Rossi, D., & Sillitti, A. Analyzing and predicting concurrency bugs in open source systems. In Neural Networks (IJCNN), 2017 International Joint Conference on (pp. 721-728). IEEE (2017).
[3] Fayollas, C., Palanque, P., Fabre, J. C., Martinie, C., & Déléris, Y. Dealing with Faults During Operations: Beyond Classical Use of Formal Methods. In The Handbook of Formal Methods in Human-Computer Interaction (pp. 549-575). Springer, Cham (2017).
[4] Lecomte, T., Deharbe, D., Prun, E., & Mottin, E. (2017, November). Applying a formal method in industry: a 25-year trajectory. In Brazilian Symposium on Formal Methods (pp. 70-87). Springer, Cham.
[5] ISO/IEC 10746-4:1998/Amd.1 Information technology — Open Distributed Processing — Reference Model: Architectural semantics — Part 4 AMENDMENT 1: Computational formalization (2001).
[6] ITU Recommendation X.904 I ISO/.IEC CD 10746-4, Open Distributed Processing - Reference Model- Part 4: Overview (1994).
[7] J.-R. Abrial: Modeling in Event-B: System and Software Engineering, Cambridge University Press New York (2010).
[8] C. Rodin, M. Jastram, and M. Butler: User's Handbook.(2011).
[9] ITU Recommendation X.903 I ISO/.IEC 10746-3: 1995, Open Distributed Processing - Reference Model- Part 3: Overview.
[10] ITU Recommendation X.902 I ISO/.IEC 10746-2: 1995, Open Distributed Processing - Reference Model - Part 2: Overview.
[11] ITU Recommendation X.901 I ISO/.IEC CD 10746-1, Open Distributed Processing - Reference Model- Part 1: Overview (1994).
[12] The international standard ISO/IEC 10746, information technology – open distributed processing – reference model: Architectural semantics, first edition (1998).
[13] Asprino, L., Nuzzolese, A. G., Russo, A., Gangemi, A., Presutti, V., & Nolfi, S. (2017). An Ontology Design Pattern for supporting behaviour arbitration in cognitive agents. Advances in Ontology Design and Patterns, 32, 85.
[14] Snook, C., Hoang, T. S., & Butler, M. Analysing security protocols using refinement in iUML-B. In NASA Formal Methods Symposium (pp. 84-98). Springer, Cham (2017).
[15] DONG, Q., JI, M. Q., ZHU, Y. F., & YANG, F. (2018). The CORAS Approach for OPM Based Risk Management. DEStech Transactions on Engineering and Technology Research, (pmsms) (2018).
[16] Jarrar, A., & Balouki, Y. Towards Sophisticated Air Traffic Control System Using Formal Methods. Modelling and Simulation in Engineering, (2018).
[17] Genilloud, G., & Wegmann, A. A foundation for the concept of role in object modelling. In Enterprise Distributed Object Computing Conference, 2000. EDOC 2000. Proceedings. Fourth International (pp. 76-85). IEEE (2000).
[18] Jarrar, A., & Balouki, Y. (2018). Formal modeling of a complex adaptive air traffic control system. Complex Adaptive Systems Modeling, 6(1), 6 (2018).
[19] Cheatham, M., Vardeman II, C. F., Karima, N., & Hitzler, P. (2017, October). Computational Environment: An ODP to Support Finding and Recreating Computational Analyses. In WOP@ ISWC (2017).
[20] Oliveira, I. L., Câmara, J. H., Torres, R. M., & Lisboa-Filho, J. Design of a Corporate SDI in Power Sector Using a Formal Model. Infrastructures, 2(4), 18 (2017).
[21] Laassiri, J. Data Security and risks for IoT in intercommunicating objects. In Proceedings of the 2nd international Conference on Big Data, Cloud and Applications(p. 3). ACM (2017).

[22] Jarrar, A., Balouki, Y., & Gadi, T. Formal Specification of QoS Negotiation in ODP System. International Journal of Electrical and Computer Engineering (IJECE), 7(4), 2045-2053 (2017).
[23] Jarrar, A., Gadi, T., & Balouki, Y. Modeling the Internet of Things System Using Complex Adaptive System Concepts. In Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems (p. 22). ACM (2017).
[24] Jarrar, A., Bellasri, O., Chougdali, S., & Balouki, Y. Formal Specification and Verification of Transmission Control Protocol. In Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems (p. 30). ACM (2017).
[25] Jarrar, A., Balouki, Y., Gadi, T., & Chougdali, S. Modeling Aircraft Landing Scheduling in Event B. In International Conference on Information Technology and Communication Systems (pp. 127-142). Springer, Cham (2017).

GPU optimized parallel implementation of NaSch traffic model

Yassine El Hafid, Abdessamad El Rharras, Mohamed Wahbi and Rachid Saadane
{yassine.el.hafid,a.elrharras,wahbi.mo,rachid.saadane}@gmail.com
SIRC/LaGeS EHTP, EHTP Casablanca, Morocco

Abstract. Traffic simulation has a practical interest for modern society, mainly in minimizing the problems of congestion, which is the main cause of road accidents. In this way, several traffic models were developed based on Cellular Automata (CA). Moreover, the simulation of traffic system requires a high computing ability. Therefore, this paper explores the ways to use the graphics-processing unit (GPU) for reducing time simulation of such applications. We propose several software implementations to maximize GPU performance for NaSch model parallelization. Compared to CPU, the simulation results show that we need optimization effort to get better performance results for GPU.

Keywords: GPGPU, CPU, CUDA, NaSch model, Road Traffic, Parallel computing optimization

1 Introduction

Road traffic is considered as an important economic lever which has a significant impact on modern society. Much research is carried out in order to reduce the problems of congestion, road accidents, or the creation of appropriate infrastructure in a given socio-economic context. Most of the information collected on traffic is based on empirical measurements by proximity sensors, traffic video or aerial and satellite images. The measurements are always done on a set of disparate vehicles in uncontrolled environments. Thus, with the development of information technology tools, the interest of scientists for numerical simulations is quickly established. Several traffic models were developed to approximate the actual drivers behavior and the conditions in which they operate. This type of simulation requires significant computing capability, causing a significant latency time ranging from several hours to several days. Additionally, new processors architectures were immersed in recent years. Designed initially for video games, graphic processors know considerable improvements in performance calculations, number of cores and memory bandwidth integrated in GPU cards. Indeed, some newer graphics cards have up to 5760 calculations cores and up to 24 GB of dedicated RAM; with a relatively affordable price for scientists. Several generations of processors were releases with significant improvements in processor architecture, power consumption and the software framework, allowing access to the features of the GPU. In this context, we took an interest in GPU to implement a basic traffic model developed by

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284233

K. Identify applicable funding agency here. If none, delete this. Nagel and M. Schreckenberg [1]. The GPU could handle a very big number of interacting cars that require memory space and high-speed calculations. Therefore, algorithm optimization to the target GPU is a first step in simulating a large number of cars drivers behaviors. Then, in the second section of this paper, we present different models and methods for road traffic especially NaSch model and we introduce the architecture of Nvidia Cuda parallel programming. The third one, we present detailed implementations of NaSch model, sequential on CPU and parallel on GPU. In the fourth part, we show the experimental results and a discussion of their interpretation. Finally, we conclude the paper and we propose future prospects of our work.

2 Introduction Methods and models for road trafic

2.1 Trafic models

To describe, understand and improve road traffic, studies provide mathematical models inspired from fluid mechanics, which describe the different interactions between the driver and his vehicle. These models include the mobile elements (vehicles, pedestrians ...) into a representation of the entire road system (roads, traffic, steering wheel ...). These studies are helpful to investment in infrastructure. The first traffic studies are born before the Second World War, they are based on an Adams approach [2] using the theory of probability, and on the work of Bruce D. Greenshields (Yale Office of Road traffic) [3] for modeling the evolution of flows, speeds and traffic at intersections. After the war, researchers were interested in tracking the vehicle [4] [5], they developed a theory of traffic flow [6] [7] and the Queueing theory [8]. Later, in the early 1970s, Payne and Whitham [9] were pioneers in proposing a new approach based on the fluid flows from fluids mechanics. More recently, while the car park grows, scientists try to improve the existing methods and propose new models more consistent with the current traffic conditions. To highlight the

Fig. 1. Cyclical road to simulate an infinite highway.

different principles and attempt a classification among all the models, Hoogendoorn and Bovy [10] propose to take the level of detail in modeling (individual element) as classification criteria to determine four groups:

- Microscopic models describe each component of the system in their behaviors and interactions in a high level of detail, for example a vehicle and its driver [4];
- Models submicroscopic, further the level of detail cited above, disclose the operated controls (indicators, speed changes ...) within the vehicle and put them in relation with the environmental conditions [10] [11];
- The mesoscopic models represent traffic in a lower level of detail by small ensembles whose behaviors are described in terms of probabilities [12] [13];
- Macroscopic models treat the traffic in a more global perspective and describe traffic in terms

of speed, flow and density; the specific characteristics of the fluid flow [6] [13].

- **2.2 NaSch model**

NaSch model [1] is a cellular automata model [14] where a section of the road is divided into cells. Each cell has a specific length of 7.5 meters, which expresses the sum of the length of most types of vehicles with inter-vehicle distances, spacing a vehicle and its two nearest neighbors in the case of a blocked traffic. This value of the minimum space occupied by a vehicle is obtained by the inverse of the blocking density 133 veh/Km, measured empirically. A cell is either empty or occupied by a single vehicle. The possibility of accidents is null and overtaking is not allowed. An iteration of NaSch corresponds to a movement of vehicles in a one-second time interval, by a four-phase algorithm. The minimum speed is 7.5 m / s (27 km / h) corresponding to a speed V =1. Moreover, other speed values are integer multiples of 7.5 m/s. As for acceleration, it takes a maximum value of 7.5 m / s (0.76 g). For the simulation, we consider a road that contains L cells, a closed circuit as illustrated in Figure 1.

- PHASE I: ACCELERATION In this phase the vehicle "i" is accelerated by a unit without exceeding the prescribed maximum speed. As well, this phase is given by equation (1).

$$V_i'(t) = \min(V_i(t) + 1, V_{max}) \quad (1)$$

- PHASE II: DECELERATION Since the vehicle "i" is not alone on the road, we must take into account the vehicle $i + 1$ in front of him, whose position is $X + 1(t)$. As well the movement of the vehicle i does not exceed the vehicle position $i + 1$ according to Equation (2).

$$V_i''(t) = \min\left(V_i'(t), d(t)\right) \ \ with\ d_i(t) = x_{i+1}(t) - x_i(t) - 1 \quad (2)$$

$$V_i''\ (t) = min(V_i'\ (t), d(t)) \text{ with } d_i(t) = x_{i_{+1}}(t) - x_i(t) - 1$$

- PHASE III: RANDOMIZATION Equation (3) allows simulating the behavior of drivers in an approximated way, and the speed is decelerated according to the probability *p*.

$$\text{If } V_i''\ (t) > 0 \quad \text{then} \quad (3)$$

$$\begin{cases} V_i(t+1) = V_i''(t) - 1 \text{ with a probability } p \\ V_i(t+1) = V_i''(t) \text{ with a probability } 1 - p \end{cases} \quad (4)$$

- PHASE IV: POSITIONS UPDATING Each vehicle is moved forward, taking into account the speed previously calculated.

$$x_i(t + 1) = x_i(t) + V_i(t + 1) \quad (5)$$

Figure 2 summarizes the four steps in Nash and Figure 3 shows the results that we

obtained by numerical simulation:

2.3 NVIDIA Parallel computing solution

1) The Architecture of NVIDIA GPU:: The first GPU card was released in 2007 called TESLA. Several generations have followed with increasing performances in: memory size, cores speed, the bandwidth of memory and finally, the pipelines configuration of tasks or threads execution. Nvidia made a classification of these processors according to their computing capabilities. There are major differences from one generation to another and even in different versions

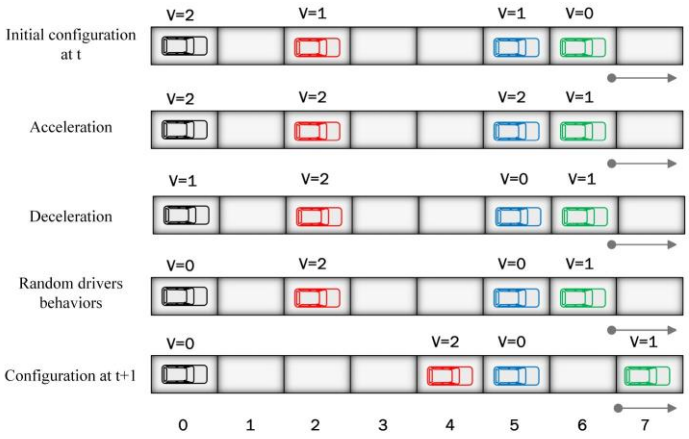


Fig. 2. Moving vehicles in the four NaSch phases.

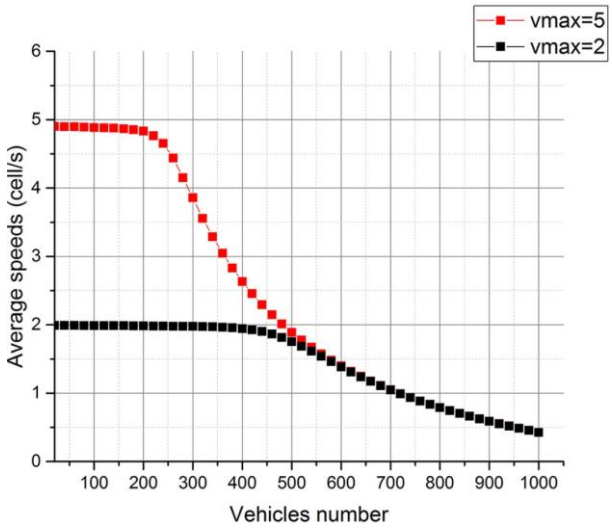


Fig. 3. Average speeds based on the number of vehicles.

of the same generation architecture. Some specific features of the processors categories gives

developers alternatives of parallelization that could reduce significantly the computation time for certain algorithms. A basic solution of GPGPU consists of a main processor CPU named Host, and an auxiliary processor GPU named Device. Each processor has its own RAM. Thus, the host has a main DDR3 RAM, and the Device has GDDR5 RAM. Communication between the two processors is done by PCIE bus. The figure 4 shows the simplified architecture of a GPU on a graphics card. The various types of memory available to the processor enable fast execution of instructions. The main computer language used is C, but it is possible to use other languages like C++, FORTRAN, Java and Phyton. However to access the functionality of the GPU we use an associated API called CUDA (Compute Unified Device Architecture). The manufacturer NVIDIA supplies this parallel computing platform. Thus, in comparison with third-party solutions, CUDA has the advantage of being improved continuously and especially to be optimized for different generations of GPUs. In a Cuda program, there are two types of code: the first, which is sequentially executed by the host, this code call device services via functions named kernels. The Device cores execute threads that are the images of these functions Kernels (second type of code) in parallel and asynchronous way. At the level of software, the threads are banded together in blocks, which their turns are grouped into a computing grid. Each grid corresponds to a kernel called by the host. The number of threads and blocks depends on the programmer choices and hardware limitations on used GPU generation. Threads can access multiple types of memory available to the GPU as shown on the figure 5.

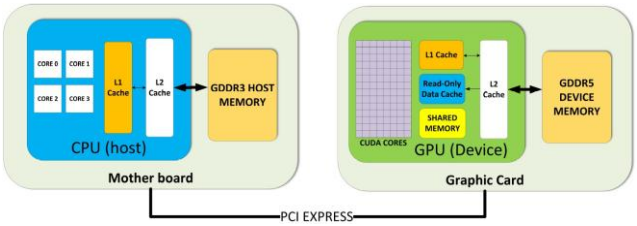


Fig. 4. Hardware architecture of GPU-based high performance computing (HPC).

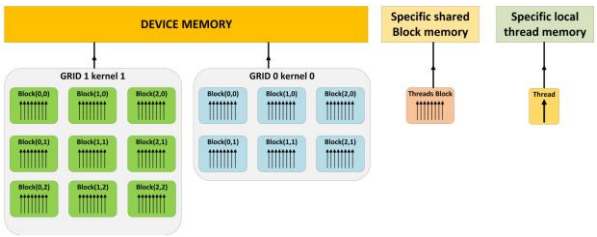


Fig. 5. Programming model organization of GPU parallel computing.

Hardware Implementation: The streaming multiprocessor (SM) is a group of computing cores; their number depends on the generation and the version of the GPU chip. It creates, manages, plans, and executes in parallel several groups of 32 threads called warp. The figure 6 illustrates an SM of Kepler architecture, which has four warps. The threads of the same warp begin together at the same program address. They have their own state registers, and their own counter instruction address, which gives them the ability to run independently. When we give to multiprocessor one or more blocks of threads to be executed, it distributes them into warps then they are organized by warps schedulers for their executions. The way a block is divided into warps is always the same; each warp contains consecutive threads, of threads whose identifier is incremented from the first warp containing the thread 0. A warp executes one common instruction at the same time; the maximum performance is achieved when all 32 threads in a warp follow the same execution path. If the execution of a warp threads diverge via a data-dependent conditional branching, then the warp runs sequentially each path branching, while the threads that do not follow the path are disabled. When all the paths are complete, the threads converge back to the same execution path. The divergence of threads execution occurs only in a warp; different warps run independently, regardless that they execute commons or severed code sequences. The threads in a warp sequentially execute certain type of instruction as atomic operations.

2 Implementation

3.1 Data model

A. Parallel Implementation

Like the sequential implementation, we adopted the same program and data structure. The major difference lies in the calculation delegation of four NaSch phases to the GPU. Moreover, Figure 8 summarizes the first implementation giving unexpected performance results. This fact led us to improve our implementation in stages. Among the experienced improvements, we use atomic operations and reduction of the requests number from CPU to the GPU (CUDA launch). It thus reached a program structure better optimized as shown in Figure 9. In addition, we test several configurations of calculations distributions cores to determine the optimization rules that will be used in future implementations. For each implementation three kernel functions are used, two of which are common to both GPU implementations: setupkernel and load configcar. A third function has two versions, and traficsimulationver2 trafic simulationver3, each applied respectively. A fourth, CalculVmoy, calculates the average speed of the vehicles and used only in the first implementation. The equivalent of this function has been completely merged in the second implementation, which in fact represent the second detailed optimization in the discussion section. The function TraficsimulationVer2 is loaded excessively by the host through the Calculpoint function. However, in traficsimulationver3 we implement Monte Carlo loop, NaSch iterations and speed average calculations within the kernel function, in this way we reduce the time cost of CUDA functions calls.

Fig. 6. Kepler SM architecture [15].

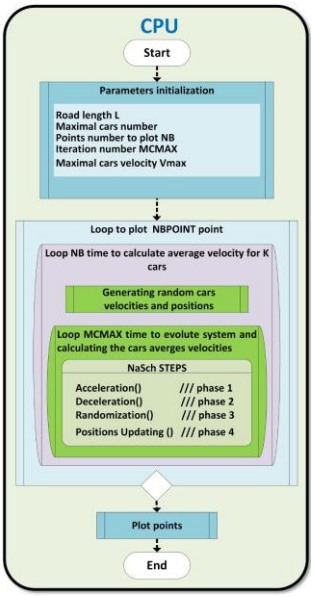


Fig. 7. Sequential implementation on CPU.

B. Sequential implementation

The sequential implementation illustrated in Figure 7 begins with the determination of the simulation parameters, in which we specify the number of cells, the maximum number of vehicles, the number of plotting points, etc. Then three nested loops execute the NaSch iterations to plot the traffic characteristic curves. Moreover, at each Monte Carlo iteration, we reinitialize position and speed parameters for the vehicles. In addition to resetting the configuration, we introduce captor table for measuring the flow in a given position depending on the number of vehicles.

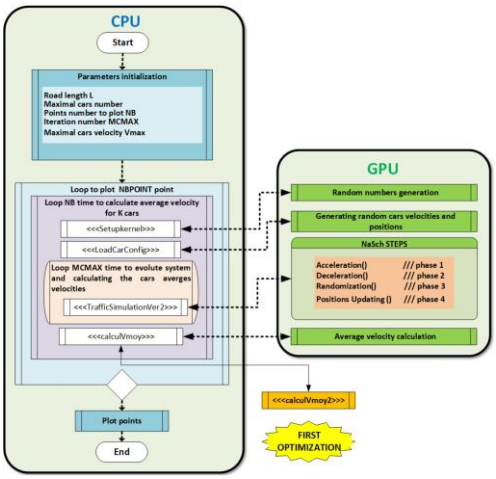


Fig. 8. The first implementation and first optimization for GPU.

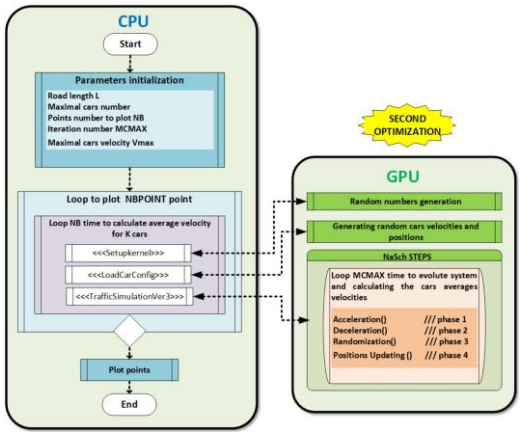


Fig. 9. The second version of implementation for GPU.

3 Result and discussion

3.1 First GPU optimization

As announced in the previous section, the results are the opposite of what we planned. Clearly seen in Figure 10, the CPU is faster compared to the GPU. Therefore, we tried to improve performance by using atomic operations. Indeed, at the beginning a kernel function called CalculVmoy() was used which calculates sequentially the average speed vehicles. The problem with this implementation is that before calculating the average, we have to wait execution of the warp containing the concerned thread. The second version of the function CalculVmoy2() uses the atomic function atomicAdd () that allows different threads from different warp to add the intrinsic speed car. Although this operation is done sequentially on the warp, deferred executions of warps based on the principle of first come, first served, reduces the relatively long waiting time of the previous method.

Moreover, we used a shared variable type to perform summation, giving more efficiency for atomic operations compared to the case of using global memory type variables; this is due to the big latency time to access to this memory type.

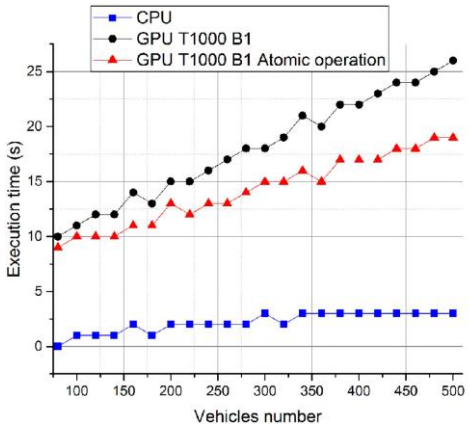


Fig. 10. Execution time CPU versus GPU Implementation 1, compared with GPU Implementation 1 using atomic operation.

A. Second GPU optimization

Despite the improvements obtained by the first implementation, the execution time remains long compared to the CPU. Thus, by using the performance analysis tool Nsight, it was noted that significant time is consumed by the Runtime API functions, specifically the CudaLaunch function to launch kernels. This is why we elaborate a second implementation.

By bringing Monte Carlo loop within the kernel function traficSimulationVer3, temporal cost of kernels launches was reduced, as shown in Figure 11. Again, the performance is not up to what is expected. We noted a significant performance gain for a relatively small number of vehicles. Nevertheless, the two curves are joined to values approximating thousand vehicles. After analyzing our implementation, we note that an atomic function is requested excessively. This function calculates the average vehicle speeds in a single iteration. To address this problem, we adopt the following idea: each vehicle (thread) accumulates its speed throughout the iterations and in the end; we make atomic summation of all values. Figure 12 reflects the improvements achieved in the second optimization. We remark the clear advantage of the GPU compared to CPU and the previous GPU implementations. Moreover, we note that the CPU and GPU curves intersect, as in the previous section. An important point to remember is that we use 1000 threads per block for time measurements in both implementations. Thereafter we explore more computing distribution configurations given the results shown in Figure 13. Thus, for 2000 vehicles, there are significant performance losses in the configuration of two threads by 1000 blocks. Although, the use of a large number of blocks means using more processing cores, so we observe a significant reduction in performance by increasing vehicles number. This could be explained by using the second atomic function, which computes the average speed based on those previously calculated by threads block. Because a block cannot access to the shared memory of its neighbors, the atomic operation is performed with global variable. The time required is therefore more gradually as the number of blocks is increased.

4 Conclusion

The general problem with GPU programming is the need to have strong technical skills to success a parallel implementation of a given algorithm. Thus, in this work we have implemented an evolving version of the NaSch algorithm. The constraints imposed by the traffic model led us to propose improvements in several stages to achieve results in line with our expectations. The implementation methods have a significant impact on computing performance. Using atomic operations and reducing kernel launching has benefit to minimize calculation latency in GPU cores. As a future work we will use our developed model the investigate applications in Deep Learning and Machine Learning [16]. An other perspectives in relation to the algorithms used are cloud and edge computing

References

[1] Nagel, K. and Schreckenberg, M.: A cellular automaton model for freeway traffic. Journal de physique I, 2(12):2221-2229, 1992. [2] William F Adams. Road traffic considered as a random series.(includes plates). Journal of the ICE, 4(1):121-130, 1936.
[3] Greenshields, BD. Channing, Ws. and Miller, Hh.: A study of traffic capacity. In Highway research board proceedings, volume 1935. National Research Council (USA), Highway Research Board.
[4] A Pipes, L.: An operational analysis of traffic dynamics. Journal of applied physics, 24(3):274-281, 1953.
[5] Robert E Chandler, Robert Herman, and Elliott W Montroll.: Traffic dynamics: studies in car following. Operations research, 6(2):165- 184, 1958.

[6] J Lighthill, M. and Beresford, Whitham G.: On kinematic waves. A theory of traffic flow on long crowded roads. In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, volume 229, pages 317-345. The Royal Society, 1955.
[7] Paul I Richards. Shock waves on the highway. Operations research, 4(1):42-51, 1956. [8] L Saaty, T.: Elements of queueing theory, volume 423. McGraw- Hill New York, 1961.
[9] J Payne, H. Models of freeway traffic and control. Mathematical models of public systems, 1971.
[10] Hoogendoorn, S. P. and HL Bovy, P.: State-of-the-art of vehicular traffic flow modelling. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 215(4):283-303, 2001.
[11] Brockfeld, E. Barlovic, R. Schadschneider, A. and Schreckenberg , M.: Optimizing traffic lights in a cellular automaton model for city traffic. Physical Review E, 64(5):056132, 2001.
[12] Maerivoet, S. and De Moor, B. Cellular automata models of road traffic. Physics Reports, 419(1):1-64, 2005.
[13] Leonard, DR. Gower, P. and Taylor. Contram, NB.: Structure of the model. Research report-Transport and Road Research Laboratory, (178), 1989.
[14] Makowiec, D.: The classification of homogeneous and symmetric cellular automata. Work, 1991, 1949. [15] NVIDIA. Nvidia kepler gk110 white paper, 2012.
[16] Abadi, M. et al. TensorFlow: A System for Large-Scale Machine Learning, 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), November 2-4, 2016 • Savannah, GA, USA.

.

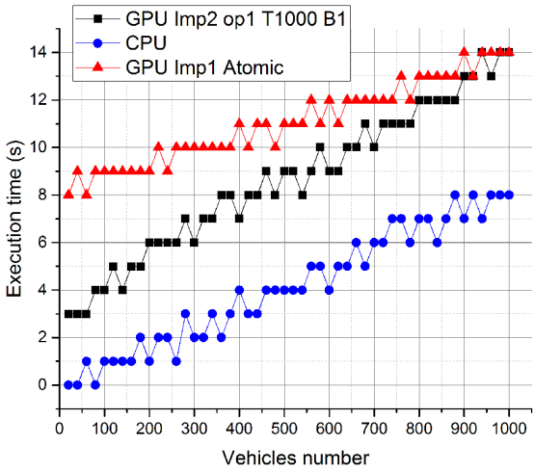


Fig. 11. Execution Time CPU versus GPU Implementation 1 atomic, compared to the second GPU optimization implementation 1.

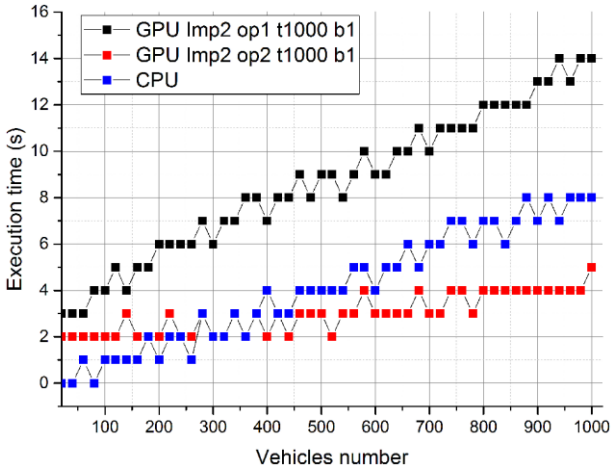


Fig. 12. Execution Time CPU versus GPU Implementation 2 optimization 1, compared to the second GPU optimization implementation.

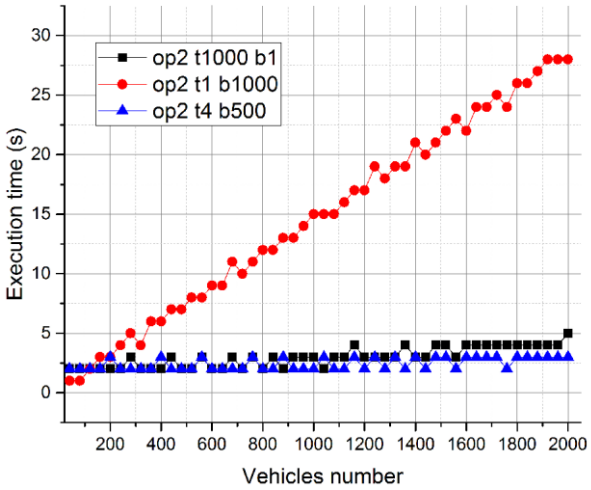


Fig. 13. Execution Time GPU Implementation 2 optimization2 with different configurations calculations distributions

Non Negative Matrix Factorization for Blind Source Separation

[1]Aoulass Nabila, [2]Chakkor Otman,
{ [1]nabilaaoulass@gmail.com, [2]o.chakkor@gmail.com }
[1]Signal and image processing, university Abdlmalek Esaadi ENSAT laboratory NTT,
[2]Telecommunication systems and networks engineering, university Abdlmalek Esaadi ENSAT

Abstract. Non negative Matrix Factorization (NMF) has been a popular representation method for pattern classification problems. It tries to decompose a non negative matrix of data samples as the product of a non negative basis matrix and a non negative coefficient matrix in NMF both supervised and unsupervised mode of operations is used. Among them supervised mode outperforms well due to the use of pre-learned basis vectors corresponding to each underlying sources. In this paper NMF algorithms such as method based in the Frobinuis norm, Kullback Leibler divergence , and an extension to NMF, by incorporating sparsity. Algorithms are used to evaluate the performance of BSS in which supervised mode is used. We further illustrate the effect of hyperparameter as the rank k let the metric chooses and the initialization of decomposition matrices, on the speed of convergence of NMF algorithm.

Keywords: Source Separation, Blind Source Separation, NMF, sparsity.

1 Introduction

The separation of sources is the operation which, from the observations, makes it possible to obtain a set of signals proportional to the sources and to identify the contribution of each sources within the observed mixture. Thus we distinguish two subproblems:

—(i) the identification of the mixture.

—(ii) the reconstruction of the sources.

This opposite problem is badly posed because without any information on the sources and on the mixture, an infinity of solutions would be admissible. It is then necessary to formulate additional hypotheses and to take into account additional information on mixing and sources.

The problem of separation sources can be approached from two points of view. The first is that the decomposition of observations on a basis of elementary signals to eliminate the redundancy of information between the different observations. So, the first methods were proposed by C. Jutten and J. Hérault [1] who realized a nonlinear (ACP) in which we can diagonalize the covariance matrix by the decomposition in eigenvalues (EVD). Due to the limitation of diagonalizable matrices, singular value decomposition (SVD) makes (PCA) always possible

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284117

based on the orthogonality constraint. It offers the least error (with respect to some measures) with the same reduced complexity, compared to other models. But it is not the only. The NMF is used in place of other low rank factorizations, such as the (SVD) [2,3], because of its two primary advantages: storage and interpretability. Due to the non negativity constraints, the NMF produces a so-called "additive parts-based" representation [2,4] of the data. One consequence of this that the factors of decomposition matrix are generally naturally sparse, there by saving a great idea of storage when compared with the (SVD)'s dense factors[5] . But is not for free. On the one hand, the decomposition of the SVD is known to have a polynomial complexity. On the other hand, it has been recently demonstrated that the factorization of NMF has a non-deterministic polynomial computation complexity (NP) [2] for which the existence of an optimal algorithm of a polynomial time is unknown. Moreover, non-orthogonal factors do not allow representation as in (PCA) but are used as a basis for unsupervised or prior modeling for supervised learning [6].

A second, more recent approach is that of the Independent Component Analysis (ICA), it will be necessary to wait for the work of P.Comon [7] to generalize this concept. The latter demonstrates, in the case of linear mixtures, that if the source signals are assumed to be mutually independent and non- Gaussian (except for at most one source), it is possible to separate these signals to a scale factor and a permutation by seeking to minimize the dependence measurements between the estimated signals at the output of the separation system. The implicit objective of the (ICA) is often to find physically significant components. However, in some field of environmental science, and using data that has the property of non-negativity, the solutions estimated by the methods based on the (ICA) lack of physical interpretability. In addition, the (ICA) can not determine the variances (energies) of the independent components as well as the order of the independent sources because the basic functions are classified by non-Gaussianities [3]. In NMF, the non-negativity constraint leads to the representation based on parts of the input mixture that helps to develop structural constraints on the source signals. NMF does not require independent evaluation and is not limited to the length of the data. It provides more important basic vectors for the reconstruction of the underlying signal than the activation vectors [3].

Among the difficulties of matrix factorization in the area of blind separation, the ratio between the number of observations and the number of sources is a problem of interest for a large number of applications and has allowed the taxonomy that we recall below. In most applications and relying on instantaneous linear mixtures, the number m of samples in X is much larger than the numbers n of observations and p of sources. We then separate the determined case p = min (n, m), over-determined $p < n$, finally underdetermined such that $p > min(n, m)$.

When a single-channel source separation problem is considered under-determined, it can not usually be solved without prior knowledge of the sources in the mixture. For this reason, the problem of estimating multiple overlapping

sources from an input mixture is unclear and complex in the (BSS) environment. But (NMF) provides a solution to this single-channel source separation problem by using its non-negativity constraint as well as a supervised mode of operation for source separation [3]. So, NMF is defined as:

$$X \approx F.G \ . \tag{1}$$

Where $X \in R_+^{n \times m}$ is the spectrogram, $F \in R_+^{n \times p}$ matrix of basis vectors (columns), $G \in R_+^{n \times m}$ is the matrix of activations (rows) of the input mixture. In NMF when the spectrogram of mixture X is given, the matrices G and F can be computed via an optimization problem by:

$$\min_{F,G} D(X||F.G) \ . \tag{2}$$

where D denotes the divergence. the reduced dimension p depends on the application and is imposed by the problem that we seek to solve. It is the same for the content of the product matrices that vary also depending on the application and the processed data and can have different physical meanings.

2 Algorithms for solving the NMF problem

2.1 NMF method based on the Frobenius norm :

The multiplicative methods can be obtained in two different ways, either by a heuristic approach, or by a Maximization-Minimization (MM) approach.
Heuristic approach : Multiplicative methods based on the Frobenius norm solve the problem (1) by rewriting it as a matrix trace, i.e.

$$J(G,F) = Tr((X - G.F)^T (X - G.F)) \ . \tag{3}$$

By developing this expression, we can show 3 functions in the expression of J(G, F), that is,

$$\begin{aligned} J(G,F) &= Tr(X^T.X) - Tr(F^T.G^T.X) - Tr(X^T.G.F) + Tr(F^T.G^T.G) \\ &= J_1 - J_2 + J_3 \ . \end{aligned} \tag{4}$$

The calculation of the gradient of each of these three functions is carried out here with respect to a matrix F, while noting that the calculation of that with respect to G is similar.

$$\frac{\partial J}{\partial F} = 2G^T.(G.F \ - X) \ . \tag{5}$$

we can identify the two non-negative functions appearing in the writing of the gradient,

$$\nabla_F^+ = 2GT.G.F \ . \quad (6)$$
$$\nabla_F^- = 2G^T.X \quad . \quad (7)$$

The heuristic approach consists in using these two non-negative functions in the update rules of F, and transpose the results for the update of G :

$$F \to F \circ \frac{\nabla_F^- J(G,F)}{\nabla_F^+ J(G,F)} \quad . \quad (8)$$

$$G \to G \circ \frac{\nabla_G^- J(G,F)}{\nabla_G^+ J(G,F)} \quad . \quad (9)$$

the multiplicative update rules for the Frobenius NMF are available in :

$$F \to F \circ \frac{(G^T .X)}{(G^T .G.F)} \ . \quad (10)$$
$$G \to G \circ \frac{(X.F^T)}{(G.F.F^T)} \quad . \quad (11)$$

The Maximization-Minimization (MM) approach :is based on two steps to obtain the update rules. The principle of this method consists of : first look for increasing the cost function by a function called auxiliary function. then in a second step, to perform the minimization of the auxiliary function.
In the problematic of the classical NMF, we can take the function J (G, F) reduced to its vector formulation and rewrite it in the form of a Taylor development in the second order, considering that , the column current of the matrix F, is the only variable of the problem

$$J(\theta) = J(\theta^k) + \nabla J(\theta^k)(\theta - \theta^k) + \frac{1}{2}(\theta - \theta^k)^T.G^T - G(\theta - \theta^k). \quad (12)$$

Where

$$J(\theta^k) = -G^T(X - G.\theta^k) \quad . \quad (13)$$

the quadratic general form of the auxiliary function H(,) is given as a function of a positive A matrix, Lee and Seung [] propose to choose the matrix A in the form:

$$\theta^{k+1} = \frac{\theta^k o\, G^T x}{G^T.G.\theta^k} \ . \quad (14)$$

The matrix formulation can be obtained by collecting the columns of the matrix F, which gives rise to the multiplicative updating rules for the Frobenius NMF problem :

$$\theta^{k+1} = \frac{\theta^k o\, G^T x}{G^T.G.\theta^k} \quad . \quad (15)$$
$$G \to G \circ \frac{(X.F^T)}{(G.F.F^T)} \quad . \quad (16)$$

2.1 NMF methods based on Kullback Leibler divergence :

The strategy is of type MM and the auxiliary function obtained is based on the concavity of the logarithmic function. The cost function to be minimized is expressed as,

$$D_{KL}(X||F.G) = \sum_{i,j}\left[X\log\frac{X}{G.F}\right] - X + G.F]_{i,j} \ . \tag{17}$$

We adopt the notation f, and x as the respective current column vectors of F and X,

$$J(f) = D_{KL}(X||F.G) = \sum_i x_i \ \log x_i \ - x_i \ + \sum_j G_{i,j} \ f_j \ - \ x_i \ \log G_{i,j} \ f_j \ \ . \tag{18}$$

Using the concavity of the logarithmic function and Jensen's inequality, the previous cost can be increased by the following auxiliary function

$$H(f, f^k) = \sum_i x_i \ \log x_i \ + \sum_j G_{i,j} \ f_j \ - \ x_i \ \ \sum_j \frac{G_{i,j} \ f_j^k}{\sum_l G_{i,j} \ f_l^k} \ (\log G_{i,j} \ f_j log \frac{G_{i,j} \ f_j^k}{\sum_l G_{i,j} \ f_l^k}) \ . \tag{19}$$

The Maximization-Minimization Theorem implies that :

$$J(f^k) > \min\left(H(f, f^k)\right) = H(f^{K+1}, f^k)) \geq J(f^{K+1}). \tag{20}$$

Minimize $H(.,f^k)$ from f:

$$\frac{\partial H}{\partial f_j} = \sum_i G_{i,j} - \frac{f_j^k}{f_j} \sum_i \ x_i \ \frac{G_{i,j} \ f_j^k}{\sum_l G_{i,j} \ f_l^k} = 0 \quad . \tag{21}$$

The minimum is then given by :

$$f^{k+1} = \frac{f^{k+1}}{G.1} \circ \left(G^T \ \frac{x^\circ 1}{G.f^k}\right) \ . \tag{22}$$

By grouping these vectors, we get the update expression of F :

$$f^{k+1} = \frac{f^{k+1}}{G.1} \circ (G^T \ \frac{x^\circ 1}{G.f^k}) \ . \tag{23}$$

By transposition, the rule for updating the matrix G is written :

$$F = \frac{F}{G^T.1}\left(G^T \frac{x^\circ 1}{G.F}\right) \ . \tag{24}$$

2.2 Sparse NMF :

The concept of sparse coding refers to a representational scheme where only a few units (out of a large population) are effectively used to represent typical data vectors [8]. In effect, this implies most units taking values close to zero while only few take significantly non-zero values. In this paper, we use a sparseness measure based on the relationship between the L_1 norm and the L_2 norm [9] :

$$sparseness(x) = \frac{\sqrt{n} - \Sigma|x_i| \Big/ \sqrt{\Sigma x_i^2}}{\sqrt{n} - 1} \quad . \tag{25}$$

where n is the dimensionality of x. Our aim is to constrain (NMF) to find solutions with desired degrees of sparseness. The first question to answer is then : what exactly should be sparse? The basis vectors F or the coefficients G? This is a question that cannot be given a general answer; it all depends on the specific application in question. Further, just transposing the data matrix gnone) must be made by the experimenter. The sparse NMF problem can be formulated as :

$$E(F, G) = \|X - F.G\|^2 \quad . \tag{26}$$

is minimized, under optional constraints :

$$\text{sparseness}(f_i) = S_f, \forall \text{ i} . \tag{27}$$
$$\text{sparseness}(g_i) = \text{Sg}, \forall \text{ i} . \tag{28}$$

Where f_i is the i-th column of F and g_i is the i-th row of G. Here, S_f and Sg are the desired sparsenesses of F and G (respectively). These parameters are set by the user.

3 uniqueness of NMF solution :

Given the formulation of the NMF problem, it is clearly to be feared the existence of invertible square matrices T such that the pairs (WT, $T^{-1}H$) are also solutions of the problem, since

$$F.G = (F.S)(S^{-1}.H) . \tag{29}$$

and that the cost of reconstruction depends only on the product F G. Given a solution (F^0,G^0), the pair(F^0.S,S^1.G) is the solution of the problem if and only if the two matrices F^0.S and (S^{-1}.G) have positive or zero coefficients. At least two types of cases can be exhibited where this is possible :

Trivial invariances : If we insist that S and its inverse S^{-1} have positive or zero coefficients, the F.S and (S^{-1}.G) products are also positive and we are in the presence of a new solution to the initial problem. In this case, it is easy to prove that such a matrix S is necessarily the product of a permutation matrix and a diagonal matrix with positive coefficients [11]. The permutation matrix introduces K degrees of invariance, but does not really change the solution; moreover, the uniqueness of a quantity is often defined at a close permutation. With regard to scale factors (diagonal matrix), the question can be solved by imposing a standardization on one of the factors F or G (in practice, one will often choose to standardize the columns of F in norm L^2). These

invariances are therefore not a real obstacle to a possible uniqueness of the solution, which will be defined by a permutation and a change of scale.
Local invariances : The product ($F^0.G^0$) can also remain invariant on points in the vicinity of the couple that makes it.Suppose that F0 and G0 have strictly positive coefficients. Given a square matrix U, not necessarily with positive coefficients, we can find ε sufficiently small such that (I + εU) is invertible. One can make the limited development : (I + εU) ' (I −εU). The matrices (I +εU) and (I +εU)−1 perform local transformations around ($F^0.G^0$); provided that this point is not situated on the edges of the positive quadrant, and that ε is chosen sufficiently small, the points (F^0(I +εU)) and ((I +εU)−1)G^0 remain in this quadrant. Thus, we obtain a new solution to the problem of NMF, the pair (F^0(I +εU), (I +εU)$^{-1}$.G^0).

4 Non-uniqueness of the general case :

We have considered only pairs of solutions expressing themselves relative to each other via a linear transformation : (F, G) and (G.S,S−1.F). In reality, there is nothing to require that two solutions producing the same product F.G be connected in this way.

$$X=\begin{pmatrix}1&1&0&0\\1&0&1&0\\0&1&0&1\\0&0&1&1\end{pmatrix}=X.I=I.X \quad . \qquad (29)$$

where I denotes the identity matrix. The matrix V is of rank 3.We can choose as factorizations (F^0 = X,G^0 = I) or (F^1 = I,G^1 = X) There is no invertible matrix S such that $F^0S = F^1$, for reasons of rank. However, if we choose K = rg (X), we can show that such counterexamples are impossible. In this case, all solutions are connected to each other by linear transformations [11].

5 The disadvantages of NMF :

Geometrically, the (NMF) consists in finding a cone belonging to the positive orthant which includes the components of the vectors of the observed data [10]. From this point of view, the cone is not always unique without additional constraints. From this geometrical interpretation, on the one hand, it appears that the (NMF) is not unique ,which poses a problem for the BSS. On the other hand, in (NMF), the criteria can be convex only according to one of the two matrices produced but not for both. The algorithms therefore only allow to converge towards a local minimum. Therefore, the convergence result strongly depends on the initialization of the algorithm. practically, it is not guaranteed that the decomposition obtained an important interpretation. To avoid this problem, it is necessary to exploit certain preliminary information or to impose certain constraints on the decomposition. For example, information from locations or special signals is used in a so-called supervised (NMF) in [3]. This method improves the accuracy of automatic transcription, but requires well-organized advance information. Another strategy is to rely on specific constraints from the characteristics of the processed signals. However, This conditions are very difficult to satisfy in the case of real data.

6 The advantages of NMF :

The NMF is used in place of other low rank factorizations, because of its two primary advantages: storage and interpretability. Due to the nonnegativity constraints, the NMF produces a so-called “additive parts-based” representation of the data. One consequence of this is that the factors F and G are generally naturally sparse, there by saving a great deal of storage. The NMF also has impressive benefits in terms of interpretation of its factors. So, the basis vectors naturally correspond to conceptual properties of the data.

7 Application :

In this section, we present the empirical evidences that support NMF as a successful document clustering and topic modeling method.consider a text processing application that requires the factorization of a term-by-document matrix X. In this case, k can be considered the number of (hidden) topics present in the document collection. In this case, F becomes a term by-topic matrix whose columns are the NMF basis vectors. The non zero elements of column 1 of F, which is sparse and nonnegative, correspond to particular terms. So we compare the clustering quality between multiplicative update (MU) NMF based on Kullback Leibler divergence and other based on the Frobenius norm , Within the sparse NMF algorithms we compare the multiplicative updating (MU), Coordinate Descent(CD) wtih defferent initializations nndsvd, nndsvdar, random.

7-1 Data Sets ”20 Newsgroups”:

This data set consists of 20000 messages taken from 20 newsgroups. One thousand Usenet articles were taken from each of the following 20 newsgroups. Approximately 4 % of the articles are crossposted. The articles are typical postings and thus have headers including subject lines, signature files, and quoted portions of other articles.

7-2 Determining a Suitable Metric When using Non-negative Matrix Factorization :

In this paper we describes algorithms for nonnegative matrix factorization (NMF) with two cost functions of divergence.This cost functions parametrized by a single shape parameter beta that takes the Euclidean distance, the Kullback Leibler divergence and the Itakura-Saito divergence (the Itakura-Saito theory is established in [2]) . So to examine the effect of different divergences on convergence speed, we test them on multiplicative update.

Table 1. the speed of convergence as a function of defferences cost function taken by fixing the number of components at 10.

Methode NMF	Frobenius	KL	Itakura-Saito
time (s)	1.509	0.277	5

7-3 The effect of inisialization NMF with Sparseness contraint :

The NMF must be initialized and the initialization selected is crucial to getting good solutions. It is well-known that good initializations can improve the speed and accuracy of the solutions of many NMF algorithms. Add to this the fact that many NMF algorithms are sensitive with respect to the initialization of one or both NMF factors, and the impact of initializations becomes very important. In this section, we compare the results of three initialization procedures for sparse NMF :Nonnegative Double Singular Value Decomposition (NNDSVD), NNDSVD with zeros filled with small random values (NNDSVDar) and random initialization :

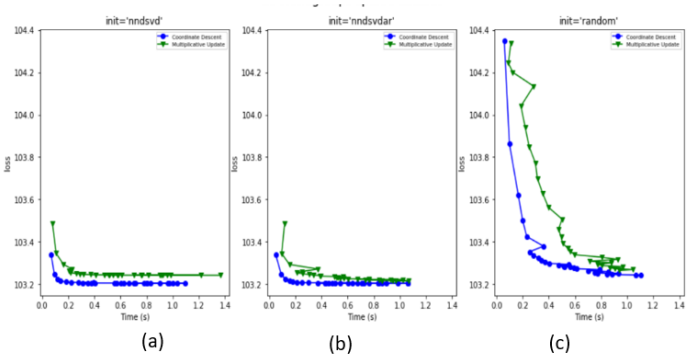


Fig. 1 graphs of different solvers taken with different inisializations: NNSDVD in (a), NNSDar in (b) and random inisialization in (c) .

.

Coordenate desent is established as the best algorithm for sparse NMF, and NN SDVD, NNSDar inisialization (graph (a), resp (b)) plays an important role to minimize convergence time and loss value. In all cases, a good initialization can improve the speed and accuracy of the algorithms, as it can produce faster convergence.

8 Conclusion :

Coordenate Non-negative matrix factorization (NMF) has proven itself a useful tool in the analysis of a diverse range of data. One of its most useful properties is that the resulting decompositions are often intuitive and easy to interpret because they are sparse. However, the sparseness achieved by NMF is not enough; in such situations like in extarction topic it might be useful to control some hyperparametres such as the inisialization of decompositions matrix that the best results are obtained by NNSDVD, NNSDar inisialisations for coordinate Descent

(CD) and suitable metric has also are markable effect on the convergence speed of the algorithm. So, we see that the best results for multiplicative update are obtained from Frobenius norm.

References

[1] Jutten and J. Herault, independent component analysis,1991.

[2] Nonnegative matrixfactorisation alorithms and applications, NGOC-DIEP HO,2008.

[3] An Experimental Survey on Non-Negative Matrix Factorization for Single Channel Blind Source Separation,Mona Nandakumar M and Edet Bijoy K,2014.

[4] Traitement des signaux parcimonieux et applications,Mohamed Aziz Sbai,2012.

[5] Algorithms, Initializations, and Convergence for the Nonnegative Matrix Factorization,Amy N. Langville ,Carl D. Meyer ,Russell Albright ,James Cox ,and David Duling,2014.

[6] NMF Factorisation par matrices non négatives,wikisat.

[7] Méthodes informées de factorisation matricielle non négative, Abdelhakim Limem,2017.

[8] O. Hoyer. Modeling receptive fields with non-negative sparse coding,2003.

[9] Non-negative Matrix Factorization with Sparseness Constraints,Patrik O. Hoyer,2004.

[10] Méthodes de séparation aveugle de sources et application à la télédétection spatiale ,Moussa Sofiane Karoui,2012 .

[11] Les factorisations en matrices non-négatives. Approches contraintes et probabilistes, application à la transcription automatique de musique polyphonique,Nancy Bertin,2010.

Reconstruction of three-dimensional scenes from a sequence of images which only one is calibrate

[1]K. Karim, [2]M. Es-sabry, N. El Akkad, M. Merras; A. Saaidi, K. Satori
{ [1]karima.karim35@gmail.com, [2]essabry.mohammed@gmail.com, [3]nabil.elakkad@usmba.ac.ma,
[4]merras.mostafa@gmail.com, [5]abderrahim.saaidi@usmba.ac.ma, [6]khalidsatori@gmail.com }

LIIAN, Department of Computer Science, Faculty of Science, Dhar El Mahraz, Sidi Mohamed Ben Abdellah, University, B.P 1796, Atlas, Fez, Morocco.

Abstract. In this work, we propose a method for reconstructing three-dimensional scenes from a sequence of images which only one image is calibrate. The system is initialized by a calibrated image. The projection matrices of the other images in the sequence are estimated from the first projection matrix (obtained by the calibration of the camera), and the matching between the images is determined by solving a system of linear equations. The intrinsic and extrinsic parameters of the camera used are estimated from the Cholesky decomposition of the projection matrix. Finally, obtaining the 3D scene is based on the Euclidean reconstruction of the detected and matched control points in the pairs of images.

Keywords: 3D reconstruction; matching; Projection matrix; Calibration and Autocalibration.

1 Introduction

3D reconstruction is a technique for obtaining a three-dimensional representation of an object from a sequence of images taken by different views ([27], [28], [29], [30], [31], [32]). Most 3D reconstruction techniques use calibration or autocalibration methods. Calibration methods use a known object (a calibration pattern 2D or 3D...) to find the cameras parameters, the projection of some points (in the image planes) which characterize the objects permit to solve a system of linear equations. Among the calibration methods, we have those presented in ([1], [2], [3]). The methods of autocalibration of the cameras are based on the estimation of the camera's parameters but without any knowledge a priori on the stage. Among these methods, we have those treated in ([4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [20], [21], [22], [23], [24], [25], [26]). The cost functions of these methods are generally non-linear, and they are formulate according to the invariants in the images and camera parameters.
In this work, we treat a three-dimensional scene reconstruction method from a sequence of images with a single image calibrated. Our system is initialized from a calibrated image.
The rest of this work is organize as follows: the second part presents some notations used in this work. The third part presents the pinhole model of the camera. Detection and matching of interest points will presented in the fourth part. The fifth part deals with the determination of the projection matrices and the autocalibration of the cameras. Experiments on real data will presented in Part 6 and the conclusion to this work is presented in Part 7.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284139

2 Notations

M_1 is a point in the 3D scene whose homogeneous coordinates are $\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$

$P_{i1}\ et\ P_{j1}$ are the projection matrices of point M_1 on the image planes respectively $i\ et\ j$

$m_{i1}\ et\ m_{j1}$ are the projections of the point M_1 on the image plane respectively $i\ et\ j$

3 Camera model

We use in this work the pinhole model of camera **Figure1.** which represents geometrically the perspective projection; this model is used to project the scene in the image plane. It is characterized by a 3 × 4 matrix

R_i : The matrix of rotation, t_i the translation vector of the camera and K_i represents the matrix of intrinsic parameters such as:

$$\begin{pmatrix} f_i & s_i = 0 & u_{0i} \\ 0 & \varepsilon_i f_i & v_{0i} \\ 0 & 0 & 1 \end{pmatrix}$$

f_i : is the focal length, ε_i is the scale factor (we consider that the pixels are square, so $\varepsilon_i = 1$) , s_i : l is the image skew, and $(u_{0i} \quad v_{0i})$: are the image coordinates of the principal point.

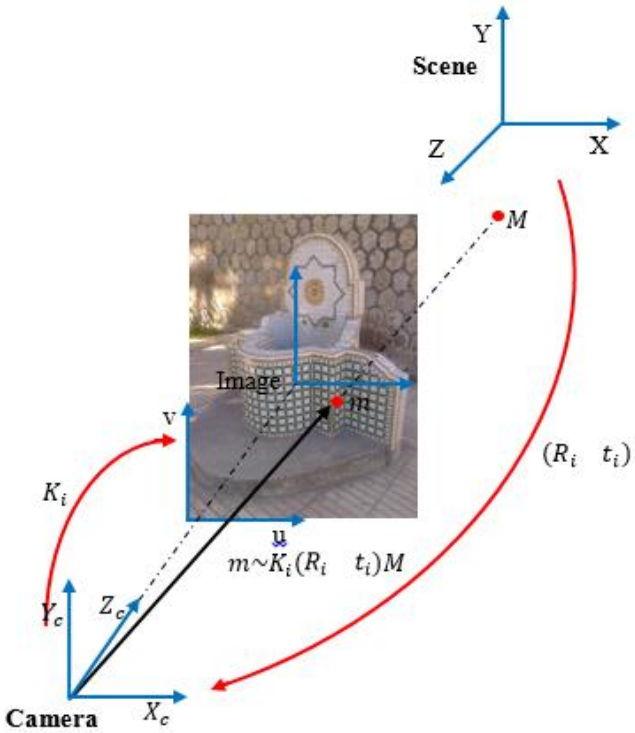


Fig. 1. Pinhole model.

4 Interest points and matching

4.1 Interest points

There are several methods to detect the interest points, among these methods we have those presented in ([15], [16]). We based in this work on the Harris detector [16]. The basic idea of Harris detector is to use the self-correlation function to determine the places where the signal changes in the two directions simultaneously.

$$G = \sum_{u,v} W(u,v) \otimes \begin{pmatrix} I_u^2 & I_u I_v \\ I_u I_v & I_v^2 \end{pmatrix} \tag{1}$$

With:

$$I_u = \frac{\partial I(u,v)}{\partial u};\ I_v = \frac{\partial I(u,v)}{\partial v} \tag{2}$$

$$W(u,v) = \frac{1}{2\pi\sigma^2} Exp\left(-\frac{u^2+v^2}{\sigma^2}\right) \quad (3)$$

Harris proposes to calculate a measure based on the determinant and the trace of G by the following formula:

$$H = \det(G) - \lambda trace(G) \quad avec\ \lambda \in [0.04 \quad 0.06] \quad (4)$$

If $H > 0$: Area containing an interest point.

4.2 Matching of interest points

The matching of interest point is an important step to calibrate the cameras‘s used; several works ([17], [18], [19]) are developed on this field. The correlation measure used in this work is the Zero mean Normalized Cross Correlation (ZNCC). [19]

$$ZNCC(s_1,s_2) = \frac{\sum_{i=-N}^{N}\sum_{j=-S}^{S}(M-E)(H-F)}{\sqrt{\sum_{i=-N}^{N}\sum_{j=-S}^{S}(M-E)^2 \sum_{i=-N}^{N}\sum_{j=-S}^{S}(H-F)^2}} \quad (5)$$

With:

$$M = I_1(u_1+i, v_1+j) \quad (6)$$

$$H = I_2(u_2+i, v_2+j) \quad (7)$$

$$E = \frac{1}{(2N+1)(2S+1)} \sum_{i=-N}^{N}\sum_{j=-S}^{S} I_1(u_1+i, v_1+j) \quad (8)$$

$$F = \frac{1}{(2N+1)(2S+1)} \sum_{i=-N}^{N}\sum_{j=-S}^{S} I_2(u_2+i, v_2+j) \quad (9)$$

And $(s_1,s_2) = ((u_1,v_1),(u_2,v_2)$ are the coordinates of the pixels considered.

5 Estimation of the projection matrices and autocalibration of camera

5.1 Estimation of the projection matrices

We know that the projection of the point $M_1$of the 3D scene on the image plane i is written in the following form:

$$m_{i1} \sim P_{i1} M_1 \quad (10)$$

Knowing that P_{i1} is obtained from calibration of the camera in the first view (image).

The previous equation contains two equations with 3 unknowns, which are the coordinates of M_1.

The projection of the point M_1 of the 3D scene on the image plane j is written in the following form:

$$m_{j1} \sim P_{j1} M_1 \tag{11}$$

The previous equation contains two equations with 14 unknowns. 11 unknowns for P_{j1} and the 3 unknowns of M_1.

To determine the coefficients of P_{j1}, we need 12 matches to generate a system of 48 linear equations with 47 unknowns.

There are 11 unknown for P_{j1} matrix and 36 unknowns for the 3D points (because we have 12 matches, therefore we have 3D points and each 3D point has 3 coordinates «unknowns »).

By solving the system, we will obtain the projection matrix P_{j1} and so on until the end of the image sequence.

5.2 Autocalibration of camera

5.2.1 Extraction of intrinsic parameters :

The projection matrix of the points of the 3D scene on the image plane is defined as follows:

$$P_j \sim (K_j R_j \quad -K_j R_j t_j) \tag{12}$$

We put $\overline{P}_j$ the 3 × 3 sub-matrix composed of the first three columns of P_j. We have:

$$\overline{P}_j \sim K_j R_j \tag{13}$$

If we multiply each side of the equation with its transpose:

$$\overline{P}_j \overline{P}^T{}_j \sim K_j R_j R^T{}_j K^T{}_j \tag{14}$$

Since R_j is orthogonal: $R_j R^T{}_j = I$. Then:

$$\overline{P}_j \overline{P}^T{}_j \sim K_j K^T{}_j \tag{15}$$

We see on the left a symmetric and positive definite matrix.

Its Cholesky decomposition gives an upper triangular matrix B_j with:

$$\overline{P}_j \overline{P}^T{}_j \sim B_j B^T{}_j \tag{16}$$

If we compare this equation with the preceding equation, while reminding us that K_j is upper triangular, we then find that B_j is indeed the calibration matrix K_j sought or almost: the element (3, 3) of K_j must be equal to 1. We then multiply B_j with the appropriate scalar in order to satisfy this constraint. Then we can extract the intrinsic parameters of B_j (or K_j).

5.2.2 Extraction of extrinsic parameters

Once the intrinsic parameters have been extracted, the calculation of the extrinsic parameters is no longer very difficult. We have :

$$P_j \sim (K_j R_j \quad -K_j R_j t_j) \tag{17}$$

Then :

$$K^{-1}{}_j P_j \sim (R_j \quad -R_j t_j) \tag{18}$$

We put A_j the 3 × 3 sub-matrix consisting of the first three columns of ${K^{-1}}_j P_j$. So we have $A_j \sim R_j$ We can make this equation exact (make disappear the symbol ~) by multiplying A_j with an appropriate scalar λ:

$$\lambda\, A_j = R_j \tag{19}$$

How to choose λ ?

The equality of the matrices implies the equality of their determinants, from which we obtain:

$$\lambda^3 \det(A_j) = \det(R_j) = 1 \tag{20}$$

So, we choose $\lambda = \sqrt[3]{1/\det(A_j)}$

Now :

$$\lambda\, {K^{-1}}_j P_j = (R_j \quad -R_j t_j) \tag{21}$$

(Note the exact equality: =). The first three columns of the left matrix then give us directly the rotation matrix R_j. Once R_j determined, the calculation of the vector t_j is trivial.

$$\lambda {R^{-1}}_j\, {K^{-1}}_j P_j = (I_{3*3} \quad -t_j) \tag{22}$$

6. Experimentations

In this phase, we took a differents views of a sequence of 19 480 × 640 images of a three-dimensional scene unknown by a CCD camera knowing that the first image is calibrate.

Figure 2.shows two of the 19 images of the used sequence.The interest points in these two images are detected by the Harris algorithm and represented in **Figure 3.** by the green color. The matching between this pair image are determined by using the ZNCC correlation measure and are shown in **Figure 4.**

Fig. 2. Two images of the three-dimensional scene

Fig. 3. The interest points detected in two images.

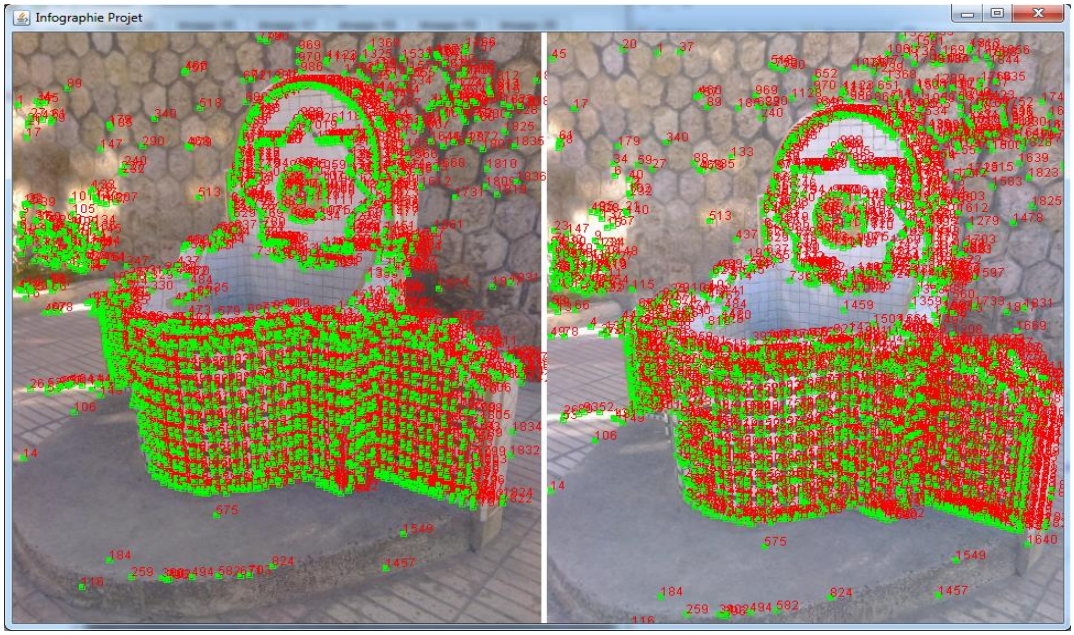


Fig. 4. The matchingsbetween the couple of two images.

The result of the reconstruction of the scene after the implementation of the different algorithms (Harris, ZNCC, Ransac ...) by the Java programming language is represented in **Figure 5.** Below.

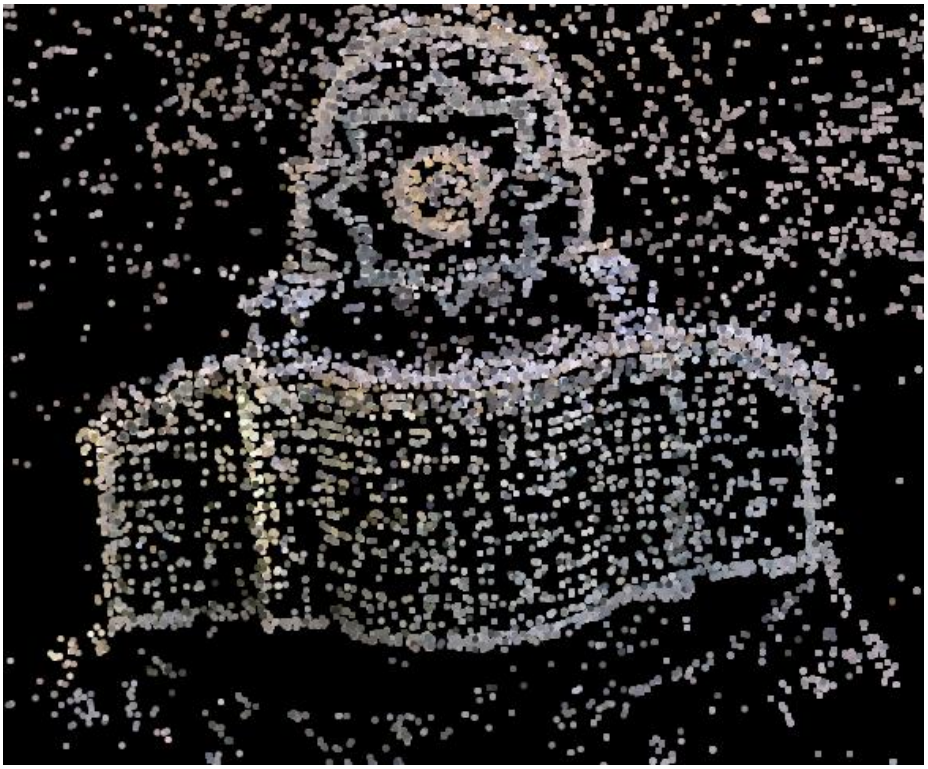


Fig. 5. Cloud points of the reconstructed 3D scene.

7. Conclusion

In this work, we have treated an approach to reconstruct three-dimensional scenes from a sequence of images which only one image is calibrate. The projection matrices of the others are obtained by solving a system of linear equations containing 16 equations at 14 unknowns. The interest points are detected by the Harris algorithm and matched by the ZNCC correlation method, and they are used with the projection matrices of the scene in the image planes to determine the point cloud coordinates , and finally to reconstruct the scene.

References

[1] Zhang, Z.: A Flexible New Technique for camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No.11, pp. 1330-1334, 2000.

[2] Gurdjos, P., Crouzil, A.,Payrissat, R.: Another Way of Looking at Plane-Based Calibration: The Centre Circle Constraint. ECCV, 2002.

[3] Boufama, B., Habert, A.: Registration and tracking in the context of augmented reality. ICGST-GVIP Journal, vol. 5 Issue 3, March 2005.

[4] Boudine, B., Kramm, S., El Akkad, N., Bensrhair, A., Saaidi, A., Satori, K.: A Flexible Technique based on Fundamental matrix for Camera Self-Calibration with Variable Intrinsic Parameters from two views. J. Vis. Commun. Image R. 39, 40–50. (2016)

[5] Manolis, I., Lourakis, A., Deriche, R.: Camera self-calibration using the kruppa equations and the SVD of the fundamental matrix: the case of varying intrinsic parameters. Technical Report 3911, INRIA, 2000.

[6] Sturm, P.: Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image and Vision Computing, 20(5-6): 415-426, 2002.

[7] Malis, E., Capolla, R.: Camera self-calibration from unknown planar structures enforcing the multi-view constraints between collineations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 4(9), September 2002.

[8] Gurdjos, P., Sturm, P.: Methods and Geometry for Plane-Based Self-Calibration. CVPR, pp. 491-496, 2003.

[9] Peijun, L., Jiaoying, S., Zhou, J., Longchao, J.: Camera Self-Calibration Using the Geometric Structure in Real Scenes. In Proceedings of the Computer Graphics International, 2003.

[10] Elsayed, E.: A Survey of Camera Self-Calibration. In Proceedings of the IEEE Conference on AVSS, 2003.

[11] Merras, M., El akkad, N., Saaidi, A., Nazih, G., Satori, K.: A New Method of Camera Self-Calibration with Varying Intrinsic Parameters using an Improved Genetic Algorithm. 8th International Conference on Intelligent Systems: SITA'13 (IEEE). , Rabat, Morocco, 08 and 09 May 2013.

[12] El akkad, N., Merras, M., Saaidi, A., Satori, K.: Camera Self-Calibration with Varying Intrinsic Parameters by an Unknown Three-Dimensional Scene. The Visual Computer (Springer).30(5), 519-530 (2014).

[13] Zetao, J.: Self-calibration of Varying Internal Camera Parameters Algorithm Based on Quasi-affine Reconstruction. Journal of Computers, Vol. 7, No. 3, March 2012.

[14] Zhao, Y., Hu, X., Lv, X., Wang, H.: Solving the Camera Intrinsic Parameters with the Positive Tri-prism Based on the Circular Points. Information Technology Journal 11(7):926-930, 2012.

[15] Trajkovic, M., Hedley, M.: Fast corner detection. Image and Vision Computing, 16:75–87, 1998.

[16] Harris, C., Stephens, M.: A combined Corner et Edge Detector. 4th Alvey vision Conference. pp. 147-151, 1988.

[17] Chambon, S., Crouzil, A.: Similarity measures for image matching despite occlusions in stereo vision. Pattern Recognit. 44(9), 2063–2075 (2011)

[18] Mori, M., Kashino, K.: Fast template matching based on normalized cross correlation using adaptive block partitioning and initial threshold estimation. In: Proceedings of IEEE International Symposium on Multimedia, pp. 196–203 (2010)

[19] Mattoccia, S., Tombari, F., Di Stefano, L.: Fast full-search equivalent template matching by enhanced bounded correlation. IEEE Trans. Image Process. 17(4), 528–538 (2008)

[20] El akkad, N., Merras, M., Baataoui, A., Saaidi, A., Satori, K.: Camera self-calibration having the varying parameters and based on homography of the plane at infinity. Multimedia Tools and Applications (Springer). DOI: 10.1007/s11042-017-5012-3, 2017.

[21] Mostafa, M., Elakkad, N., Saaidi, A., Gadhi Nazih, A., Satori, K. Robust method for camera calibration with varying parameters using hybrid modified genetic simplex algorithm. Journal of Theoretical and Applied Information Technology. Vol. 51 No.,pp. 363-373, 2013.

[22] Baataoui, A., El akkad, N., Saaidi, A., Satori, K., Masrar, M.: Spherical Self-Calibration of Cameras Characterized by Varying Intrinsic Parameters. International Journal of Software Engineering and Its Applications. Vol. 8, No. 7, pp. 23-40, 2014.

[23] Baataoui, A., El akkad, N., Saaidi, A., Satori, K., Masrar, M.: Robust Method for Camera Self-Calibration by an Unkown Planar Scene. Machine Graphics & Vision. Vol. 23, No. 1, pp. 37-57, 2014.

[24] El akkad, N., Saaidi, A., Satori, K.: Self-calibration based on a circle of the cameras having the varying intrinsic parameters. The 3rd International Conference on Multimedia Computing and Systems: ICMCS'12 (IEEE). Tangier, Morocco, May 10-12, 2012.

[25] El akkad, N., Merras, M., Saaidi, A., Satori, K.: Camera self-Calibration with Varying Parameters from Two views. Wseas Transactions on Information Science and Application, Vol. 10, No. 11, pp. 356-367, 2013.

[26] El akkad, N., Merras, M., Saaidi, A., Satori, K.: Robust Method for Self-Calibration of Cameras Having the Varying Intrinsic Parameters. Journal of Theoretical and Applied Information Technology. Vol. 50, No. 1, pp. 57-67, 2013.

[27] Merras, M., Saaidi, A., El akkad, N., Satori, K.: Multi-view 3D reconstruction and modeling of the unknown 3D scenes using genetic algorithms. Soft computing (Springer). DOI : 10.1007/s00500-017-2966-z, 2017.

[28] El hazzat, S., Merras, M., El akkad, N., Saaidi, A., Satori, K.: 3D reconstruction system based on incremental structure from motion using a camera with varying parameters. The Visual Computer (Springer). DOI : 10.1007/s00371-017-1451-0, 2017.

[29] El akkad, N., El hazzat, S., Saaidi, A., Satori, K.: Reconstruction of 3D scenes by camera self-calibration and using genetic algorithms. 3D Research (Springer). Vol. 7, No. ¬6, pp. 1-17, 2016.

[30] KARIM, K., El akkad, N., Satori, K.: Reconstruction of the 3D scenes from the matching between image pair taken by an uncalibrated camera. The 3rd International Conference on Big Data, Cloud and Applications: BDCA'18. Kenitra, Morocco, (Springer) doi.org/10.1007/978-3-319-96292-4_35, April 04-05, 2018.

[31] Merras, M., El akkad, N., El hazzat, S., Saaidi, A., Nazih, G., Satori, K.: Method of 3D Mesh Reconstruction from Sequences of Calibrated Images. The 4th International Conference on Multimedia Computing and Systems. ICMCS'14 (IEEE). Marrakesh, Morocco, April 14-16 2014.

[32] El hazzat, S., Merras, M., El akkad, N., Saaidi, A., Satori, K.: Enhancement of sparse 3D reconstruction using a modified match propagation based on particle swarm optimization. Multimedia Tools and Applications (Springer). DOI: https://doi.org/10.1007/s11042-018-6828-1, 2018.

An approch for FPGA implementation of LST-SW algorithm

Assaad El Makhloufi, Samir El Adib , Naoufal Raissouni
{ assaadelmakhloufi@gmail.com, adibsamir@gmail.com, naoufal.raissouni.ensa@gmail.com }

Remote Sensing & Geographic Information System Unit (RS&GIS)
University Abdelmalek Essaadi, ENSA Tetouan, Morocco

Abstract. Satellite remote sensing (SRS) needs to make use of up-to-date technology. The trend in SRS missions has always been towards using hardware devices with smaller size, lower cost, more flexibility, and higher computational power. Therefore, Field Programmable Gate Array (FPGA) technology is highly used by the scientific community and hardware developers to implement different SRS algorithms. Thus, FPGAs offer well-suited architectural features; large number of programmable logic elements, distributed and block RAMs, DSP and register slices and look up tables. This paper describes an approach to the implementation of the Land Surface Temperature Split-Window (LST-SW) algorithm structure based on FPGA technology.

Keywords: FPGA, Remote sensing, Virtex, VHDL, LST-SW.

1 Introduction

Recently, Field Programmable Gate Array (FPGA) technology has become a viable target for the implementation of different algorithms in different fields. These developments are offering new opportunities especially in satellite remote sensing. Digital sensors mounted onboard the remote sensing satellites scan vast areas of the earth's surface during day and night and beam this data to the satellite ground stations for further processing and usage in different applications: Digital Signal Processing (DSP), video processing, control systems engineering, bioinformatics, aerospace and defense systems. With their continuous development and improvement, remote sensing satellites [1] are increasingly being utilized and use of the latest available technological components. The emergence of peripherals such as Field Programmable Gate Arrays (FPGAs) [2], which can bridge the gap in on-board and real-time analysis of remote sensing data. Some FPGAs incorporate a large number of arithmetic blocks that can be low-complexity blocks such as simple multipliers or can be relatively more complex like the Digital Signal Processing (DSP) units which consist of combinations of various components like multipliers, adders, accumulators, shift registers. A DSP unit significantly accelerates the FPGA's performance and allows achieving greater productivity and flexibility, while decreasing cost and power consumption [3].

FPGAs are fully reconfigurable, a technological feature that combines the flexibility of traditional microprocessors with the performance of Application Specific Integrated Circuits (ASIC) devices [4]. Also, the reconfigurable hardware was certified by international remote sensing agencies, which was frequently used in remote sensing missions and space-borne earth observation missions [5]. FPGA also offers exact results with compact size, and low cost which makes the reconfigurable system interesting for onboard data processing.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284220

The FPGA architecture makes it possible to implement any combinational and sequential circuit, which can range from a simple logic function to a high-end soft-processor [6]. In the past, FPGAs have been mainly used in signal processing and network packet analysis. However, thanks to the high-speed embedded resources included in the FPGA such as DSP slices and fast memories, they are now also utilized for algorithm acceleration either as coprocessors or standalone systems, System-On-Chips (SoCs). Moreover, several vendors reserve space in the FPGAs to custom soft hard-IP (HIPs) processors.

Recently, FPGAs have become a viable target technology for implementation different algorithm in remote sensing such us: A Real-Time Marker-Based Visual Sensor Based on a FPGA and a Soft-Core Processor [7], FPGA Implementation of an Algorithm for Automatically Detecting Targets [8], FPGA Implementation of the N-FINDR Algorithm [9], On-Board Ortho-Rectification for images Based on an FPGA [10], FPGA Implementation of JPEG-LS Remote Sensing Image Coding Algorithm [11]. These computing systems combine the flexibility of general-purpose processors with the speed of application-specific processors. The objective of this work is to develop an FPGA-based hardware version of the Land Surface Temperature Split-Window (LST-SW) algorithm structure [12] on FPGA.

As image sizes and bit depths grow larger, software has become less useful in the images processing realm. Recently developed hybrid FPGAs, such as the Xilinx Virtex-5 [13], offer the versatility of running diverse software applications on embedded processors while at the same time taking advantage of reconfigurable hardware resources, all on the same chip package. In fact, radiation hardened FPGAs are in great demand for military and space applications [14] [15]. Xilinx FPGAs have been used in more than 50 missions. In this work, we use a Xilinx Virtex-5 XC5VLX50T FPGA as a baseline architecture since it is similar to the existing FPGAs [8] that have been certified by several international agencies for remote sensing applications. They are based on the same architecture so we could immediately implement our design on them.

2 Role of FPGA in satellite remote sensing applications

The trend in remote sensing missions has always been towards using hardware devices with smaller size, lower cost, more flexibility, and higher computational power [16] [17]. Reconfigurable Hardware (RH) provides a flexible medium to implement hardware circuits.

The RH resources are reconfigurable post-fabrication, allowing a single-based hardware design to implement a variety of circuits. FPGAs can be reconfigured to avoid hardware faults [18], whether they result from fabrication or the environment.

Reconfigurable hardware offers the necessary flexibility and performance with reduced energy consumption compared to other high-performance processors. By mapping functionality to FPGAs, the computer designer can optimize the hardware for a specific application resulting in acceleration rates of several orders of magnitude over general purpose computers. Furthermore, these devices are characterized by lower form/wrap factors compared to parallel platforms and by higher flexibility than ASIC solutions. Moreover, satellite-based remote sensing instruments can only include chips that had been certified for space conditions. This is because space-based systems must operate in an environment in which radiation effects

have an adverse impact on integrated circuit operation [19]. Ionizing radiation can cause soft-errors in the static cells used to hold the configuration data. This will affect the circuit functionality and can cause system failure. Therefore, it requires special FPGAs that provide on-chip reconfiguration error-detection and/or correction circuitry. High-speed, radiation-hardened FPGA chips with million gate densities have recently emerged can support the high throughput requirements for the remote sensing applications. Radiation-hardened FPGAs are in great demand for military and space applications. For instance, industrial partners such as Microsemi Corporation or Xilinx (www.xilinx.com) have been producing radiation-tolerant antifuse FPGAs for several years for high-reliability space-flight systems. Microsemi FPGAs have been on board more than 100 launches and Xilinx FPGAs have been used in more than 50 missions.

3 Description of LST Split-Window algorithm

Land Surface Temperature (LST) has a significant role in the land surface characters on local and global scale and his one of most key parameters in the biophysics of land surface processes [20]. LST plays an important role in the cycle of the natural ecosystem evolution and global change. LST determination from satellite data is mainly influenced by the atmosphere and surface emissivity.

LST is defined as the surface radiometric temperature corresponding to the instantaneous field-of-view of the sensor or more precisely [21], as the `ensemble directional radiometric surface temperature [22]. The term "ensemble" depicts the bulk contribution of an inhomogeneous pixel. For a given sensor viewing direction, LST depends on the distribution of temperature and emissivity within a pixel and the spectral channel of measurement [23].

In our case, the LST Split Window (LST-SW) algorithm proposed by (Sobrino J. & Raissouni N) has been used for implementation purpose on FPGA (see eq. 1):

$$LST = T_4 + 1.40\,(T_4-T_5) + 0.32\,(T_4-T_5)2 + 0.83 + (57-5W)\,(1-\varepsilon) - (161-30W)\,\Delta\varepsilon\,. \quad (1)$$

where T4 and T5 in (K) are the brightness temperatures measured in the Advanced Very High-Resolution Radiometer (AVHRR) Channels 4 and 5 on board the National Oceanic and Atmospheric Administration (NOAA) satellite series, W is the total amount of atmospheric water vapor in (g cm-2). ε and $\Delta\varepsilon$ are, respectively, the average effective emissivity in both channels and the spectral variation of emissivity. Also, T_4, T_5, W, ε are binary images contains pixels of images coded in 16 bits.

4 Proposed Hardware FPGA architecture for the LST-SW algorithm

The architecture proposed of the hardware used to implement the LST algorithm (see **Figure 1)**, in our implementation we have devices our algorithm in two parts (see equations (2) and equations (3)):

$$\text{Part 1(LST1)} = T_4 + 1.40\,(T_4-T_5) + 0.32\,(T_4-T_5)^2 + 0.83\,. \quad (2)$$

Part 2(LST2) = (57-5W) (1- ε) -(161-30ε) Δε . (3)

LST = LST1 + LST2. . (4)

For data input, we use a hard disk (PC) to store our binary data. In addition, our architecture contains two FIFO: the first FIFO to transmit the binary images T_4, W and also the result of the first part. The second FIFO to transfer the binary images T_5, ε and the result of the second part. Finally, we transmit each part in the two FIFOS: FIFO 1 (part1) and FIFO 2 (part2). The LST module is used to implement our version of the LST algorithm.

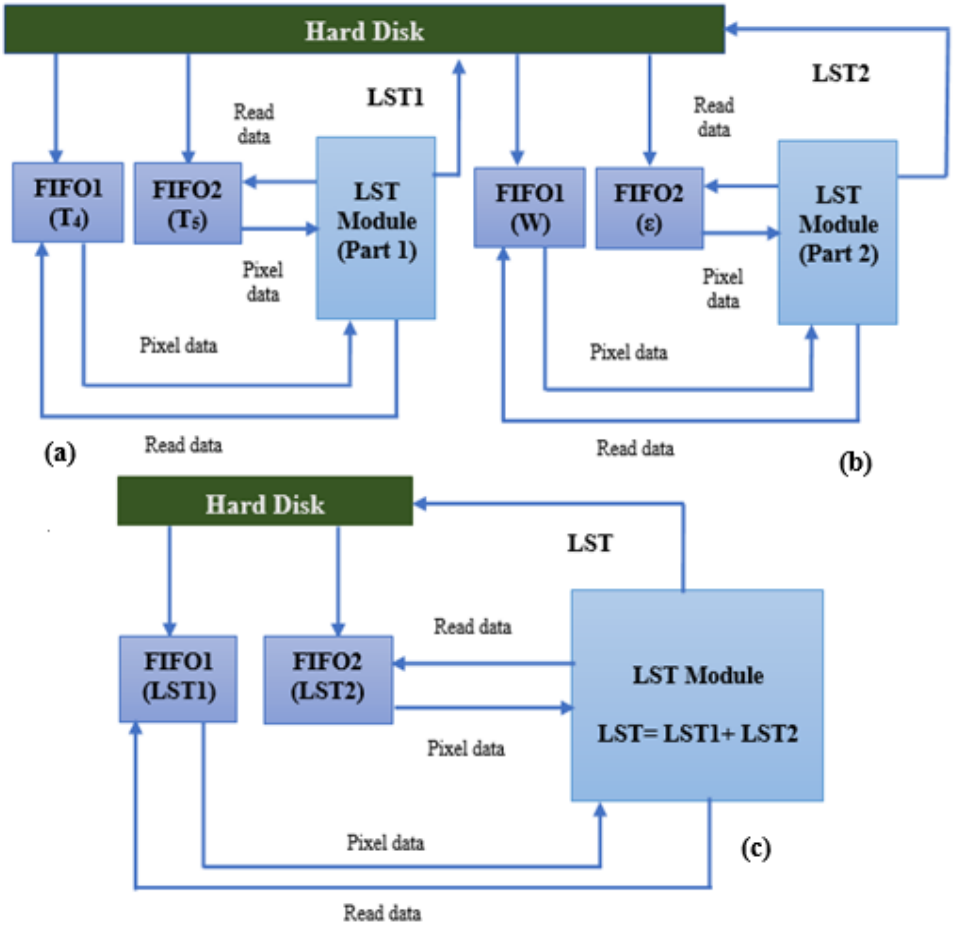


Fig. 1.Hardware architecture proposed for implementation of the LST S-W algorithm

To calculate LST1 we send two images: T_4 through FIFO1 and T_5 into FIFO 2 then the calculation happened in the LST module (see **Figure 1. [a]**). The last result of part1 stored in the hard disk. For LST2 we send two images: W through FIFO1 and ε into FIFO 2 then the calculation happened in the LST module (see **Figure 1. [b]**). The last result of part2 stored in the hard disk. Finally, to calculate LST we send both parts that stored in memory through the

two FIFOs: LST1 through FIFO1 and LST2 into FIFO2 (see **Figure 1. [c]**) then we add this result and send it to the hard disk.

LST hardware module is the heart of our system, and we took special care in its definition using VHDL as the hardware description Language. This module has set as top module of our system, because he contains different hardware part such us, registers, arithmetic-logic units. Also, all the calculations of our algorithm are done in this module: addition, subtraction, multiply. In our implementation, we multiply 100 with the coefficient of the algorithm (see equation (4) and (5)) to avoid the floating-point and their difficulty in their implementation, and when we got the result, we divided the coefficients by 100 to get back to our equations.

$$\text{Part 1(LST1)} = 100 \times T_4 + 140\,(T_4 - T_5) + 32\,(T_4 - T_5)^2 + 83\,. \quad (5)$$

$$\text{Part 2(LST2)} = 100 \times (57 - 5W)\,(1 - \varepsilon) - 100 \times (161 - 30\varepsilon)\,\Delta\varepsilon\,. \quad (6)$$

The LST-SW algorithm has been described using VHDL (VHSIC Hardware Description Language), and the design has been synthesized and implemented using Xilinx ISE 9.2i tool. As we said, T_4, T_5, W, ε are binary images generated by ENVI software in form of. Png files, but our Xilinx ISE cannot support this type of files. A program was written using MATLAB to support design verification through file manipulation. Therefore, we have written a program to convert this png files to text file format which was read by the VHDL.

5. Implementation results

5.1 Input Data

In our proposed implementation we have used 4 images (format: 512 columns, 512 rows). we chose to work with images similar to the real satellite images (T_4, T_5, W, ε) to check the operation of our system Then we have converted this image to text file contains pixels of images coded in 16 bits.

5.2 Test bench generation

The most time-consuming part of a system design of this nature is verifying that the VHDL code operates as intended. This is done by executing the device code in a simulator. The simulation of the device is only meaningful if the inputs to the system are manipulated in a way that causes the device to operate in its intended manner. This is easily done using VHDL to create a test bench for the entity under test. One of the major improvements made as test bench development progressed was the adoption of file I/O for loading processing parameters and image data into the system.

We have creating a test bench to circulate satellite data into system, which are the latter text format files contains pixel values that what we called the file Input/Output.

5.3 Register Transfer Level (RTL) schematic

Figure 2 show the Register Transfer Level of the first part and second part and the result of our algorithm, as we can see this figures shows the input signal which are the outlar

(T_4), outlar 2(T_5), outlar 3(W),outlar 4(ε), Also the LST1 and LST2 and they are data satellite as text file contain pixel coded in 16 bits, and the output that contain the result of the first part LST 1 and second part LST 2 and the LST result.

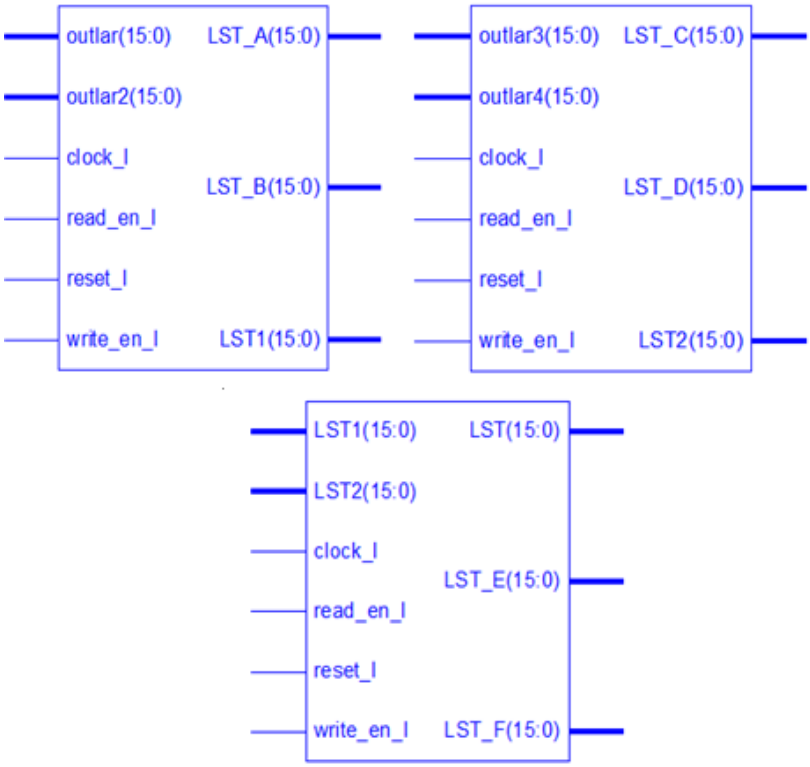


Fig. 2. Register Transfer Logic for the system

The **Figures 3,4** and **5** shows the simulation results of the proposed implementation of the LST-SW algorithm in each part. In this simulation we see all the data that be sending into system as binary number coded in 16 bits using the test bench for reads images file and it was written specifically to replace the acquisition of images. This data buffering by FIFO, it is a structure used in hardware application when we need to buffer a data and it's like a pipe where the first element entering into the pipe is the first element that output from the pipe. Therefore, for LST1 (see Figure 3), we send the first element of our image T4 (outlar) into FIFO 1 and the first element of image T5 (outlar2) into FIFO 2 and we calculated the first element it means the first pixel in each data together then we stored this result in the first line, this operation continue until the last element.

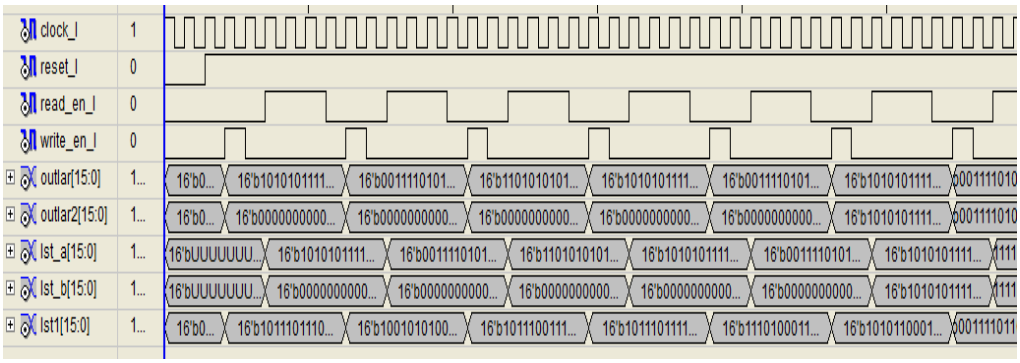


Fig. 3. Simulation Results for the LST1

The same way, to calculated LST2 (see Figure 4), we send the first element of our image W (outlar3) into FIFO 1 and the first element of ε(outlar4) into FIFO 2 and we calculate the first element it means the first pixel in each data together then we stored this result in the first line, this operation continues until the last element.

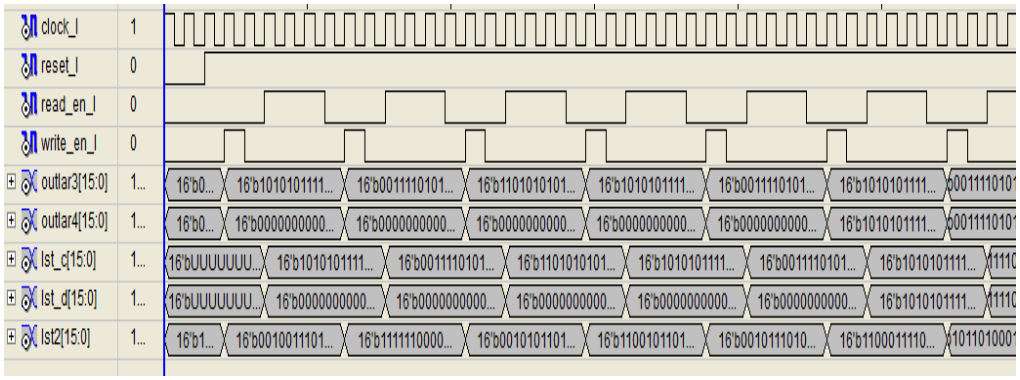


Fig. 4. Simulation Results for the LST2

Figure 5 shows the result of the addition of the two parts, with the same way, we transfer the result of the LST1 into FIFO 1 and the result of LST 2 into FIFO 2, this calculation is done like the earlier ones, we add the first element pixel in each FIFO and we memorized, this operation continues until the last pixel.

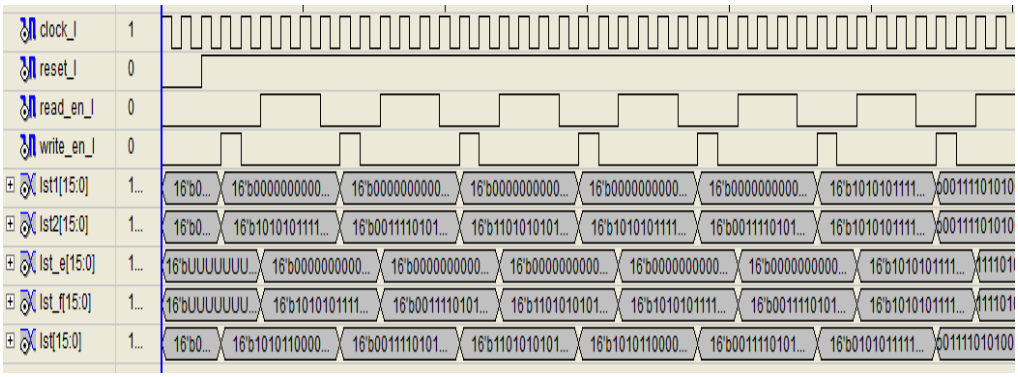


Fig. 5. Simulation Results for the LST

5.4 LST-SW FPGA implementation results

In this subsection, we conduct an experimental result of the computational performance of the proposed FPGA implementation. **Table 1** shows the resources used for our hardware implementation of the proposed LST algorithm design, conducted on the Virtex-5 LX50T FPGA. This FPGA has a total of 4706 slices, 28800 slice registers, and 28800 four input look-up tables available. In addition, the FPGA includes some heterogeneous resources, such as 48 DSP48s.In our implementation, we took advantage of these resources to optimize the design, The Block Slice Registers are used to implement the FIFOs without using block RAMs, with others slices are used for the implementation of the LST-SW algorithm with the DSP48s multiplier.

Table 1. Summary of resource utilization for the FPGA-based implementation of the LST algorithm.

Target FPGA	**Virtex-5 LX50T**		
Device utilization	**Used**	**Total**	**%**
Number of Slice Registers	3306	28800	11
Number of Slice LUTs	2231	28800	7
Number of Slices	222	7200	3
Number of bonded IOBs	84	480	17
Number of DSP48Es	27	48	56
Maximum Frequency	324.657 Mhz		

Logically, the number of DSP is higher because of the existence of the calculation in each part (LST1 and LST2) that consumes much of the resources of the FPGA.

6. Conclusion

The choice of an architecture for the embedded images processing is not trivial and that a compromise must certainly be found between flexibility, consumption, performance, cost and speed of design in terms of TTM (Time-To-Market). Also, the number of remote sensing applications requiring fast response of algorithm analysis has been growing exponentially in

recent years. In this paper, we have proposed an architecture for the FPGA implementation of the Land Surface Temperature Split Window (LST-SW) algorithm, one of the most well-known approaches for determination of Land Surface Temperature from satellite data in the remote sensing community. The system was designed using VHDL in a Hight level design method. All part of the design has been simulated and implementing using Xilinx tools with target FPGA Virtex-5 LX50T.

References

[1] Roy,S., Behera,D,M.,Srivastav,K,S.: Satellite Remote Sensing: Sensors, Applications and Techniques. Vol. 87, pp 465–472. Sciences, INDIA (2017)

[2] Yang,C. , Li,B., Chen ,L., Wei,C.: A Spaceborne Synthetic Aperture Radar Partial Fixed-Point Imaging System Using a Field- Programmable Gate Array—Application-Specific Integrated Circuit Hybrid Heterogeneous Parallel Acceleration Technique. Sensors (MDPI), (2017)

[3] Yaseer,D.,Teresa,R.: Power estimation technique for DSP architectures. Digital Signal Processing. Vol 19, pp. 213-219 (2009)

[4] Singh,G.: Reconfigurable Computing A review of the technology and its architecture. Journal of VLSI and Signal Processing. Vol 3, pp. 08-14 (2013)

[5] Gonzalez,C., Resano,J., Mozos,D., Plaza,A., and Valencia,D.: FPGA Implementation of the Pixel Purity Index Algorithm for Remotely Sensed Hyperspectral Image Analysis. Journal on Advances in Signal Processing, Vol 10, pp. 13 (2010)

[6] García,G., Jara,C., Pomares,J., Alabdo,A., Poggi,L., Torres,F. A Survey on FPGA-Based Sensor Systems: Towards Intelligent and Reconfigurable Low-Power Sensors for Computer Vision, Control and Signal Processing. Sensors (MDPI), (2014)

[7] Tayara, H., Ham,W., Chong,K.: A Real-Time Marker-Based Visual Sensor Based on a FPGA and a Soft-Core Processor.Sensors(MDPI). Vol 16, (2016)

[8] Gonzalez,C., Bernabé,S., Mozos,D., Plaza,A.: FPGA Implementation of an Algorithm for Automatically Detecting Targets. IEEE journal of selected topics in applied earth observations and remote sensing.Vol 9,pp .13 (2016)

[9] Gonzalez,C., Resano,J., Mozos,D., Plaza,A.: FPGA Implementation of the N-FINDR Algorithm for Remotely Sensed Hyperspectral Image Analysis. IEEE transactions on geoscience and remote sensing.Vol 50, (2012)

[10] Zhou,G., Zhang,R., Liu,N., Huang,J., Zhou,X.: On-Board Ortho-Rectification for Images Based on an FPGA. Remote Sensing (MDPI). Vol 9, (2017)

[11] Wang,H.: FPGA Implementation of JPEG-LS Remote Sensing Image Coding Algorithm. Advances in Engineering Research.Vol 150,(2017)

[12] Sobrino,J, Raissouni,N.: Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco.International Journal of Remote Sensing. Vol 21,(2000)

[13] Bulens,P.,Standaert,F.,Pellegrin,P.:Implementation of the AES-128 on Virtex-5 FPGAs.Progress in cryptology.pp. 16-26 (2008)

[14] Kramer,H., Cracknell, A.: An overview of small satellites in remote sensing, International Journal of Remote Sensing, Vol 29, (2008)

[15] Lysaght, P., Blodget, B., Mason, J., Young, J., Bridgford, B.: Enhanced architectures, design methodologies and CAD tools for dynamic reconfiguration of Xilinx FPGAs. Proc. Int. Conf. Field Programmable Logic Appl. pp. 1–6 (2006)

[16] Heinz, D., Chang,C.:,Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery. IEEE Trans. Geosci. Remote Sensing. Vol. 39, pp. 529–545 (2001)

[17] Plaza, A., Martinez, P., Perez, R., Plaza, J.: A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE Trans. Geosci. Remote Sensing. vol. 42, pp. 650–663 (2004)

[18] Laplante, P.: Computing requirements for self-repairing space systems, Journal of Aerospace Computing, Information and Communication. vol 2, pp. 154–169 (2005)
[19] Neil, B., Dawood, A.:Reconfigurable computers in space: Problems solutions and future directions.In Proc. Military Aerosp. Appl. Programmable Logic Devices Conf., pp. 1–2 (1999)
[20] Morrow, N., Friedl, MA.: Modelling biophysical controls on land surface temperature and reflectance in grasslands.Agricultural and Forest Meteorology. Pp. 147-161 (1998)
[21] Prata,F., Sobrino,J.: Thermal Remote Sensing of Land Surface Temperature From Satellites: Current Status and Future Prospects. Remote Sensing Reviews. Pp. 175-224 (1995)
[23] Norman,J.,Becker,F.: Terminology in thermal infrared remote sensing of natural surfaces Reviews'.pp. 159-173 (1995)
[23] Becker,F.,Li,Z.: Surface temperature and emissivity at various scales: Definition, measurement and related problems. Remote Sensing Reviews.vol 10, pp. 225-253 (1995)

Hybrid Beamforming for Relay-Aided mmWave Backhaul links

[1]Mostafa Hefnawi, [2]Esmael Yahya, [3]Jamal Zbitou, [4]Mohamed Aboulfatah, [5]Hassan Abdelmounim

{ [1]hefnawi@rmc.ca, [2]zbitou.jamal@yahoo.fr, [4]mohamed.aboulfatah@uhp.ac.ma, [5]hassan.abdelmounim@hotmail.fr }

[1,2]Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON, Canada

[3,4,5] LMEET FST of Settat, University Hassan 1st, Settat, Morocco

Abstract. Massive multiple-input multiple-output (MIMO) and millimeter-wave (mmWave) technologies have emerged as a promising solution to enhance the backhaul wireless link in 5G Heterogeneous networks (HetNets). These mmWave backhaul links, however, are very susceptible to a significant path loss due to the blockage of the line-of-sight and massive antenna arrays may not be sufficient to alleviate such losses. To this end, relays are usually deployed to provide alternative routes that help boost links with high path loss. In this paper, therefore, we consider using relay base stations (RBS) in mmWave backhaul links between small cell base stations (SBS) and a macro-cell base station (MBS). It is assumed that the SBSs, the RBSs, and the MBS are all equipped with massive antenna arrays employing hybrid analog and digital beamforming. The analog beamformers are based on the selection of fixed multi-beams using a constrained eigenbeamforming scheme while the digital beamformers are based on the maximum ratio transmission and maximum ratio combining (MRT/MRC) schemes that maximize the transmit and receive SINRs of the effective channels created by the actual channel and the analog beamformer. The performance evaluation in terms of the beampatterns and the ergodic channel capacity shows that the proposed HBF scheme achieves near-optimal performance with only 4 RF chains and requires considerably less computational complexity.

Keywords: Hybrid beamforming, HetNets, Relays, Massive MIMO, mmWaves.

1 Introduction

Recently, millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) technologies [1],[2],[3],[4],[5][6] have emerged as a promising solution to enhance the backhaul wireless link in 5G Heterogeneous networks (HetNets) [5],[6],[7],[8],[9],[10],[11]. On the one hand, the mmWave backhaul links can provide the Gigahertz bandwidth that can be achieved by conventional optical fiber link without the restriction of deployment and installation of small cells. On the other hand, because of the small wavelength of mm-waves, a large number of antennas can be deployed and can provide a high gain to compensate for the pathloss of the mmWaves. However, mmWave massive antenna arrays work better in the presence of line-of-sight (LoS) and may not be sufficient to alleviate the severe losses due to the blockage of these LoSs. To overcome this problem, relay assisted backhaul link can be incorporated to efficiently transmit the signals between the small cell base stations (SBS) and the macro-cell base station (MBS). In this paper, therefore, we consider using relay base stations (RBS) in mmWave backhaul links between the SBSs and the MBS, where the SBSs, the RBSs, and the MBS are all

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284218

equipped with massive antenna arrays. Moreover, in order to reduce the number of RF chains required by fully-digital beamforming massive arrays, we employ a combination of RF analog beamformers and baseband digital beamformers, known as hybrid beamforming (HBF) [12],[13],[14],[15],[16],[17],[18]. In such a hybrid configuration, the analog RF beamforming matrix, built from analog hardware like phase-shifters, is used to connect M_a antenna elements to N_{RF} RF chains, where $N_{RF} < M_a$. Previous studies on hybrid massive MIMO mainly focused on single-user systems and exploited the sparse nature of the mmWave to develop low-complexity hybrid precoding algorithms [12],[13],[14]. MU-MIMO cases were studied in [15],[16],[17]. In [15] a scheme called "Joint Spatial Division Multiplexing" (JSDM) was proposed to create multiple "virtual sectors" which reduces signaling overhead and computational complexity of downlink training and uplink feedback. In [16],[17] it was shown that the required number of RF chains only needs to be twice the number of data streams in order to achieve the same performance of any fully-digital beamforming scheme. These studies, however, did not consider HBF in the context of HetNets and focused primarily on macro-cellular systems. In this paper, we propose to extend HBF to relay-assisted backhaul links where the SBSs, the RBSs, and the MBS are all equipped with massive hybrid antenna arrays. On the one hand, the analog beamformers are based on the creation of the best fixed multi-beams by eigendecomposition of the backhaul channels. On the other hand, the digital beamformers are based on the maximum ratio transmission and maximum ratio combining (MRT/MRC) schemes [19] that maximizes the transmit and receive SINRs of the effective channels created by the cascade of the analog beamforming weights and the actual channel.

2 System Model

We consider the backhaul uplink in the HetNet of Figure 1, where K SBSs are connected to the MBS through an RBS using a two-hop relaying path. It is assumed that the RBS, the SBSs, and the MBS are equipped with $M_a - element$ transmitting/receiving massive hybrid antenna arrays. For the SBSs-to-relay link, it is assumed that the number of transmit/receive RF chains is identical and is equal to the number of data streams N_{RF}. On the other hand, it assumed that for the relay-to-MBS link, the number of transmit/receive RF chains is identical and is equal to the number of SBSs K.

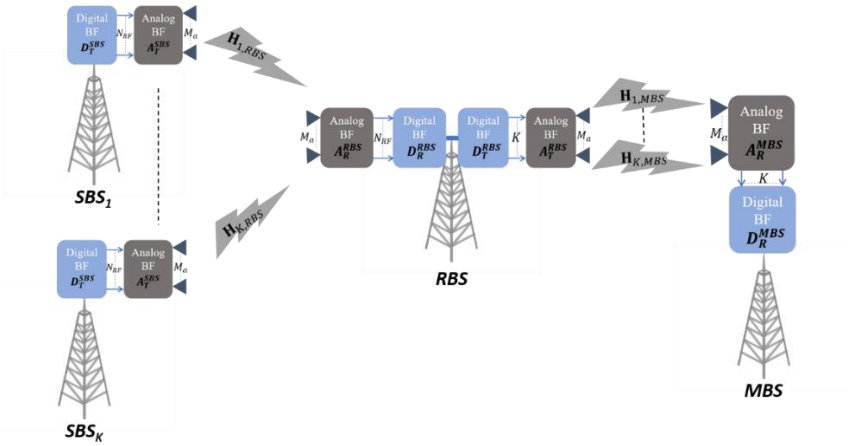


Fig. 1. System model: *K* SBSs connected to an MBS through two-hop relaying links.

2.1 SBSs-to-Relay Link

The k^{th} SBS applies its signal $\boldsymbol{s}_k^{SBS}$ of N_{RF} data streams to an $N_{RF} \times N_{RF}$ diagonal transmit digital beamforming weight matrix $\boldsymbol{D}_{T,k}^{SBS}$ followed by an $M_a \times N_{RF}$ transmit analog beamforming matrix $\boldsymbol{A}_{T,k}^{SBS}$. If we denote the combined digital-analog transmit beamformer for the k^{th} SBS as $\mathbf{w}_{T,k}^{SBS} = \boldsymbol{A}_{T,k}^{SBS}\boldsymbol{D}_{T,k}^{SBS}$, then the $M_a \times 1$ transmitted signal $\boldsymbol{x}_{T,k}^{SBS}$ at the output of the antenna array of the k^{th} SBS can expressed as

$$\boldsymbol{x}_{T,k}^{SBS} = \mathbf{w}_{T,k}^{SBS}\boldsymbol{s}_k^{SBS}, \tag{1}$$

and the array output of the RBS can be written as

$$\boldsymbol{y}_{RBS} = \sum_{k=1}^{K} \mathbf{H}_{k,RBS}\mathbf{w}_{T,k}^{\boldsymbol{SBS}}\boldsymbol{s}_k^{SBS} + \mathbf{n}_{RBS}, \tag{2}$$

where $\boldsymbol{y}_{RBS}$ is the $M_a \times 1$ vector containing the outputs of the $M_a -$ element antenna array at the RBS, $\mathbf{H}_{k,RBS}$ is the $M_a \times M_a$ channel matrix representing the transfer functions from the $M_a -$element antenna array of the k^{th} SBS to the $M_a -$element antenna array of the RBS, and $\mathbf{n}_{RBS}$ is the received $M_a \times 1$ complex additive white Gaussian noise vector at the RBS.

The RBS detects the k^{th} SBS signal by applying the output of the array $\boldsymbol{y}_{RBS}$ to the $N_{RF} \times M_a$ receiving analog weight matrix, $\boldsymbol{A}_{R,k}^{RBS}$, followed by a diagonal $N_{RF} \times N_{RF}$ receive digital beamforming weight matrix $\boldsymbol{D}_{R,k}^{RBS}$. If we denote the combined digital-analog receive beamformer for the k^{th} SBS as $\mathbf{w}_{R,k}^{RBS} = \boldsymbol{A}_{R,k}^{RBS}\boldsymbol{D}_{R,k}^{RBS}$, then the detection of the the k^{th} SBS signal by the RBS can be expressed as

$$\widehat{\boldsymbol{x}}_{k,RBS} = \left(\mathbf{w}_{R,k}^{RBS}\right)^H \boldsymbol{y}_{RBS} = \mathbf{S}_k^{RBS} + \mathbf{S}_{I_k}^{RBS} + \mathbf{N}_{RBS}, \tag{3}$$

where $\mathbf{S}_k^{RBS} = \left(\mathbf{w}_{R,k}^{RBS}\right)^H \mathbf{H}_{k,RBS}\mathbf{w}_{T,k}^{SBS}\boldsymbol{s}_k^{SBS}$ is the k^{th} SBS received signal, $\mathbf{S}_{I_k}^{RBS} = \left(\mathbf{w}_{R,k}^{RBS}\right)^H \sum_{i=1,i\neq k}^{K} \mathbf{H}_{i,RBS}\,\mathbf{w}_{T,i}^{SBS}\boldsymbol{s}_i^{SBS}$ is the multiple-access interference (MAI) from the $K-1$ other SBSs, and $\mathbf{N}_{RBS} = \left(\mathbf{w}_{R,k}^{RBS}\right)^H \mathbf{n}_{RBS}$ is the noise signal at the array output of the RBS.

Assuming that $\boldsymbol{s}_k^{SBS}$ are complex-valued random variables with normalized unit power, i.e., $\mathbb{E}\left[\boldsymbol{s}_k\,\boldsymbol{s}_k^H\right] = \mathrm{I}_{N_{RF}}$, we can express the SINR at the RBS for the k^{th} SBS as

$$\begin{aligned}\gamma_k^{RBS} &= \frac{\left(\boldsymbol{D}_{R,k}^{RBS}\right)^H\left(\boldsymbol{A}_{R,k}^{RBS}\right)^H \mathbf{H}_{k,RBS}\boldsymbol{A}_{T,k}^{SBS}\boldsymbol{D}_{T,k}^{SBS}\left(\boldsymbol{D}_{T,k}^{SBS}\right)^H\left(\boldsymbol{A}_{T,k}^{SBS}\right)^H \mathbf{H}_{k,RBS}^H\boldsymbol{A}_{R,k}^{RBS}\boldsymbol{D}_{R,k}^{RBS}}{\left(\mathbf{w}_{R,k}^{RBS}\right)^H \mathbf{B}_{RBS}\left(\mathbf{w}_{R,k}^{RBS}\right)} \\ &= \frac{\left|\left(\boldsymbol{D}_{R,k}^{RBS}\right)^H \boldsymbol{\mathcal{H}}_{k,RBS}\boldsymbol{D}_{T,k}^{SBS}\right|^2}{\left(\mathbf{w}_{R,k}^{RBS}\right)^H \mathbf{B}_{k,RBS}\left(\mathbf{w}_{R,k}^{RBS}\right)},\end{aligned} \tag{4}$$

where $\boldsymbol{\mathcal{H}}_{k,RBS} = \left(\boldsymbol{A}_{R,k}^{RBS}\right)^H \mathbf{H}_{k,RBS}\left(\boldsymbol{A}_{T,k}^{SBS}\right)$ represents the effective channel and $\mathbf{B}_{k,RBS} = \sum_{i=1,i\neq k}^{K} \mathbf{H}_{i,RBS}\mathbf{w}_{T,i}^{SBS}\left(\mathbf{w}_{T,i}^{SBS}\right)^H \mathbf{H}_{i,RBS}^H + \sigma_n^2\mathbf{I}_{M_a}$ is the covariance matrix of the interference-plus-noise at the RBS.

2.2 Relay-to-MBS Link

The RBS applies the received k^{th} SBS signal, $\widehat{\boldsymbol{x}}_{k,RBS}$, to the k^{th} selected beam port of the transmit hybrid beamformer. For simplicity, we will assume that each SBS signal is forwarded to the MBS using a separate beam (i.e., a separate RF chain). Thus, if we reorganize the *K* SBSs'

signals into a vector as $\hat{\boldsymbol{x}}_{RBS} = [\hat{\boldsymbol{x}}_{1,RBS}, \hat{\boldsymbol{x}}_{2,RBS}, \cdots, \hat{\boldsymbol{x}}_{K,RBS}]$ and we denote the RBS transmit analog beamformer as $\boldsymbol{A}_T^{RBS} = \left[\boldsymbol{a}_{T,1}^{RBS}, \boldsymbol{a}_{T,2}^{RBS}, \cdots, \boldsymbol{a}_{T,K}^{RBS}\right]$ and the digital beamformer as $\boldsymbol{D}_T^{RBS} = diag\left[d_{R,1}^{MBS}, d_{R,2}^{MBS} \ldots\ d_{R,K}^{MBS}\right]$, then the $M_a \times K$ transmitted signal, $\boldsymbol{s}_T^{RBS}$, at the output of the RBS antenna array can be expressed as

$$\boldsymbol{s}_T^{RBS} = \boldsymbol{A}_T^{RBS}\boldsymbol{D}_T^{RBS}\hat{\boldsymbol{x}}_{RBS}\,, \tag{5}$$

and the received signal at the array output of the MBS can be written as

$$\boldsymbol{y}_{MBS} = \mathbf{H}_{MBS}\boldsymbol{A}_T^{RBS}\boldsymbol{D}_T^{RBS}\hat{\boldsymbol{x}}_{RBS} + \mathbf{n}_{MBS}\,, \tag{6}$$

where $\boldsymbol{y}_{MBS}$ is the $M_a \times 1$ vector containing the outputs of the $M_a - \text{element}$ antenna array at the MBS, $\mathbf{H}_{MBS}$ is the $M_a \times M_a$ channel matrix between the RBS and the MBS, $\mathbf{n}_{MBS}$ is the received $M_a \times 1$ complex additive white Gaussian noise vector at the MBS.

The output of the array $\boldsymbol{y}_{MBS}$ is applied to the $M_a \times K$ receiving analog weight matrix of the MBS, $(\boldsymbol{A}_R^{MBS})^H$, followed by the $K \times K$ receive digital beamforming weight matrix $(\boldsymbol{D}_R^{MBS})^H$, then the detection of the K SBSs' signals by the MBS can be expressed as

$$\begin{aligned}\hat{\boldsymbol{x}}_{MBS} &= (\boldsymbol{D}_R^{MBS})^H(\boldsymbol{A}_R^{MBS})^H\boldsymbol{y}_{MBS} \\ &= (\boldsymbol{D}_R^{MBS})^H(\boldsymbol{A}_R^{MBS})^H\mathbf{H}_{MBS}(\boldsymbol{A}_T^{RBS})^H(\boldsymbol{D}_T^{RBS})^H\hat{\boldsymbol{x}}_{RBS} + (\boldsymbol{D}_R^{MBS})^H(\boldsymbol{A}_R^{MBS})^H\mathbf{n}_{MBS}\,,\end{aligned} \tag{7}$$

which results in the detection of the k^{th} SBS signal being expressed as

$$\hat{\boldsymbol{x}}_{k,MBS} = \left(d_{R,k}^{MBS}\right)^*\left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{H}_{MBS}\boldsymbol{a}_{T,k}^{RBS}\left(d_{T,k}^{RBS}\right)\hat{\boldsymbol{x}}_{k,RBS} + \left(d_{R,k}^{MBS}\right)^*\left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{n}_{MBS}\,, \tag{8}$$

Using (3), and denoting $\boldsymbol{\mathcal{H}}_{k,MBS} = \left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{H}_{MBS}\boldsymbol{a}_{T,k}^{RBS}$ as the effective channel, $\hat{\boldsymbol{x}}_{k,MBS}$ and the SINR of the k^{th} SBS at the MBS can be expressed, respectively, as

$$\begin{aligned}\hat{x}_{k,MBS} = \left(d_{R,k}^{MBS}\right)^*\left(d_{T,k}^{RBS}\right)\boldsymbol{\mathcal{H}}_{k,MBS}\left(\mathbf{S}_k^{RBS} + \mathbf{S}_{I_k}^{RBS} + \mathbf{N}_{RBS}\right) \\ + \left(d_{R,k}^{MBS}\right)^*\left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{n}_{MBS}\,,\end{aligned} \tag{9}$$

$$\gamma_k^{MBS} = \frac{\left|\left(d_{R,k}^{MBS}\right)^*\left(d_{T,k}^{RBS}\right)\boldsymbol{\mathcal{H}}_{k,MBS}\mathbf{S}_k^{RBS}\right|^2}{\left(d_{R,k}^{MBS}\right)^*\mathbf{B}_{k,MBS}d_{R,k}^{MBS}} \tag{10}$$

where $\mathbf{B}_{k,MBS}$ is the covariance matrix of the interference-plus-noise at the MBS and is given by $\mathbf{B}_{k,MBS} = \mathbf{B}_{I_k} + \mathbf{B}_N$, with $\mathbf{B}_N = \left|\left(d_{T,k}^{MBS}\right)\right|^2\boldsymbol{\mathcal{H}}_{k,MBS}\mathbf{N}_{RBS}\mathbf{N}_{RBS}^H\boldsymbol{\mathcal{H}}_{k,MBS}^H + \sigma_{n_{MBS}}^2\left(\boldsymbol{a}_{k,R}^{MBS}\right)^H\boldsymbol{a}_{k,R}^{MBS}$ and $\mathbf{B}_{I_k} = \left|\left(d_{T,k}^{MBS}\right)\right|^2\boldsymbol{\mathcal{H}}_{k,MBS}\mathbf{S}_{I_k}^{RBS}\left(\mathbf{S}_{I_k}^{RBS}\right)^H\boldsymbol{\mathcal{H}}_{k,MBS}^H$.

Assuming that the transmit and receive digital beamformer are identical, (10) can be simplified as

$$\gamma_k^{MBS} = \mathbf{B}_{k,MBS}^{-1}\left|\boldsymbol{\mathcal{H}}_{k,MBS}\left(\boldsymbol{D}_{R,k}^{RBS}\right)^H\boldsymbol{\mathcal{H}}_{k,RBS}\boldsymbol{D}_{T,k}^{SBS}\right|^2\,, \tag{11}$$

2.3 Channel Model

For the two-hop relaying links, we consider mmWave propagation channels with limited scattering, which can be modelled by the narrowband clustered channel representation, based

on the extended Saleh-Valenzuela model [13]. We assume a scattering environment with N_{cl} scattering clusters randomly distributed in space and within each cluster, there are N_{ray} closely located scatterers.

The channel matrix between the k^{th} SBS and the RBS and between the RBS and the MBS can be expressed, respectively, as

$$\mathbf{H}_{k,RBS} = \sqrt{\frac{M_a^2}{N_{cl}N_{ray}}} \sum_{i}^{N_{cl}} \sum_{j=1}^{N_{ray}} \alpha_{ij}\, \boldsymbol{a}_{RBS}(\emptyset_{i,j}^{r}, \theta_{i,j}^{r}) \boldsymbol{a}_{k,SBS}^{*}(\emptyset_{i,j}^{t}, \theta_{i,j}^{t}) \,, \tag{12}$$

$$\mathbf{H}_{MBS} = \sqrt{\frac{M_a^2}{N_{cl}N_{ray}}} \sum_{i}^{N_{cl}} \sum_{j=1}^{N_{ray}} \alpha_{ij}\, \boldsymbol{a}_{MBS}(\emptyset_{i,j}^{r}, \theta_{i,j}^{r}) \boldsymbol{a}_{RBS}^{*}(\emptyset_{i,j}^{t}, \theta_{i,j}^{t}) \,, \tag{13}$$

where α_{ij} are the complex gains of the j^{th} ray in the i^{th} scattering cluster and are assumed i.i.d $\mathcal{CN}(0, \sigma_{\alpha,i}^2)$ with $\sigma_{\alpha,i}^2$ representing the average power of the i^{th} cluster, $\emptyset_{i,j}^{r}\ and\ \emptyset_{i,j}^{t}$ are the azimuth angles of arrival and departure respectively, $\theta_{i,j}^{r}\ and\ \theta_{i,j}^{t}$ are the elevation angles of arrival and departure respectively, $\boldsymbol{a}_{RBS}(\emptyset_{i,j}^{r}, \theta_{i,j}^{r})$, $\boldsymbol{a}_{MBS}(\emptyset_{i,j}^{r}, \theta_{i,j}^{r})$, and $\boldsymbol{a}_{k,SBS}(\emptyset_{i,j}^{t}, \theta_{i,j}^{t})$ represent the normalized array response vectors of the RBS, MBS, and the k^{th} SBS respectively.

It is assumed that the N_{ray} azimuth and elevation angles, $\emptyset_{i,j}^{r,t}\ and\ \theta_{i,j}^{r,t}$ are randomly distributed with a uniformly-random mean cluster angle of $\emptyset_{i}^{r,t} and\ \theta_{i}^{r,t}$ respectively, and a constant angular spread of $\sigma_{\emptyset^{r,t}}\ and\ \sigma_{\theta^{r,t}}$ respectively.

3 Proposed Hybrid Beamforming

The proposed hybrid beamforming is performed in two stages. First, the analog beamformers at the SBSs-to-RBS and RBS-to-MBS links select a set of beams using eigenbeamforming and imposing the phase-only constraint on each selected eigenvector. Beam selection can be realized by a network of RF switches that feed the data streams to the best ports (selected eigenvectors) of a Butler matrix. Once the analog beamformer is known, the transmit and receive digital weight vectors are obtained using the SINR-based MRT/MRC schemes.

3.1 SBSs-to-RBS

The transmit analog weight vectors of the k^{th} SBS are based on eigen-beamforming scheme and are given by

$$\begin{gathered} \boldsymbol{A}_{T,k}^{SBS} = \left[\boldsymbol{a}_{T,k,1}^{SBS}, \boldsymbol{a}_{T,k,2}^{SBS}, \cdots, \boldsymbol{a}_{T,k,L_d}^{SBS}\right] \\ subject\ to \left|\boldsymbol{A}_{T,k}^{SBS}(i,j)\right|^2 = 1 \end{gathered}, \tag{14}$$

where $\boldsymbol{a}_{T,k,i}^{SBS}$ denote the i^{th} selected $N_{RF} \times 1$ eigenvector corresponding to the i^{th} maximum eigenvalue of $\boldsymbol{H}_{k,\text{RBS}}^{H}\boldsymbol{H}_{k,\text{RBS}}$.

Assuming channel reciprocity, the receive analog weight vectors of the MBS could be chosen as $\boldsymbol{A}_{R,k}^{MBS} = \boldsymbol{A}_{T,k}^{SBS}$.

For fixed analog beamforming weights, $\boldsymbol{A}_{T,k}^{SBS}$ and $\boldsymbol{A}_{R,k}^{MBS}$, the transmit optimal digital weight vector of the k^{th} SBS, $\boldsymbol{D}_{T,k}^{SBS}$, and the receive optimal digital weight vector of the MBS, $\boldsymbol{D}_{R,k}^{MBS}$, are obtained by the MRT/MRC scheme that maximizes (4) and are given by

$$\boldsymbol{D}_{T,k}^{SBS} = \boldsymbol{D}_{R,k}^{MBS} = \boldsymbol{B}_{k,RBS}^{-1}\boldsymbol{\mathcal{H}}_{k,RBS}\boldsymbol{V}_{BL}\,, \tag{15}$$

where $\boldsymbol{V}_{BL}$ is the eigenvector corresponding to the maximum eigenvalue of $\left(\boldsymbol{\mathcal{H}}_{k,RBS}\right)^{\boldsymbol{H}}\boldsymbol{\mathcal{H}}_{k,RBS}$

3.2 RBS-to-MBS

For the RBS-to-MBS link, the transmit analog weights of the RBS and the receive analog weight vectors of the MBS are based on the singular value decomposition (SVD) of the channel matrix, $\mathbf{H}_{MBS}$:

$$\mathbf{H}_{MBS} = \boldsymbol{U}_{MBS}\,\boldsymbol{\Sigma}\,\boldsymbol{U}_{RBS}^{H} \tag{16}$$

where $\boldsymbol{U}_{MBS} \in \mathbb{C}^{M_a \times K}$ and $\boldsymbol{U}_{RBS} \in \mathbb{C}^{M_a \times K}$ are semi-unitary matrices and $\boldsymbol{\Sigma}$ is an $K \times K$ diagonal matrix with the largest K singular values $\sigma_1, \cdots, \sigma_K$ on its diagonal.

The transmit and receive analog weight matrices of the RBS, and the MBS can then be expressed, respectively, as

$$\begin{aligned} \boldsymbol{A}_T^{RBS} &= \left[\boldsymbol{a}_{T,1}^{RBS}, \boldsymbol{a}_{T,2}^{RBS}, \cdots, \boldsymbol{a}_{T,K}^{RBS}\right] = \boldsymbol{U}_{RBS}\ , \\ \boldsymbol{A}_R^{MBS} &= \left[\boldsymbol{a}_{R,1}^{MBS}, \boldsymbol{a}_{R,2}^{MBS}, \cdots, \boldsymbol{a}_{R,K}^{MBS}\right] = \boldsymbol{U}_{MBS}\ , \\ subject\ to\ &|\boldsymbol{A}_T^{RBS}(i,j)|^2 = |\boldsymbol{A}_R^{MBS}(i,j)|^2 = 1\ . \end{aligned} \tag{17}$$

Using (7) and (8), $\hat{\boldsymbol{x}}_{MBS}$ and $\hat{\boldsymbol{x}}_{k,MBS}$ can be expressed as

$$\hat{\boldsymbol{x}}_{MBS} = (\boldsymbol{D}_R^{MBS})^H\,\boldsymbol{\Sigma}\,\boldsymbol{D}_T^{RBS}\hat{\boldsymbol{x}}_{RBS} + (\boldsymbol{D}_R^{MBS})^H(\boldsymbol{A}_R^{MBS})^H\mathbf{n}_{MBS}\,, \tag{18}$$

$$\begin{aligned} \hat{x}_{k,MBS} &= \sigma_k\left(d_{R,k}^{MBS}\right)^*\left(d_{T,k}^{RBS}\right)\hat{\boldsymbol{x}}_{k,RBS} + \left(d_{R,k}^{MBS}\right)^*\left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{n}_{MBS} \\ &= \sigma_k\left(d_{R,k}^{MBS}\right)^*\left(d_{T,k}^{RBS}\right)\left(\mathbf{S}_k^{RBS} + \mathbf{S}_{I_k}^{RBS} + \mathbf{N}_{RBS}\right) + \left(d_{R,k}^{MBS}\right)^*\left(\boldsymbol{a}_{R,k}^{MBS}\right)^H\mathbf{n}_{MBS} \end{aligned}\ , \tag{19}$$

The SINR of the k^{th} SBS at the MBS, given in (11), can then be simplified as

$$\gamma_k^{MBS} = \mathbf{B}_{k,MBS}^{-1}\left|\sigma_k\left(\boldsymbol{D}_{R,k}^{RBS}\right)^H\boldsymbol{\mathcal{H}}_{k,RBS}\boldsymbol{D}_{T,k}^{SBS}\right|^2\,, \tag{20}$$

where $\mathbf{B}_{k,MBS} = \sigma_{\text{k}}^2\left|\left(d_{T,k}^{RBS}\right)\right|^2\left(\mathbf{S}_{I_k}^{RBS}\left(\mathbf{S}_{I_k}^{RBS}\right)^H + \mathbf{N}_{RBS}\mathbf{N}_{RBS}^H\right) + \sigma_{n_{MBS}}^2\left|\left(d_{T,k}^{RBS}\right)\right|^2\left(\boldsymbol{a}_{k,R}^{MBS}\right)^H\boldsymbol{a}_{k,R}^{MBS}$,

Note that the SINR, γ_k^{MBS}, given in (20) is independent of the digital beamformers, $\boldsymbol{D}_T^{RBS}$ and $(\boldsymbol{D}_R^{MBS})^H$, of the RBS-to-MBS link. This property enables us to choose the optimal digital beamformers that satisfy $\boldsymbol{D}_T^{RBS}(\boldsymbol{D}_R^{MBS})^H \propto \mathrm{I}_K$ or simply choose $\boldsymbol{D}_T^{RBS} \propto \mathrm{I}_K$ and $\boldsymbol{D}_R^{MBS} \propto \mathrm{I}_K$.

The ergodic channel capacity for each user l_s is given by [20],

$$\mathcal{C} = \mathbb{E}[log_2\{1 + \gamma_k^{MBS}\}]\,, \tag{21}$$

where $\mathbb{E}$ [.] denotes the expectation operator.

4 Simulation Results

In our simulation setups, we consider six SBSs (K=6) connected to one macro-cell through one relay station. The SBSs, the RBS, and the MBS use the same number of antennas, $M_a = 64$. The number of RF chains for SBSs-to-RBS links is $N_{RF} = 2 \; or \; 4$. We assume QPSK modulation.

Figure 2 shows the beampattern of the proposed HBF with four RF chains and the optimal fully-digital one for the SBSs-to-RBS links. The optimal beamformer has about four dominant beams that are similar to the selected beams of the proposed HBF, which means that near-optimal performance could be achieved by transmitting data streams through those four beams. Figure 3, on the other hand, compares the ergodic channel capacity of the proposed HBF and the optimal fully-digital one. It is observed that as we increase the number of RF chains, the performance gap between the two schemes is reduced, and the near-optimal solution was achieved by the proposed HBF using four RF chains. Compared to the single cell MU-MIMO case presented in [12],[13],[14], near-optimal performance was obtained with only five RF and for the MU-MIMO case in [16],[17] it was shown that the required number of RF chains could be reduced to two to achieve the fully-digital beamforming performance. However, unlike our case, where we have focused on the backhaul link and assumed a two-hop relay link that connects multiple small cells to a macro cell, these studies focused primarily on macro-cellular systems.

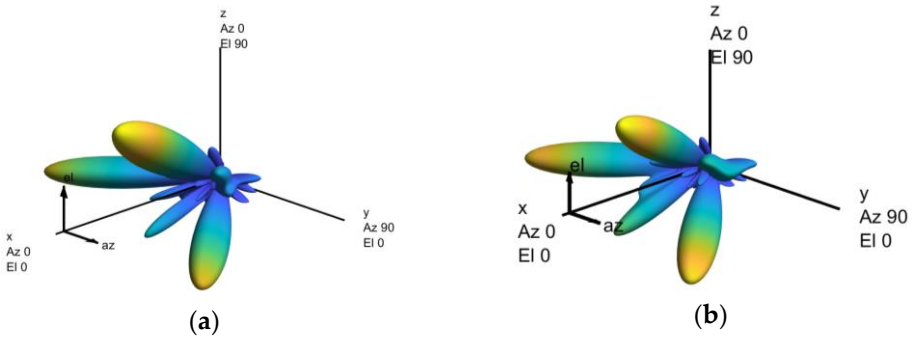


Fig. 2 Beampattern of the access link: (**a**) Proposed HBF, 4 RF chains; (**b**) Optimal beamforming.

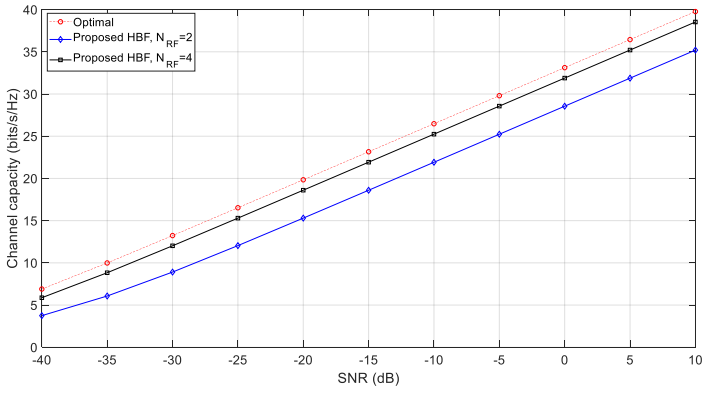


Fig. 3. Channel capacity for a different number of RF chains: Proposed HBF vs optimal

5 Conclusion

In this paper, we extended hybrid beamforming to relay-aided mmWave backhaul links where multiple small cell base stations (SBS) are connected to a macro-cell base station (MBS) through a two-hop backhaul with manageable interference between the SBSs. The performance evaluation in terms of the beampatterns and the ergodic channel capacity shows that the proposed HBF scheme achieves near-optimal performance with only four RF chains and requires considerably less computational complexity.

Acknowledgments. The author would like to thank the Canadian Microelectronics Corporation (CMC) for providing the Heterogeneous Parallel Platform to run the computationally-intensive Monte-Carlo Simulations.

References

[1] Marzetta T. L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Tran. Wirel. Comm. 2010, 9(11), 3590 – 3600.

[2] Rusek F., Persson D., Lau B., Larsson E., Marzetta T. L., Edfors O., Tufvesson F. Scaling up MIMO: Opportunities and challenges with very large arrays. IEEE Signal Process. Mag. 2013,30,1,40–60.

[3] Ngo H. Q., Larsson E. G., and Marzetta T. L. Energy and spectral efficiency of very large multiuser MIMO systems. IEEE Trans. Commun. 2013, 61(4),1436–1449.

[4] Hoydis J., ten Brink S., Debbah M. Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?. IEEE J. Sel. Areas Commun. 2013, 31(2), 160–171.

[5] Siddique, U.; Tabassum, H.; Hossain, E.; Kim D. I. Wireless backhauling of 5G small cells: Challenges and solution approaches. IEEE Wireless Communications, Special Issue on "Smart Backhauling and Fronthauling for 5G Networks. 2015, 22, 22-31.

[6] Gao, Z.; Dai, L.; Mi, D.; Wang, Z.; Imran, M. A.; Shakir, M. Z. MmWave Massive MIMO Based Wireless Backhaul for 5G Ultra-Dense Network. IEEE Wireless Communications. 2015, 22, 13-21.

[7] Tabassum H., Hamdi S. A., and Hossai E. Analysis of massive MIMO-enabled downlink wireless backhauling for full-duplex small cells. IEEE Transactions on Communications. 2016, 64(6), 2354-2369.

[8] Shariat M., Pateromichelakis E., Quddus A., Tafazolli R. Joint TDD Backhaul and Access Optimization in Dense Small-Cell Networks. IEEE Transactions on Vehicular Technology. 2015, 64(11), 5288-5299.

[9] ElSawy H., Hossain E., Kim D. I. HetNets with cognitive small cells: User offloading and Distributed channel allocation techniques. IEEE Communications Magazine, Special Issue on "Heterogeneous and Small Cell Networks (HetSNets). 2013, 51(6), 28-36.

[10] Yan Z., Zhou W., Chen S., Liu H. Modeling and Analysis of Two-Tier HetNets with Cognitive Small Cells. IEEE Access. 2016, 5, 2904-2912.

[11] Hefnawi M. Capacity-Aware Multi-User Massive MIMO for Heterogeneous Cellular Network. In Proceedings of the IEEE International Conference on Selected Topics in Mobile and Wireless Networking, Tangier, Morocco, 20-22 June 2018.
[12] Sohrabi F. , Yu W. Hybrid beamforming with finite-resolution phase shifters for large-scale MIMO systems. In Proceedings of the IEEE Workshop Signal Process. Adv. Wireless Commun., Stockholm, Sweden, 28 June-1 July 2015.
[13] El Ayach O., Rajagopal S., Abu-Surra S., Pi Z., and Heath R. Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wireless Commun. 2014, 13(3), 1499-1513.
[14] Alkhateeb A., El Ayach O., Leus G., Heath R. Channel estimation and hybrid precoding for millimeter wave cellular systems. IEEE J. Sel. Topics Signal Process. 2014, 8(5), 831-846 .
[15] Sohrabi F. , Yu W. Hybrid digital and analog beamforming design for large-scale MIMO systems. In Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Brisbane, Australia, Apr. 2015.
[16] Liang L., Dai Y., Xu W., Dong X. How to approach zero-forcing under RF chain limitations in large mmwave multiuser systems?. In Proc. IEEE/CIC Int. Conf. Commun. , Shanghai, China , 2014.
[17] Liang L., Xu W., Dong X. Low-complexity hybrid precoding in massive multiuser MIMO systems. IEEE Wireless Commun. 2014.
[18] S. Han, I C.-L., Xu Z., and Rowell C.. Large-scale antenna systems with hybrid precoding analog and digital beamforming for millimeter wave 5G. IEEE Commun. Mag. 2015, 53(1), 186 –194.
[19] Kang M. A comparative study on the performance of MIMO MRC systems with and without cochannel interference. IEEE Transactions on Communications. 2004, 52(8), 1417 – 1425.
[20] Sulyman A. I. , Hefnawi M. Adaptive MIMO Beamforming Algorithm Based on Gradient Search of the Channel Capacity in OFDMSDMA System. IEEE Commun. Letters. 2008, 12(9), pp. 642-644.

Noise Reduction using DUET algorithm for dual-microphone mobile station

[1]Aoulass Nabila, [2]Chakkor Otman,
{ [1]nabilaaoulass@gmail.com, [2]o.chakkor@gmail.com }

[1]Signal and image processing, university Abdlmalek Esaadi ENSAT laboratory NTT
[2]Telecommunication systems and networks engineering, University Abdlmalek Esaadi ENSAT

Abstract. This paper presents a novel technique which can improve the quality of signal used for mobile station using dual-microphone. Our method is based on the degenerate unmixing estimation technique (DUET). This technique can separate any number of sources using only two mixtures. The DUET algorithm is used as a tool to denoise a received signals. Multiple sources are acquired by two sensors. The received signals are called mixtures, witch they processed in order to recover only one source signal, whereas others are considered as noise.

Keywords: Blind Separation of sources, Degenerate Unmixing Estimation Technique, histogram, time-frequency masking.

1 Introduction

Blind source separation (BSS) consists of estimating a set of latent signals from a set of observed signals, called mixtures. Typically, the observations are obtained at the out put of a set of sensors, where each sensor receives a different combination of the source signals. The adjective ”blind” high lights two facts[1] :

- The source signals are not observed.
- No information is available on the mixing system.

The BSS is a very active research subject because of its many practical applications. In the field of audio signal separation, we find the ”cocktail-party” problem, that we encounter when we record several speakers who speak simultaneously, and that we want to isolate a speaker. Radar and sonar when signals from multiple transmitters interferes at the receiver antenna, and in many other areas. The lack of prior knowledge of the source signals and the mixing system is always compensated by assumptions that unknown sources should answer[1].

Some common assumptions are that the sources must be statistically independent. The implementation of ICA algorithm is the Fast Independent Component Analysis (FastICA). The main problem using FastICA is that it assumes instantaneous mixing without time delay in the recording process.

As we know, many techniques based on ICA and among the recently proposed methods are found the DUET (Degenerate Unmixing Technical Estimation) wich is one of the demixing

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284118

algorithms in the field of BSS. It can separate any number of sources using only two mixtures, such this method is based on W-disjointorthogonal sources.

There are several algorithms for signal separation that use similar or even completely different concepts.

DUET is a solution for separating for signals using two microphones locatd in a mobile station(MS). This method is suited for audio source separation. To realize this separation we will implement the DUET solution for speech signals in real evironnement with a noisy background.

The DUET algorithm is effective for sources with parsimony property in the time-frequency domain, such as the voice signal, that is, the target speech signal in a noisy environment can be effectively recognized. However, in many cases, there are more sources than mixtures, so we refer to a case as degenerate[5]. In the degeneration of the blind source, separation poses a challenge because the mixing matrix is not reversible.

Basically, the traditional method such as the ICA of the demixing by estimating the inverse mixture matrix of which the DUET algorithm operates in the frequency domain, no inverse matrix is calculated and this is one of the reasons why it shows very good performance. Another advantage over ICA is that the number of sources can be greater than the number of mixtures, infact it can be used for an arbitrary number of sources.

ICA methods have several drawbacks. Often, it is required that the number of source signals is known in advance, which is not our case.

When a mobile user talked in a public place e.g. office, markets, universities, a speech noise is created. Then, the corrupted speech needs to enhance in a minimum quality so that the user in opposite side can understand. Speech enhancement algorithms are, therefore, developed and improved depending on cost effectiveness and maintaining the minimum quality multiple unwanted speakers.

2 Degenerate Umnixing Estimation Technique

2.1 Introduction

The Degenerate Unmixing Estimation Technique (DUET) is an algorithm that allows us to blindly separate an arbitrary number of sources with only two sensors. The received signals are called mixtures that provide the time- frequency representations of the sources. DUET assumes the sources are already separate in the time-frequency plane.

The sources are disjoint they do not overlap too much, which is ideal for speech[5]. This method applied when the sources are W-disjoint orthogonal, when the time-frequency representations of any two signals in the mixtures are disjoint sets.

The principle of this method is to perform a search for amplitude-Delay mix parameters, and simultaneously construct time-frequency bit masks that are used to partitionone of the mixtures to recover the original source signals[5].

The blind separation of the audio signals using the DUET algorithm requires following a set of successive steps in order to estimate the sources:

2.2 Creation of mixing model

Firstly we need to establish a model to describe the mixed signals. In our case, we have two microphones that can capture the k sources that are represented by a number of people who speak in the scene, are supposed to be on different locations.

The mixed signals x_1 and x_2 can be described as follows:

$$x_1(n) = \sum_{j=1}^{k} s_j(n) \quad . \tag{1}$$

$$x_2(n) = \sum_{j=1}^{k} a_j(n).s_j(n - \delta_j) \quad . \tag{2}$$

where a_j is a relative attenuation factor, and δ_j is the arrival delay between the sensors.

2.3 Hamming window and FFT

In order to avoid spectral leakage, a Hamming window can be applied to a signal in the frequency domain[7]. The Hamming window has relatively low values, close to zero,at the ends and values up to one, in the center of the window.

When the Hamming window is applied to a signal,the amplitudes of the frequencies in the highest and lowest part of the frequency spectrum are attenuated[7]. The amplitude of the frequencies in the middle of the spectrum remains essentially the same[7].

The x_1 and x_2 mixtures are sampled and divided into blocks of length N with overlap. These sample blocks are multiplied with the Hamming W windowing function and the n the Fast Fourier transform is applied.

$$x^{\theta}_{1,2}(n) = W(n)x_{1,2}(n) \quad . \tag{3}$$

$$X_{1,2}(f) = \sum_{n=0}^{N-1} x^{\theta}_{1,2}(n).e^{\frac{-i2\pi nf}{N}} \quad . \tag{4}$$

2.4 Estimation of amplitude-delay parameter

So if we will refer to the Fourier transform of sources $S_j(f)$ For a given source j we can describe their FFT as follows:

$$\begin{bmatrix} X_1(f) \\ X_2(f) \end{bmatrix} = \begin{bmatrix} 1 \\ a_j e^{-i2\pi f \delta_j} \end{bmatrix} S_j(f) \quad . \tag{5}$$

The DUET algorithm is based on the property of W-disjoint orthogonality between the sources, and each source has a sparse frequency spectrum for a given time compared to the other such that[8]:

$$S_j(f).S_i(f) = 0, \quad \forall i, j \quad i \neq j \quad . \tag{6}$$

They then assume that the sources are orthogonal W-disjoint in time-frequency plane associated with the transform considered with the window W[8]. This strong hypothes is means that the sources have no overlap in the time frequency domain. The associated amplitude and delay to a source are then given by:

Where:

$$a_j, \delta_j = \left(\| \frac{X_2(f)}{X_1(f)} \|, Im(\log(\| \frac{X_2(f)}{X_1(f)} \|))/2\pi f \right) \quad . \tag{7}$$

$$\frac{X_2(f)}{X_1(f)} = a_j e^{-i2\pi f \delta_j} \quad . \tag{8}$$

After we added the time dimension, and then realize a time-frequency study, leading to: Using the hypothesis that the sources are Wdisjoint-orthogonal, and performs an amplitude/delay histogram of all the values of $a_j;\delta_j$ encountered of which this histogram has peaks around the values corresponding to the parameters of the mixing matrix seen in the time-frequency plane.

2.5 Separation via binary masking and IFFT

This step consists in generating a bit mask for each pair of parameters (a_j,δ_j) of mixture corresponding to the time-frequency points which produce particular mixing parameters then the sources are separated by multiplying each mask by a mixtures[8].

The mask is defined as:

$$M_j(t, f) = \begin{cases} 1 & si \quad J(t, f) = j \\ 0 & elsewhere \end{cases} \quad . \tag{9}$$

As J is the objective function obtained by maximum likelihood ML.

Essentially,weas sign each time-frequency point to the pair of mix parameters,which is the best explains the mixes at that particular timefrequency point. We denoise the signals via masking and the ML combination:

$$J(t,f) = ArgMin_j \frac{| a_j e^{-i2\pi f \delta_j} X_1(t,f) - X_2(t,f) |^2}{1 + a_j^2} \quad . \tag{10}$$

We can then reconstruct the sources from their timefrequency representations by converting them back to the time domain using the inverse Fast Fourier transform(IFFT).

3. Simulation

We implement blind signal reduction using DUET algorithm. We use two mixed signals x_1 and x_2 from five sources. It means that the mobile station use two microphones, a first one is used for a short acoustic path between mouth,which ensures less reverberation and high direct path power. A second microphone is placed on the back of the mobile station in order to capture the speech signal with a lower sound pressure level.

Steps to proceed for noise reduction are summarized as follows:

• **Calculation of FFT and Hamming window**:

The first step is to apply FFT and then build a time frequency representation of the two mixtures, so we are interested in signals of limited length, and a real signal can only have a limited duration. This can only be done on a finite number of points using the hamming window

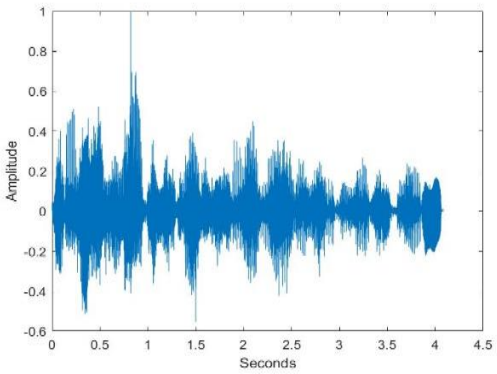


Fig. 1. Ar Mixed signal x_1.

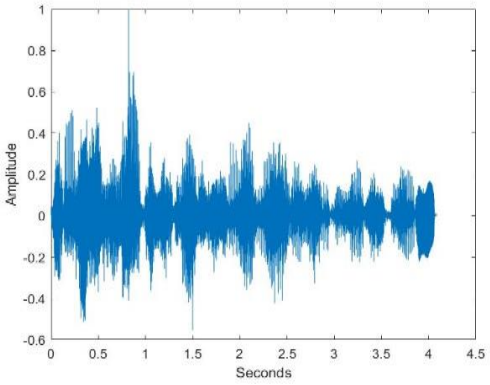


Fig. 2. Mixed signal x_2.

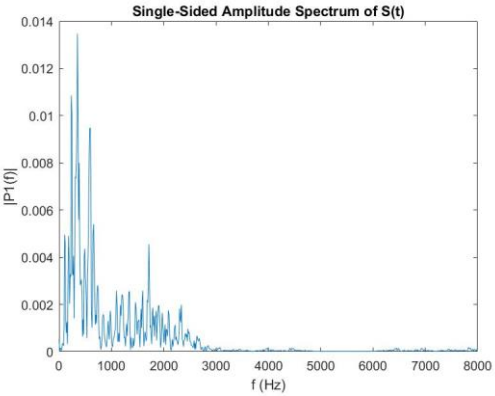


Fig. 3. x_1 after FFT a short term.

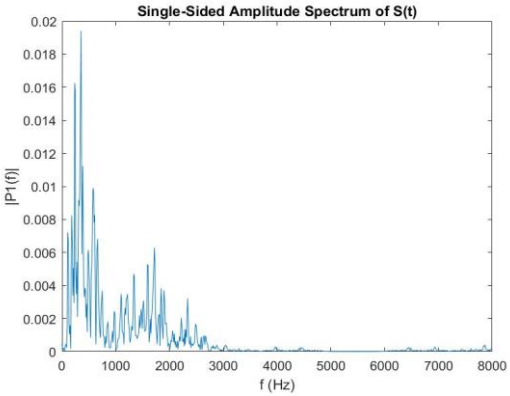


Fig.4 . x_2 after FFT a short term.

• **Calculation of attenuation parameters and the weighted histogram:**

For each time-frequency point, the phase and amplitude will be calculated, then we build a 2D histogram (one dimension is the phase, the other is the amplitude). This histogram present speaks around the values corresponding to the parameters of the mixing matrix in the time-frequency plane.The parameters p and q represent the weight of the histogram. The histogram in the figure 5 shows five peaks associated with the two mixtures(x_1 and x_2). The peaks are exactly the pairs of mixing parameters for each of the five sources. Using the calculated symmetric attenuation $\alpha_j = a_j - 1/a_j$: It is weighted in terms of two-dimensional cross power (p = 1,q = 0),and pairs of two mixtures (x_1 and x_2)of five sources.

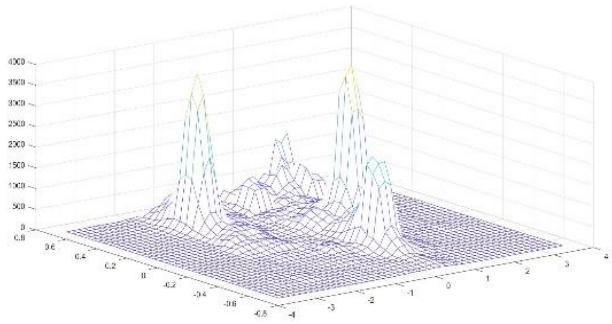


Fig. 5. calculation of histogram

• **Separation of sources:**

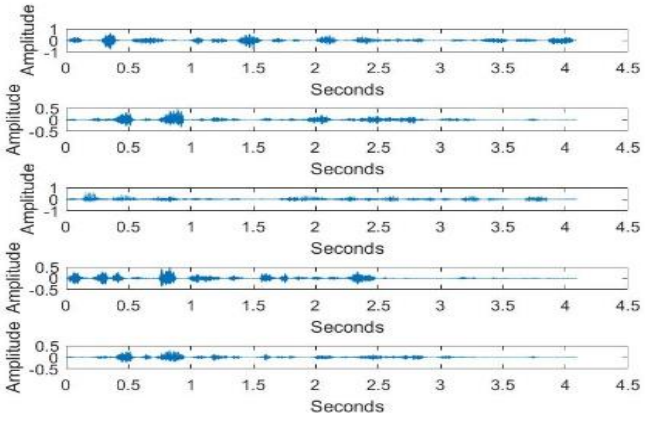


Fig. 6. estimated sources

We generate a binary mask for each pair of mixing parameters at time-frequency points, which produce special mixing parameters, and then separate the sources by multiplying each mask by one of the mixtures, then weapply IFFT to represent the estimated sources in the time domain,figure 6.

3. Conclusion

We propose a solution for noise reduction for speech enhancement in dual-microphone mobile phones. The solution is based on the DUET blind signal separation. The principle of DUET is based on the W-disjoint orthogonal assumption of source signals, which is verified for mobile station conversations. The simulations with recorded signals shows a good performance with low complexity of the system, making it possible for an easy integration in to mobile stations.

References

[1] K.Itoh and M. Mizushima.Environmental noise reduction based on speech/non-speech identifiation for hearing aids. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing(ICASSP),vol.1,pp.419-422,1997.

[2] Zbynek Koldovsky, Petr Tichavsky, David Botka. Noise reduction in dual-microphone mobile phones using a bank of pre-measured target-cancellation filters. Acoustics, Speech and Signal Processing (ICASSP),2013IEEEInternationalConferenceon,26-31May2013.

[3]Alexander A SGunawana, Albert Stevelino, Heri Ngarianto, Widodo Budiharto, Rini Wongsob. Implementation of Blind Speech Separation for Intelligent Humanoid Robot using DUET Method. Procedia Computer Science,Volume116,2017,Pages87-98.

[4]BELL, A. J., SEJNOWSKI, T. J An information maximization approach to blind separation and blind deconvolution. Neural Computation,1995,pp.1129-1159.

[5] CARDOSO, J.-F Blind signal separation: statistical principles. Proceedings of the IEEE, 1998, vol.86, no.10, pp.2009-2025.

[6] M.Billette Analyse en composantes indépendantes avec une matrice demélange sparse. Université de Montréal des études supérieure sen vue de l'obtention du grade de Maitres sciences(M.Sc.) en Statistique, 20juin2013.
[7] Yang Wang, Orgür Yilmaz, Zhengfang Zhou. Phase aliasing correction for robust blind sources eparation using DUET.IEEE transaction on signal processing.
[8] Scott Rickard, S. Makino et al.The DUET blind source separation algorithm in Blind speech separation, Springer,pp.217-241,2007.
[9] Alexander Jourjini, Scott Rickard, Orgür Yilmaz, Blind source separation of disjoint orthogonal signal: Demixing N sources from two mixtures". Proc ICASSP2000, vol. 15, pp. 2985-2988, June 5–9, 2000.
[10] ScottRichard, The DUET Blind Source Separation Algorithm. University College Dublin 2007 Springer.
[11] N. Yousefian and P. C. Loizou, A dual-microphone algorithm that can cope with competing-talker scenarios, IEEE Trans. on Audio, Speech,and Language Processing,vol.21,no.1,pp.143–153,2013.

Sharing a Secret Image Based in CMY model using RSA Method in Visual Cryptography

[1]Asma Hilmi, [2]Soufiane Mezroui, [3]Ahmed El Oualkadi
{ [1]asmaa00hilmi@gmail.com, [2]mezroui.soufiane@yahoo.fr, [3]eloualkadi@gmail.com }

Abdelmalek Essaadi University. National school of applied sciences of Tangier (ENSAT), Laboratory of Information and Communication Technologies (LabTIC), ENSA Tanger, Route Ziaten, BP 1818, Tanger principale, Morocco

Abstract. Many techniques in visual cryptography are proposed for sharing a secret message. Among the methods cited in the literature the one of creation the secure shares based in RGB color, using AES (Advanced Encryption Standard),and ECC (Elliptical Curve Cryptography) algorithms. Our proposed technique consists to share the secret image in multiple shares. First the original image is decomposed to the CMY colors, after we created the multiple shares. This multiple shares are used to transfer the secret image by using the encryption and decryption RSA algorithm. The principle consists the created the number of shares based on the number of basics matrices, after we apply the RSA algorithm in the encryption process for each share and decryption process.

Keywords: visual cryptography, XOR, share, RSA, encryption, decryption

1 Introduction

In the face of cyber theft, internet thread and hacker, there is a big demand to organisation for realising a transmission secure and realable of these informations and properties. Based on Naor and Shamir method [1], visual cryptographic scheme (VCS) is a special type of secret sharing scheme, where the secret is a black and white image. A VCS is able to code a secret image into k shadow image, the general principle that with k or more than k we reconstruct the secret image , but we cannot have any information with less than k. All share is printed on a distinct transparency [2], encryption and decryption method are done for each share.When all shares are superimposed , the original image should be visible. color visual cryptography is presented [3], which can encrypt the color secret messages into RGB color or with the complementary color CMY (Cyan, Magenta, Yellow).Two mathematical operations used: the booleen OR and the booleen XOR [4]. However the schemes based on operation OR (OVCS) suffer from the huge size of sharing (reflected by the pixel expansion) and poor quality (reflected by contrast) of the recovered secret image. It is for this reason that the schemes based on operation XOR (XVCS) has been studier to achieve advanced properties such as good contrast and resolution.
This paper is organized as follows. A review of some related works is given in Section II. Section III the proposed methodology and gives an overview about RSA cryptosystem. Section IV presents the implementation of the proposed method scheme, the original image,

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284210

the shares and the encrypted and decrypted shares. Finally, the conclusion is given in Section V.

2 Related Works

In [5], Shankar K and Eswaran created a modern cryptographic technique of sharing a secret with encapsulated shares. They created a separate matrix Ri, Gi, Bi from the RGB color .The basic matrices of RGB color are obtained by dividing each and every value in Ri, Gi and Bi by 2. Once the shares are created, it is encrypted separately by using Advanced Encryption Standard (AES) algorithm.

In [6,7], the proposed method of Shankar K and Eswaran is used to create the shares from their pixel value of RGB, the extracted pixel values are used to create the multiple shares (share1, share2,….share n) and then the image shares are divided into blocks. The blocks of each share are encrypted by using ECC (Elliptical curve cryptography) method and the encrypted image is decrypted by using the decryption of ECC method.

3 Proposed methodology

As we reworked in related work, the methods proposed in the state of art consist in creating the shares of each color R, G and B, after encrypting and decrypting them. In our proposed methodology, we have tried to create the shares from the three colors, ie; each share will contain Ci + Mi + Yi.

The proposed methodology is used to send the original image from the sender to the receiver. The figure 1 shows the block diagram of the proposed method.

From the original image we based on the model CMY to present the colors, because (R, G, B) and (C, M, Y) are complementary colors. In fact the two models colors (R, G, B) and (C, M, Y) have the following relationships: C = 255-R; M = 255-G; Y = 255-B; white (0, 0, 0) and complete black (255,255,255).

The pixel values are taken into consideration to create the basic matrices (Ci,Mi,Yi). The basic matrices C1,C2,C3, M1,M2,M3 and Y1,Y2,Y3 are obtained by using modulo 3 in Ci,Mi and Yi. After we created the shares, in this case we have 2^3 shares for each Ci,Mi,Yi colors. The multiple shares are encrypted by using the RSA method, and the encryption shares are decrypted by using RSA method. The public and private keys in encryption and decryption are using randomly. The final image will be compared to the original image.

2.1 Shares creation

The original image is extracted in three pixel color CMY and given by,

$$PIXEL = \sum C + M + Y \tag{1}$$

A pixel value is considered for each color. With these pixel value, we crate the basics matrices., these values are separated in three basic matrices according to the following relations:

$$C1 = M1 = Y1 = 1 \bmod 3 \quad (2)$$

$$C2 = M2 = Y2 = 2 \bmod 3 \quad (3)$$

$$C3 = M3 = Y3 = 1 \bmod 3 \quad (4)$$

We have three basics matrices (i=3), then the number of shares is equal 2^3 =8.

The basic matrices constructed from the pixel value of the C,M, Y and the shares are created according to the following :

$$Cs1 = C1 \oplus Km \oplus 128 \quad (5)$$

$$Cs2 = Cs1 \oplus C2 \quad (6)$$

$$Cs3 = C \oplus 128 \quad (7)$$

$$Cs4 = Cs3 \oplus Km \quad (8)$$

$$Cs5 = Cs1 \oplus C \quad (9)$$

$$Cs6 = C \oplus C2 \oplus Km \oplus 128 \quad (10)$$

$$Cs7 = C \oplus C3 \oplus Km \oplus 128 \quad (11)$$

$$Cs8 = Cs6 \oplus Km \quad (12)$$

Csi, i=1...8, denote the shares, Ci, i=1..3, denote the basic matrices, C is the Cyan color and Km is the generated random key, its size is that of the original image.

We repeat the same process to two other pixels to get the eight shares of Magenta color; Ms1,Ms2....Ms8, and the eight shares of yellow color; Ys1, Ys2....Ys3.

2.2 RSA cryptosystem

RSA Cryptosystem is one of the first public-key cryptosystems and is widely used to secure data transmission. In such a cryptosystem, the encryption key is public and it is different from the decryption key which is kept secret (private). [8].

The RSA public-key cryptosystem was developed by R.L. Rivest, A. Shamir, and L. Adleman in 1978 [9]. The RSA cryptosystem is simply the modular exponentiation.

The modulus n is the product of two large prime's p and q, public and private keys are obtained by:

$$ed = 1 \ mod \ \varphi(n) \quad (13)$$

The encryption operation is performed using the public key n and e as follows:

$$C = M^e \ (mod \ n) \quad (14)$$

Where M is the plaintext such that 0<M<n and C is the ciphertext which can be decrypted using the private key n and d as follows:

$$M = C^d \ (mod\ n) \tag{15}$$

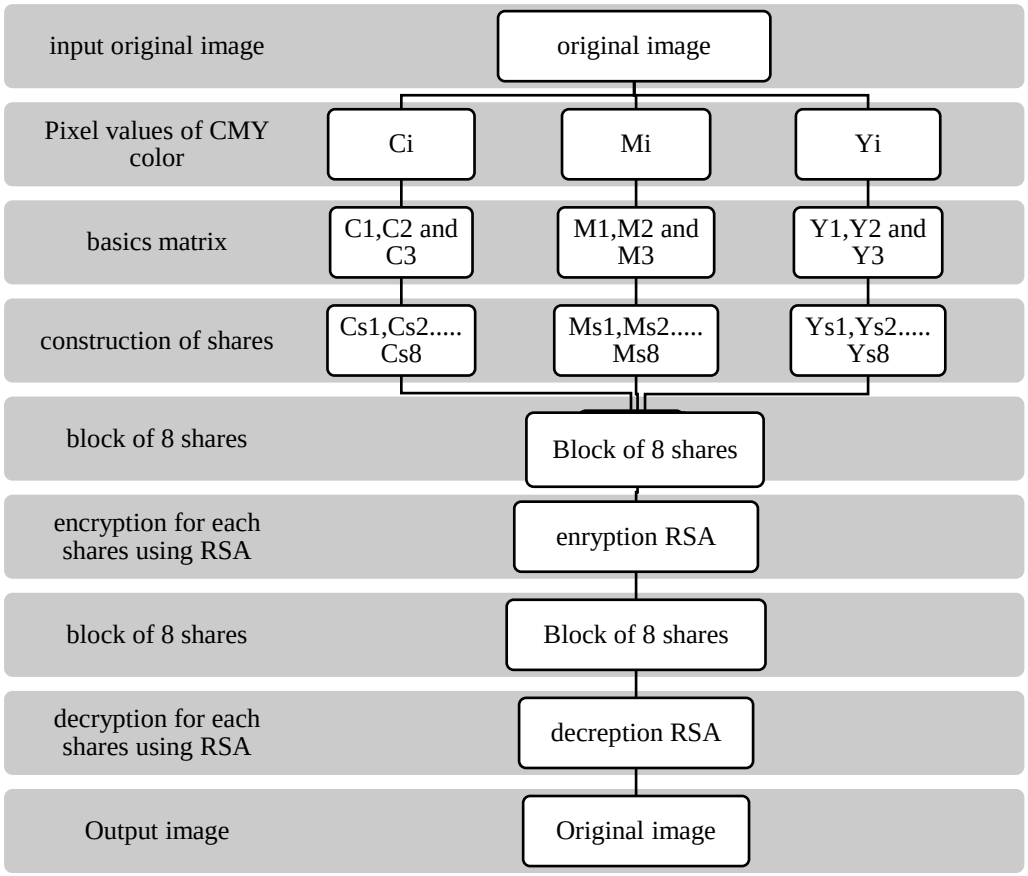


Fig. 1. Block diagram of the proposed method.

3 Results and discussions

In this paragraph we have presented the methods to realize our proposition, as well as a recapulative table which presents the shares, encrypted and decrypted shares.

3.1 Pseudo code of proposed methodology

In Step1, we input the original image, after we extract the pixel values from the input image to create the basic matrices. One key matrix is generate randomly with size of the original image. The multiple shares are created from CMY pixel values where $s=2^i$. The shares for each color are created using the XOR function according to the algorithm above.

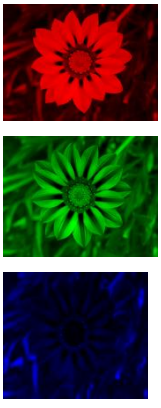
SHARE1= Cs1+Ms1+Ys1.
SHARE2= Cs2+Ms2+Ys2.
SHARE3= Cs3+Ms3+Ys3.
SHARE4= Cs4+Ms4+Ys4.
SHARE5= Cs5+Ms5+Ys5.
SHARE6= Cs6+Ms6+Ys6.
SHARE7= Cs7+Ms7+Ys7.
SHARE8= Cs8+Ms8+Ys8.

We apply the encryption method based on RSA method for each shares using the public key, in the decryption method private key is used to decrypt the 8 encrypt shares. All decrypt shares are stacked together to get the original image.

3.2 Experimental results

The proposed method is implemented by using MATLAB Simulator. The experimental result obtained from the implementation of the proposed method scheme, the original image, the shares and the encrypted and decrypted shares are represented in table2.

Table 1. Decomposition of image in RGB model and CMY model.

Image RGB	**Decomposition color RGB**	**Convert to CMY**	**Image CMY**
			

.

Table 2. Proposed method the secret image, the shares, encrypted, decrypted and final image.

Input image	Eight shares	Encrypted shares	Decryprted shares	Output image	Convert to RGB

3.3 Discussions and Comparisons

The simulation gives three finals images; the first we apply the XOR with the image given after decryption and the model RGB, the second is the image realized after decryption, the third we apply the XOR with the image given after decryption and the model CMY.

The table 2 summarizes the results of our work, also it compares our results with the case of Shankar and Erswran.

The SHANKAR and ERSWAN method is working on the ordinary decomposition color RGB. In our methodology we work with the CMY model, which presents more details of the colors. The choice of RSA cryptosystem is based on the type of encryption that is encrypted by block of characters, which prevents a cryptanalyst from using the frequency analysis to try to break the keys, Moreover, with this method; the encryption is asymmetric which solves the problem of key communication. However AES is a symmetric encryption and ECC needs more computation during encryption and signature verification.

The figure 2 presents a comparison the different steps of construction the two methods. SHANKAR and ERSWAN method create 4 shares for each color, which gives 12 shares to encrypt and decrypt, that increases the processing time of encryption and decryption, on the other hand in our methodology, we have encrypted and decrypted only 8 shares. The algorithm proposes in our method to mix the three colors in each share increases the level of security and make the deciphering for a hacker a difficult thing.

Table 3. Table presents a comparison to results of our methodology and Shankar and Eswran methodology.

	Color decomposition	Basics matrix	Number of shares	Encryption and Decryption method
Our Methodology	**CMY**	**3**	**8**	**RSA**
SHANKAR,ERSWAN Methodology[5,6,7]	**RGB**	**2**	**4**	**AES, ECC**

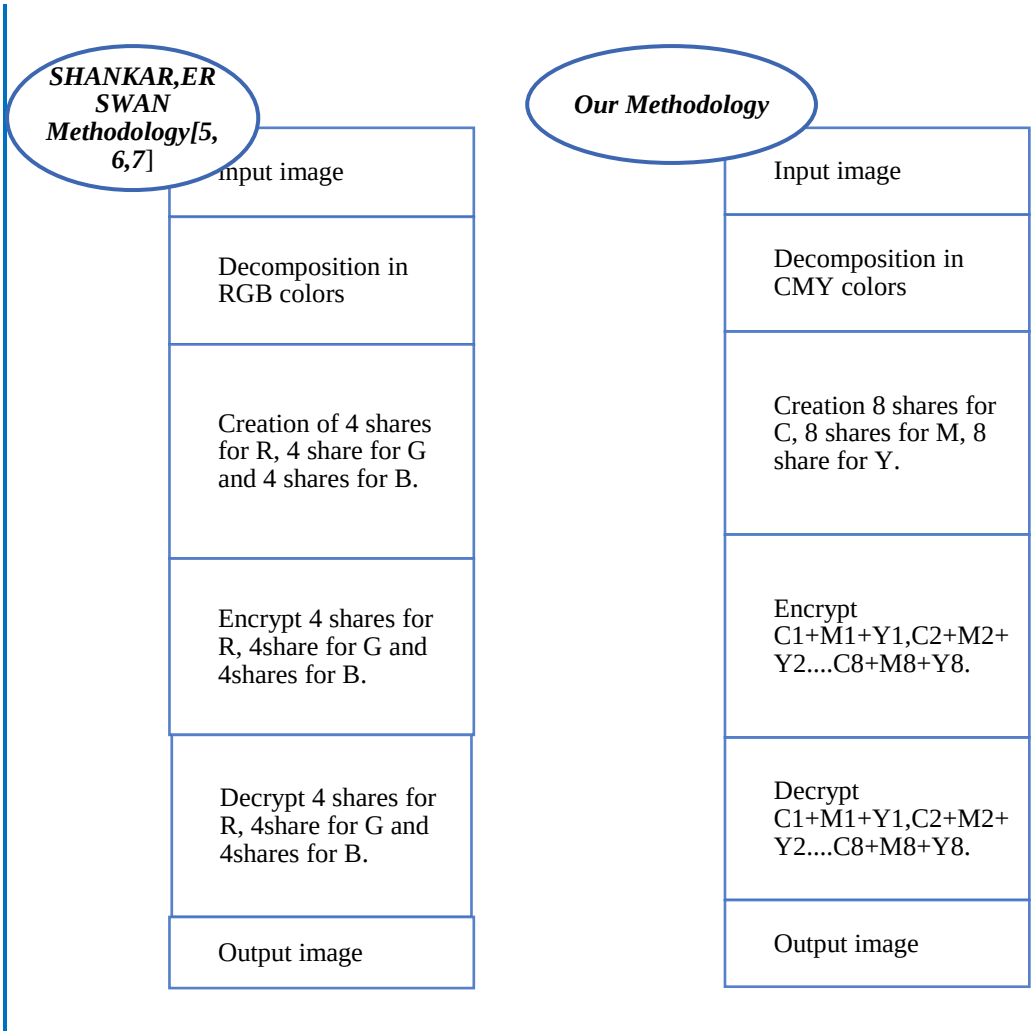


Fig. 2. Comparison between our methodology and SHANKAR,ERSWAN methodology.

4 Conclusion

In this paper, the original image is decomposed in CMY colors, the shares of the secret image are created by using an algorithm of decomposition the basics matrices, after this shares is encrypted and decrypted using the RSA method. This type of work will increase the security of document sharing, it is a method that will be added to the work of shankar and

eswran. in our methodology, we have to increase the number of basic matrices and the number of shares to augment the degree of security and to make the decryption for an attacker difficult.

References

[1] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology — EUROCRYPT'94.*, 1994.

[2] *W.Q. Yan, D. Jin and M. S. Kankanhalli, "Visual cryptography for print and scan applications," in IEEE International Symposium on Circuits and Systems, Vancouver, 2004.*

[3] Y.C. Hou, C.Y. Chang, and SF Tu, "Visual cryptography for color images based on halftone technology," in *Proc. of International Conference on Information Systems,Analysis and Synthesis,World Multiconference onSystemics, Cybernetics and Informatics*, 2001.

[4] A. Adhikari, "Linear algebraic techniques to construct monochrome visual cryptographic schemes for general access structure and its applications to color images,"Des. Codes Cryptogr., vol. 73, pp. 865–895, December 2014.

[5] K. Shankar† and P. Eswaran "Sharing a Secret Image with Encapsulated Shares in Visual Cryptography" *4th International Conference on Eco-friendly Compiting and comminication Sysrem,ICECCS, 2015.*

[6] K. Shankar† and P. Eswaran "RGB-Based Secure Share Creation in Visual Cryptography Using Optimal Elliptic Curve Cryptography Technique" *Journal of Circuits, Systems, and Computers Vol. 25, No. 11 (2016) 1650138 (23 pages),2016.*

[7] K. Shankar† and P. Eswaran "RGB-Based multiple Share Creation in Visual Cryptography with Aid of Elliptic Curve Cryptography" *Network and Security, China Communication , February 2017.*

[8] Sami A. Nagar ; Saad Alshamma "High speed implementation of RSA algorithm with modified keys exchange *2012, 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012.*

[9] *R. L. Rivest, A. Shamir and L. Adleman "A method for obtaining digital signatures and public- key cryptosystems" Communications of the ACM, vol. 21, pp. 120-126, 1978.*

A Wavelet-Based ECG Delineation and Automated Diagnosis of Myocardial Infarction in PTB Database

[1]Rachid Haddadi, [2]Elhassane Abdelmounim, [3]Mustapha El Hanine, [4]Abdelaziz Belaguid
{ [1]haddadirachid@hotmail.com, [2]hassan.abdelmounim@hotmail.fr, [3]melhanine@yahoo.fr }

[1,2,3]Laboratory of Systems Analysis and Information Processing, Faculty of Sciences and Technics, Hassan First University, Settat, Morocco

[4]Laboratory of physiology, Mohammed-V University Rabat, Morocco

Abstract. In this work, we present an ECG delineation and the automated diagnosis of coronary artery disease in the electrocardiogram (ECG). In preprocessing stage, the baseline wander (BLW) and 60 Hz power line interference (PLI) were removed using discrete wavelet transform (DWT). The QRS detection is carried out using Daubechies (Db4) DWT. Feature extraction and classification is done using a convolutional neural network (CNN) containing three convolutional layers, three max-pooling layers, and three fully connected layers. The standard 12 lead ECG signals of 50 healthy subjects and 50 myocardial infarction subjects (MI) of one minute are obtained from the Physikalisch-Technische Bundesanstalt (PTB) database. We achieved an accuracy of 94.83%. sensitivity of 94.75%, and specificity of 94.93% on PTB database.

Keywords: Electrocardiogram, DWT, QRS complex, Convolutional neural network.

1 Introduction

Cardiovascular disease is the leading cause of death in the world. According to the World Health Organization (WHO), more than 17.7 million people die each year due to cardiovascular disease, representing 31% of total global mortality [1]. This number is in permanent growth. The main risk factors are obesity, tobacco, high blood pressure, diabetes and genetic factors. Ischemic heart disease happens when coronary arteries get narrower and reduce the blood flow to the heart. This is also called coronary heart disease (CHD). This can ultimately lead to the heart attack [2]. The electrocardiogram (ECG) is the most used cardiological examination at hospitals and cardiology offices to diagnose heart disease. It is easy to set up, painless and inexpensive.

As illustrated in Figure 1, an ECG is a series of waves and deflections recording the heart's electrical activity from a certain view. Each view called a lead, monitor voltage changes between electrodes placed in different positions on the body. Leads I, II, and III are bipolar leads. Leads aVR, aVL, aVF, V1 through V6 are unipolar leads. An ECG tracing looks different in each lead because the recorded angle of electrical activity changes with each lead [3]. The ECG signal is characterized by a recurrent wave sequence of P, QRS, and T-wave associated with each beat. Each wave of the ECG represents a phase of heart functioning. Each deformation detected on the duration and the forms of these waves can be considered as an indicator of cardiac abnormalities.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284216

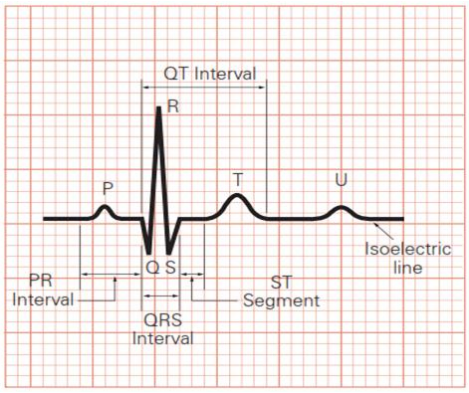


Fig.1. ECG segments and time intervals

Ischemia and Myocardial Infarction (MI) are among the most serious cardiac abnormalities. These anomalies appear on the ECG signal as ST segment or T wave changes. The ST segment elevation and depression indicate the myocardial ischemia. ST segment elevation is measured from the isoelectric line level (the flat part of the ECG between the P wave and the QRS complex) to the J point. The J point is the "junction" point between the end of QRS complex and the onset of ST segment.

There have been several methods dealing with automated characterization of Coronary Artery Disease (CAD). Sun et al. [4] proposed myocardial infarction detection, in 12 ECG leads, using multiple instance learning technique. Authors uses Support vector machine (SVM), NN and k-nearest neighbor (KNN) classifiers. They achieved MI detection sensitivity 92.3% and specificity 88.1%. Harldsson et al [5] used Hermite expansions with NN technique to classify MI in 12 leads ECG. They achieved ROC area 83.4%. Lahiri et al. [6] proposed MI detection in 12 leads ECG using Phase space fractal dimension of ECG and Artificial neural network. They obtained sensitivity 96,55% and specificity 95,24%.

Sharma et al. [7] proposed MI classification in 12-lead ECG using wavelet transform and SVM. Over 549 records of PTB databases, they achieved MI detection sensitivity 93%, specificity 99% and accuracy 96%. In Ref. [8], Arif et al. proposed KNN classifier to detect MI in 12-lead ECG. They used 3200 Normal beats and 16,960 MI and obtained sensitivity 99.97% and specificity 99.9%. Zheng et al. [9] used Naïve Bayes, SVM and Random Forest classifiers to detect the presence of old myocardial infarction in body surface potential maps. They achieved MI detection accuracy 81.9% with Naïve Bayes, 82.8% with SVM and 84.5% with Random Forest classifier. In Ref. [10] Acharya et al. proposed KNN classifier for automated detection and localization of myocardial infarction in 12-lead ECG. They used 125,652 normal beats and 485,753 MI from PTBDB and obtained sensitivity 99.45%, specificity 96.27% and accuracy 98.80%. Acharya et al. [11] used KNN and decision tree (DT) classifiers to detect (CAD) in 12-leads ECG. With Bispectrum-KNN, they achieved accuracy 98.2% sensitivity 94.8%, and specificity 99.3%. Pyakillya et al. [12] proposed Deep Learning algorithm for ECG Arrhythmia Classification using convolutional neural network (CNN). They used 1-dimentional (1D) CNN architecture comprising 7 convolutional layers with filter width 5 and 128 neurons + max-pooling and dropout after every layer + Global Average Pooling + 3 FCN layers with (256/128/64) neurons + dropout after every layer + softmax layer with 4 outputs (4 classes: normal sinus rhythm, AF, other rhythm and noisy). The best resulted accuracy on validation data is about 86%. Andrew Ng's scientific group (Pranav Rajpurkar, Awni Y. Hannun,

Masoumeh Haghpanahi, Codie Bourn) [13] proposed 34-layer CNN architecture for Arrhythmia Detection. They exceed the average cardiologist performance in both recall (sensitivity) and precision (positive predictive value). Liu et al. [14] proposed Real-time Multilead CNN (ML-CNN) for MI detection. To make the ML-CNN suitable for multilead ECG processing, they use 1-D kernels shared among the different leads to generate local optimal features. they achieved sensitivity 95.40%, specificity 97.37% and accuracy 96.00%. T. Reasat et al. [15] used shallow convolutional neural networks to detect Inferior Myocardial Infarction (IMI) in 3 leads ECG signal; lead II, III and aVF. The performance of the model is evaluated on IMI and healthy signals obtained from PTB database. They achieved an accuracy, sensitivity and specificity of 84.54%, 85.33% and 84.09%, respectively.

In this work, we propose an deep convolutional neural networks (CNN) in the classification of two ECG classes; normal and IMI.

2 Methodology

2.1 Data collection

ECG signals used in this study are taken from the Physikalisch-Technische Bundesanstalt (PTB) database [16]. The database contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects).

Each subject is represented by one to five records. There are no subjects numbered 124, 132, 134, or 161. Each record includes 15 leads: the conventional 12 leads (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6) together with the 3 Frank lead ECGs (Vx, Vy, Vz). Each signal is digitized at 1000 samples per second, with 16-bit resolution [17].

In this work, we used ECG signal from lead II, III and aVF and focused on inferior myocardial infarction (IMI) and healthy signals.

2.2 Preprocessing

In the preprocessing stage, the Baseline wander (BW) and 60 Hz power line interference (PLI) must be removed. The removal of this disturbance is an important step in ECG signal analysis, to produce a stable signal for subsequent automatic processing [18].

The method that we propose to eliminate the baseline is based on DWT decomposition up to level 8, which generates a set of approximation coefficients(C8), and eight sets of detail coefficients(d1,…,d8). By cancellation of approximations, the filtered signal is recovered from the details only [19]. Some results in the removal of baseline wanders is shown in Figure 2. For removal of 60Hz PLI interference, we propose to add a sine wave of 60 Hz frequency to the ECG test signals, to exemplify the original ECG signal as contaminated with PLI. The signal is recovered from all coefficients except details D(7,58), D(7,59) and D(7,60). This cancels all frequencies in the band 60 ± 1.5 Hz [20]. The results in removal of 60Hz PLI interference are shown in Figure. 3. [20].

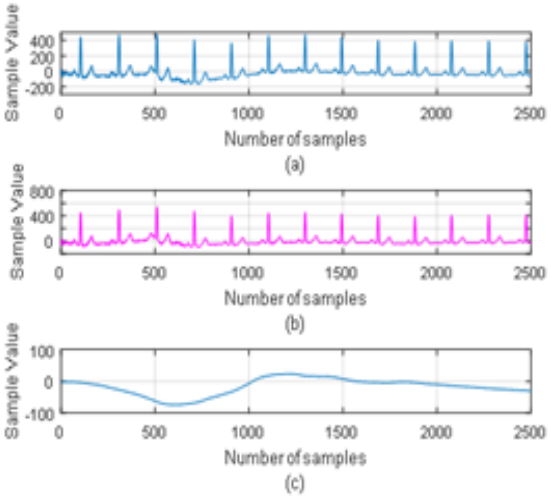


Fig. 2. Results for removal of baseline wanders: (a) noised ECG: (b) Baseline free ECG; (c) Removed baseline

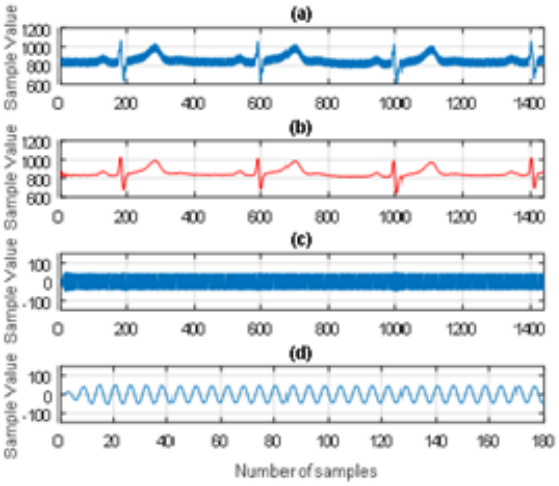


Fig. 3. Results in removal of PLI interference: (a) noised ECG; (b) Filtered ECG; (c) 60 Hz power line noise; (d) 60 Hz noise zoomed in 0.5s.

3 Detection of QRS complexes

The QRS complex reflects the electrical activity within the heart during the ventricular contraction, the time of its occurrence, as well as its shape, provide much information about the current state of the heart. The QRS detection is the most important step for almost all ECG analysis algorithms. Pan and Willis J.Tompkins were the first to develop a real-time QRS detection algorithm that uses an automatically adjustable threshold [21]. In this work, the decomposition of the ECG signal is made using Daubechies (Db4) DWT up to level 8. To choose the best coefficient, we compare the correlation coefficient between all the detail coefficients individually with the original ECG signal. The coefficient d4 [8.125Hz-16.25Hz] has the highest cross-correlation compared to other coefficients. Therefore, in the time domain, d4 is highly correlated with the original signal. The ECG signal is reconstructed with the coefficient d4. To detect the R Peaks, we use a hard thresholding method [19]. Figure 4 shows the original ECG signal with R peaks localization. The algorithm is also evaluated, in 12 leads PTB database, using Two performance measures : sensitivity (Se) and Positive predictivity (+P) defined respectively as:

$$= \frac{}{+} \tag{1}$$

$$+ = \frac{}{+} \tag{2}$$

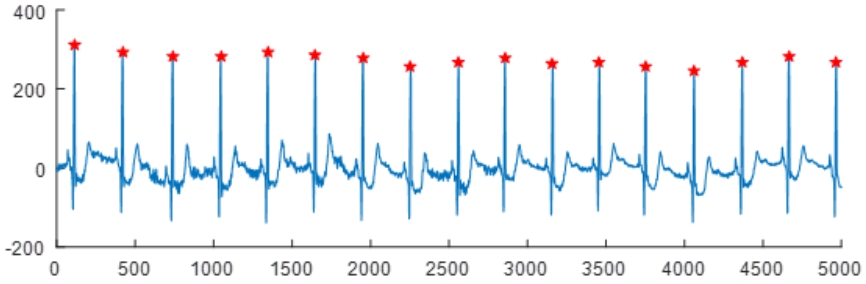


Fig. 4. Original ECG signal with R peak detection.

Where TP means a true positive, FN means a false negative and FP denotes a false positive detection.

Table 1 illustrates the one-minute results for three patients from PTB database:

Table 1. Results for 1 minute R peaks detection from PTBDB

Patient	Records	TP	FN	FP	Se (%)	+P (%)
PTB P.013	s0045lrem	980	5	4	99.49	99.59
	s0051lrem	957	5	5	99.48	99.48
	s0072lrem	1032	5	5	99.52	99.52
PTB P.019	s0058lrem	1032	4	7	99.61	99.33
	s0070lrem	986	4	5	99.60	99.50
	s0077lrem	847	5	6	99.41	99.30
PTB P.021	s0065lrem	852	5	5	99.42	99.42
	s0073lrem	1041	5	3	99.52	99.71
	s0097lrem	1064	4	5	99.63	99.53
PTB P.117	s0291lrem	780	5	6	99.36	99.24
	s0292lrem	790	4	5	99.50	99.37
Total		**10361**	**51**	**56**	**99.51**	**99.46**

4 CNN based Classifier

Convolutional neural networks (CNN) have emerged from the study of the visual cortex of the brain, and they have been used in image recognition since the 1980s "the neocognitron" [22], which has progressively evolved into what we today call the convolutional neural networks. An important step was taken in 1998 by Yann LeCun, Bottou, Yoshua Bengio and Patrick Haffner, who introduced the famous LeNet-5 architecture, widely used to recognize handwritten check numbers.

4.1 Pre-processing

Each ECG beat consists of a maximum of 670 samples (fs = 1KHz); considering the maximum values of the PR and QT intervals (PRmax + QTmax = 200ms + 470ms = 670ms). This value will be the size of the input vector (m = 670). To make sure that the P, QRS and T waves are included, we use the positions of the R peaks.

To solve the problem of amplitude scaling and eliminate the shift effect before introducing the ECG segments into the one-dimensional CNN, each segment of the ECG signals is normalized with the Z-score normalization method.

4.2 The architecture

In this work, we use a CNN architecture comprising 3 convolutional layers, 3 subsampling layers (Max pooling) and 3 fully connected layers. Each convolutional layer is followed by LeakyRelu [23] activation function to recover the non linearity into the data [24]. The output layer of the last fully connected network is a Softmax layer (layer 9) with a two-dimensional "vector" output since this is a two-class problem (normal and IMI). Figure 5 illustrates the overall structure of our proposed CNN. Table 2 summarizes the details of this CNN structure.

The number of neurons at the output of each channel was evaluated using this formula:

$$= \lfloor \frac{+2\ -}{\ } \rfloor + 1 \tag{3}$$

where n the length of input vector, m the Kernel size, p the padding (p=0) and s the Stride.

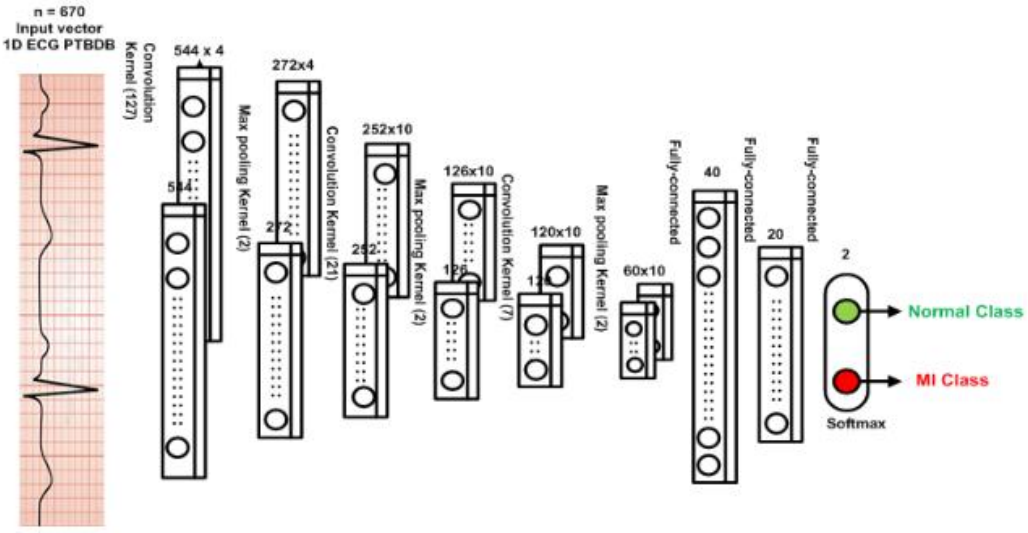


Fig. 5. The proposed CNN architecture

Table 2. Details of the CNN structure

Input vector size (670)			
Layers	Kernel size	stride	Number of neurons for each channel
(1) Convolution	127	1	544
(2) Max pooling	2	2	272
(3) Convolution	21	1	252
(4) Max pooling	2	2	126
(5) Convolution	7	1	120
(6) Max pooling	2	2	60
(7) Fully connected			40
(8) Fully connected			20
(9) Fully connected			2

4.3 Experiments and Results

In this study, we use a backpropagation, based on stochastic gradient descent (SGD), to adjust the momentum, regularization and learning rate. The best performances were obtained by adjusting the Momentum at 5x10-3, the learning rate 0.6, and the Regularization at 0,1. We performed several epochs of training and testing. We validates the CNN model for each epoch. The classification error is stabilized and has converged after 45 epochs.

To validate our model, we used 10-fold cross validation. The performances (Accuracy, sensitivity, and specificity) are evaluated at each iteration. The overall performance is the average performance recorded in the 10 iterations. To make a better comparison between various performance indices, the confusion matrix presented in Table 3 is used. Overall performance with optimal settings is shown in Figure 6.

Table 3. Confusion matrix of the CNN classifier

	Predicted	
	Normal	IMI
Normal	1780	95
IMI	1895	105

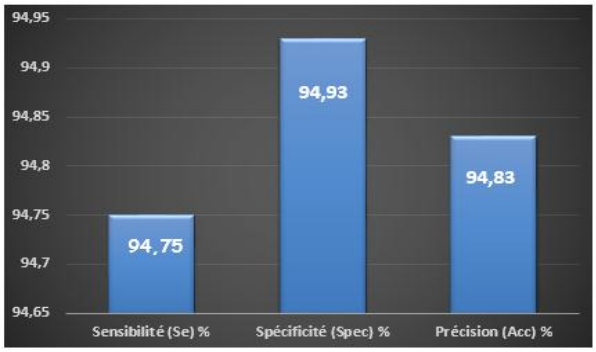


Fig. 6. Results of the CNN classifier

4.4 Discussion

Table 4 shows some techniques used in the literature for the automated classification of myocardial infarction validated on the same database (PTBDB).

Table 4. Comparison of our MI classifiers with other methods in PTB database

Author, Year	Ref	Classifier	Performance		
			% Se	% Spec	% Acc
Sun et al., 2012	[25]	SVM	92,60	82,4	-
Safdarian et al., 2014	[26]	Naïve Bayes	-	-	94,74
Lu et al. 2000	[27]	Fuzzy ANN	84,60	90	-
Liu et al., 2017	[28]	CNN	95,40	97,37	96
Our work		CNN	94,75	94,93	94,83

The performance of our system is not the best, but it exceeds several works cited in the table above. Our precision is close to that of Liu et al. [28] obtained with a CNN network. This last work uses a larger data size than ours, with 122 training records and 125 test records, for a total of 34769 12-lead beats. In our study, we used only 100 ECG recordings of one minute, on three leads DII, DIII and aVF; commonly used derivations for the diagnosis of inferior infarction. This choice is justified by a hardware limitation since we use the CPU version of Keras-tensorflow. To reduce the calculation time and increase the data size, we have to use the GPU version.

5 Conclusion and future work

In this paper, we present an algorithm of automatic detection of inferior myocardial infarction based on convolutional neural network (CNN), using leads II, III and aVF from PTB database. In the preprocessing stage, discrete wavelet transform (DWT) was used to remove the baseline wander and power line interference. DWT was also used for QRS detection. The filtered signals are normalized and sent directly to the CNN network. We achieved an average sensitivity, specificity and accuracy of 94,75%, 94,93% and 94,83% respectively.

In this study, we used a binary classification that does not provide information about the location of myocardial infarction. Indeed, there are 10 types of infarction in the PTB database, classified according to the location in the myocardium. In the future, our CNN classifier could be extended to an 11-class classifier to be able to recognize the different types of infarction.

References

[1] World Health Organization (WHO), "Cardiovascular Diseases," (2017)

[2] http://www.heart.org/HEARTORG/Conditions/HeartAttack/PreventionTreatmentofHeartAttack/Silent-Ischemia-and-Ischemic-Heart-Disease_UCM_434092_Article.jsp#.WYmRzITyjIU.

[3] Shirley A. Jones, ECG Success Exercises in ECG Interpretation, MS Ed, MHA, EMT-P (2008)

[4] L. Sun, Y. Lu, K. yang, S. Li, "ECG analysis using multiple instance learning for myocardial infarction detection,"IEEE Trans. Biomed. Eng. 9 (12), pp. 3348–3356 (2012)

[5] H. Haraldsson, L. Edenbrandt, M. Ohlsson, “Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks, ” Artif. Intell. Med, vol. 32, pp. 127–136 (2004)
[6] T. Lahiri, U. Kumar, H. Mishra, S. Sarkar, A.D. Roy, “Analysis of ECG signal by chaos principle to help automatic diagnosis of myocardial infarction, ” J. Scient. Industr. Res, vol.68, pp. 866–870 (2009)
[7] L.N. Sharma, R.K. Tripathy, S. Dandapat, “Multiscale energy and eigenspace approach to detection and localization of myocardial infarction, ” IEEE Trans. Biomed. Eng., vol. 62 (7), pp. 1827–1837 (2015)
[8] M. Arif, I.A. Malagore, F.A. Afsar, “Detection and localization of myocardial infarction using k-nearest neighbor classifier, ” J. Med. Syst., vol. 36 (1), pp. 279–289 (2012)
[9] H. Zheng, H. Wang, C.D. Nugent, D.D. Finlay, “Supervised classification models to detect the presence of old myocardial infarction in body surface potential maps, ” In: Proc. of Computers in Cardiology, pp. 265–268 (2006)
[10] R. Acharya et al.” Automated detection and localization of myocardial infarction using electrocardiogram: a comparative study of different leads, ” Knowledge-Based Syst., vol. 99, pp. 146–156 (2016)
[11] U.R. Acharya et al. “Application of higher-order spectra for the characterization of coronary artery disease using electrocardiogram signals, ” Biomed. Signal Process. Control, vol. 31, pp. 31–34 (2017)
[12] B. Pyakillya, N. Kazachenko, and N. Mikhailovsky, "Deep Learning for ECG Classification," IOP Conf. Series: Journal of Physics: Conference Series, vol. 913, no. 1, pp. 1-5 (2017)
[13] Rajpurkar P, Hannun Awni Y, Haghpanahi M, Bourn C and Ng Y. Ng, “A 2017 Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks,” (Preprint arXiv:1707.01836v1).
[14] Wenhan Liu et al., "Real-time Multilead Convolutional Neural Network for Myocardial Infarction Detection," IEEE Journal of Biomedical and Health Informatics, vol. 99, pp. 1-12 (2017)
[15] T. Reasat, C. Shahnaz, “Detection of Inferior Myocardial Infarction using Shallow Convolutional Neural Networks”, In: Proc. of 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), 21 - 23 Dec 2017, Dhaka, Bangladesh, pp. 718-721 (2017)
[16] Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals”, Circulation, 101(23): e215-e220
[17] https://physionet.org/physiobank/database/ptbdb
[18] Bailey et al. “Recommendations for the Standardization and Specifications in Automated Electrocardiography: Bandwidth and Signal Processing”, Circulation, vol.81, 730-739 (1990)
[19] R.Haddadi, E.Abdelmounim, M.El Hanine, and A.Belaguid, "Discrete Wavelet Transform Based Algorithm for Recognition of QRS Complexes," World of Computer Science and Information Technology Journal (WCSIT), Vol. 4, No.9, pp. 127-132, (2014)
[20] R. Haddadi, E. Abdelmounim, M. El Hanine, and A. Belaguid, "ST Segment Analysis Using Wavelet Transform," IJCSNS International Journal of Computer Science and Network Security , vol. 17, no. 9, pp. 102-107, September (2017)
[21] J. Pan, J. Tompkins, “A Real-Time QRS Detection Algorithm, ” IEEE Trans. Biomed. Eng. BME, vol. 32 (3), pp. 230–236 (1985)
[22] K. Fukushima, "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected," Biol. Cybernetics, vol. 36, no. 4, pp. 193-202, (1986).
[23] He.K. Zhang, X. Ren, S. Sun, Delving deep into rectifiers: surpassing human-level performance on image net classification, 1026–1034, (2015)
[24] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 436–444(2015)
[25] Li Sun, Yanping Lu, Kaitao Yang, and Shaozi Li, "ECG Analysis Using Multiple Instance Learning for Myocardial Infarction Detection", IEEE Transactions on Biomedical Engineering, vol. 59, no. 12, pp. 3348 - 3356, (2012)

[26] N. Safdarian, N. J. Dabanlo, and G. Attarodi, "A New Pattern Recognition Method for Detection and Localization of Myocardial Infarction Using T-Wave Integral and Total Integral as Extracted Features from One Cycle of ECG Signal", Journal of Biomedical Science and Engineering, vol. 7, no. 10, pp. 818-824 (2014)

[27] H.L. Lu, K. Ong, and P. Chia, "An automated ECG classification system based on a neuro-fuzzy system", Computers in Cardiology 2000, vol. 27, pp. 387–390 (2000)

[28] Wenhan Liu et al., "Real-time Multilead Convolutional Neural Network for Myocardial Infarction Detection", IEEE Journal of Biomedical and Health Informatics, vol. 99, pp. 1-12 (2017)

A Layout Investigation of Large Wind Farm in Akhfennir using Real Coded Genetic Algorithm

Mohammed Amine Hassoine[1], Fouad Lahlou[2], Adnane Addaim[3], Abdessalam Ait Madi[4]
romema@yahoo.fr[1], lahloufouad@hotmail.com[2], adnane.addaim@uit.ac.ma[3], abdessalam.aitmadi@uit.ac.ma[4]
Faculty of Sciences[1,2], National School of Applied Sciences[3,4], Ibn Tofail University, 14000 Kénitra, Morocco

Abstract. The objective of this study is to evaluate the effect of wind turbine spacing in large wind farm on the total energy loss of the wind farm, the power loss is due to the wake effect between wind turbines, on a site gathering several wind turbines, if the wind turbines are too close, the loss of power increases with the wake effect. This paper presents an investigation into optimal wind farm layout in 88 wind farm configurations of a hypothetical WF in Tarfaya, to search the optimal number of Wind Turbines (WTs), the wind farm are Installed on an area of 35 km2 (7000m×5000m), with the aim to maximize the electrical power generated by all WTs and grows the annual economic profitability of the WF, in order to approve the result of this investigation a new approach based on the division of the wind farm in the sub domain method is proposed, to search the optimal location of Wind Turbines by mean an RCGA (Real Coded Genetic Algorithm). This new proposed approach is promising in terms of the applicability in large wind farm. It is also more suitable when performing the wind farm layout assessment in WPP (Wind Power Project).

Keywords*:* Wind Farm, Power, layout, Optimization, Cost, Energy, RCGA, Genetic Algorithm.

1 Introduction

The development of wind farms (WFs) requires optimal placement of the wind turbines (WTs) in relation to each other in order to maximize the production of the wind farm. If the turbines are too close, the wake effects prevent to provide the full power If they are too far away, the wind resource is not exploited optimally. The wake power losses currently measured in the wind farm can reach 20% [1]. Thus, optimizing the placement of wind turbines becomes a real challenge. The selection of an appropriate site is essential to the success of a wind farm (WF) project it plays a crucial role in financial returns, a good arrangement of wind turbines in a wind farm can improve the performance of the wind farm (WF). Many studies have been carried out by different researchers to evaluate the effects of wind turbine spacing (WTS) on the performance of the wind farm [2], [3], [4],[5], along and across the prevailing wind direction by varying wind turbine spacing in prevailing wind direction (PWD) and crosswind direction (CWD), the wind turbine spacing was expressed as a multiple of the rotor diameter of wind turbine. Mosetti et al. [2] modeled an wind farm of 2 km × 2 km by 5D in CWD and 5D in PWD, Archer et al. [3] studied The most efficient arrangement of the Lillgrund offshore wind farm, close to Sweden, which includes 48 turbines with six different configurations, 3.2D to 6.4D in CWD and 4.3D to 8.6D in PWD, Gao et al. [4] investigated the potential for offshore wind farm by 5D to 15D along the PWD and 5D to 12D in CWD, García et al. [5] conducted an experimental wind tunnel (1D to 18D) and concluded in the distance between 6D and 15D, the wake shows a similar profile of velocity behind the WT. In this work, we investigated regular layout and irregular layout of wind turbines within a large wind farm. In order to perform this investigation, we consider a hypothetical large wind farm (7000m×5000m) under 88 configurations witch the number of wind turbines is varied between 30 and 176 and arranged in regular layout, wind turbines were separated by 5D to 15D along the PWD and 5D to 12D in CWD. The best regular layout is achieved by comparing the ratio of cost per total power generation for each configuration. With the aim to find the best irregular layout of large wind farm in-terms of minimum unit cost of energy produce, we proposed a new method based on the division of the large wind farm (LWF)in the sub domain, to search the optimal location in irregular layout of Wind Turbines by mean an RCGA (Real Coded Genetic Algorithm).

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284188

The organization of this paper is presented as follows. Section 2 presents a brief presentation of the wind farms model. Section 3 gives the optimization method. The results and their interpretations are discussed in section 4. The conclusion of this proposed work is given in section 5.

2 MODELS OF WIND FARM

2.1 Wind farm Site location and Data Source description

The **Figure 1**-(a) gives a map of the Tarfaya territory. Tarfaya is a Moroccan coastal town. It is located at the edge of the Atlantic Ocean, about 890 km southwest of the capital Rabat, 100 km north-north-east of Laayoune. Its latitude and longitude Coordinates are respectively 27° 54′ 48″ N and 12° 55′ 54″ W . For most of the year, the wind blows so hard in Tarfaya city. Consequently is known for its innovative economic projects, such as that of the largest park in Africa, called the Tarfaya wind farm (Akhfennir WF). As seen in **Figure 1**-(b), the location of a hypothetical WF is selected near of the Tarfaya WF. The lifetime of an wind farm is assumed to be 20 years [6], the wind data of twenty years (1998-2018), which are used here, are extracted from the MERRA-2 [7], [8].

a b

Fig. 1. Tarfaya territory map and location of the hypothetical wind farm: (a) Tarfaya territory map; (b) Akhfennir wind farm layout (117 WTs) and hypothetical wind farm.

Table 1 shows latitude and longitude coordinates of MERRA-2 point and those of Tarfaya WF and hypothetical WF.

Table 1. MERRA-2 grid point, Akhfennir wind farm and hypothetical wind farm coordinates.

Reference	Latitude N	Longitude W
Akhfennir WF	27 °57′3.956″N	11°59′52.058″W
Tarfaya hypothetical WF	28° N	11° 52' 29.999' W
MERRA-2	28°N	11° 52' 29.999' W

2.2 Wind Characteristics and Analytical Model

Weibull distribution and Cumulative distribution function are been used to express the wind speed frequency distribution. We used here graphic method [9] in which Weibull parameters have been estimated by using linear LSM (Least-Squares regression Method) [19] with regress variable taken as wind speed to compute the best fitted line. The hourly wind speed is extrapolated to the hub height of the wind turbines by using the power law. Here we use the hourly wind data (20 years) from the MERRA-2 point. **Figure 2**-(a) shows the hourly mean wind speed at MERRA-2 site A (Tarfaya hypothetical WF) over the last twenty years at 80 meters, with an average of 7.37 m/s indicated by the green line. **Figure 2**-(b) shows Wind speeds distributions at 80 m above ground and the Weibull distributions of Tarfaya

hypothetical WF. the Plot of twice logarithm of Cumulative distribution function should yield a straight line. The gradient of the line is k (shape parameter) and the intercept with the y-axis is - k ln (c) [10], F**igure 3-**(a) shows the numerically linearized data of twice logarithm of Cumulative distribution function and graphically data. The values of shape parameter *k* and scale parameter *c* (obtained graphically) are nearly similar to the numerical. In order to give a view of how wind speed and direction are distributed. The wind rose are used to find the prevailing wind direction in location of Tarfaya hypothetical WF site. Predominant wind direction in this region is calculated over a 20-year period (1998-2018), which shows that the wind direction is predominately generated from the ENE (East North-East) and E (East), See **Figure 3**-(b).

a b

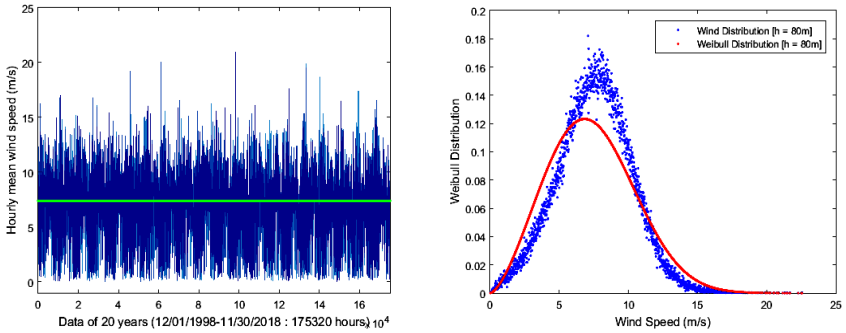


Fig. 2. Wind Characteristics: (a) Wind data in the Site A, from 1998 to 2018; (b) Wind speeds distributions at 80 m and the Weibull distributions of Tarfaya hypothetical WF

a b

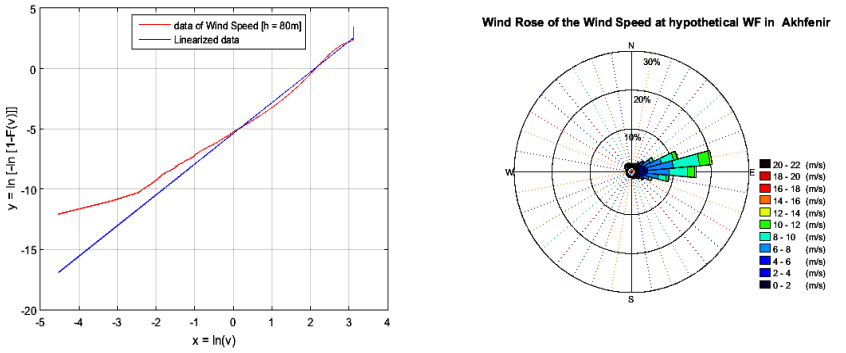


Fig. 3. Wind Characteristics: (a) linearized data; (b) Wind rose of wind speed at hypothetical WF.

The weibull parameters and mean speed are represented in Table 2.

Table 2. Weibull parameters and mean speed.

Site	Hub Height (m)	Weibull K	Weibull C	Mean speed (m/s)
MERRA-2	50	2.5452	7.7114	6.8452
Tarfaya HWF	80	2.5452	8.3137	7.3798

2.3 Wake effect model

When wind flows through the rotor of a WT behind another WT, the wake expands with down-stream distance. The model used here is a Jensen wake model [11]. This model assumed that the wake expands linearly with down-stream distance. The velocity deficit is a loss due to the interaction between two WTs placed one behind the other, In multiple wakes of N turbines, assuming that the kinetic energy deficit of a mixed wake is equal to the sum of the energy deficits [12].

The power generated by the wind farm P_{wf}, expressed by equation 1, is the total power for all the wind turbines in the wind farm.

$$P_{wf} = \sum_{i=1}^{N} P_i \tag{1}$$

where, N is the number of WT and P_i is the power of WT i .

The efficiency of the WF is the ratio of the total power produced by the WF to the sum of the power for each individual WT. The efficiency of the farm can be expressed as shown in [13], by equation 2.

$$\eta_{wf} = \frac{P_{wf}}{\sum_{i=1}^{N} P_{si}} \tag{2}$$

where P_{si} is the power of WT i if it is functioning as a single turbine.

The AEP (Annual Energy Production) for all wind turbines within wind farm can be calculated using the following expression [14]:

$$AEP = T \sum_{i=1}^{directions} \sum_{j=1}^{speed} \sum_{k=1}^{turbines} Frequency_{ijk}\, Power_{ijk} \tag{3}$$

where, $Frequency_{ijk}$ is the frequency or probability of wind coming from direction i, with wind speed j on to the wind turbine k, T is the number of hours in one year (365×24) and $Power_{ijk}$ is the power (in kilowatts) generated by that turbine for the same wind speed and direction .

2.4 Cost model and objective function

In this article the objective function will be used, this function serves as a criterion for determining the best arrangement of the WF. The goal is then to minimize this function up to the optimum, this function will be used. As shown by Grady [15]. This objective function is expressed by the following equation 4.

$$\text{Objective} = \frac{\text{Cos}\,t}{p_{wf}} \tag{4}$$

Where Cost is cost function, which is a non-dimensional function of WTs number (N) given by the equation below.

$$Cost = N\left(\frac{2}{3} + \frac{1}{3}e^{-0.00174N^2}\right) \quad (5)$$

3 OPTIMIZATION PROCESS

We consider 7km×5km wind farm that is divided into a grid of (nD×mD) under 88 configurations (n=5..12, m=5..15), see **Table 3**. with the number of wind turbines is between 30 and 176 see **Table 4**. The search for the optimal configuration of WTs within wind farm (regular layout) is carried out by an iterative approach, the CWD distance remains constant between the wind turbines and the distance PWD varies from 5D to 15D then the CWD distance is incremented up to 12D. When the optimal solution is obtained, the number of WTs and the distances separation between wind turbines along CWD and PWD are known. After The large wind farm (LWF) is divided into 4 sub-wind farms (SWF1, SWF2, SWF3, SWF4) area of 2km × 3km, a new search for the optimal configuration of WTs within LWF (irregular layout) is achieved by using RCGA, the distance between the 4 sub-wind farms is 12D along CWD and 15D along PWD. To execute our approach, we have developed a MATLAB software code using the RCGA (Real Coded Genetic Algorithm) with 4 subdomains (4 sub wind farms). The output computing results are wind speed, WF power generation, and the cost. The program will be stopped provided that the best fitness stills the same without any change in 400 iterations. The specifications of wind farm parameters are represented in **Table 5.**

Table 3. Configuration number (CWD-PWD).

PWD	CWD							
	5D	6D	7D	8D	9D	10D	11D	12D
5D	1	12	23	34	45	56	67	78
6D	2	13	24	35	46	57	68	79
7D	3	14	25	36	47	58	69	80
8D	4	15	26	37	48	59	70	81
9D	5	16	27	38	49	60	71	82
10D	6	17	28	39	50	61	72	83
11D	7	18	29	40	51	62	73	84
12D	8	19	30	41	52	63	74	85
13D	9	20	31	42	53	64	75	86
14D	10	21	32	43	54	65	76	87
15D	11	22	33	44	55	66	77	88

Table 4. Wind turbines number (CWD-PWD).

PWD	CWD							
	5D	6D	7D	8D	9D	10D	11D	12D
5D	176	144	128	112	96	96	80	80
6D	143	117	104	91	78	78	65	65
7D	121	99	88	77	66	66	55	55
8D	110	90	80	70	60	60	50	50
9D	99	81	72	63	54	54	45	45
10D	88	72	64	56	48	48	40	40
11D	77	63	56	49	42	42	35	35
12D	77	63	56	49	42	42	35	35
13D	66	54	48	42	36	36	30	30
14D	66	54	48	42	36	36	30	30
15D	66	54	48	42	36	36	30	30

Table 5. specifications of wind farm parameters.

Parameters	Specifications
Wind turbine model	SWT-2.3 MW -93 [16]
Rated power (kW)	2300
Hub height of WTs, Z (m)	80
Wind turbine rotor radius, Rr (m)	46.5
Turbine thrust coefficient, C_T	0.88
Roughness length of ground, Z_0 (m)	0.1
The entrainment constant, α	0.0748
The axial induction factor, a	0.3268

4 Results and discussion

In the present study, 88 configurations of wind farm layout is considered, by varying the spacing between wind turbines along and across the prevailing wind direction, 8 different WTs separations in CWD (5D to 12D) with 11 different WTs separations in PWD (5D to 15D) were allowed to evolve over 88 wind farm layout. After the execution of the program, the best solution for 42 WTs from configuration number 63 (42C63) placement in WF is achieved with a best fitness value of 0.000330009. The comparison results between the top five best layout of wind farm (42C63, 42C43, 42C44, 40C83 and 42C42) is presented in **Table 6**. For 42 WTs from configuration number 63 (42C63), the total power is 86816.79 kW, the fitness value is 0.000330009 and the efficiency is 89.87%. The best solutions of 42 WTs layout within wind farm are depicted in **Figures 12**-(a) and **12**-(b), the wind turbines numbers and Total power evolution by configuration and by spacing are depicted respectively in **Figure 4** and **Figure 5.**

Table 6. The comparison results between the top five best layout

Number of WTs	Number of WTs	Total power (kW)	Fitness value	Efficiency (%)	spacing (CWD)	spacing (PWD)
63	42	86816.79498	0.000330009	89.87245857	10D	12D
43	42	86613.90335	0.000330782	89.66242583	8D	14D
44	42	86573.65257	0.000330936	89.62075835	8D	15D
83	40	82634.95925	0.000332675	89.82060788	12D	10D
42	42	86118.73973	0.000332684	89.14983409	8D	13D

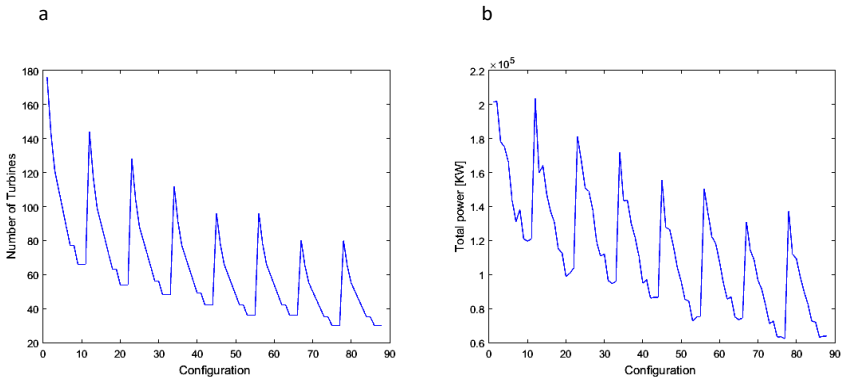


Fig. 4. WT numbers and Total power evolution by configuration: (a) WT numbers evolution; (b) Total power evolution

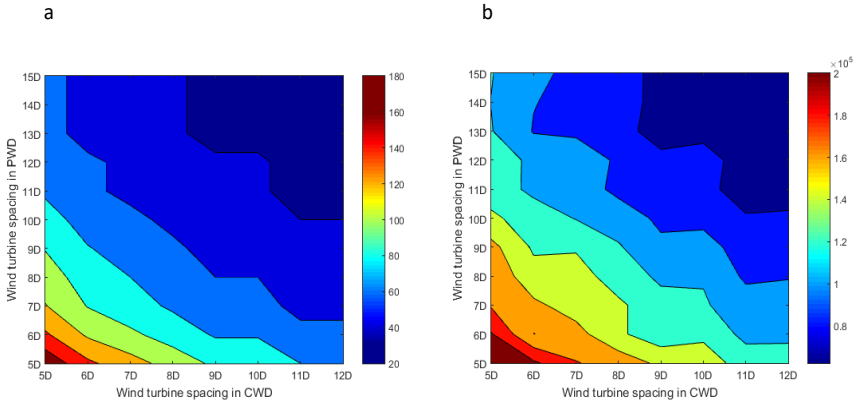


Fig. 5. WT numbers and Total power evolution by spacing: (a) WT numbers evolution; (b) Total power evolution

Fitness and Efficiency evolution by configuration are depicted in **Figure 6.** Fitness evolution by spacing are depicted in **Figure 7**-(a). Fitness versus Total power evolution by configuration are depicted in **Figure 7**-(b). these figures indicate that the best Fitness value corresponding at configuration number 63 with 10 D in CWD and 12 D in PWD.

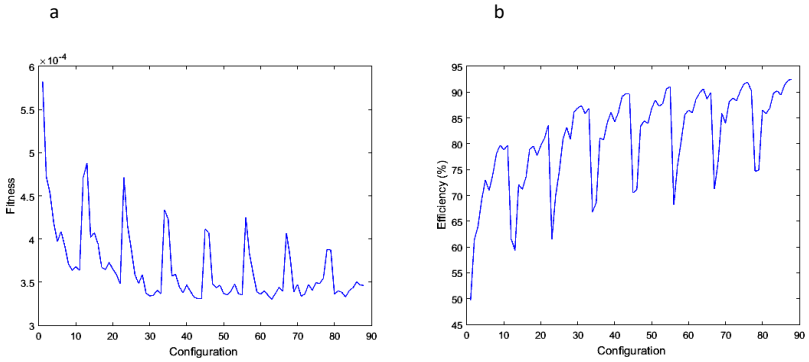


Fig. 6. Fitness and Efficiency evolution by configuration: (a) Fitness evolution; (b) Efficiency evolution

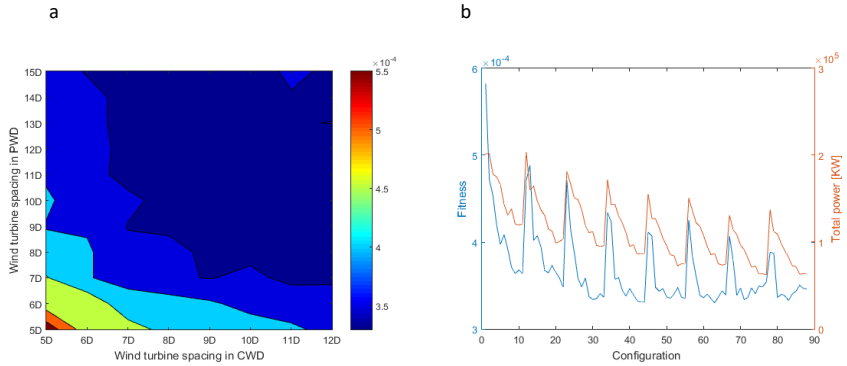


Fig. 7. Fitness versus Total power evolution: (a) Fitness evolution by spacing; (b) Fitness evolution versus Total power evolution by configuration

In the present study, 45 individuals (Probability of crossover value of 0.9, Probability of mutation value of 0.1) were allowed to evolve over 1000 iterations. After the execution of the

RCGA program for 600 iterations, the best solution for 11 WTs placement in SWF and 42 WTs placement in LWF is achieved with a best fitness value of 0,000325853 (42 WTs). For 42 WTs in LWF (irregular layout), the total power is 87924,02 kW, the fitness value is 0,000325853 and the efficiency is 91,018 %. The optimized result obtained for 42 WTs (irregular layout) shows that the optimal arrangement for 42 WTs achieved by using RCGA gives a lower fitness value than 42 WTs (regular layout), See **Table 6.** Moreover, the efficiency evolution, the total power evolution and the fitness evolution of 11 WTs (SWF) and 42 WTs (LWF) over the searching period are depicted respectively in **Figure 8**, **Figure 9** and **Figure 10**. **Figure 11**-(b), and **Figure 11**-(a), shows respectively the optimal layouts of SWF and LWF (irregular layout).

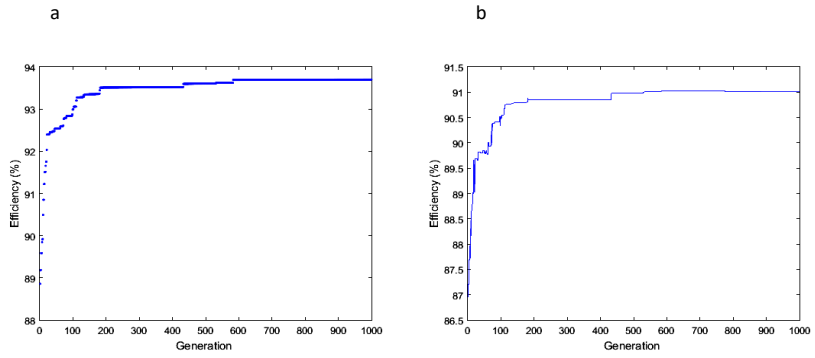


Fig. 8. Efficiency evolution (SWF vs LWF): (a) Efficiency evolution of 11 WTs; (b) Efficiency evolution of 42 WTs.

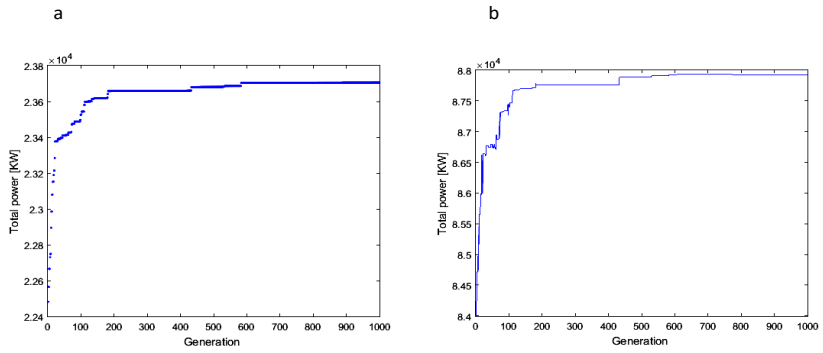


Fig. 9. Total power evolution (SWF vs LWF): (a) Total power evolution of 11 WTs. ; (b) Total power evolution of 42 WTs.

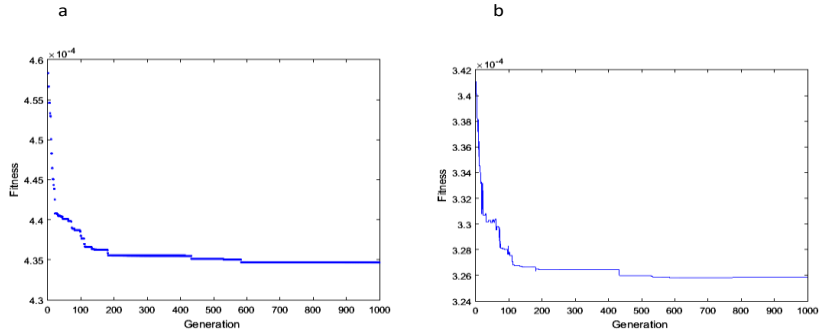


Fig. 10. Fitness evolution (SWF vs LWF): (a) Fitness evolution of 11 WTs; (b) Fitness evolution of 42 WTs

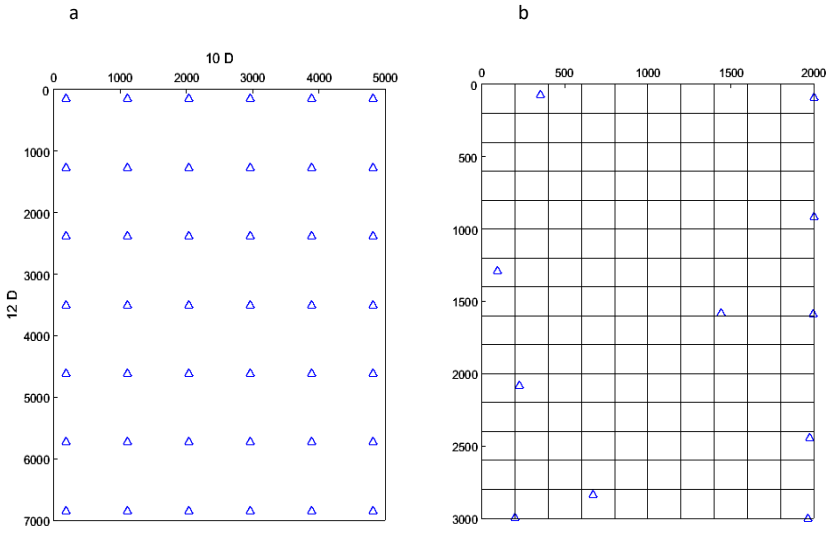


Fig. 11. Wind farm layouts: (a) optimal layout of large wind farm (42 WTs, regular layout) ; (b) optimal layout of Sub wind farm1 (11 WTs, irregular layout) .

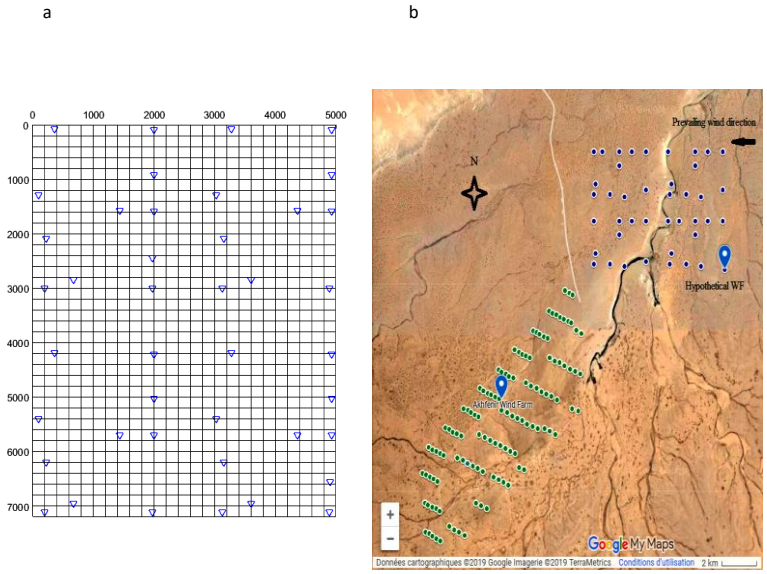


Fig. 12. Wind farm layouts: (a) optimal layout of large wind farm (42 WTs, irregular layout) ; (b) GoogleMyMaps: Optimal layout of hypothetical wind farm (placed perpendicular to prevailing wind direction).

5 Conclusions

In this article, we investigated regular layout and irregular layout of wind turbines within a large wind farm, we have proposed and applied a new approach based on the division of the large wind farm in the sub domain (Sub wind farm), and we have applied real coded genetic algorithm (RCGA) approach to achieve the best placement of WTs in order to get the most out of the power production. The carried out results from Matlab Software showed that the proposed approach is promising in terms of the applicability in large wind farm. It is also more suitable when performing the wind farm layout assessment in WPP (Wind Power Project). In the ongoing research, we will take into consideration the real wind condition of the WFs by using the more complex wake model .

References

[1] Barthelmie, R. et al.: Modelling and measuring flow and wind turbine wakes in large wind farms offshore. Wind Energy. 12, 5, 431-444 (2009).

[2] Mosetti, G. et al.: Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. Journal of Wind Engineering and Industrial Aerodynamics. 51, 1, 105-116 (1994).

[3] Archer, C. et al.: Quantifying the sensitivity of wind farm performance to array layout options using large-eddy simulation. Geophysical Research Letters. 40, 18, 4963-4970 (2013).

[4] Gao, X. et al.: Investigation into the optimal wind turbine layout patterns for a Hong Kong offshore wind farm. Energy. 73, 430-442 (2014).

[5] García, L. et al.: Experiments in the wind turbine far wake for the evaluation of an analytical wake model. Journal of Physics: Conference Series. 854, 012015 (2017).

[6] Tremeac, B., Meunier, F.: Life cycle analysis of 4.5MW and 250W wind turbines. Renewable and Sustainable Energy Reviews. 13, 8, 2104-2110 (2009).

[7] "Modern-Era Retrospective analysis for Research and Applications, Version 2" https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/. [Accessed: 12- Dec- 2018].

[8] "Data Collections" https://disc.sci.gsfc.nasa.gov/datasets?keywords=%22MERRA-. 2%22&page=1&source=Models%2FAnalyses%20MERRA-2. [Accessed: 12- Dec- 2018].

[9] Deaves, D., Lines, I.: On the fitting of low mean windspeed data to the Weibull distribution. Journal of Wind Engineering and Industrial Aerodynamics. 66, 3, 169-178 (1997).

[10] Lun, I., Lam, J.: A study of Weibull parameters using long-term wind observations. Renewable Energy. 20, 2, 145-153 (2000).

[11] N. O., Jensen.: A note on wind generator interaction. Roskilde, Denmark: Risø National Laboratory. November (1983).

[12] Katić, I., J, Højstrup, J., Jensen, N.: A simple model for cluster efficiency. European wind energy association conference and exhibition. October (1986).

[13] Chowdhury, S. et al.: Unrestricted wind farm layout optimization (UWFLO): Investigating key factors influencing the maximum power generation. Renewable Energy. 38, 1, 16-30 (2012).

[14] Mittal, P. et al.: A novel hybrid optimization methodology to optimize the total number and placement of wind turbines. Renewable Energy. 86, 133-147 (2016).

[15] Grady, S. et al.: Placement of wind turbines using genetic algorithms. Renewable Energy. 30, 2, 259-270 (2005).

[16] "SWT-2.3-93 (Siemens)" https://www.thewindpower.net/turbine_en_22_siemens_swt-2.3-93.php. [Accessed: 01- Dec- 2018].

New Model and Sliding mode control of a three phase grid connected photovoltaic system without DC-DC Converter

[1]Hicham Bahri, [2]Mohamed Aboulfatah, [3]M'hamed Guisser, [4]Elhasan Abdelmounim, [5]Mohammed El Malah, [6]Adekanle Simon

{ [1]hbahri.inf@gmail.com, [2]mohamed.aboulfatah@uhp.ac.ma, [3]mguisser@gmail.com, [4]hassan.abdelmounim@hotmail.fr, [5]medelmalah@gmail.com, [6]adekanlesimon@gmail.com }

Department of Physics, Laboratory ASTI, FST-Settat, University Hassan I, Settat, Morocco

Abstract. This paper presents a new model and an advanced control strategy of a three phase grid connected photovoltaic system without DC-DC converter. The novel model is characterized by decoupled inputs and high relative degree of the outputs. A sliding mode control without chattering phenomena based on this new model is applied in order to control the system to track the maximum power point produced by the photovoltaic generator PVG, convert all this power to active power and compensate the harmonic components and reactive power caused by nonlinear load if it is connected to the PV system. This controller is able to realize all these objectives with a high dynamic performance in presence of changing atmospheric conditions, system's disturbance and with any kind of load or without load. Simulation results demonstrate efficient performance and robustness of the proposed strategy.

Keywords: Inverter, Maximum Power Point Trackers, Unity power factor, Photovoltaic Generator, Power Conversion Harmonics, Pulse Width Modulation Converters ,Reactive power compensation, Nonlinear load.

1 Introduction

Recently, due to environmental pollution and depletion of fossil fuels, renewable energy sources have gained more attention of researchers and investors, and in particular, research on photovoltaic systems is constantly being conducted [1]. Especially, about the photovoltaic system connected to the utility grid, that produces electricity at a site close to customers or that is tied to an electric distribution system [2], this system can be realized by using a DC-DC converter and inverter (Tow-Stage topology), or just an inverter (Single-stage topology). The second topology is more economical owing to the elimination of DC-DC converter. This also decreases the number of the sensors needed by the controller and thereby reduces measurement errors. This topology has gained attention, especially in low voltage application [3]. This work is devoted to optimize the operation of a three phase grid single stage connected photovoltaic system by controlling the voltage source inverter VSI in order to realize the following objectives:

- Tracking maximum active power MPPT;
- Generate maximum available active power to feed a load and supply the grid;
- Compensate harmonic components and reactive power caused by nonlinear load.

Many control strategies have been designed to achieve these objectives [4]-[8], but they aren't able to control the system to track precisely the maximum power produced by the

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284092

photovoltaic generator PVG or they are not good compensators of current's harmonics components caused by nonlinear load.

This paper presents, for the first time, a new model of the studied system. This model decouples the control input of the inverter and increases the relative degree of the first output selected to be controlled in this work, which enhances the dynamic performance of the controller in the regulation of this output. After that, the paper proposes a new control strategy of the studied system based on the new model in dq0 axis applying a sliding mode technique [9], which is characterized by its efficient and robust dynamic performance. Furthermore, the reactive power needed or caused by nonlinear load is calculated by using a new simplify mathematic expression in dq0 axis. As results of this work, the elaborated controller achieved all the objectives with a fast response time and very good precision in presence of changes in atmospheric conditions (irradiation and temperature) and system perturbation (parameters disturbance and harmonics pollution caused by nonlinear load). This later achieved performance is so attractive; it demonstrates the robustness of this strategy in front of systems disturbance compared to the other control strategies already deployed in the literature. The supplementary advantage of this controller is that it performs efficiently in presence of any kind of load or even without load in the case where the system will just supply the grid with the achievement of the unity power factor UPF.

Thereafter, the system description and modeling is presented in section 2. The sliding mode control is developed in section 3. In section 4 the simulation results are illustrated using inductive load in order to demonstrate the performance of the controller. Finally, the paper presents a short conclusion..

2 System description and modelling analysis

The electrical schema of the studied system is shown in fig. 1. It consists of a PV Generator PVG coupled to a three phase grid and or without three phase load via a three phase Voltage source inverter VSI and a simple low pass filter (L,r). The capacitor C_p is used to stabilize the input voltage of the inverter.

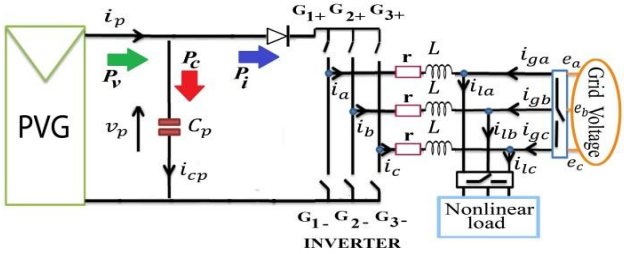


Fig. 1. Proposed topology for a PVG system connected to a three phase grid and any kind of load.

Where i_p , v_p are the PVG current and voltage, P_v is the PVG power, P_i is the inverter active power. i_a, i_b, i_c are the inverter currents. e_a, e_b, e_c are the grid voltages, C_p is the input capacitor of the inverter, R_l is the load resistance, L_l is the load inductance, i_{ga}, i_{gb}, i_{gc} are the grid currents, i_{la}, i_{lb}, i_{lc} are the load currents, $G_{1\mp}, G_{2\mp}, G_{3\mp}$ are the switching signals of the inverter switchs.

The mathematical model of the global system is given by the following equations:

$$\begin{cases} \frac{dv_p}{dt} = \frac{1}{C_p}\frac{P_v}{v_p} - \frac{1}{C_p}\frac{P_i}{v_p} \\ \frac{di_a}{dt} = -\frac{r}{L}i_a - \frac{1}{L}e_a + \frac{1}{L}v_a \\ \frac{di_a}{dt} = -\frac{r}{L}i_b - \frac{1}{L}e_b + \frac{1}{L}v_b \\ \frac{di_a}{dt} = -\frac{r}{L}i_c - \frac{1}{L}e_c + \frac{1}{L}v_c \end{cases} \quad (1)$$

Wherve v_a, v_b, v_c are the inverter voltages

The park's transformation matrix used in this paper is given by:

$$M_{abc}^{dq0} = \frac{2}{3}\begin{pmatrix} \sin(\theta) & \sin(\theta - \frac{2\pi}{3}) & \sin(\theta - \frac{4\pi}{3}) \\ \cos(\theta) & \cos(\theta - \frac{2\pi}{3}) & \cos(\theta - \frac{4\pi}{3}) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad (2)$$

θ represents the phase angle of the grid calculated by PLL technique [10].

By using this matrix, the expressions of the active power and reactive power in dq0 frame can be simplified by [11]:

$$P_i = \frac{3}{2}E_d i_d \quad (3)$$

$$Q_i = -\frac{3}{2}E_d i_q \quad (4)$$

Using park's transformation, this model can be transformed into d-q axis and expressed in the following form:

$$\begin{cases} \frac{dv_p}{dt} = \frac{1}{C_p}\frac{P_v}{v_p} - \frac{1}{C_p}\frac{1}{v_p}\frac{3}{2}E_d I_d \\ \frac{dI_d}{dt} = -\frac{r}{L}I_d + wI_q - \frac{1}{L}E_d + \frac{1}{L}V_d \\ \frac{dI_q}{dt} = -\frac{r}{L}I_q - wI_d - \frac{1}{L}E_q + \frac{1}{L}V_q \end{cases} \quad (5)$$

Where:

$$\begin{pmatrix} E_d \\ E_q \\ E_0 \end{pmatrix} = M_{abc}^{dq0}\begin{pmatrix} e_a \\ e_b \\ e_c \end{pmatrix}; \begin{pmatrix} I_d \\ I_q \\ I_0 \end{pmatrix} = M_{abc}^{dq0}\begin{pmatrix} i_a \\ i_b \\ i_c \end{pmatrix}; \begin{pmatrix} V_d \\ V_q \\ V_0 \end{pmatrix} = M_{abc}^{dq0}\begin{pmatrix} v_a \\ v_b \\ v_c \end{pmatrix}.$$

$(V_d\ V_q)$ is the input control vector considered in this work.

The advantage of this new model is that it decouples the inputs V_d and V_q such as V_d control the direct component I_d and V_q control the quadratic component I_q of the inverter output current, which enhances the stability of the global system. Moreover, the PVG output voltage v_p isn't directly in function of the control inputs, which increases the relative degree of the first output chosen in this work. That enhances the stability degree of this output and eliminates the chattering phenomena of the sliding mode control law produced if the output has a relative degree less strictly than 2. The chattering phenomena can saturate the inverter switches.

After having modelled the system, the next section presents the development of the control strategy.

3 Control strategy

The goal of this work is to elaborate a control strategy able to realize the following

objectives:

- Achieve the MPPT with excellent precision and very good response time in presence of irradiation and temperature changes.
- Convert the maximum power generated by the PVG to active power.
- Compensate the reactive power and harmonic components caused by load.
- Guarantee the global asymptotic stability of the system

All these objectives must be realized with a high performance even in the presence of system perturbation. To achieve these objectives with the fixed performance, the sliding mode method is selected to develop the controller which is characterized by high dynamic performance and robustness.

The sliding mode control is described in the next sub sections.

3.1 MPPT and active power control

To achieve the MPPT and convert all the available PVG power to active power, the following output is selected to be controlled:

$$Y_1 = \frac{\delta P_v}{\delta v_p} = i_p + v_p \frac{\partial i_p}{\partial v_p} \tag{6}$$

By controlling this output to converge to its reference $Y_{1ref} = 0$, the system will track the MPP with good precision [12].

The sliding mode control law of the first output is developed as follows:

The relative degree of the output Y_1 is 2. Therefore, the sliding mode surface must be defined as follows:

$$\sigma_1 = \dot{\varepsilon}_1 + \gamma_1 \varepsilon_1 \tag{7}$$

Where γ_1 is a positive constant.

$$\varepsilon_1 = Y_1 - Y_{1ref} = i_p + v_p \frac{\partial i_p}{\partial v_p} \tag{8}$$

And

$$\dot{\varepsilon}_1 = \left(2\frac{\partial i_p}{\partial v_p} + v_p \frac{\partial^2 i_p}{\partial {v_p}^2}\right)\dot{v}_p \tag{9}$$

The first Lyapunov function can is defined by:

$$V_1 = \frac{1}{2}{\sigma_1}^2 \tag{10}$$

So, its derivative is calculated as:

$$\dot{V}_1 = \sigma_1 \dot{\sigma}_1 \tag{11}$$

$\dot{V}_1$ must be forced to be negative, in order to stabilize ε_1 to zero. For that $\dot{\sigma}_1$ can be selected as this form:

$$\dot{\sigma}_1 = -K_1 sing(\sigma_1) \tag{12}$$

Where K_1 is a positive constant.

In other hand, $\dot{\sigma}_1$ can be expressed by using (7).

$$\dot{\sigma}_1 = \left(2\frac{\partial^2 i_p}{\partial {v_p}^2} + v_p \frac{\partial^3 i_p}{\partial {v_p}^3}\right)\dot{v}_p \left[\frac{P_v}{C_p v_p} - \frac{3E_d I_d}{2C_p v_p}\right] + \left(2\frac{\partial i_p}{\partial v_p} + v_p \frac{\partial^2 i_p}{\partial {v_p}^2}\right)\left[\frac{1}{C_p}\frac{\dot{P}_v v_p - P_v \dot{v}_p}{{v_p}^2}\right] - \left(2\frac{\partial i_p}{\partial v_p} + v_p \frac{\partial^2 i_p}{\partial {v_p}^2}\right)\frac{3E_d}{2C_p v_p}\left[-\frac{r}{L}I_d + wI_q - \frac{1}{L}E_d + \frac{1}{L}V_d\right] + \left(2\frac{\partial i_p}{\partial v_p} + v_p \frac{\partial^2 i_p}{\partial {v_p}^2}\right)\frac{3E_d I_d}{2C_p {v_p}^2}\dot{v}_p + \gamma_1\left(2\frac{\partial i_p}{\partial v_p} + v_p \frac{\partial^2 i_p}{\partial {v_p}^2}\right)\left[\frac{P_v}{C_p v_p} - \frac{3E_d I_d}{2C_p v_p}\right] \tag{13}$$

By exploiting the equation (12) and (13) of $\dot{\sigma}_1$, V_d can be calculated as follows:

$$V_d = -\frac{2C_p v_p L}{3E_d(2\frac{di_p}{dv_p}+v_p\frac{\partial^2 i_p}{\partial {v_p}^2})}[-K_1 sing(\sigma_1) - \left(2\frac{\partial^2 i_p}{\partial {v_p}^2} + v_p\frac{\partial^3 i_p}{\partial {v_p}^3}\right)\dot{v}_p\left[\frac{P_v}{C_p v_p} - \frac{3E_d I_d}{2C_p v_p}\right] - \left(2\frac{\partial i_p}{\partial v_p} + v_p\frac{\partial^2 i_p}{\partial {v_p}^2}\right)\left[\frac{1}{C_p}\frac{\dot{P}_v v_p - P_v \dot{v}_p}{{v_p}^2}\right] - \left(2\frac{\partial i_p}{\partial v_p} + v_p\frac{\partial^2 i_p}{\partial {v_p}^2}\right)\frac{3E_d I_d}{2C_p {v_p}^2}\dot{v}_p - \gamma_1\left(2\frac{\partial i_p}{\partial v_p} + v_p\frac{\partial^2 i_p}{\partial {v_p}^2}\right)\left[\frac{P_v}{C_p v_p} - \frac{3E_d I_d}{2C_p v_p}\right] + \left(2\frac{\partial i_p}{\partial v_p} + v_p\frac{\partial^2 i_p}{\partial {v_p}^2}\right)\frac{3E_d}{2C_p v_p}\left[-\frac{r}{L}I_d + wI_q - \frac{1}{L}E_d\right]] \quad (14)$$

This control law will force the Lyapunov function to be negative, which guarantees asymptotic convergence of the first output Y_1 to zero; achieving the MPPT with a high precision and fast response time. This performance can be regulated by using the controller parameters K_1 and γ_1. Furthermore, v_p is stabilized at the optimum voltage corresponding to the MPP P_{vmax}. Therefore, the capacitor power P_C is null. In fact,

$$P_C = i_{cp} v_p = C_p v_p \frac{dv_p}{dt} \quad (15)$$

which demonstrates that:

$$P_a = P_{vmax} \quad (16)$$

signifying that the inverter will convert all the available PVG power P_{vmax} to active power.

3.2 Reactive power control

In order to guarantee the Unity Power Factor and the quality of the power in the level of the power flow between the system and the utility grid, the reactive power must be controlled to be null if there is a linear load or there isn't a load connected to the system ,or equal to reactive power needed or caused by a nonlinear load. This reactive power is calculated by using a new mathematical expression in dq0 axis as follows:

$$Q_L = -\frac{3}{2}E_d I_{lq} \quad (18)$$

Where:

$$\begin{pmatrix} I_{ld} \\ I_{lq} \\ I_{l0} \end{pmatrix} = M_{abc}^{dq0}\begin{pmatrix} i_{la} \\ i_{lb} \\ i_{lc} \end{pmatrix}$$

So, the inverter must generate a reactive power Q_i such that:

$$Q_i = Q_L \quad (19)$$

The second output, which has been chosen to be controlled is:

$$Y_2 = I_q \quad (20)$$

Using (4) the reference of this output is given by:

$$Y_{2ref} = I_{qref} = -\frac{2}{3E_d}Q_L \quad (21)$$

The sliding mode control of the second output Y_2 is developed thereafter:
Firstly let's define the following tracking error:

$$\varepsilon_2 = Y_2 - Y_{2ref} = I_q - I_{qref} \quad (22)$$

The degree relative of Y_2 is 1
So, the sliding surface can be defined by:

$$\sigma_2 = \varepsilon_2 \quad (23)$$

The second Lyapunov function is expressed by:

$$V_2 = \frac{1}{2}{\sigma_2}^2 \tag{24}$$

So, its derivative is:

$$\dot{V}_2 = \sigma_2 \dot{\sigma}_2 \tag{25}$$

$\dot{V}_2$ must be negative under any condition. For that, let $\dot{\sigma}_2$ be:

$$\dot{\sigma}_2 = -K_2 sat(\sigma_2) \tag{26}$$

Where K_2 is a positive constant and Sat is a function used to eliminate the chattering phenomena, because the second output has a degree relative equal to 1. This function is defined by:

If absolute value of σ_2 <=Δ then

Sat= $\frac{\sigma_2}{\Delta}$;

Else

Sat= sign(σ_2);

End;

Where Δ is a setting parameter.

The derivative of σ_2 can be calculated by using (23) as follows:

$$\dot{\sigma}_2 = -wI_d - \frac{R}{L}I_q + \frac{1}{L}V_q - \frac{1}{L}E_q - \dot{I}_{qref} \tag{27}$$

Therefore, the quadratic control law is given by:

$$V_q = L\left[-k_2 sat(\sigma_2) + wI_d + \frac{r}{L}I_q + \frac{1}{L}E_q + \dot{I}_{qref}\right] \tag{28}$$

This control input V_q will drive I_q to converge asymptotically to I_{qref}, which permits the compensation of reactive power and harmonic component caused by the load with a high performance.

To summarize, fig. 2 shows a block diagram of the PVG system and the proposed controller.

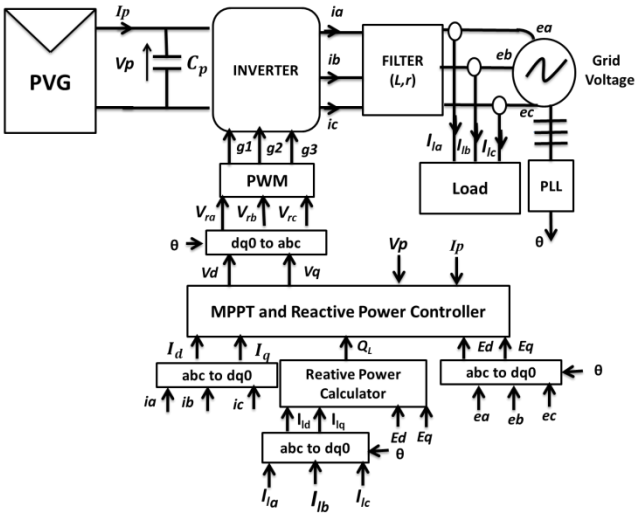


Fig. 2. Control block diagram of the PVG system and the proposed controller.

v_{ar}, v_{br}, v_{cr} are the PWM voltage references and g_1, g_2, g_3 are the switching signals of the inverter. The controller setting parameters are K_1, γ_1, K_2 and Δ.

The control input V_d and V_q will force the derivative of the Lyapunov function $\dot{V}_1$ and $\dot{V}_2$ to be negative, which guarantees the asymptotic stability of the global system.

All this performance of the elaborated controller will be demonstrated in the next section by using the simulation results under MATLAB/SIMULINK.

4 Simulation results

In order to demonstrate the performance of the proposed controller, the paper presents the simulation results of the PVG system controlled by the sliding mode control simulated under Simulink/Matlab platform. The scenario used in this simulation is shown in Fig. 3. It presents the variation of climatic conditions (solar irradiation and temperature) at specific time. To demonstrate the robustness of the controller, the disturbance has been introduced at 12s by changing the value of the input capacitor C_p amount by 20% of its original values and at 15s, the load current will be polluted by about 10% of its fundamental value.

The PVG used in this work is formed by 50 Photovoltaic modules in parallel; every module consists of 20 Siemens SM55 panels in series. The SM55 generate a 55W at 1000 $\mathrm{W/m^2}$ AND 25°C. So, this PVG has a peak power of about 50KW at standard atmospheric conditions.

The nonlinear load used in this simulation is an inductive load; which could have been any kind of load. In fact, the controller is developed independently of the kind of load.

Table 1 presents the parameters of the system and the controller.

Table 1. Parameters of the System and the Controller

System Parameters	Controller Parameters
Input capacity: $\mathrm{C_P} = 4700$ µF	$\mathrm{K_1} = 80000$
	$\gamma_1 = 70$
Filter inductance: $L_f = 3$ mH	$\mathrm{K_2} = 110000$
Filter resistance: $r = 0.002\Omega$	$\Delta = 30$
Load inductance: $L = 20$ mH	PWM Frequency: 10 KHz
Load resistance: $R = 10\Omega$	
Grid voltage: 380V	
Grid frequency: 50 Hz	

Fig. 4 shows the control inputs V_d and V_q. Its waveform doesn't contain distortion or fluctuation, which could have saturated the switching of the inverter by generating a high frequency PWM output control signal.

Fig. 5 shows the behavior of the first output $Y_1 = \frac{\delta P_v}{\delta v_p}$. It converges to zero with a very fast response and very good precision under climatic condition changes and system disturbance. Moreover the MPPT is achieved with the same performance under any condition as shown in fig. 6, which permits to exploit the maximum power produced by the PVG. Furthermore, all the available PVG power is converted to active power as proved by fig. 7. In fact, the inverter active power tracks the PVG power efficiently with good robustness.

Fig. 8 shows the power flow of the inverter output active power, grid active power and load active power. So, part of the inverter power is used to feed the load and the rest is supplied to the grid.

Now, let's presents the simulation results corresponding to the reactive power control. The behavior of the second output $Y_2 = I_q$ is shown in fig. 9. So, I_q converges with a high precision and very fast response to its reference I_{qref} , which compensates the reactive power caused by the load under atmospheric condition changes and system perturbations as

demonstrated in fig. 10. Fig. 11 shows that the reactive power injected into the grid converges to zero with negligible response time. Moreover, the unity power factor is achieved in the level of the power flow between the utility grid and the system. So, the grid current and voltage are in phase under PVG power variation, system parameters disturbance and load harmonic pollution as shown in fig. 12, fig. 13, and fig. 14.

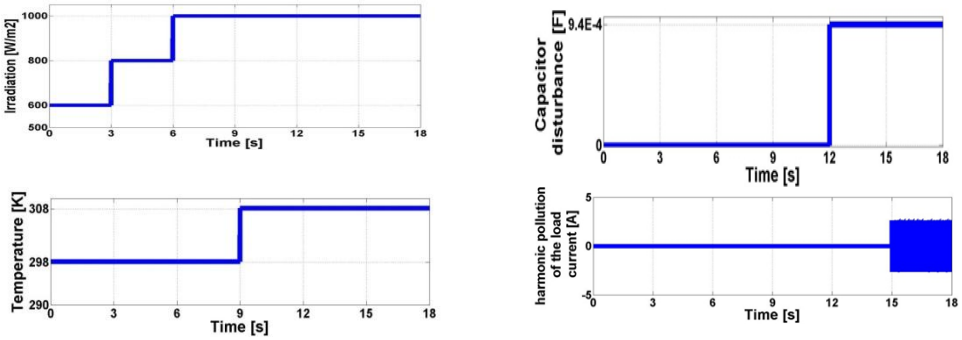


Fig. 3. Irradiation and temperature changes and disturbance values of the inverter input capacitor C_p and harmonic pollution signal of the load current.

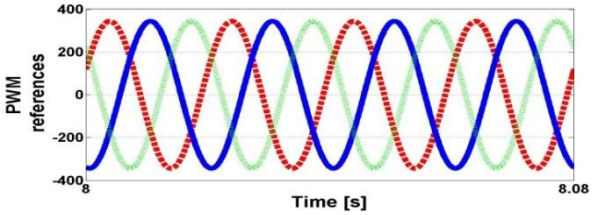


Fig. 4. The inverter three phase references output voltages.

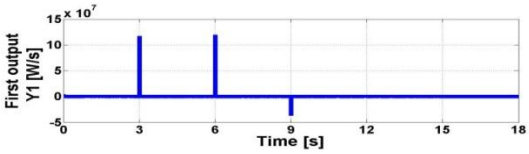


Fig. 5. The derivative of the PVG power P_v with respect to PVG voltage v_p.

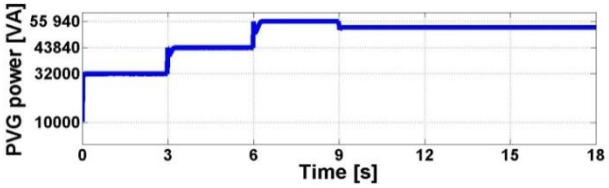


Fig. 6. MPPT achievement under abrupt variation of irradiation and temperature and in the presence of system parameter disturbance and load current harmonic pollution.

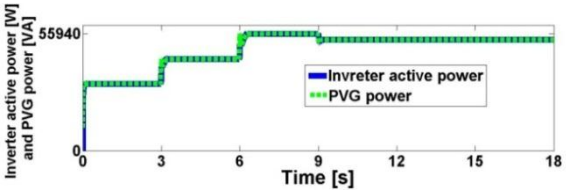


Fig. 7. The evolution of the inverter active power in function of the variation of PVG power.

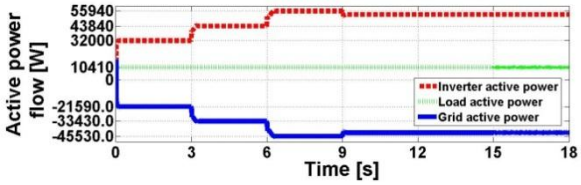


Fig. 8. Active power flow between the inverter and the grid with the active power consumed by the load.

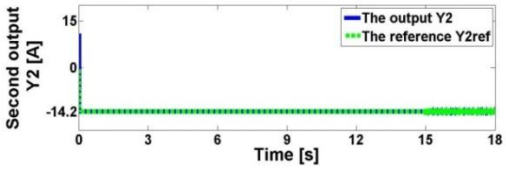


Fig. 9. The behavior of the quadratic component current of the inverter I_q.

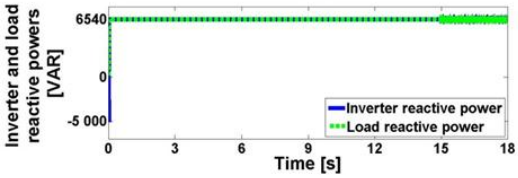


Fig. 10. Compensation of the load reactive power by the inverter output reactive power.

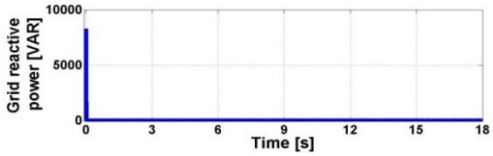


Fig. 11. Grid reactive power.

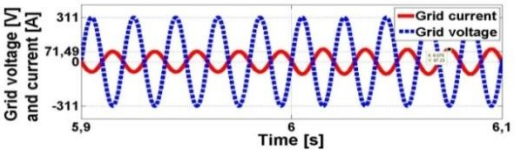


Fig. 12. Phase 'a' of the grid current and voltage (achievement of UPF under fast variation of the PVG power with a negligible response time).

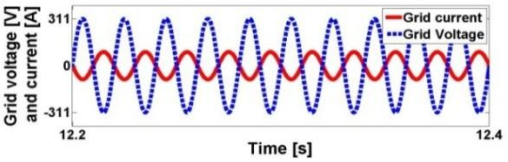


Fig. 13. Phase 'a' of the grid current and voltage (achievement of UPF in presence of inverter input capacitor disturbance).

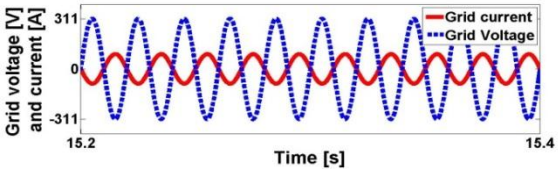


Fig. 14. Phase 'a' of the grid current and voltage (achievement of UPF during harmonic distortion of load current).

5 Conclusion

This paper presented a new model and advanced control strategy of a three phase grid connected photovoltaic system with or without load. The new model has decoupled the input control, which simplifies the operation and enhances the performance of the controller. The controller has been developed by using a sliding mode control without chattering phenomena based on dq0 transformation of the new model; in order to realize the main objectives with a high performance under climatic condition changes and system disturbance. These objectives are tracking the Maximum Power Point and achieving the UPF in the level of the grid power in the presence of load or without load. Mathematical analysis has been used to demonstrate the asymptotic stability of the controlled system and simulation results proved the very good performance of the controller.

References

[1] Jinsoo Kim and Oh yang, Design of a High-Efficient Grid-Connected Three-Phase Three-Level T-Type PV Power System, *International Journal of Control an Automation, Vol. 7, No. 10,pp. 65-78,* 2014.

[2] J. P. Lyons and V. Vlatkovic, Power Electronics and Alternative Energy Generation, Proceedings of PESC 04, pp.16-21, Germany, June 2004.

[3] Yousry atia, M.M.Salem, Aref Eliwa and Mohamed Zahran, Simulation and Implementation of Power Control Strategy Grid Connected Photovoltaic System, WSEAS TRANSACTIONS on SYSTEMS and Control,Vol. 10, E-ISSN. 2224- 2856, 2015.

[4] M.R. Bengourina, M. Rahli and L. Hassaine, Direct power control of a grid connected photovoltaic system associated with an active power filter, Revue des Energies Renouvelables, Vol. 20 N°1 pp. 99 – 109, 2017.

[5] Ehsan Afshari, Babak Farhangi, Frede Blaabjerg, Control Strategy for Three-Phase Grid

[6] Connected PV Inverters Enabling Current Limitation under Unbalanced Faults, IEEE Transactions on Industrial Electronics, pp. 8908 – 8918, 31 July 2017.

[7] Mateus F. Schonardie and Denizar C. Martins, Three-Phase Grid-Connected Photovoltaic System With Active And Reactive Power Control Using dq0 Transformation, pp. 1202-1207, IEEE 2008.

[8] C. Cecati, A. Dell'Aquila, M. Liserre. A novel three-phase single stage distributed power inverter. IEEE Transactions on Power Electronics, Vol. 19, Issue. 5, pp. 1226 – 1233, Sept. 2004.

[9] S. Nonaka. A novel three-phase sinusoidal PWM voltage source inverter and its application for photovoltaic power generation system, Power Conversion Conference - Nagaoka 1997, Vol. 2, pp. 755 – 758, 3-6 Aug 1997.

[10] M'hammed Guisser, Elhassane Abdelmounim, Mohammed Aboulfatah, Abdesselam EL-Jouni, Nonlinear Observer-Based Control for Grid Connected Photovoltaic System, IOSR Journal of Electrical and Electronics Engineering, Vol. 9, e-ISSN: 2278-1676, p-ISSN: 2320-3331, Issue. 5, PP 40-52, Sep – Oct. 2014.

[11] Evgenije Adzic, Vlado Porobic, Boris Dumnic, Nikola Celanovic, Vladimir Katic, PLL Synchrnization in Grid connected Converters, International Conference on Engineering and Technology (ICET-2013), pp. 15, Serbia May 15-17,2013.

[12] Farhan Mahmood, Luigi Vanfretti and Hossein Hooshyar, Modeling of a Detailed Photovoltaic Generation System for EMT-Type Simulation, ENERGYCON 2014, Dubrovnik, Croatia, May 13-16, 2014.

[13] M. Guisser, M. Aboulfatah, E. Abdelmounim, H. Medromi,J. Saadi. Nonlinear Control Design for Maximum Power Point Tracking and Unity Power Factor of a Grid-Connected Photovoltaic Renewable Energy Systems, IOSR Journal of Electronics and Communication Engineering, vol.9, no.5, pp.62-71, 2014.

Power Control of a Wind Energy Conversion System based on a Doubly Fed Induction Generator

Bouchaib Rached, Mustapha Elharoussi, Elhassane Abdelmounim
{ bouchaib.rached@gmail.com, m.elharoussi@gmail.com, hassan.abdelmounim@hotmail.fr }

ASTI Laboratory, FST Hassan University in Settat, Morocco

Abstract. Over the last years, there has been a strong penetration of renewable energy resources into the power supply network. Wind energy generation has played and will continue to play a very important role in this area for the coming years. Currently, the wind using a doubly-fed induction generator « DFIG» are the more used for the production of the electric energy. Our work consists of modeling of a chain of conversion of the wind energy where the doubly-fed induction generator operates at variable speed. thereafter, we consider the regulating of the active and reactive powers in order to ensure optimum operation. To this effect, we have established a vectorial control to ensure a decoupling between the electromechanical variables. The command is developed and validated using Matlab/Simulink in order to analyze by simulation the behavior of the chain in the fields of possible function.

Keywords: Wind energy conversion system (WECS); Vector control; Doubly Fed Induction Generator (DFIG).

1 Introduction

To increase the operational speed range of wind turbines and eliminate soft starter and grid-side reactive power compensator, a variable speed turbines based on doubly fed induction generator (DFIG) were developed. The configuration of Wind Energy Conversion System (WECS) using a DFIG and power electronic converter is shown in Figure 1 [1][2]. As the name implies, power from the doubly fed induction generator is fed to the grid through both stator and rotor windings.

The power converter is usually made of two-level voltage source converters connected in a back to back configuration. The machine-side converter, also called a rotor-side converter (RSC), controls the generator torque/speed or active/reactive power, whereas the grid side converter (GSC) controls the net DC-bus voltage[3][4] .

The advantages of the DFIG are that rotor power allows variable speed, high starting torque capability, and flexible reactive power control.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284101

This work presents a control chain conversion based on a (D.F.I.G). To improve the transient and steady state performance and the power factor of generation, a stator flux oriented vector control scheme is used in this work. The vector control structure employs conventional PI controllers for the decoupled control of the stator side active and reactive power. The whole system is modeled and simulated using Matlab/Simulink and the results are analyzed.

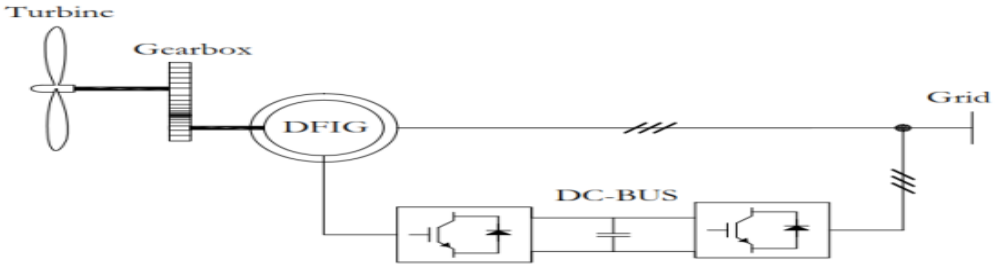


Fig. 1. Variable speed wind turbine.

2 WIND TURBINE MODELING

The power contained in the wind is given by the kinetic energy of the flowing equation [5][2][6]:

$$P_{aero} = \frac{1}{2} \cdot C_P(\beta,\lambda) \cdot \rho \cdot \pi \cdot R_{Pale}^2 \cdot V_{wind}^3 \quad (1)$$

Where ρ is the air density (kg/ m³), R_{Pale} is the blade radius (m), V_{wind} is the wind speed, $C_P(\beta,\lambda)$ is the turbine power coefficient, and λ is the tip speed ratio, which is defined by:

$$\lambda = \frac{\Omega_t \cdot R_{Pale}}{V_{wind}} \quad (2)$$

Where Ω_t is the Rotational speed of the turbine.

The aerodynamic torque is given by:

$$T_{aero} = \frac{P_{aero}}{\Omega_t} = \frac{1}{2} \cdot C_P(\lambda,\beta) \cdot \rho \cdot S \cdot \frac{V_{wind}^3}{\Omega_t} \quad (3)$$

Where S is the area swept by the pales of the turbine.

3 DFIG MODELING AND POWER CONTROL

3.1 DFIG model

For a doubly fed induction machine, the Concordia and Park transformation's application to the traditional a,b,c model allows to write a dynamic model in a d-q reference frame as follows by [7][8][9]:

Electrical equations

$$\begin{cases} V_{Sd} = R_S I_{Sd} + \dfrac{d}{dt}\Phi_{Sd} - \omega_S \Phi_{Sq} \\ V_{Sq} = R_S I_{Sq} + \dfrac{d}{dt}\Phi_{Sq} + \omega_S \Phi_{Sd} \\ V_{rd} = R_r I_{rd} + \dfrac{d}{dt}\Phi_{rd} - \omega_r \Phi_{rq} \\ V_{rq} = R_r I_{rq} + \dfrac{d}{dt}\Phi_{rq} + \omega_r \Phi_{rd} \end{cases} \tag{4}$$

Stator and rotor flux equations

$$\begin{cases} \Phi_{Sd} = L_S I_{Sd} + MI_{rd} \\ \Phi_{Sq} = L_S I_{Sq} + MI_{rq} \\ \Phi_{rd} = L_r I_{rd} + MI_{Sd} \\ \Phi_{rq} = L_r I_{rq} + MI_{sq} \end{cases} \tag{5}$$

Stator and rotor active and reactive powers

$$\begin{cases} P_S = v_{Sd} \cdot i_{Sd} + v_{Sq} \cdot i_{Sq} \\ Q_s = v_{Sq} \cdot i_{Sd} - v_{Sd} \cdot i_{Sq} \end{cases} \tag{6}$$

$$\begin{cases} P_r = v_{rd} \cdot i_{rd} + v_{rq} \cdot i_{rq} \\ Q_r = v_{rq} \cdot i_{rd} - v_{rd} \cdot i_{rq} \end{cases} \tag{7}$$

The electromagnetic torque

$$T_{em} = pM\left(I_{rd} I_{Sq} - I_{rq} I_{Sd}\right) \tag{8}$$

3.2 Establishment of the Control Strategy

Through aligning the direct axis of the reference frame along the stator flux vector position (FOC) [10][11][12], the following relationships can be derived from Eq. (5):

$$\phi_{Sq} = 0 \Leftrightarrow \phi_S = \phi_{Sd} \tag{9}$$

$$\begin{cases} V_{Sd} = 0 \\ V_{Sq} = \omega_S \Phi_{Sd} = V_s \end{cases} \tag{10}$$

So, we can write:

$$\begin{cases} \Phi_{Sd} = L_S I_{Sd} + MI_{rd} \\ \Phi_{Sq} = L_S I_{Sq} + MI_{rq} = 0 \end{cases} \tag{11}$$

The adaptation of the relations (6), (7), and (11) to the chosen system of axes and the simplifying hypotheses considered in our case gives:

$$\begin{cases} P_S = -V_S \dfrac{M}{L_S} I_{rq} \\ Q_S = \dfrac{V_S \Phi_S}{L_S} - \dfrac{V_S M}{L_S} I_{rd} \end{cases} \tag{12}$$

$$\begin{cases} P_r = g V_S \dfrac{M}{L_S} I_{rq} \\ Q_r = g \dfrac{V_S M}{L_S} I_{rd} \end{cases} \tag{13}$$

Therefore, the d and q axis components of the rotor current have to be manipulated to control the stator reactive and active power respectively. The current control loop is based on the rotor circuit equations of the DFIG.

$$\begin{cases} v_{rd} = R_r I_{rd} - L_r \omega_r \sigma I_{rq} + L_r \sigma \dfrac{dI_{rd}}{dt} \\ v_{rq} = R_r I_{rq} + L_r \omega_r \sigma I_{rd} + \dfrac{M \omega_r}{L_S} \Phi_{sd} + L_r \sigma \dfrac{dI_{rq}}{dt} \end{cases} \tag{14}$$

The control variables Vdr and Vqr of the rotor voltage can be obtained from eq.(14). The influence of the cross-coupling between the d-q axis components of rotor current on system performance is small, which can be eliminated by adopting some control law. The model of the vector control of the rotor-side converter obtained from the above analysis is shown in Fig. 2.

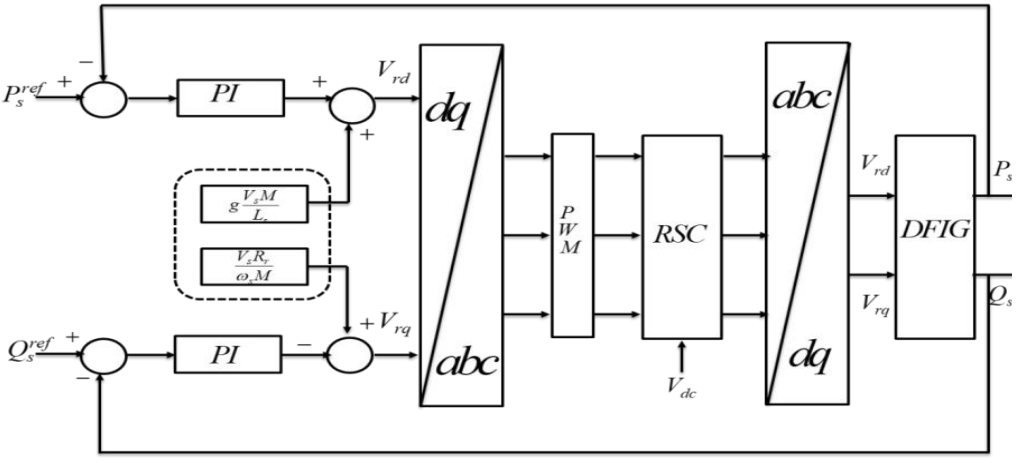


Fig. 2. Block Diagram of vector control design

4 SIMULATION RESULTS

The DFIG connected directly to the grid through the stator, and its speed is controlled via a back-to-back PWM converter. The parameters of the DFIG are given in Table 1.

Table 1. Detailed parameters of the DFIG.

Rated power	1.5 MW
ρ	1.225 Kg/m3
Blade Radius R	35.25 m
Stator resistance RS	0.012Ω
Rotor resistance Rr	0.021Ω
Stator inductance LS	0.0137 H
Rotor inductance Lr	0.0136 H
Mutual inductance M	0.0135H
Moment of inertia J	0.175 kg.m²
Damping coefficient f	0.0024 N/rd/s
Number of pole pairs p	2
C	2.2 mF

Figures 3 and 4 illustrate respectively the stator active power and reactive power. We can see the robustness of the power control of the DFIG. Figures 5 and 6 show the rotor and stator currents waveforms. The frequency of these current, vary according to the slip s.

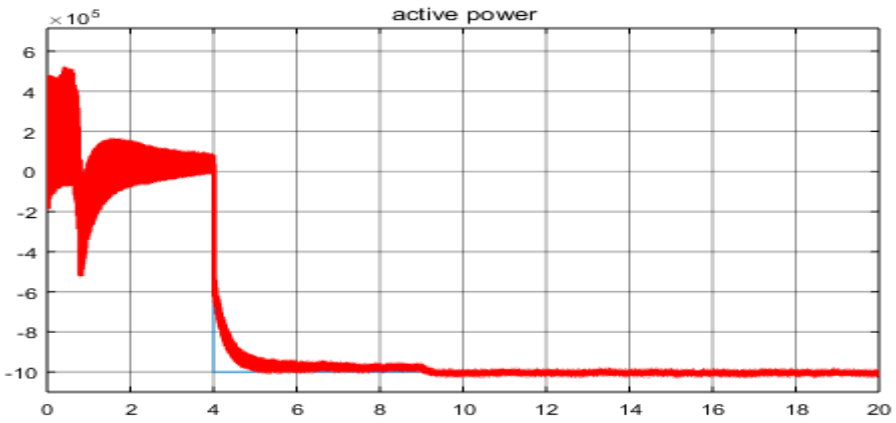


Fig. 3. Stator active power response

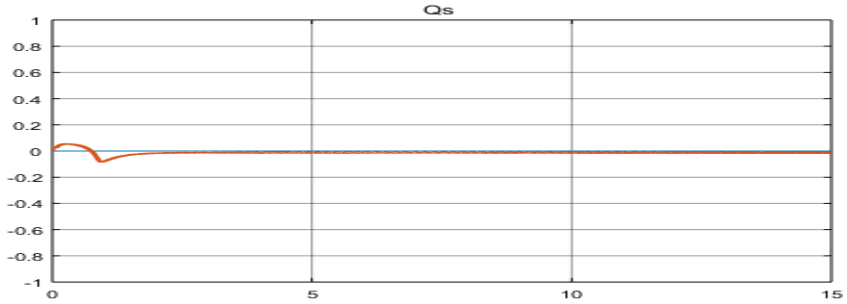


Fig. 4. Stator reactive power response

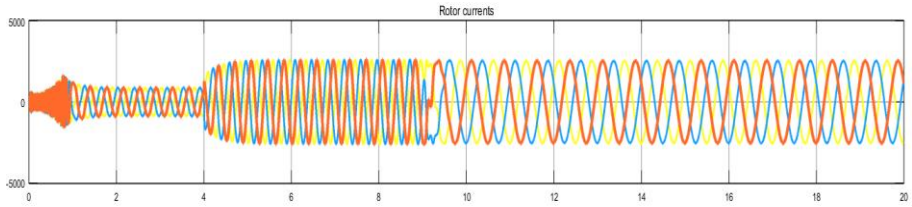


Fig. 5. rotor currents response

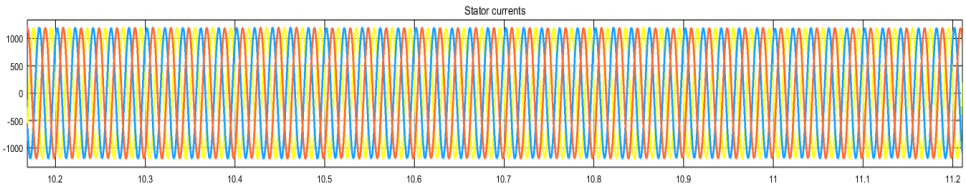


Fig. 6. Stator currents response

The active power of DFIG follows the variation of the setpoint and the reactive power is mainted to zero, which signified the reactive power output is not affected. The simulation result indicates that the active and reactive power decoupled control is achieved and the performance is good.
According to the results found in literature[13][14], which show the results of the implementation of PI controllers on wind power systems, and by observing the previous results, we see that the controller presented in this paper has a fewer oscillations during setpoint changes, and the peaks of disturbances are much less significant.

5 CONCLUSION

The control of the doubly fed induction generator, used in variable-speed wind power generation, was considered. And a control structure using standard proportional integral PI controller and a field-oriented control strategy is described and simulated.
The simulation result indicates that the active and reactive power decoupled control is achieved and the performance is good.

References

[1] I. Munteanu, Optimal control of wind energy systems : towards a global approach. Springer, 2008.

[2] G. Abad, Doubly fed induction machine : modeling and control for wind energy generation applications. Wiley-Blackwell Pub, 2011.

[3] J. A. Baroudi, V. Dinavahi, and A. M. Knight, "A review of power converter topologies for wind generators," International Journal of Renewable Energy, vol. 32, no. 14, pp. 2369–2385, 2007.

[4] Z. Zhu and J. Hu, "Electrical machines and power-electronic systems for high-power wind energy generation applications: Part I – market penetration, current technology and advanced machine systems,"COMPEL:The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 32, no. 1, pp. 7–33, 2013.

[5] P. M. M. Bongers, "Modeling and Identification of Flexible Wind Turbines and A Factorizational Approach to Robust Control Design." 1994.

[6] B. Beltran et al., "Sliding Mode Power Control of Variable Speed Wind Energy Conversion Systems To cite this version : Sliding Mode Power Control of Variable Speed Wind Energy Conversion Systems," 2010.

[7] G. S. Kaloi, J. Wang, and M. H. Baloch, "Active and reactive power control of the doubly fed induction generator based on wind energy conversion system," Energy Reports, vol. 2, pp. 194–200, Nov. 2016.

[8] C. Hamon, K. Elkington, and M. Ghandhari, "Doubly-fed induction generator modeling and control in DigSilent PowerFactory," in 2010 International Conference on Power System Technology, 2010, pp. 1–7.

[9] A. Boyette, "Contrôle-commande d'un générateur asynchrone à double alimentation avec système de stockage pour la production éolienne," Dec. 2006.

[10] S. Khojet El Khil, I. Slama-Belkhodja, M. Pietrzak-David, and B. de Fornel, "Power distribution law in a Doubly Fed Induction Machine," Math. Comput. Simul., vol. 71, no. 4–6, pp. 360–368, Jun. 2006.

[11] M. Elhaissouf, E. Lotfi, B. Rached, M. Elharoussi, and A. Barazzouk, "DSP Implementation in the Loop of the Indirect Rotor Field Orientation control for the Three-Phase Asynchronous Machine," in Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems - ICCWCS'17, 2017, pp. 1–7.

[12] E. Lotfi, B. Rached, M. Elhaissouf, M. Elharoussi, and E. Abdelmounim, "DSP Implementation in the Loop of the Vector Control Drive of a Permanent Magnet Synchronous Machine," in Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems - ICCWCS'17, 2017, pp. 1–7.

[13] A. Dekhane, S. Lekhchine, T. Bahi, S. Ghoudelbourg, and H. Merabet, "DFIG modeling and control in a wind energy conversion system," 2012 1st Int. Conf. Renew. Energies Veh. Technol. REVET 2012, pp. 287–292, 2012.

[14] S. P. Uma and S. Manikandan, "Control technique for variable speed wind turbine using PI controller," in 2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN), 2013, pp. 640–643.

Boost converter system modelling and incremental conductance algorithm for photovoltaic system via Matlab/Simulink

Benlafkih Abdessamad and Chafik Elidrissi Mohamed

{ ab.benlafkih@gmail.com, labogem15@gmail.com }

Laboratory of Engineering Energy and Materials, Faculty of Sciences
University Ibn Tofail, Kenitra, Morocco

Abstract. This paper presents a method of modeling and simulation of photovoltaic panel in MATLAB/Simulink using solar cell block from SimElectronics library. The method is used to determine the characteristic of a particular photo voltaic cell panel I-V & P-V, and to study the influence of different values of solar radiation at different temperatures concerning performance of photo voltaic cells and to study the influence of different values of temperatures at different solar radiation. Also from the simulation it is inferred how the maximum power point is tracked using Incremental conductance algorithm to maximize the power output of the PV panel.
Keywords: System modeling, solar cell, MPPT, photovoltaic panel, boost converter, Incremental conductance algorithm, simulation, Matlab/Simulink.

1 Introduction

Renewable energy sources offer clean production of electrical power using sunlight, wind, biomass, tidal waves etc. Renewable energy generation have grown greatly due to the concerns of climate change and the increase in oil prices. The growth in renewable energy has been very consistent in the last two decades. Not only the increasing concerns regarding climate change and the soaring of oil prices but also the great support by renewable energy legislation and incentives with a close to 150 billion US Dollars in 2007 have brought this alternative source of electrical power generation to the big picture [1]. Photovoltaic (PV) systems are one of the most popular renewable energy sources. It is an interesting energy source as it is not only renewable but also inexhaustible and non-polluting unlike the conventional fossil fuels such as coal, oil and gas. These unique features have made power generation through Photovoltaic sources one of the most popular renewable energy sources in last decade [2]. Photovoltaics convert sunlight into electrical energy using photoelectric effect. Sun's radiation is converted directly into usable electricity by the photovoltaic systems. Photovoltaic (PV) systems are made of photovoltaic modules which are semiconductor devices that convert the incident solar radiation directly into electrical energy. PV power depends greatly on solar irradiation and temperature. As a result the power generated by solar systems are not constant. Apart from this clean conversion of solar energy into electrical energy, one thing which is holding back the photovoltaic systems is their lack of reliability. Depending on different solar irradiation levels and temperatures, their production rate varies. Therefore by adding a supplemental source of power, Solar Power‘s reliability can be greatly increased where this supplement energy source will work as a backup

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284106

energy source. Whenever the load demand will not be fully met by the primary solar energy source it will be supported by the backup energy source. And on the other hand when the demand will be less than the generation, the primary solar source will energize the backup source.

2 PV MODEL

Figure 1 presents the widely used single-diode equivalent circuit model to represent the characteristics of the PV panel for analyzing and simulating the PV system. The circuit model represents the PV panel as a current source I_{ph} in parallel with a single diode and a shunt resistor R_P as well as a series resistor R_S . The current I_{PV} and voltage V_{PV} from the PV panel is characterized by the following Equations (1) and (2):

$$I_{PV} = I_{ph} - I_D - \frac{V_{PV}+R_S I_{PV}}{R_S} \quad (1)$$

$$I_D = I_0 \left[exp\left(\frac{q(V_{PV}+R_S I_{PV})}{nAKT}\right) - 1 \right] \quad (2)$$

where I_{ph} is the current generated by the incident light, I_D is the diode current, I_0 is the reverse saturation current, q is the electron charge (1.602 × 10−19 C), K is the Boltzmann constant (1.38 × 10−23 J/K), T is the operating temperature of the cell in Kelvin (K), A is the diode ideality constant, and *n* is the number of diodes in series to form the single-diode model. The current source Iph mainly depends on the irradiation level of sun light as well as the operation temperature of the solar panel.

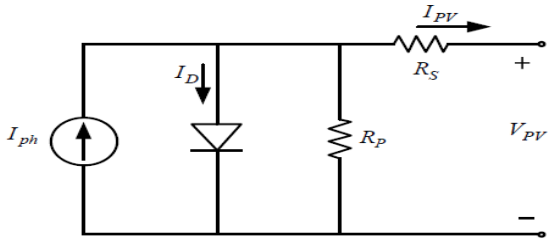


Fig. 1. Single-diode equivalent circuit model

In this study, the PV module was selected to conduct the design and circuit simulation; the table 1 present the PV model characteristics at (25 C° and 1000 W/M^2).

Table 1. The PV MODEL characteristics at (25 C° and 1000 W/M^2)

Paramerter	Value
Maximum Power	**61.24 W**
Tension at Pmax	**17.32 V**
Current Pmax	**3.5 A**
Open circuit Voltage	**21.02 V**
Short circuit Current	**3.8 A**

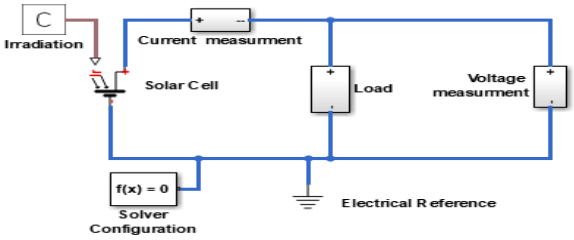


Fig. 2. PV Model Simulation

2.1 Current -Voltage and Power-Voltage characteristics

One way of studying the consistency of the model is developed to study the shape of the current-voltage characteristics I-V, Fig. 3 and P-V, power-voltage, Fig. 4, was obtained using the equations the electrical model (1) and (2)

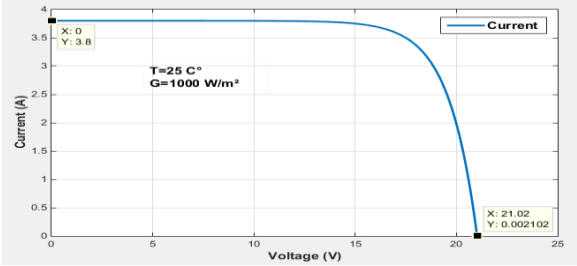


Fig. 3. Single diode equivalent circuit Model

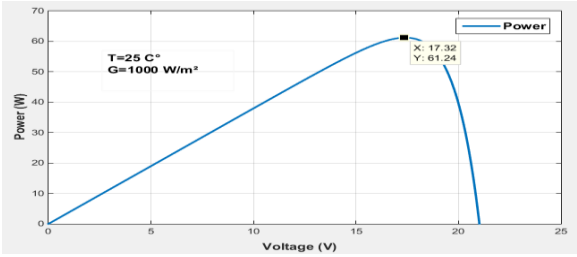


Fig. 4. P-V Characteristics curve of PV panel Simulation

2.2 Influence of temperature

To characterize PV cells, we used the model Fig.2, to provide the values of voltage, current product and the power generated. We present the I-V and P-V characteristics in figures 5 and 6 respectively of PV panel, for G = 1000 W/m^2 given, and for different values of temperature. If the temperature of the photovoltaic panel increases, the open circuit voltage Voc decreases. The temperature increase is also reflected by the decrease of the maximum power.

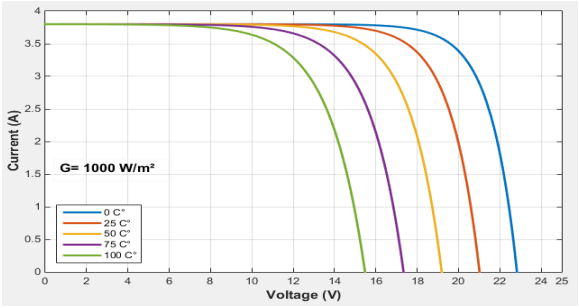


Fig. 5. Simulated I-V Curves of PV Model influenced by tepmerature

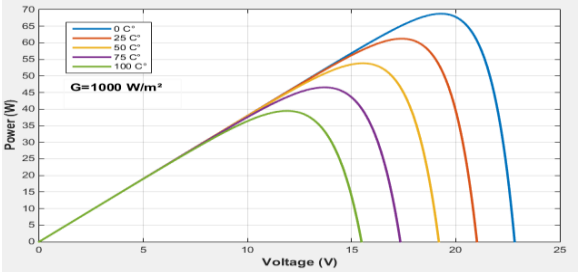


Fig. 6. Power versus Voltage P-V curves influenced by tempurature

2.3 Influence of irradiation

Now, we present the I–V and P–V characteristics in figures 7 and 8 respectively of the PV photovoltaic model at a given temperature T = 25°C for different solar illumination levels.

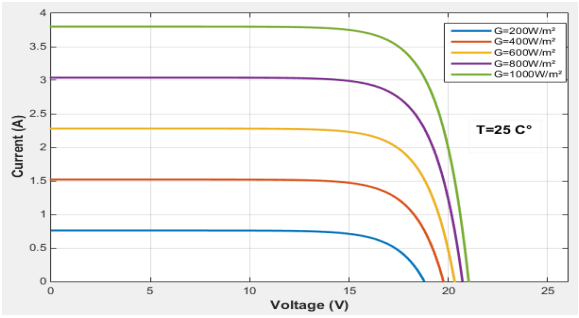


Fig.7. Simulated I-V curves of PV model influenced by solar illumination

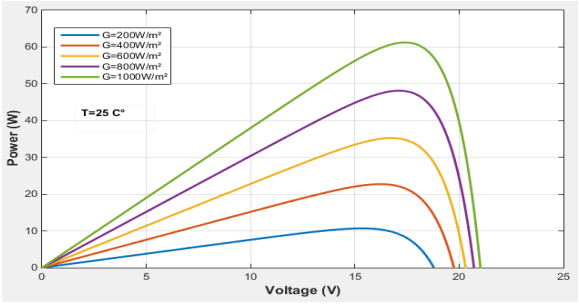


Fig. 8. Power versus voltage P-V curves influenced by solar illumination

2.4 DC/DC boost converter

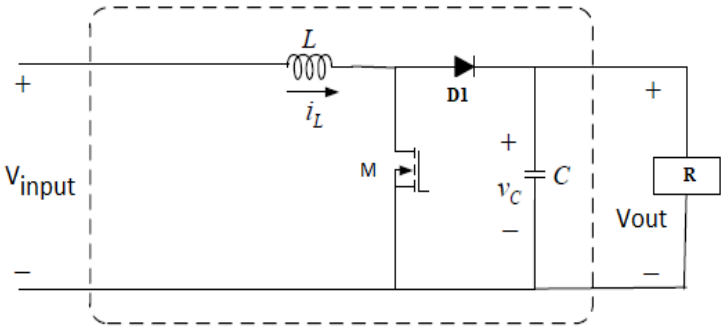


Fig.9. DC/DC boost converter

The boost DC/DC converter offers a greater level of capability, it as expected it extra components may be required to provide the level of functionality needed [3], [4], [5].
In Fig. 9 a DC/DC Boost converter is shown. The switching period is $T$and the duty cycle is D. Assuming continuous conduction mode of operation, during T_{ON} the period of the cycle, switch M is ON and the input voltage is impressed across the inductor. Since the load current is instantaneously provided by the output capacitor during this interval, the capacitor voltage (output voltage) decreases, the state space equations are given by Eq.(3).

$$\begin{cases} \frac{di_L}{dt} = \frac{1}{L}\,[v_{in}] \\ \frac{dv_c}{dt} = -\frac{1}{c}\left(\frac{v_c}{R}\right) \end{cases} \tag{3}$$

During the other interval of the switching period (T_{OFF}) , Diode D1 is turned ON and the inductor energy is transferred to the output, providing both the load current and also charging the output capacitor, the equations are given by Eq.(4).

$$\begin{cases} \frac{di_L}{dt} = \frac{1}{L}\left[v_{in} - v_c\right] \\ \frac{dv_c}{dt} = \frac{1}{c}\left(i_L - \frac{v_c}{R}\right) \end{cases} \quad (4)$$

2.5 MPPT Control algorithm

The configuration of MPPT controller is shown in Fig.10. The inputs of MPPT controller are voltage and current of the PV model, while the output is PWM (pulse width modulation) for controlling the duty cycle of the boost converter. The system is simulated using Matlab/Simulink. Many MPPT techniques have been proposed in the literature; examples are the Perturb and Observe (P&O), Incremental Conductance (IC), Fuzzy Logic, and so forth. The P&O algorithm is very popular and simple [6], [7], [8], [9], [10].

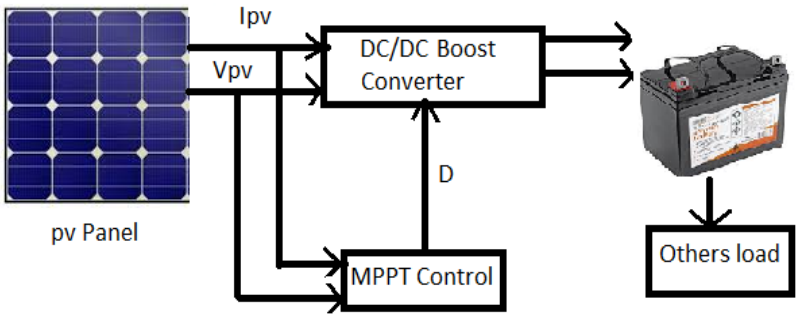


Fig.10. PV System with Power conveter and MPPT Control

2.6 Incremental conductance method

This method is based on the fact that the slope of the power curve of the panel is zero at the MPP, positive to the left and negative to right. This method is based on the fact that the slope of the power curve of the panel is zero at the MPP, positive to the left and negative to right [11], [12], [13], [14]. Since :

$$\frac{dP}{dV} = \frac{d(IV)}{dV} = I + V\frac{dI}{dV} = I + V\frac{\Delta I}{\Delta V} \quad (5)$$

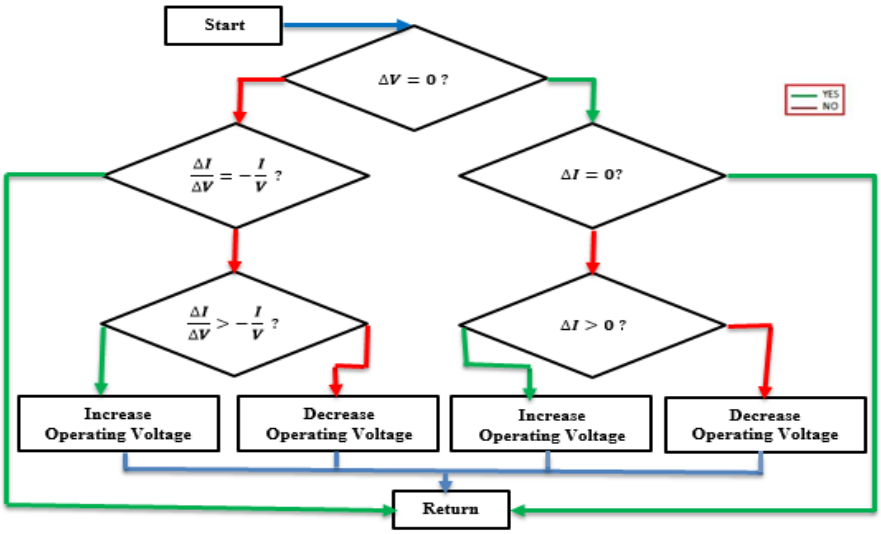


Fig.11.Organigram of incremental conductance algorithm

$$\begin{cases} \frac{\Delta I}{\Delta V} = -\frac{I}{V} & at\ the\ MPP \\ \frac{\Delta I}{\Delta V} > -\frac{I}{V} & left\ of\ the\ MPP \\ \frac{\Delta I}{\Delta V} < -\frac{I}{V} & right\ of\ the\ MPP \end{cases} \quad (6)$$

The MPP can be tracked by comparing the instantaneous conductance to the incremental conductance, as shown in the organogram of Fig.11. The detailed Simulink model is shown in Fig.12. The Vpv and Ipv are taken as the inputs to MPPT unit, duty D cycle is obtained as output.

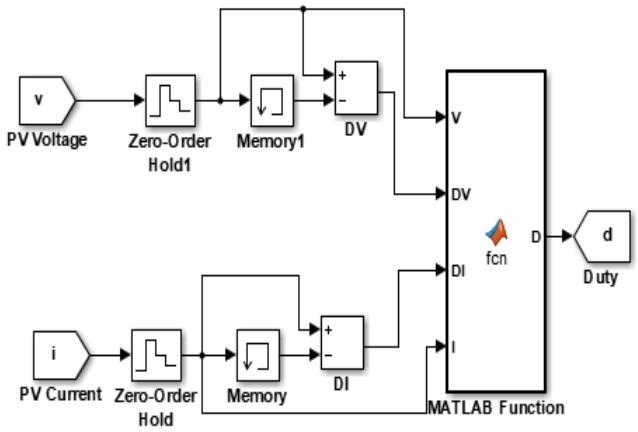


Fig.12. Maximum power point tracking by incremental conductance method

3 Simulation model of PV system with incremental conductance algorithm

The model shown in Fig.13 represents a model of PV solar panel connected to resistive load through a DC/DC boost converter with Incremental Conductance Algorithm. The boost system specifications are given as follows:

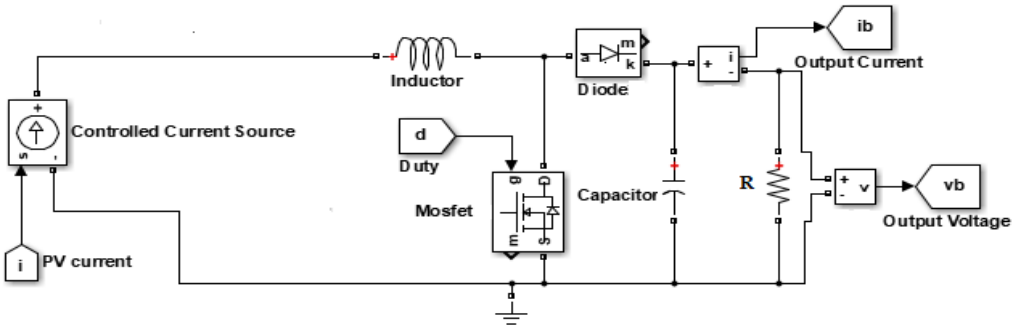


Fig.13 Boost converter circuit with PV input

Figure 14 presents how the irradiance that falls on PV solar panel is changing. The voltage and the current vary depending on irradiance. The curve of variable irradiance is plotted using a signal builder, where the irradiance is not very realistic, that will be equivalent to do very fast cloud moving for example, what allowing to the sun changing instantaneous which is not happen, but allow to give an idea of measure of how fast the controller responds [15].

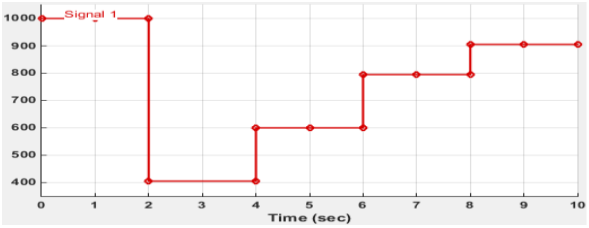


Fig.14. Variation of solar radiation

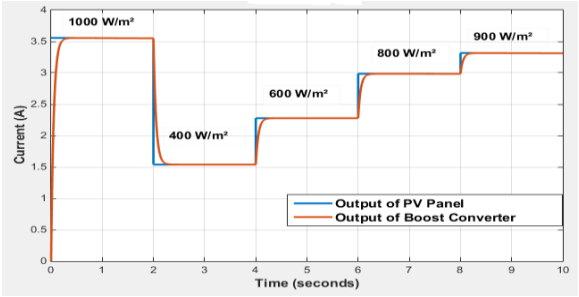


Fig.15. Output current of the Boost converter and output current of the PV panel

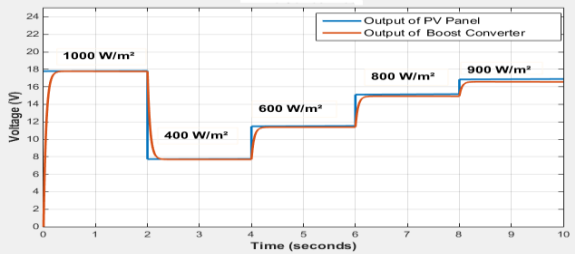


Fig.16. Output Voltage of the Boost converter and output Voltage of the PV panel

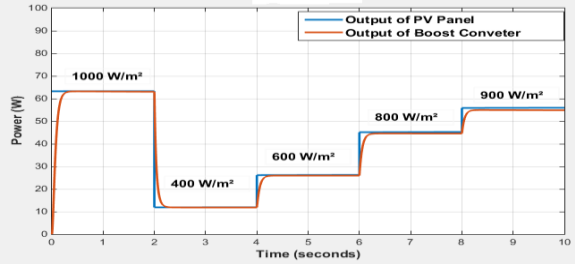


Fig.17. Output power of the Boost converter and output power of the PV panel

Figure 15 to 17 presents the results of the simulation model PV panel. The current, voltage and power output of PV panel and at the output of circuit connected to the photovoltaic panel. The Irradiance is variable, passing successively through the following values: 1000, 400, 600, 800 and 900 W/m^2. To test the operation of the system, the change of solar radiation was modeled. The temperature is fixed at 25 °C and the level of solar radiation is varied with five levels. The first level of illumination is set at 1000 W/m² then at 2 second the solar irradiation level passed abruptly at the second level 400 W/m^2, next at 4 second, it passed at the third level 600 W/m^2 , after that at 6 second, it passed at the 800 W/m^2, Finally at 8 second, it passed at the last level G=900 W/m^2. An illustration of the relationship between the irradiation and the output power of PV panel is shown in figure 15 to 17 to explain the effectiveness of the algorithm mentioned. According to the simulation results presented above, all quantities to regulate, and converge well to references, and after a few time acceptable response = 0.1s These results show the effectiveness of the algorithm and the relationship between the illumination and the output power of the PV panel, and show the operation of the boost converter.

4 Conclusion

In this paper, a standalone PV system has been simulated by Matlab/Simulink. Incremental conductance algorithm has been used for maximum power point tracking. Simulation results show that the system operates in the maximum power point. This technique has an advantage over the perturb and observe method because it can determine when you reach the MPP without having to oscillate around this value. It can also perform MPPT under rapidly increasing and decreasing irradiance conditions with higher accuracy than the perturb and observe method.

References

[1] A. A bete, E .Barbis io, F.Cane, and P. Demartini, ―Analysis of photovoltaic modules with protection diodes in presence of mismatching,‖ in Photovoltaic Specialists Conference, 1990., Conference Record of the Twenty First IEEE, pp. 1005–1010 vol. 2 (1990)

[2] Jie Shi ; Wei-Jen Lee ; Yongqian Liu ; Yongping Yang ; Wang, Peng ‖Forecasting power output of photovoltaic system based on weather classification and support vector machine‖ Industry Applications Society Annual Meeting (IAS), 2011 IEEE DOI: 10.1109/IAS.2011.6074294 Publication Year: 2011 Y.T. Tan, D.S. Kirschen and N. Jenkins, 'A Model of PV Generation Suitable for Stability Analysis', IEEE Transactions on Energy Conversion, Vol. 19, N°4, pp. 748 – 755, (2004).

[3] A. Brambilla, M. Gambarara, A. Garutti and F. Ronchi, 'New Approach to Photovoltaic Panels Maximum Power Point Tracking', 30th Annual IEEE Power Electronics Specialists Conference, PESC'99, Vol. 2, pp. 632 - 637, (1999).

[4] J.A. Gow and C.D. Manning, 'Development of a Photovoltaic Panel Model for Use in Power Electronics Simulation Studies', IEE Proceedings on Electric Power Applications, Vol. 146, N°2, pp. 193 - 200, March (1999).

[5] E. Solodovnik, S. Liu and R. Dougal, 'Power Controller Design for Maximum Power Tracking in Solar Installations', IEEE Transactions on Power Electronics, Vol. 19, N°5, pp. 1295 - 1304, (2004).

[6] J.A. Gow and C.D. Manning, 'Development of a Photovoltaic Panel Model for Use in Power Electronics Simulation Studies', IEE Proceedings on Electric Power Applications, Vol. 146, N°2, pp. 193 , March.(2000)

[7] M. Nabipour, M. Razaz, S.GH Seifossadat, S.S. Mortazavi "A new MPPT scheme based on a novel fuzzy approach" Renewable and Sustainable Energy Reviews 74 (2017) 1147–1169-journal homepage: www.elsevier.com/locate/rser.(1999)

[8] Serkan Öztürka, Polat Poşpoşa, Volkan Utalayb, Atila Koçb, Muammer Ermişc, Işık Çadırcıa,"Operating principles and practical design aspects of all SiC DC/AC/DC converter for MPPT in grid-connected PV supplies" Solar Energy 176 380–394.journal homepage: www.elsevier.com/locate/solener. (2018)

[9] Buvana Lefevrea, Bert Herteleera, Sven De Breuckerb, Johan Driesena "Bayesian inference based MPPT for dynamic irradiance conditions" Solar Energy 174 1153–1162, journal homepage: www.elsevier.com/locate/solener. (2018)

[10] A. Chermitti, O. Boukli-Hacene and S. Mouhadjer, 'Design of a Library of Components for Autonomous Photovoltaic System under Matlab/Simulink', International Journal of Computer Applications, (0975 – 8887), Vol. 53, N°14, (2012).

[11] T.C. Yu and T.S. Chien, 'Analysis and Simulation of Characteristics and Maximum Power Point Tracking for Photovoltaic Systems', In Proceedings of the International Conference on Power electronics and Drive Systems (PEDS '09), pp. 1339 - 1344, January (2009).

[12] M.A.S. Masoum, H. Dehbonei and E.F. Fuchs, 'Theoretical and Experimental Analyses of Photovoltaic Systems With Voltage-and Current-Based Maximum Power-Point Tracking', IEEE Transactions of Energy Conversion, Vol. 17, N°4, pp. 514 – 522, (2002)

[13] F. Adamo, F. Attivissimo, A. Di Nisio and M. Spadavecchia, 'Characterization and Testing of A Tool for Photovoltaic Panel Modeling', IEEE Transactions on Instrumentation and Measurement, Vol. 60, N°5, pp. 1613 - 1622, May (2011).

[14] I.H. Altas and A.M. Sharaf, 'A Photovoltaic Panel Simulation Model for Matlab-Simulink GUI Environment', International Conference on Clean Power, pp. 341 – 345, (2007).

Miniature 2.45 GHz Patch Antenna with Defected Ground Structure for Rectenna Applications

[1]Abdellah Taybi, [2]Abdelali Tajmouati, [3]Jamal Zbitou, [4]Ahmed Errkik, [5]Larbi El Abdellaoui, 6Mohamed Latrach

{ [1]taybi.abdellah@gmail.com, [2]tajmoua@gmail.com, [3]zbitou.jamal@yahoo.fr, [4]ahmed.errkik@uhp.ac.ma, [5]elabdellaoui20@hotmail.com, [6]mohamed.latrach@eseo.fr }

[1,2,3,4,5] LMEET, FST of Settat, Hassan 1st University, Settat, Morocco
[6]Microwave Group ESEO Angers, France

Abstract. In the context of miniaturization of microwave systems that this paper is aimed at. We have designed a novel miniaturized 2.45 GHz antenna dedicated to operate for wireless power transmission applications. At first, the antenna was radiating at 5.8 GHz with the optimal dimentions, then, and by conserving the same size, we have switched the operating frequency to 2.45 GHz by using a novel configuration of Defected Ground Structure (DGS). The final structure have been designed and optimized by using Computer Simulation Technology (CST) Microwave Studio .The obtained results confirm the proper functioning of the antenna.

Keywords: Wireless Power Transmission, DGS, Rectenna, CST Microwave Studio.

1 Introduction

The wave of innovations continues to improve day by day and particularly in the field of telecommunications, which are in a period of globalization. Therefore, studies have focused on significant miniaturization, especially for the antenna part to meet modern telecommunications needs. The antenna is an essential part of any wireless device, especially RECTENNA (RECtifying anTENNA) systems. Figure 1 shows a basic block diagram of rectenna system. A rectenna is a device that collects incident electromagnetic radiation and converts it to DC power. The rectenna consists of four elements: a receiving antenna, an input filter, a diode or combination of diodes [1] that performs wave rectification, and an output filter that makes DC energy available at the load. The possibility of using Rectennas for wireless power transmission [2] purposes have been initiated and experimented by W.C Brown [3] by using a model of helicopter powered by microwaves transmited from the ground and received by an attached rectenna. A variety of rectenna designs optimized for conversion efficiency in the industrial, scientific, medical band (ISM) reaching important efficiencies for different values of input power have been reported [4,6].

There are several types of antennas depending on geometry, gain, beam shape and bandwidth. Patch antennas [7,9] (Figure 2) are widely used for microwave applications mainly for those involving wireless power transmission. It has the advantage of low weight and is made by engraving a printed circuit board. Due to their manufacturing technology, they can be integrated as close as possible to electronic circuits, occupying a reduced volume and conforming to different types of surfaces. Their main advantage is their low manufacturing cost.

Otherwise, the operating frequency has a significant weightiness on the antenna's size. Several operating frequencies of the rectenna have been considered and investigated.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284085

Components of microwave power transmission have traditionally been focused on 2.45 GHz and recently moving up to 5.8 GHz, which has a smaller antenna aperture area than that of 2.45 GHz.

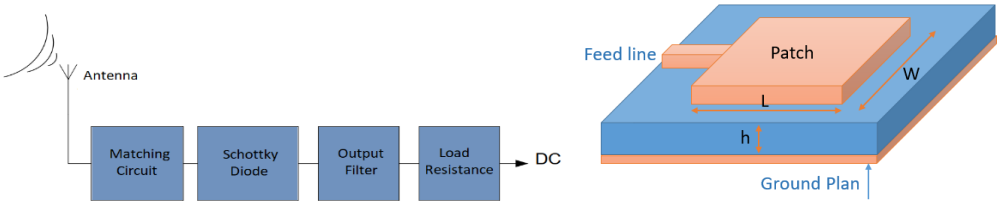


Fig. 1. (a) Block diagram of rectenna circuit (b) Rectangular microstrip antenna configuration

Many research topics have been devoted to benifits from small size and performant antennas, hence, Defected Ground Structure (DGS) [10,12] is becoming an effective and efficient technology for the miniaturization of patch antennas. DGS is a periodic motif engraved on the ground plan. The name of this technique simply means that a defect is placed on the ground plane which is generally considered an anomaly of a current well.

In this paper, a novel DGS configuration have been employed on a 5.8 GHz antenna's ground plane so that the frequency could be switched to 2.45 GHz without losing in terms of antenna's performances.

2 5.8 GHz initial antenna analysis and design

Owing to technological advances in telecommunications and continuous scientific research concerning printed antennas, and given the multiple requirements in the field of communication, the use of conventional microwave antennas is no longer able to meet these requirements. For this reason, printed antennas replace conventional ones in a number of applications, including rectenna systems. In general, a printed antenna (Fig. 1(b)) consists of a ground plane, one or more substrate layers that may have the same or different permittivities. The surface contains a radiating element of any geometry (rectangular, circular, slit, or more elaborate shapes). Among the many advantages: A low manufacturing cost, lightweight and space-saving, the possibility of networking to improve directivity and the compatibility with hybrid and MMIC (Microwave Monolithic Integrated Circuit) circuits.

2.1 Calculation of Width (W), Length (L) and Effective Dielectric Constant ε_{reff}

- For an effective radiator, practical width of the patch antenna that leads to good radiation efficiencies is given by [13]:

$$W = \frac{C_0}{2F_0}\sqrt{\frac{2}{1+\varepsilon_r}} \qquad (1)$$

Where C_0 is the free-space velocity of light i.e. 3×10^{-8} m/s and ε_r is the dielectric constant of material.

- The length of the patch determines the resonance frequency thus it is a critical factor for narrowband patch. Since it is not possible to accurately account the fringing field the results are not definite. Below is the equation to calculate the length of the patch :

$$L = \frac{\lambda_{eff}}{2} - 2\Delta L \qquad (2)$$

Where ΔL is the length extension because of fringing field, which can be calculated as follow:

$$\Delta L = 0.412\, h\, \frac{(\varepsilon_{reff}+0.3)\left(\frac{W}{h}+0.264\right)}{(\varepsilon_{reff}-0.258)\left(\frac{W}{h}+0.8\right)} \qquad (3)$$

- The value of effective dielectric constant is less than dielectric constant of the substrate, because of the fringing fields are not confined in dielectric substrate around the periphery of the patch only, but is also spread in the air. The value of this effective dielectric constant is given by [13]:

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-1/2} \qquad (4)$$

2.2 Initial antenna design

The development of rigorous techniques to solve maxwell's equations has introduced and imposed computer tools in electromagnetism. These tools are increasingly used in the analysis and design of microwave devices used in microwave and wireless communications applications. Digital electromagnetic calculation tools are therefore becoming an essential element in the design of a patch antenna due to rapid growth of the computer's processing power. The choice of simulation software therefore depends on the problem to be solved, for example the size, type and material of the structure, the complexity of the geometry. In our case, we have chosen the transient solver of Computer Simulation Technolgy (CST) Microwave Studio [14] which is an electromagnetic simulation software of 3-D passive structure based on the resolution of Maxwell's equations according to the technique of integral equations (Finite Technical Integration).

The proposed patch antenna is depicted in figure 2.

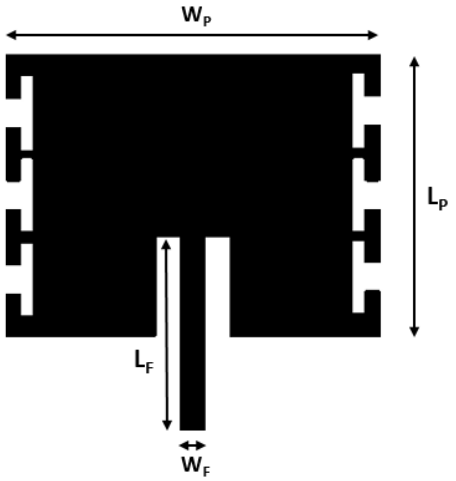


Fig. 2. The proposed patch antenna.

The FR-4 substrate have been used with dielectric permittivity constant 4.4, thickness of 1.6mm and loss tangent of 0.025. We have also chosen to feed our patch antenna with the microstrip-line method since it is easy to fabricate, simple to match by controlling the inset position and rather simple to model. The following figure shows the reflection coefficient of the proposed antenna at the resonance frequency 5.8 GHz. The antenna is well matched at the desired frequency.

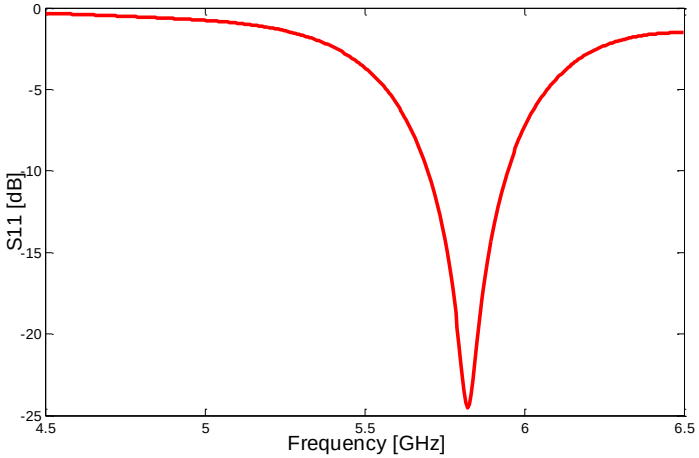


Fig. 3. Simulated reflection coefficient versus frequency.

3 Defected Ground Structure Design

In order to improve the performance of a patch antenna, several processes and techniques have been used, including the DGS technique, which stands for "Defected Ground Structure". This technique consists of applying more or less regular deformations (periodic or not) to the structure of the ground plane. Current research has allowed a greater degree of freedom in the design and location chosen on a ground plan. Whatever these deformations, they have a direct influence on the distribution of electrical charges and currents, which greatly modifies the characteristics of the antenna: the bandwidth, gain, directivity, radiation pattern, etc.. Many forms of DGS unit cells have been proposed in the literature for antenna design. Several types of patterns are illustrated in Figure 4.

Figure 5 shows the structure of the proposed Dumbbell shaped DGS with a ring slot between the two edges of the dumbell. Once investigating the whole parameters of the DGS, with a series of optimizations and simulations, the usage of the structure can be summarized and then etched on the metallic ground plane of the afromentionned microstrip antenna, with the dimentions shown in figure 5: L_{DGS1}, L_{DGS2}, L_1, L_2, L_3, W_1 ,W_2 ,W_3.

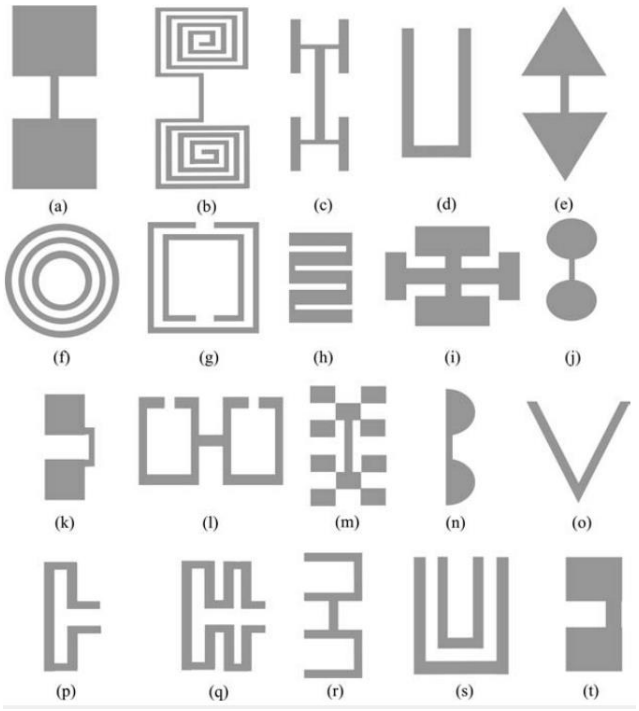


Fig. 4. Different DGS geometries: (a) dumbbell-shaped (b) spiral-shaped (c) H-shaped (d) U-shaped (e) arrow head dumbbell (f) concentric ring shaped (g) split-ring resonators (h) interdigital (i) cross-shaped (j) circular head dumbbell (k) square heads connected with U slots (l) open loop dumbbell (m) fractal (n) half-circle (o) V-shaped (p) L-shaped (q) meander lines (r) U-head dumbbell (s) double equilateral U (t) square slots connected with narrow slot at edge. [15]

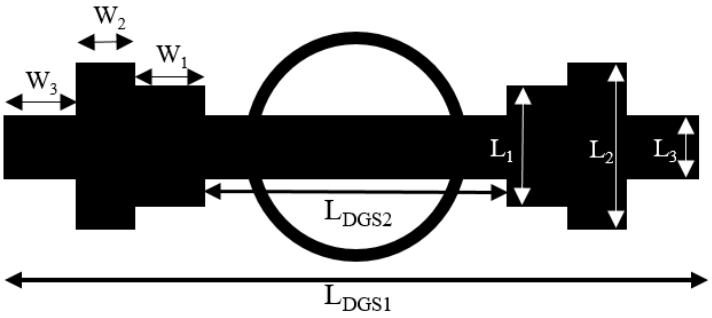


Fig. 5. The proposed Defected Ground Structure.

4 Final Antenna Design with DGS

The aim of this reseach study being the miniaturization of the 2.45 GHz patch antenna, we have assembled the designed 5.8 GHz antenna and the proposed new configuration of Defected Ground Structure on an FR4 substrate with dielectric permittivity constant 4.4, thickness of 1.6mm and loss tangent of 0.025 as shown in the figure 6, the whole system was simulated by using the Transient solver of CST Microwave Studio. The Defected Ground Structure have been placed in an optimized way permitting to obtain the desired results . Figure 7 shows the simulated reflection coefficient of the system, we can conclude from the graph that the resonating frequency is well shifted from 5.8 GHz to 2.45 GHz of the Industrial Scientific Medical band (ISM) with a return loss of -34 db and bandwidth of 260 Mhz. The 2D radiation pattern is given by Figure 7 in the E-plane, which show a stable and bi-directional radiation pattern for the resonant frequency band. The miniaturisation have been achieved with a size reduction of more than 80%.

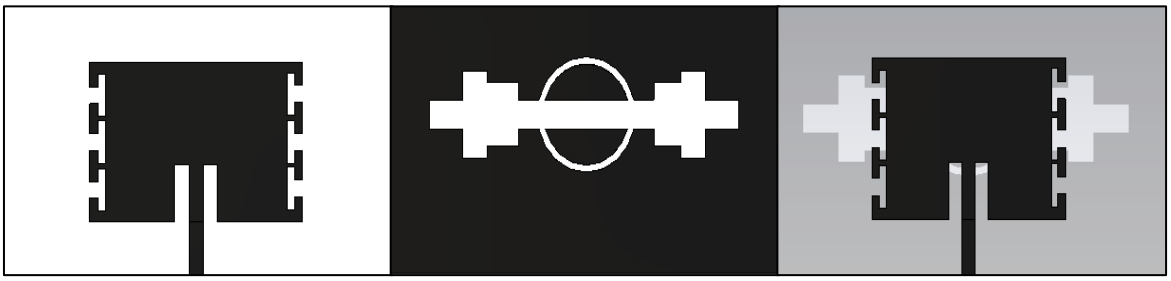


Fig. 6. Final Antenna Design

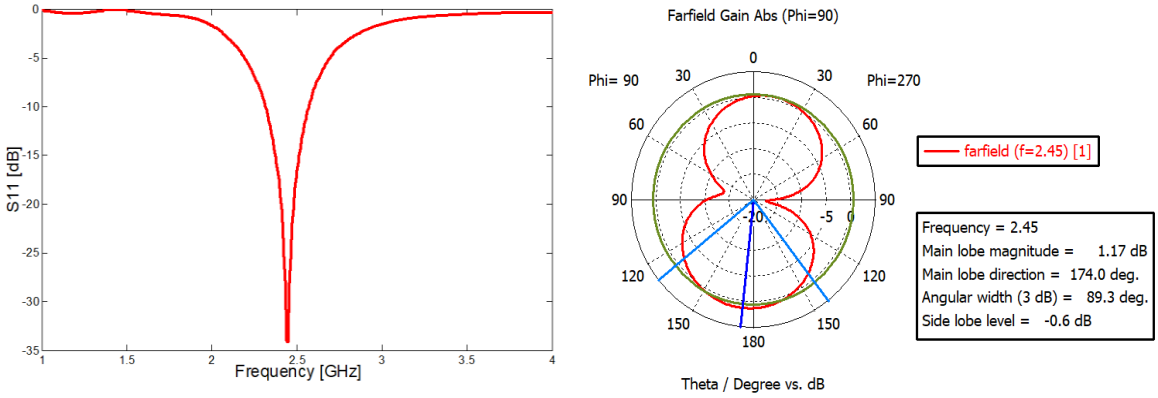


Fig. 7. Simulation results (a) Reflection coefficient versus frequency (b) Antenna 2D Radiation pattern in E-Plane at 2.45 GHz

5 Conclusion

This paper reports the design and simulation of a novel patch antenna with an integrated Defected Ground Structure (DGS) etched in the ground plan. This configuration leads us to ensure a good size reduction compared to a conventional antenna radiating at 2.45 GHz The CST Microwave Studio have been used to design, optimize and simulate the whole system and an FR-4 substrate was adopted. Interresting results have been ontained.

References

[1] "Surface mount RF Schottkybarrier diode HSMS-282X series," AgilentTechnol., Palo Alto, CA, Tech. Data, 2001.

[2] K. M. Z. Shams and M. Ali, "Wireless power transmission to a buried sensor in concrete "IEEE Sensors Journal, vol. 7, no. 12, pp. 1573–1577, 2007.

[3] W. C. Brown. The history of power transmission by radio waves. IEEE Trans. Microw. Theory Techn. Sep. 1984; MTT-32(9): 1230-1242.

[4] Jamal Zbitou, Mohamed Latrach "Hybrid Rectenna and Monolithic Integrated Zero-Bias Microwave Rectifier,"IEEE Transactions On Microwave Theory And Techniques, Vol. 54, No. 1, January 2006

[5] A. Taybi, A. Tajmouati, J. Zbitou, A. Errkik, M. Latrach, L. El Abdellaoui. A New Configuration of a High Output Voltage 2.45 GHz Rectifier for Wireless Power Transmission Applications. Telecommunication, Computing, Electronics and Control Journal (TELKOMNIKA). October 2018; 16(5): 1939-1946.

[6] Seung-Tae Khang, Dong-Jin Lee, In-Jun Hwang, Tae-Dong Yeo, and Jong-Won Yu, Microwave Power Transfer With Optimal Number of Rectenna Arrays for Midrange Applications. IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, VOL. 17, NO. 1, pp. 155-159 (2018).

[7] Kai-Fong Lee , Shing Lung Steven Yang, Ahmed A. Kishk. Dual- and Multiband U-Slot Patch Antennas. IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, VOL. 7, pp. 645-647 (2008)

[8] Fu-Jhuan Huang, Tzong-Chee Yo, Chien- Ming Lee, and Ching Hsing Luo, "Design of Circular Polarization Antenna With Harmonic Suppression for Rectenna Application," IEEE Antennas Wireless Propag. Lett., Vol. 11, pp. 592-595,2012.

[9] C. Y. D. Sim, M. H. Chang, and B. Y. Chen. Microstrip-fed ring slot antenna design with wideband harmonic suppression. IEEE Trans. Antennas Propag. Sep. 2014; 62(9): 4828-4832.

[10] Lin Dang, Zhen Ya Lei, Yong Jun Xie, Gao Li Ning, and Jun Fan. A Compact Microstrip Slot Triple-Band Antenna for WLAN/WiMAX Applications. IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, VOL. 9, 2010.

[11] Debatosh Guha, Senior, Chandrakanta Kumar, and Surendra Pal. Improved Cross-Polarization Characteristics of Circular Microstrip Antenna Employing Arc-Shaped Defected Ground Structure (DGS). IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, VOL. 8, 2009

[12] Chandrakanta Kumar, Senior Member, IEEE, Mahammad Intiyas Pasha, and Debatosh Guha, Senior Member, IEEE. Defected Ground Structure Integrated Microstrip Array Antenna for Improved Radiation Properties. IEEE Antennas and Wireless Propagation Letters. (2016).

[13] C. A. Balanis, "Antenna Theory, Analysis and Design,"John Wiley & Sons. Third Edition, New York, 2005.

[14] Computer Simulation Technolgy (CST) Microwave Studio. www.cst.com

[15] Debatosh Guha, Yahia M.M. Antar. MICROSTRIP AND PRINTED ANTENNAS NEW TRENDS, TECHNIQUES AND APPLICATIONS. John Wiley and sons Ltd.

Doubly Fed Induction Generator Wind Turbine Control Based on a Novel MRAS Observer

[1]Oluwaseun Simon Adekanle, 1M. Guisser, [1]E. Abdelmounim, [1]M. Aboulfatah, [1]H. Bahri,
[2]Adeola Balogun
{ adekanlesimon@gmail.com, balog975@yahoo.com }

[1]Laboratory of System Analysis and Information Processing, Université Hassan 1er Settat, Morocco
[2]Department of Electrical and Electronic Engineering University of Lagos, Lagos, Nigeria

Abstract. This paper proposes a novel Doubly Fed Induction Generator (DFIG) control approach using an optimized set of observers. Firstly, an Adaptive Disturbance Rejection Scheme (ADRS) Model Reference Adaptive System (MRAS) observer is applied to estimate generator speed. Secondly, a Fuzzy Logic MPPT controller is designed to optimized power conversion. Finally to achieve the overall sensor-less control, speed, current and DC-voltage control are achieved using a novel version of perturbation observer while taking into account the power loss in the system. The technique is evaluated for a 1.5MW DFIG under variable wind speed, measurement noise, grid voltage dip and parameter variation and compared with the conventional PI-controller.

Keywords: ADRS, Integral Backstepping Control, DFIG Control, MRAS, Fuzzy Logic, MPPT, Extended State Observer, Crowbar, DC-chopper.

1 Introduction

The purpose of Maximum power Point Tracking (MPPT) is to optimize the operation of the wind turbine to capture the maximum power possible. The first approach is the technique using relative velocity, which aims to control the speed of rotation of the generator to keep the relative speed at its optimum value [1]. This method requires a perfect knowledge of the parameters of the turbine and the measurement of the wind speed. Meanwhile, the use of a single anemometer leads to a local measure of wind speed that is not representative of the effective wind speed incident on the blades. Thus, an erroneous measurement of the wind speed leads to a degradation of the extracted power. To solve this problem of wind speed measurement, some researchers have worked on the estimation of wind speed using different techniques. Song et al in [2] proposed and compared two methods of wind speed estimation using different versions of the Kalman filter. The inconvenience of this methods resides in the fact that Kalman filters are complex and require high number of tuning parameters. In addition the stability of these estimator is not verified in close loop with controllers to achieve the complete control of the DFIG wind turbine. Belmokhtar et al in [3] developed a Fuzzy Logic based technique to estimate

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284107

wind speed and control the DFIG in close loop. However, stability analysis is not given. The second method for extracting the maximum wind power is by the Perturbation and Observation technique which is often used in the photovoltaic system [4]. It is easy to implement, less expensive and requires neither the knowledge of turbine parameters nor the measurement of wind speed [5]. However, this algorithm does not give good convergence because we have to make a compromise between the speed of convergence and performance. Other versions were then proposed such as neural network, gradient methods to give better performances but are still difficult to implement [1]. Optimal Torque Control (OTC) is a method which consists in adapting the electromagnetic torque (or power) to its optimum value without the need to measure wind speed. The disadvantage of this OTC technique is that characterization tests are required off-line to design the look-up table. In general, the OTC method is unsuitable for medium and large wind turbines because their inertia slows down the response during abrupt or rapid changes in wind speed [6]. We propose therefore a method based on Fuzzy Logic maximum power point search algorithm which is independent of wind speed measurement and which indirectly takes power loss in the system into consideration. It depends solely on the measurement of generator speed and output power.

Many controllers like PI-controller [7], Sliding Mode Controller [8], Backstepping [9], Model Predictive Control [10], and Fuzzy Logic Controller [11] have been successfully designed to control generator speed and currents in order to extract maximum power from the DFIG turbine through Maximum Power Point Tracking (MPPT) Techniques. They are equally employed to optimize the exchange of electrical energy between the turbine and the utility grid through Unity Power Factor (UPF) control. These controllers are synthesized assuming that all variables needed by the control algorithm are measurable from material and economic point of view. However, using mechanical, current and voltage sensors increases the cost of both equipment and maintenance, and reduces the reliability of the overall wind turbine system because they are prone to measurement noise and faults. Several techniques to observe one or more variables of the wind turbine system have been proposed in the literature. The authors in [12] proposed a Luemberger observer to estimate rotor currents from voltage, speed and power measurements. Only the RSPC control strategies is presented and the measurement of generator speed reduces the efficiency of the overall system. In [13] a modified PLL is employed to estimate generator speed but presented results show slow convergence of the estimator-controller. A Fuzzy-MRAS generator speed observer is elaborate in [3] but the stability and convergence of the Fuzzy adaptive mechanism with 25 rules are not proven.

In this paper an ADRT-based MRAS observer through modified Nonlinear Extended State Sliding Mode Observer and Lyapunov stability theory is proposed to observe generator speed. The observer takes into account only stator voltage measurement and observed rotor currents.

The control of the DFIG by stator flux orientation[14] stator voltage [15] or air gap flux [16] is a technique synthesized by assuming that the grid voltage is ideal and the stator flux is constant or vary very slowly. However, stator flux introduces high nonlinearity into the system during network fault. In fact, the variation of the stator flux during grid voltage drop depends on the value of the stator flux at the instance of occurrence and an exponential term [17]. Other authors like *Bhattarai et al.* [18] have proposed the use of stator and current measurements to estimate the flux in open loop. However, the accuracy of the flux estimation depends strongly on the machine parameters. If grid voltage dip occurs, the control laws cannot therefore respond directly and quickly to contain the disturbances generated by flux estimation incertitude and nonlinearity during fault moments. This causes degradation of the performance of these controllers during the voltage drop.

In this paper, an adaptive technique based on disturbance rejection is proposed to control the speed of the DFIG, rotor current-, filter current and the DC bus voltage. It consists of estimation followed by compensation, in real time, of network disturbances, parameter variation, modeling error and the decoupling terms. It is a simple and unique controller design method that has been proposed by Han [6] and improved by Gao [7]. This observer-controller has shown an apparent efficiency and a robust performance even without the exact knowledge of the system parameters.

2 ADRS control

Fig. 1 illustrates the complete proposed observer-based Fuzzy MPPT disturbance rejection control of the DFIG wind turbine. Fuzzy MPPT controller extracts optimal generator reference speed that leads to the maximization of extracted power. Two control objectives are intended to be carried out by the ADRS technique: tracking of the reference speed and desired stator reactive power. The MRAS observer estimates the actual generator speed from current measurement and supplies it to the controller. At the GSPC, the control objective is to keep the DC bus voltage constant at its nominal value and annul the reactive power generated by the filter.

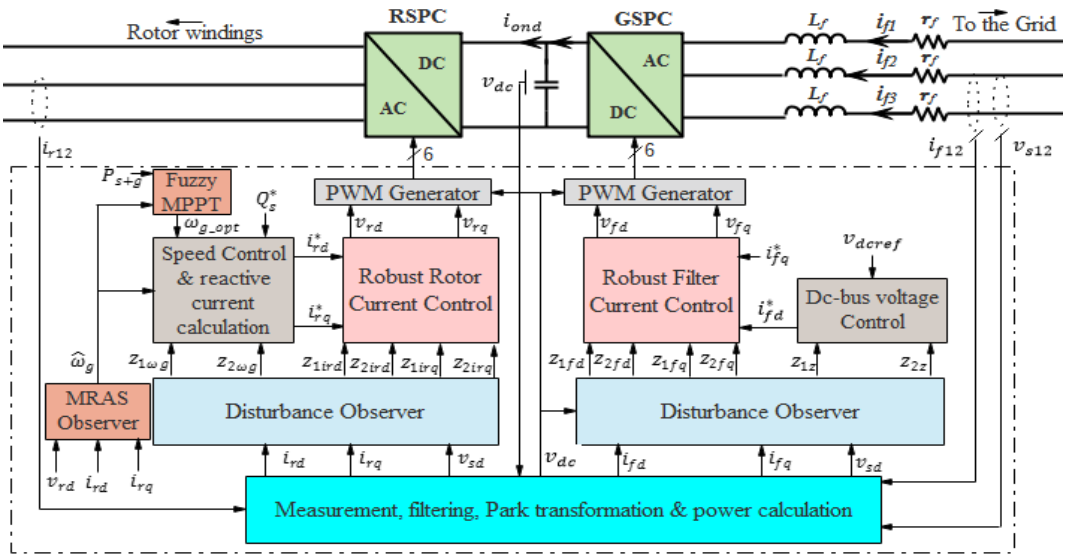


Figure 1 : Synoptic diagram of the overall control structure.

2.1 MRAS Speed Observer

The proposed MRAS speed estimator consists of a reference model (available variable) and an adaptive model (equation of the available model with $\widehat{\omega}_g$ as variable to be estimated) as illustrated in fig. 2. The error between both models is driven to zero by an adaptation mechanism based on a simple Adaptive Disturbance Rejection Scheme (ADRS). The output of the reference model v_{rd} is easily accessible at the output of RSPC controller. The adaptive model under stator voltage orientation control is obtained from voltage and flux equations [19] and expressed in (1).

$$\hat{v}_{rd} = \sigma L_r \frac{di_{rd}}{dt} + R_r i_{rd} - \sigma L_r (\omega_s - \hat{\omega}_g) i_{rq} - (\omega_s - \hat{\omega}_g) \frac{L_m \omega_s \varphi_{sq}}{L_s} \quad (1)$$

Fig. 3 describes the ADRS adaptive mechanism technique applied to reduce the error between the reference and adaptive models. If the parameters of the controller are carefully chosen, the error between the reference v_{rd} and adaptive $\hat{v}_{rd}$ models converges to zero and the controller output is the estimate of the DFIG rotational speed. The main advantage of using ADRS as adaptation mechanism is its robustness and high precision. This technique is chosen because it does not require knowledge of the dynamics of the adaptive model. It is based solely on the consideration of the adaptive model as having a general form described by (2), where θ is an unknown variable, b is a control constant and $\hat{\omega}_g$ is the variable to be observed.

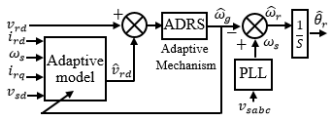


Figure 2 : MRAS speed observer technique.

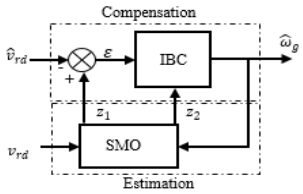


Figure 3 : MRAS Adaptive mechanism.

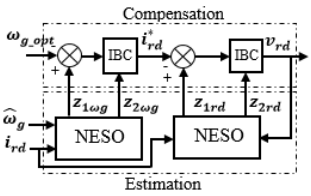


Figure 4 : Disturbance observer and compensation for speed and current control.

$$\frac{dv_{rd}}{dt} = b\hat{\omega}_g + \theta(\omega_s, \varphi_{sq}, i_{rd}, i_{rq}) \quad (2)$$

The first step in the convergence of $\hat{v}_{rd}$ to v_{rd} is then the determination of the unknown variable θ in order to be able to synthesize an appropriate control law. The solution then is to develop an observer to construct θ from v_{rd}. An observer for the system described by (2) is the optimized Sliding Mode Observer (SMO) presented in (3).

$$\begin{cases} \dot{z}_1 = z_2 + b\hat{\omega}_g + \rho_1 \left(|e|^{\alpha_1} sign(e) + |e|^{\beta_1} sign(e) \right) + k_1 sign(e) \\ \dot{z}_2 = \rho_2 \left(|e|^{\alpha_2} sign(e) + |e|^{\beta_2} sign(e) \right) + k_2 sign(e) \\ e = \hat{v}_{rd} - z_1 \end{cases} \quad (3)$$

where : z_1 and z_2 are the estimations of $\hat{v}_{rd}$ and θ respectively. The constants ρ_1, ρ_2, α_1, α_2, β_1, β_2, k_1, k_2 are the adjustment parameters of the observer. Equation (4) is substituted into (3) in order to avoid chattering.

$$sign(e) = \begin{cases} 2\left(\frac{1}{1 + exp^{-\tau e}} - \frac{1}{2}\right), & |e| \leq \delta \\ sign(e), & |e| > \delta \end{cases} \quad (4)$$

Once the estimate of the unknown variable is available, it remains to synthesize a control technique for the tracking of the reference model. After the observation, the dynamics of $\hat{v}_{rd}$ becomes (5). Application of Integral Backstepping Control thus gives (6).

The slip speed ω_r is calculated from the stator angular speed ω_s extracted by the PLL and the observed rotor speed $\widehat{\omega}_g$ as ($\omega_r = \omega_s - \widehat{\omega}_g$). A simple integral action on the observed rotor angular speed is applied to obtain the angle for Park transformation of stator variables.

$$\dot{z}_1 = z_2 + b\widehat{\omega}_g \tag{5}$$

$$\widehat{\omega}_g = \frac{1}{b}\left(-k\varepsilon - z_2 + \dot{v}_{\mathrm{rd}} - k_i(z_1 - v_{\mathrm{rd}})\right) \tag{6}$$

2.2 Non Linear Extended State Observer

Here we propose a real time estimation and compensation of the coupling terms as well as other disturbances that may occur in the system due to parameter variation, measurement noise or grid voltage dip. To proceed, we take the disturbance into account in speed, current and DC voltage dynamics as expressed in (7). The disturbance, in form of parametric incertitude, flux nonlinearity or modelling error, is represented as a lumped parameter $\emptyset$ in speed, rotor, filter and DC bus mathematical models. The lumped parameter, the coupling terms as well as the nonlinearity in the models are considered as a generalized disturbance ($z_{2\omega g}$, z_{2rd}, z_{2rq}, z_{2z}, z_{2fd}, z_{2fq}). Expressions of rotor, filter and DC bus models are hence reduced and presented in (7) where ($z_{1\omega g}$, z_{1rd}, z_{1rq}, z_{1z}, z_{1fd}, z_{1fq}) represent (ω_g, i_{rd}, i_{rq}, z, i_{fd}, i_{fq}).

$$\begin{cases} \dfrac{dz_{1\omega g}}{dt} = uz_{1rd} + z_{2\omega g} \\ \dfrac{dz_{1rd}}{dt} = b_r v_{rd} + z_{2rd} \\ \dfrac{dz_{1rq}}{dt} = b_r v_{rq} + z_{2rq} \end{cases} \qquad \begin{cases} \dfrac{dz_{1fd}}{dt} = b_f v_{fd} + z_{2fd} \\ \dfrac{dz_{1fq}}{dt} = b_f v_{fq} + z_{2fq} \\ \dfrac{dz_{1z}}{dt} = b_z z_{1fd} + z_{2z} \end{cases} \tag{7}$$

The generalized disturbances in form of extended states are:

$$\begin{cases} u = -\dfrac{3}{2}p\dfrac{L_m}{JL_s}\varphi_{sq} \\ z_{2\omega g} = \left(C_m - f_v\omega_g + \emptyset_{\omega g}\right)/J \\ z_{2rd} = -\dfrac{R_r i_{rd}}{\sigma L_r} + \omega_r i_{rq} + \dfrac{L_m\omega_r\varphi_{sq}}{\sigma L_r L_s} + \emptyset_{rd} \\ z_{2rq} = -\dfrac{R_r i_{rq}}{\sigma L_r} - \omega_r i_{rd} + \emptyset_{rq} \\ b_r = \dfrac{1}{\sigma L_r} \end{cases} \qquad \begin{cases} z_{2fd} = -\dfrac{R_f}{L_f}i_{fd} + \omega_s i_{fq} + \dfrac{v_{sd}}{L_f} + \emptyset_{fd} \\ z_{2fq} = -\dfrac{R_f}{L_f}i_{fq} - \omega_s i_{fd} + \emptyset_{fq} \\ z_{2z} = \left(-P_{l_{filtre}} - P_{l_{CPCR}} - v_{dc}.i_{ond}\right)/C \\ \quad +\emptyset_z \\ b_f = \dfrac{-1}{L_f} \\ b_z = \dfrac{3v_{sd}}{2C} \end{cases}$$

2.2.a Disturbance Observers

A Nonlinear Extended State Observer (NESO) for the system (dynamiccs of $z_{1\mathrm{rd}}$ for example) describe by (7) can be expressed as:

$$\begin{cases} \dot{\hat{z}}_{1\mathrm{rd}} = \hat{z}_{2\mathrm{rd}} + bv_{rd} + \rho_1\left(|e|^{\alpha_1}sign(e) + |e|^{\beta_1}sign(e)\right) + k_1 sign(e) \\ \dot{\hat{z}}_{2\mathrm{rd}} = \rho_2\left(|e|^{\alpha_2}sign(e) + |e|^{\beta_2}sign(e)\right) + k_2 sign(e) \\ e = z_{1\mathrm{rd}} - \hat{z}_{1\mathrm{rd}} \end{cases} \quad (8)$$

If the observer parameters are judiciously chosen, e converges to zero therefore $\hat{z}_{1\mathrm{rd}}$ and $\hat{z}_{2\mathrm{rd}}$ are estimates of $z_{1\mathrm{rd}}$ and $z_{2\mathrm{rd}}$ respectively. In order to reduce the chattering phenomenon of the Sliding Mode Observer caused by the sign function, a super twisting algorithm is substituted into (8) where the sign function is replaced with the expression in (9). τ is a parameter inversely proportional to δ.

$$sign(e) = \begin{cases} 2\left(\frac{1}{1+exp^{-\tau e}} - \frac{1}{2}\right), & |e| \le \delta \\ sign(e), & |e| > \delta \end{cases} \quad (9)$$

A careful study of the observer in (8) by [20] proposed a technique for determining the parameters of the NESO and it is applied in this paper as follows $1 < \rho_1, \rho_2 < +\infty$, $0.5 < \alpha_1 < 1$, $\alpha_2 = 2\alpha_1 - 1$, $\beta_1 = 1/\alpha_1$, $\beta_2 = \beta_1 + (\alpha_1 - 1)$ and $k_2, k_2 > 0$.

2.2.b Disturbance compensation

A simple error feedback control can be used to converge the current to its reference value as used in [20]–[22]. The disadvantage of this method is that the observation error is not taken into account in the command. This increases the convergence time or introduces a static error between the current and its desired value. To improve on this, we will use a more robust control technique that will ensure quick convergence to desired values.
Reference speed value ω_{g_opt} is computed online by the Fuzzy MPPT controller and its tracking by the controller gives the d-axis rotor current reference i_{rd}^*. The desired DC-bus voltage V_{dcref} is the nominal value of 1150V (1 p.u) and its tracking by the controller gives the d-axis filter current reference i_{rq}^*. This consequently implies that $z_{ref} = V_{dcref}^2/2 = 0.5\mathrm{p.u}$ due to the change of variable between V_{dc} and z.

Step 1: to control the speed of rotation of the generator and the DC bus voltage to follow the optimal rotational speed ω_{g_opt} generated by the MPPT command and the reference voltage z_{ref} of the DC bus respectively, we define the tracking errors by (10).

$$\varepsilon_{\omega g} = \hat{z}_{1\omega g} - \omega_{g_opt} + k_{i\omega}\int\left(\hat{z}_{1\omega g} - \omega_{g_opt}\right)dt \quad (10)$$

$$\varepsilon_{rd} = \hat{z}_{1rd} - i_{rd}^* \quad (11)$$

For easy readability, we replace i_{rd}^* and $i_{fd}^*$with α_{rd} and α_{fd} respectively. Following the Integral Backstepping controller method procedures presented in [19], the final control variable that guarsnttes the convergence of the speed to its optimal value ω_{g_opt} is:

$$v_{rd} = \frac{1}{b_r}\left(-k_{rd}\varepsilon_{rd} - u\varepsilon_{\omega g} - \hat{z}_{2rd} + \dot{\alpha}_{rd}\right) \quad (12)$$

Fig. 4 illustrates the technique of disturbance observation and compensation by the Extended State Observer and the Integral Backstepping respectively for rotational speed tracking.

3 Simulation results

To verify the performance of the proposed technique (observer-controller in close-loop), we will present the simulation results under measurement noise and network voltage drop. Fig. 5 shows the wind speed profile with which the turbine blades are driven.

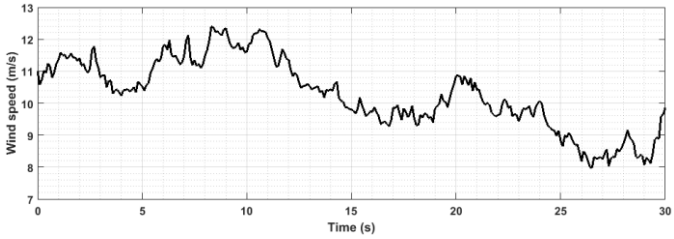


Figure 5 : Wind speed profile.

3.1 Simulation under Measurement Noise

In this paper, the effectiveness of the proposed technique is evaluated in the presence of measurement noises often introduced into the system by sensors. To achieve this, noise has been added, without filtering, to the measured variables that are not observed in this work. The distorted measurements (fig. 6) are then introduced to the control algorithm to synthesize the controller, the observers and the Fuzzy MPPT. For three-phase variables, only the first phases are presented for clarity but in simulation, the same noise is added to all the three phases.

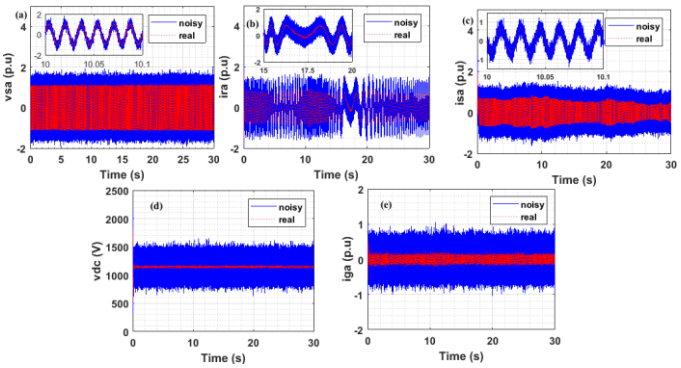


Figure 6 : System inputs with noise introduced.

Fig. 7.a shows the rotational speed, its estimate and the reference generated by the Fuzzy Logic MPPT controller. It is clear that despite the noisy inputs, the adaptive speed observer (MRAS), using the ADRS technique as an adaptation mechanism, rejects the noise and observes the speed with very good accuracy. We observe the same robustness on the normalized DC bus voltage (Fig. 7.b). The noises are rejected and the real state converges towards the reference (0.5 p.u). Fig. 7.c-f present the measured currents (currents coming from the noisy three-phase currents),

reference currents and the currents actually transmitted to the control algorithm by the observer. It is obvious that the disturbance observer rejects the noise and eliminates its effects on the control as the noisy currents are filtered by the observer. Reference currents are also noisy because they are generated from noisy quantities. However, the currents track their references while at the same time rejecting the noises.

Fig. 8 shows the output power, power coefficient and output currents of the wind turbine. We note that the active and reactive powers evolve without being disturbed by measurement noise. Likewise, the power coefficient evolves around the maximum possible efficiency of 50%. The currents follow wind speed profile without any overshoot above the nominal value of 1 p.u.

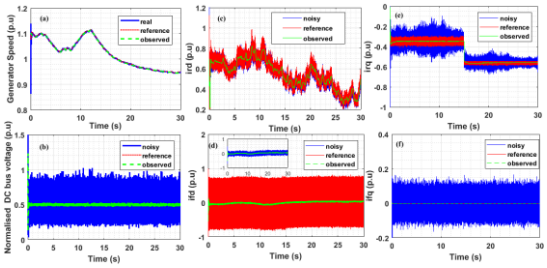


Figure 7 : Influence of observers on measurement noise

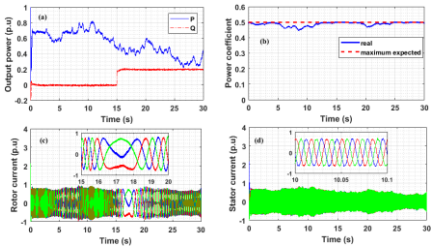


Figure 8 : DFIG output under noisy sensors.

3.2 Simulation under Deep Grid Voltage Dips

Simulations in this section are performed to verify and compare the capability of the proposed ADRS and PI-controller to satisfy the strictest scenario of LVRT requirement of the German grid code. According to the code, wind turbine generators are required to remain connected to the grid for at least 0.625s at the occurrence of grid voltage dip of up to 85%. A three-phase short-circuit fault is introduced to the grid at at the point of DFIG connection to the grid. The three phases of the grid voltage drop to 5% (95% dip) of their nominal values at t=10s and the dip lasts till t=10.625s.

Two DFIG protection schemes are included in the DFIG set-up. The first is the active crowbar which is activated when rotor current or DC bus voltage exceed their respective thresholds to limit inrush rotor current during fault. The second is the DC-chopper connected in parallel with the DC-link capacitor to limit DC voltage swell. It is activated at the same instance as the active crowbar. The deactivation of both protection circuits occurs simultaneously when both DC voltage and rotor current return below their thresholds. 1.35p.u and 1.2p.u are chosen as thresholds for DC voltage and rotor current respectively.

Fig. 9 shows that grid voltage is the same under both controllers, dips to 0.05p.u and its recovery is rapid with neither overshoot nor oscillations.

According to fig. 10.a, stator current overshoot poses no threat as it is limited below 1p.u at the instance of voltage rising under both controllers. However, significant stator current surge of 125% can be observed during voltage falling under PI-controller, which far exceeds admissible stator current. Under ADRS technique, stator current is limited around its admissible value of 1.2p.u during voltage rising.

In fig. 10.b, rotor current surge and settling are admissible during grid voltage recovery under both controllers as it is limited within the safe operating zone of the back-to-back converters. Unfortunately, rotor current evolution during voltage falling is identical to that of the stator current, it surges far above the save operating limit of the power converters up to 2.75p.u for the

PI-controller. The proposed approach using ADRS controller was able to keep the inrush current at 1.25p.u around the save operating area of the converters.
Comparing the DC voltage under PI and ADRS controllers, fig. 11 shows that the DC bus experienced a voltage swell during the network voltage drop under both controllers. Under control by the PI technique, the overvoltage exceeds 1.35p.u admissible for the safety of the DC bus and reaches 1.75p.u whereas under the control by ADRS, the surge reaches only 1.3p.u, which is just below the admissible voltage across the DC capacitor.

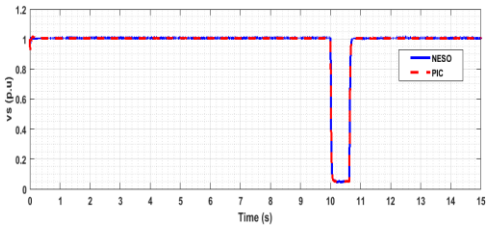


Figure 9: Grid voltage under short-circuit fault.

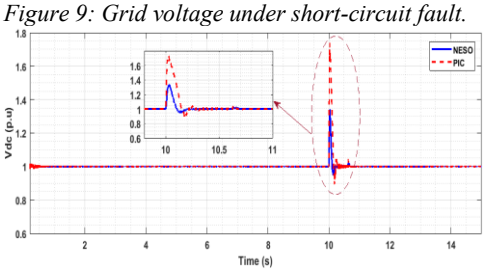


Figure 11 : DC-bus voltage under grid voltage dip.

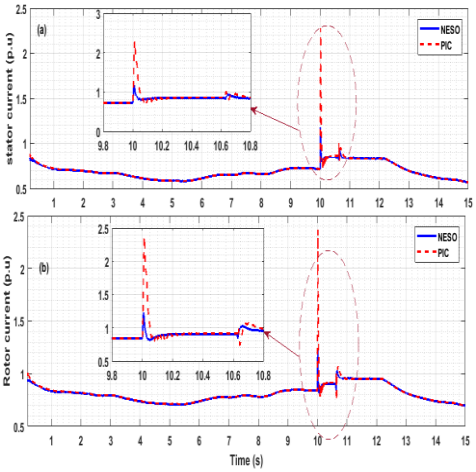


Figure 10: Stator and rotor currents under grid voltage dip.

Conclusion

The application of a novel MRAS observers in this paper does not only reduce the cost of installation, it also increases the reliability of wind turbine installations, eliminates measurement noise and reduces the complexity of control circuit setup

A precise measurement of the decoupling terms of the direct and quadrature components of the rotor currents and the filter is necessary for the good performance of the wind turbine control techniques. We have estimated decoupling terms by treating them as another state of the system. An IBC based on Lyapunov stability theory taking into account the integral of the error is then used for tracking current references.

The robustness the ADRS to measurement noise and grid fault has been demonstrated. The 1.5MW wind turbine connected to the infinite network is controlled by the PI and ADRS controllers. The comparative results under MATLAB / SIMULINK show that the ADRS is better able to reduce wind turbine rotor and stator overcurrent, thus increasing the safety level of bidirectional converters with respect to overcurrent caused by network voltage drop. In addition, the DC bus voltage is limited below its safe operating limit under control by ADRS, protecting the DC bus from overvoltage.

References

[1] M. A. Abdullah, A. H. M. Yatim, C. W. Tan, and R. Saidur, "A review of maximum power point tracking algorithms for wind energy systems," *Renew. Sustain. Energy Rev.*, vol. 16, no. 5, pp. 3220–3227, Jun. 2012.

[2] D. Song, J. Yang, M. Dong, and Y. H. Joo, “Kalman filter-based wind speed estimation for wind turbine control,” *Int. J. Control. Autom. Syst.*, vol. 15, no. 3, pp. 1089–1096, 2017.
[3] K. Belmokhtar, M. L. Doumbia, and K. Agbossou, “Novel fuzzy logic based sensorless maximum power point tracking strategy for wind turbine systems driven DFIG (doubly-fed induction generator),” *Energy*, vol. 76, pp. 679–693, 2014.
[4] C.-M. Hong, T.-C. Ou, and K.-H. Lu, “Development of intelligent MPPT (maximum power point tracking) control for a grid-connected hybrid power generation system,” *Energy*, vol. 50, pp. 270–279, Feb. 2013.
[5] L. Qu and W. Qiao, “Constant Power Control of DFIG Wind Turbines With Supercapacitor Energy Storage,” *IEEE Trans. Ind. Appl.*, vol. 47, no. 1, pp. 359–367, Jan. 2011.
[6] J. S. Thongam, M. Tarbouchi, R. Beguenane, A. F. Okou, A. Merabet, and P. Bouchard, “An optimum speed MPPT controller for variable speed PMSG wind energy conversion systems,” in *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, 2012, pp. 4293–4297.
[7] J. Ouyang, T. Tang, Y. Diao, M. Li, and J. Yao, “Control method of doubly fed wind turbine for wind speed variation based on dynamic constraints of reactive power,” *IET Renew. Power Gener.*, 2018.
[8] F. Zargham and A. H. Mazinan, “Super-twisting sliding mode control approach with its application to wind turbine systems,” *Energy Syst.*, no. 209, 2018.
[9] B. Bossoufi, M. Karim, A. Lagrioui, M. Taoussi, and A. Derouich, “Observer backstepping control of DFIG-Generators for wind turbines variable-speed: FPGA-based implementation,” *Renew. Energy*, vol. 81, pp. 903–917, 2015.
[10] A. Bektache and B. Boukhezzar, “Nonlinear predictive control of a DFIG-based wind turbine for power capture optimization,” *Int. J. Electr. Power Energy Syst.*, vol. 101, pp. 92–102, Oct. 2018.
[11] W. Ayrir, M. Ourahou, B. El Hassouni, and A. Haddi, “Direct torque control improvement of a variable speed DFIG based on a fuzzy inference system,” *Math. Comput. Simul.*, May 2018.
[12] R. Bhattarai, N. Gurung, A. Thakallapelli, and S. Kamalasadan, “Reduced-Order State Observer-Based Feedback Control Methodologies for Doubly Fed Induction Machine,” *IEEE Trans. Ind. Appl.*, vol. 54, no. 3, pp. 2845–2856, 2018.
[13] A. Giannakis, A. Karlis, and Y. L. Karnavas, “A combined control strategy of a DFIG based on a sensorless power control through modified phase-locked loop and fuzzy logic controllers,” *Renew. Energy*, vol. 121, pp. 489–501, Jun. 2018.
[14] N. Sarma, P. M. Tuohy, J. M. Apsley, Y. Wang, and S. Djurović, “DFIG stator flux-oriented control scheme execution for test facilities utilising commercial converters,” *IET Renew. Power Gener.*, vol. 12, no. 12, pp. 1366–1374, 2018.
[15] G. Calderón, J. Mina, J. C. Rosas-Caro, M. Madrigal, A. Claudio, and A. López, “Simulation and Comparative Analysis of DFIG-based WECS Using Stator Voltage and Stator Flux Reference Frames,” *IEEE Lat. Am. Trans.*, vol. 15, no. 6, pp. 1052–1059, 2017.
[16] S. Jing, N. Xiaohong, and L. Kean, “Air gap field-oriented vector control strategy for high-power electrically excited synchronous motor based on full-order flux linkage observer,” *Sci. Res. Essays*, vol. 9, no. 9, pp. 367–373, May 2014.
[17] M. Rahimi and M. Parniani, “Grid-fault ride-through analysis and control of wind turbines with doubly fed induction generators,” *Electr. Power Syst. Res.*, vol. 80, no. 2, pp. 184–195, Feb. 2010.
[18] R. Bhattarai, S. Member, N. Gurung, S. Member, and S. Ghosh, “Parametrically Robust Dynamic Speed Estimation Based Control for Doubly Fed Induction Generator,” *IEEE Trans. Ind. Appl.*, vol. PP, no. c, pp. 1–8, 2017.
[19] O. S. Adekanle, M. Guisser, E. Abdelmounim, and M. Aboulfatah, “Robust Integral Backstepping Control of Doubly Fed Induction Generator Under Parameter Variation,” *Int. Rev. Autom. Control*, vol. 10, no. 2, 2017.
[20] S. Xiong, W. Wang, X. Liu, Z. Chen, and S. Wang, “A novel extended state observer,” *ISA Trans.*, vol. 58, pp. 309–317, 2015.
[21] S. Gao, C. Mao, D. Wang, and J. Lu, “Dynamic performance improvement of DFIG-based WT using NADRC current regulators,” *Int. J. Electr. Power Energy Syst.*, vol. 82, pp. 363–372, 2016.

[22] Z. Song, T. Shi, C. Xia, and W. Chen, “A novel adaptive control scheme for dynamic performance improvement of DFIG-Based wind turbines,” *Energy*, vol. 38, no. 1, pp. 104–117, 2012.

A New Modeling of the Junction Metal Semiconductor for Rectenna System

Abdelhak Hamid allah[1], Ahmed Errkik[2], Jamal Zbitou[3] ,Larbi El Abdellaoui[4],Abdelali Tajamouati[5],Mohamed Latrach[6]
{abdelhakhamidallah@gmail.com[1], ahmed.errkik@uhp.ac.ma[2], zbitouj@gmail.com[3]}

LMEET Laboratory FST of Settat Hassan 1st University Morocco[1,2,3,4,5], Microwave group ESEO Angers France[6]

Abstract. This article presents a new design and optimization of the junction Metal-Semiconductor (Au/n-GaAs) optimized in performances.This study has as objective the design of a new Schottky diode which is dedicated to the design of new RF-DC rectifiers in order to increase and to improve the performances of rectenna, and the microwave applications in general. The modelling is based initially on the determination of the breakdown voltage and then the calculation and optimization of all the principal parameters of the Schottky diode. For validation, the proposed study is compared with the commercial diodes.

Keywords: Junction Metal-Semiconductor, GaAs Schottky Diode, Schottky Barrier, Rectenna, Diode Modeling .

1 Introduction

The semiconductor devices are in the foreground of modern technology. Each day, the powerful semiconductor devices are developed and contribute to the spectacular results got in various spheres of the scientific and technology activity. The structures metal semiconductors (Schottky diode) are components widely used in microelectronics and microwaves applications for the manufacturing of integrated circuits on one hand and as a basic competent in commutation, on the other hand. In this context the Schottky diode is the most used element for the manufacturing and the design of the rectifier rectenna. Among the semiconductor with direct band gap and higher electronic mobility, the gallium arsenide (GaAs) [1], [2]. It's a semiconductor which starts to become a serious competitor for silicon (Si). Moreover, it can easily be doped n or p, and its operating temperature reaches at least 600°C-700°C [3], [4]. Indeed, the characteristic of the metal-semiconductor junction is nonlinear when it functions at high powers and its performances on the conversion operation are very limited. The objective is the design of a diode Au/n-GaAs Zero biased, which works with low, high power levels and having a good detection sensitivity and good performance in terms of rectifying operation. The study begins with the fixation of a breakdown voltage which is suitable for high levels of power, then the optimization of the parameters and properties of the diode using the methodology detailed in **Figure 1** and in the following section.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284081

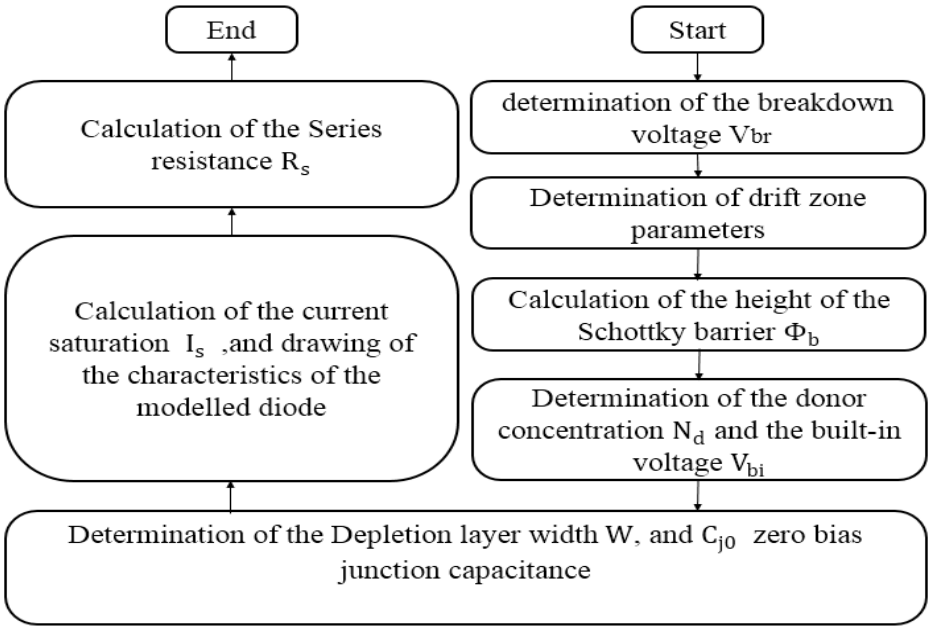


Fig. 1. Diagram of the calculation and modeling carried out in Matlab.

2 Modeling and optimized parameters

According to [5] a study conducted to evaluate the conversion efficiency of a rectenna. The breakdown voltage V_{br} , the zero-biased junction capacitance C_{j0} and the series resistance R_s Are the principal elements which determine the efficiency of RF-DC conversion for a rectenna and the global performances of commercial diodes, those parameters are related to the properties of selected materials [1]. In addition, the maximum power of a rectenna or the critical input power where the output voltage DC becomes limited, is given by the following equation $P_{RF} = \frac{{V_{br}}^2}{4R_L}$. Therefore, this power is limited by the value of the breakdown voltage thus also limiting the efficiency and the performance of the rectenna. For this reason, we decided to begin this modeling with the determination of the breakdown voltage V_{br} and then to extract and modeled all principal parameters of the diode.

2.1 Determination of drift region parameters

The breakdown voltage we had chosen for this study is $V_{br} = 100\,V$. According to [6] the relationships between the breakdown voltage, the drift region width, and drift region concentration or (drift region doping) are expressed by the equations (1) and (2).

$$N_d{}^+ = 2 \times 10^{18} \times (V_{br}{}^{-4/3})\ cm^{-3} \quad (1)$$
$$W_D = 2.67 \times 10^{10} \times (N_d{}^+)^{-7/8}\ cm \quad (2)$$

With $N_d{}^+$ is the drift region concentration and W_D is the drift region width, to see **Figure 2**.

Fig. 2. Structure metal semiconductor.

The value of the breakdown voltage decreases notably when the concentration increases, and conversely for the width the voltage increases. The variation between dimensions of the drift region and the breakdown voltage is traced in **Figure 3.a** and **Figure 3.b**.

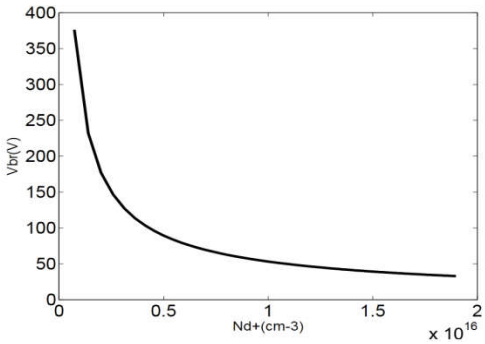


Fig. 3.a. Variation of the breakdown voltage V_{br} vs. Drift region concentration $N_d{}^+$.

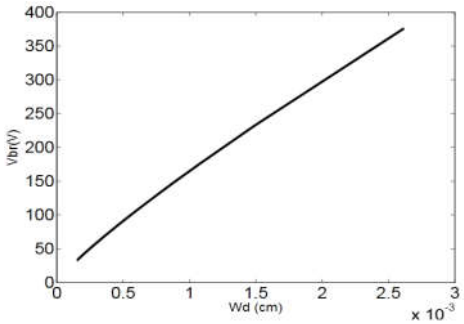


Fig. 3.b. Variation of the breakdown voltage V_{br} vs. Drift region concentration $N_d{}^+$.

2.2 Schottky barrier height Φ_b

The barrier height can be expressed as (3)

$$\Phi_b = \Phi_m - X_s \quad (3)$$

With Φ_m work function of the metal (gold for our study), and X_s is the electronic affinity of semiconductor (GaAs).

For our study, the work function of Au is $\Phi_m = 5.06$ eV [7],[8], and the electronic affinity of GaAs is $X_s = 4.07$ eV[1], therefore the barrier height is $\Phi_b = 0.98$ eV. **Figure 4** is traced using the equation $\Phi_b = 0.12\Phi_m + 0.37$ [7], representing the variation between a barrier height and a work function.

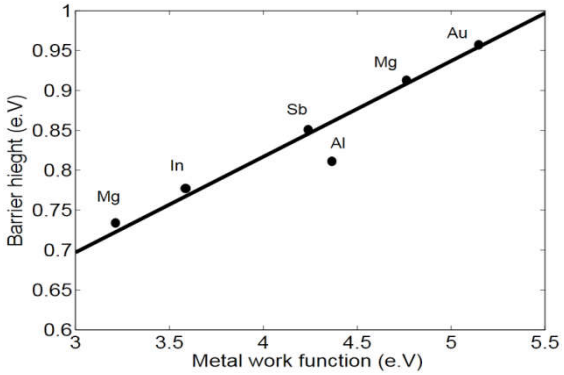


Fig. 4. Barrier height values for different metals plotted against the metal work function.

2.3 Donor concentration N_d and the built-in voltage V_{bi}

To calculate the built-in voltage, we substitute the equations (3) and (5) into (4), the curve that models the evolution of V_{bi} is given in the **Figure 5.a**. Regarding the breakdown voltage V_{br}, we replace the relations (4), (5) and (6) in the equation (7). The curve in **Figure 5.b** shows the variation between the donor concentration N_d and voltage V_{br}, we note we have a high donor concentration the breakdown voltage decreases [1], [9], [10].

$$V_{bi} = \Phi_b - \frac{KT}{q} \times \ln\left(\frac{N_c}{N_d}\right) \tag{4}$$

$$N_c = 4.7 \times 10^{17} \times \left(\frac{T}{300}\right)^{3/2} \tag{5}$$

$$E_{cr} = \frac{2.49 \times 10^6}{1 - 0.25 \times \log(N_d \times 10^{-16})} \tag{6}$$

$$V_{br} = \frac{\varepsilon_s \times (E_{cr})^2}{2qN_d - V_{bi}} \tag{7}$$

Where V_{bi} is the built-in voltage and, N_c the density of states at the conduction band edge, E_{cr} is the critical electric field, ε_s is the permittivity of the semiconductor.

The donor concentration and the built-in voltage we got for $V_{br} = 100\ V$, Are $N_d = 5 \times 10^{22} m^{-3}$, and $V_{bi} = 0.99\ V$.

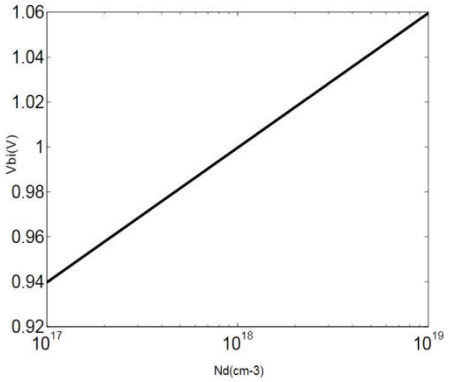


Fig. 5.a Variation of the built-in voltage V_{bi} vs.Donor concentration N_d.

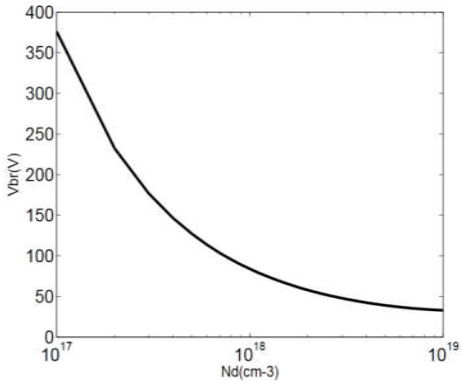


Fig. 5.b Variation of breakdown voltage V_{br} vs. Donor concentration N_d

2.4 Depletion layer width W, and C_{j0} zero bias junction capacitance

The equation to calculate W is given in the following form (8).

$$W = \sqrt{\frac{2\varepsilon_s}{qN_d}(V_{bi} - V - kT/q)} \quad (8)$$

Where V is the applied bias, q the elictronic charge, T is the temperature (K), with $kT/q \approx 26$ mV for T=25°C.?

The results of the variation of depletion layer width W versus donor concentration and the applied bias are given within the **Figure 6**. The condition to respect between depletion region width W and diode thickness d is defined by $W < d$. The presence of the depletion region generates a static capacity C_{j0} or (zero bias junction capacitance) is defined by the relation (9)

$$C_{j0} = \sqrt{\frac{q\varepsilon_s N_d S^2}{2V_{bi}}} \quad (9)$$

2.5 The saturation current current I_s , (I-V) characteristics of the modeled diode

The saturation current is expressed by the equation (10). On the other hand, in the bibliographical support [7] the ideality factor or the emission constant is $n = 1.01$ for a metal function work $\Phi_m = 5.06$. Concerning how to trace the characteristic (I-V-T), we use the Shockley equation (11). The **Figure 7** represents the characteristic in forward bias and under linear scale.

$$I_s = S \times A \times T^2 \times \left[\exp\left(\frac{-q\Phi_b}{KT}\right)\right] \quad (10)$$

With S the diode area, $A = 8.2 \times 10^4 A/m^2K^2$ the effective Richardson constant $K = 38 \times 0.6285\ 10^{-4}$ is Boltzmann's constant.

$$I = I_s \times \left[\exp\left(\frac{qV}{nKT}\right) - 1\right] \quad (11)$$

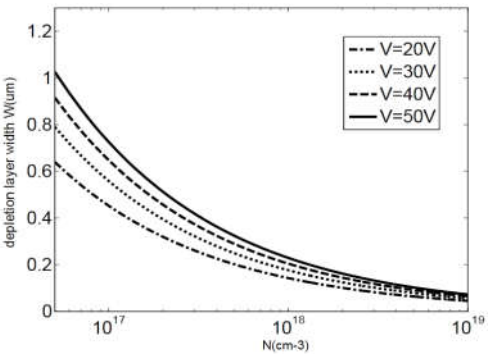


Fig. 6. Depletion layer width W vs. Donor concentration N_d, and V.

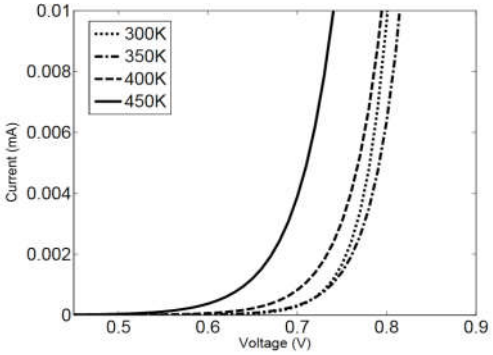


Fig. 7. Linear scale characteristic (I-V-T).

2.6 Series resistance R_s

The series resistance consists of three resistances $R_s = R_g + R_{ch} + R_c$ where R_g, R_{ch}, R_c are respectively the junction resistance, the canal resistance, and contact resistance; the three resistances are given by equations (12), (13), (14). After the calculation and the simplification between all equations, the value obtained for the series resistance is $R_s = 4.6\ \Omega$, the variation of the latter is shown in **Figure 8**.

$$R_g = \frac{\rho_m D}{3tW} \quad (12)$$

$$Rch = \frac{R_1(R_2 + R_3)}{R_1 + R_2 + R_3} \quad (13)$$

$$\begin{cases} R_1 = \dfrac{1100L}{aN_d{}^{0.82}A} & (13a) \\ R_2 = \dfrac{1100L}{a^+(N_d{}^+)^{0.82}A} & (13b) \\ R_3 = \dfrac{1100La}{N_d{}^{0.82}W_D} & (13c) \end{cases}$$

$$R_c = \frac{2100}{P \times \sqrt{(a^+)}(N_d{}^+)^{0.66}} \quad (14)$$

$$P = 2D + 2L + W \quad (14a)$$

Where $\rho_m = 2.5 \times 10^{-8}\Omega.m$ [4] is the metal resistivity, W_D is the drift region width $N_d{}^+$ is the drift region concentration; all the other parameters are detailed in **Figure 9.a** and **Figure 9.b**.

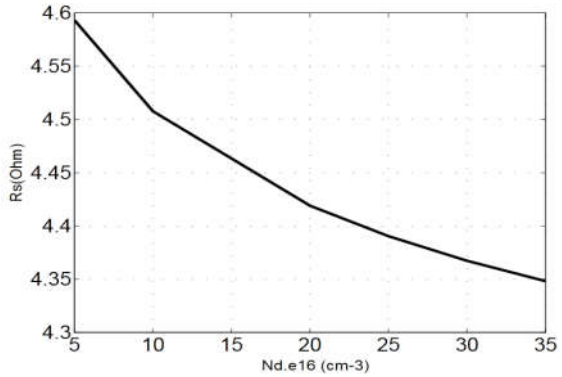


Fig. 8. Variation of series resistance R_s vs. Donor concentration N_d.

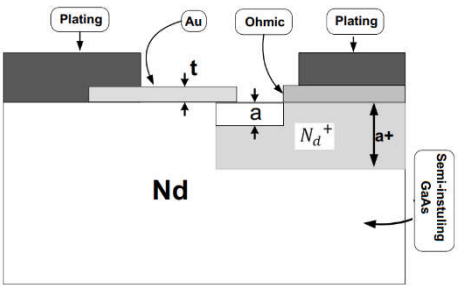


Fig. 9.a.

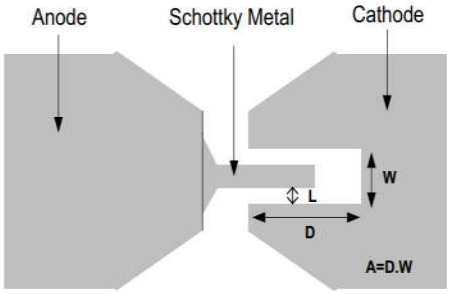


Fig. 9.b.

Fig. 9. a and **b** : the structure of metal-semiconductor detailed.

2.7 Electron mobility μ and cutoff frequency f_c

The electron mobility is a concept used in physics to characterize the conducting mediums of the electric current; the gallium arsenide is a semiconductor with higher electronic mobility the maximum electronic mobility for GaAs is approximately $\mu_{max} = 9200\ cm^2/V.s$ at $T = 300K$. According to the supports [11], [12], there exist two equations to calculate the electronic mobility (15), (16). The variation of the mobility versus of temperature and donor concentration is presented in **Figure 10.**

$$\mu(N_d) = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + (N_d/Nref)^{\sigma}} \quad (15)$$

$$\mu(T) = \mu(300K) \times \left(\frac{300K}{T}\right)^{\alpha} \quad (16)$$

Where μ_{min} , μ_{max} maximum and minimum mobility $Nref = 1.26 \times 10^{17} cm^{-3}$[1]. Reference doping concentration of GaAs, $\sigma = 2.15, \alpha = 0.61$ are two constants.

For traditional high-frequency model (**Figure 11**) of a Schottky diode, the cutoff frequency is given by the equation (17). For this study we find $f_c = 55$ GHz.

$$f_c = \frac{1}{2\pi R_s C_{j0}} \quad (17)$$

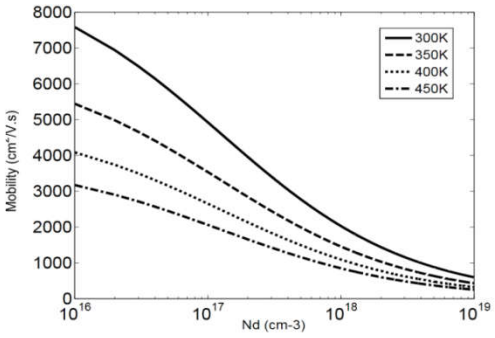


Fig. 10. Variation of electronic mobility vs. Temperature T and donor concentration N_d.

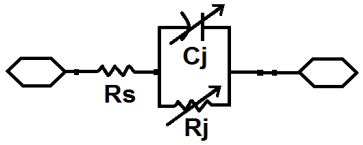


Fig. 11. Traditional model of a Schottky diode with R_s (series resistance), C_j (junction capacitance), R_j (junction resistance).

Table 1. Comparison between our work and the commercial diode used in other works.

Diode	Reference	V_{br} (V)	R_s (Ω)	C_{j0} (pF)	I_s (A)
BAT-03	[13]	4.6	5.0	0.1385	$13 .10^{-8}$
KA-BAND	[5]	9.0	4.85	0.13	$5.24 .10^{-15}$
HSMS-2860	[14]	7.0	5.0	0.18	$5 .10^{-8}$
HSMS-2820	[15]	15.0	6.0	0.70	$2.2 .10^{-8}$
HSMS-2850	[16]	7.0	25.0	0.18	$5.0 .10^{-8}$
MA4E1317	[17]	7.0	4.0	0.02	-
Optimized Diode	This work	100	4.6	0.6291	$6 .10^{-15}$

Table 2. Parameters of optimized diode.

Parameters	Symbol	Value	Unit
Breakdown voltage	V_{br}	100	V
Series resistance	R_s	4.6	Ω
Zero bias capacitance	C_{j0}	0.6291	pF
Saturation current	I_s	6	fA
Built-in voltage	V_{bi}	0.9	V
Barrier Height	Φ_b	0.98	eV
Ideality factor	n	1.01	-
Cutoff frequency	f_c	55	GHz

3 Conclusion

We showed through this study a new modeling of metal semiconductor in Au/n-GaAs technology, in which we calculated all the principal parameters of the junction. Compared to the parameters of commercialized Schottky diode reported in previous work (**Table 1**), we can conclude that the modeled diode (**Table 2**) is more performing in several parameters and more

precisely in the parameters which determine the efficiency of diode, and this was the global goal of our study. Moreover, the proposed study has a several advantages, firstly this study can be used as the guideline for diode conception of a rectifier design, for microwaves applications. Secondly, the study has enabled us to conceive adaptable diodes to our application. Finally, the study allows a high efficiency of microwave applications on which the diodes are the principal element like a rectenna.

References

[1] S. M. Sze.: Physics of Semiconductor Devices. John Wiley & sons, New York. London. Sydney. Toronto.1969.

[2] Sheng S. LI.: Semiconductor Physical Electronics. 2nded, Springer, USA, 1969.

[3] Shur M.: Gallium Arsenide versus Silicon — Applications and Modeling. In: Snowden C.M. (eds) Semiconductor Device Modeling. Springer, London,1989.

[4] Winfried M.: Electronic properties of semiconductor interfaces. Springer-Verlag Berlin Heidelberg ,2004.

[5] Yoo, T.-W., and K. Chang.: Theoretical and Experimental Development of 10 and 35 GHz Rectennas. IEEE Transactions on Microwave Theory and Techniques 40(6): 1259–1266.

[6]Baliga, B. Jayant, 2009.: Chapter 2 Advanced Power Rectifier Concepts. New York: Springer.

[7] Tyagi, M. S., 1977,: Surface States and Barrier Height at Metal-GaAs Interface. Japanese Journal of Applied Physics 16(S1): 333.

[8] Sachtler, W.M.H., G.J.H. Dorgelo, and A.A. Holscher, 1966.: The Work Function of Gold. Surface Science 5(2): 221–229.

[9] Seong-Jin Kim, 2006.: Breakdown Voltage Characteristics of SiCSchottky Barrier Diode with Aluminum Deposition Edge Termination Structure. Journal of the Korean Physical Society, Vol. 49, December 2006, pp. S768-S773.

[10] Hudait, M.K, and S.B Krupanidhi, 2001.: Doping Dependence of the Barrier Height and Ideality Factor of Au/n-GaAs Schottky Diodes at Low Temperatures .Physica B: Condensed Matter 307(1–4): 125–137.

[11] Christou, A., and W.M. Webb, 2006.: GaAs MMIC Reliability - High Temperature Behavior. Center for Risk and Reliability. https://books.google.co.ma/books?id=3jqGxfdr_04C.

[12] Shur, Michael, Sergey L. Rumyantsev, and M. E. Levinshteĭn, eds., 2006, Page 97.: SiC Materials and Devices. Selected Topics in Electronics and Systems, v. 40. New Jersey ; London: World Scientific.

[13] Ou, Jun-hui, Shao Yong Zheng, Andrey S. Andrenko, Yuanxin Li, and Hong-Zhou Tan, 2018.: Novel Time-Domain Schottky Diode Modeling for Microwave Rectifier Designs . IEEE Transactions on Circuits and Systems I: Regular Papers 65(4): 1234–1244.

[14] Hakim Takhedmit, 2010.: Modélisation et conception de circuits de réception complexes pour la transmission d'énergie sans fil à 2.45 GHz.Autre. Ecole Centrale de Lyon, 2010 .

[15] Zbitou,J.,Latrach, M. and Toutain, S. 2006. : Hybrid rectenna and monolithic integrated zero-bias microwave rectifier. IEEE Transactions on Microwave Theory and Techniques. 54, 1 (Jan.2006), 147–152.

[16] Z. Harouni, L. Cirio, L. Osman, A. Gharsallah, and O. Picon, : A dual circularly polarized 2.45-GHz rectenna for wireless powertransmission, IEEE Antennas Wireless Propag. Lett., vol. 10, pp. 306–309, 2011.

[17] Ladan, Shabnam & Babu Guntupalli, Ajay & Wu, Ke. (2014). : A High-Efficiency 24 GHz Rectenna Development Towards Millimeter-Wave Energy Harvesting and Wireless Power Transmission. Circuits and Systems I: Regular Papers, IEEE Transactions on. 61. 3358-3366. 10.1109/TCSI.2014.2338616.

Valorization of by-products of fish by reactive extrusion to produce biofuel

Dounia MAKOURE [1,2], Abdellah ARHALIASS [1], Adil ECHCHELH [2], Jack LEGRAND [1]
{ dounia.makoure@etu.univ-nantes.fr[1,2], Abdellah.Arhaliass@univ-nantes.fr[1],
adilechel@gmail.com[2]}

[1] University of Nantes, Laboratoire GEPEA, UMR-CNRS 6144, CRTT, Boulevard de l'Université, 44600, Saint Nazaire, France
[2] University Ibn Tofail, b.p 133 av. université, 14000, Kenitra, Maroc

Abstract. : In this paper we present the reactive extrusion process to extract the lipids from by-products of fish. The study focuses on the technical feasibility of using these purified lipids as an additive (between 5 and 40% dilution) in a diesel fuel for use in an engine. The parameters of extrusion process were optimized in order to extract the most percentage of oil and the results of lipids extraction were compared with other techniques like batch process. The physico-chemical characteristics (viscosity, density, calorific value, acidity index) of extracted and transesterified oil are compared with diesel characteristics. Further tests will concern the production of mixtures of the obtained oil and the diesel oil at different dilutions.

Keywords: Reactive extrusion,valorization, by-products, transesterification, biodiesel

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284111

1 Introduction

The depletion of oil resources of the planet and the enhancement of laws and regulations promoting “green” fuels oriented the scientific research toward the discovery of new sources of energy. The activities of smoking fish and canning industry generate every year several tons of by-products representing the residue of the fish and crustaceans manufacturing. The valorization of these by-products (residue biomass) constitutes an important issue to allow a more profitable and more sustainable exploitation of the marine resource. It is therefore important to develop new, high value-added ways of promoting these co-products. The main process used is the reactive extrusion process to extract the lipids used as a biofuel. Moreover, the residue marine biomass represents an alternative source for biofuel production. Biofuel represents an alternative to diesel derived from fossil energy. It is a clean energy source that emits less greenhouse gases.,Its combustion is also less dangerous and non-toxic. Biofuel can be used in pure form B100, B20 is a mixture of 80% diesel and 20% biodiesel, B5 contains 95% diesel and 5% biodiesel [1]. Biodiesel is characterized by physico-chemical properties, the most interesting are viscosity, calorific value, flash point, chemical composition, acidity. The most common method for producing biodiesel is the transesterification of vegetable oils and animal fats[2]–[5].

2 Materials and methods

2.1 Raw material

The fish co-products used in this study were collected in fish shops. The origin of these co-products are two types of fatty and medium-fat fish such as salmon, trout and sardine. The enzymes used for hydrolysis are Neutrase and Protamex. they are industrial enzymes produced by genetic engineering by Novozyme AS. The optimal conditions for Protamex are: pH between 5.5 and 7.5 and temperature between 35 and 60°C. Protamex can be inactivated by heating to 85°C for 10 min. For Neutrasse, its optimal pH is between 5.5 and 7.5 and its optimal temperature is between 35 and 50°C. The Neutrasse inactivation is done by heating to 90°C for 5 min.

2.2 Batch enzymatic hydrolysis

The first enzymatic hydrolysis tests were carried out in a double jacketed reactor to optimize the conditions of the enzymatic hydrolysis reaction (reaction temperature, enzyme type, enzyme/substrate ratio and reaction time).The reactor is composed of a thermal regulation system, pH and temperature probe to control the reaction parameters and a pneumatic stirrer to homogenize the mixture see Figure1.

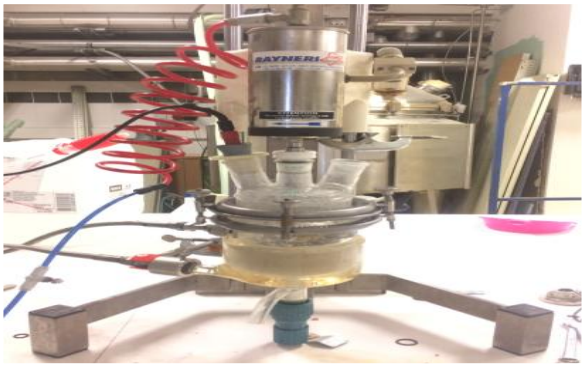


Fig. 1. Batch test equipment

2.3 Enzymatic hydrolysis by reactive extrusion

The reactive extrusion is a process that combines the strength of the extruder's mixture with the chemical industry. The extrusion consists in forcing a product to flow into a screw/barrel system called an extruder. The material moves under the action of high pressures of one or more screws that rotate in a barrel. These screws will also compress and shear the material. All the tests were carried out by using corotative twin screw extruder type Clextral BC21 see figure2.

The extruder is composed of the following components:

-A solid feed hopper equipped with a dosing pump,
-two parallel, co-rotating and interpenetrating screws,
-a thermally regulated barrel and a water circulation cooling system,
-a variable speed motor.

Fig. 2. twin screw extruder type Clextral BC21

2.4 Centrifugation

Centrifugation is a process of separating mixture compounds due to the difference in densities of each compound by centrifugal force. This step is performed after the enzyme inactivation by using an Awel centrifuge - model CF 20/20-R. The hydrolysate is centrifuged at 4400 rpm (3500g) for 30 minutes at 20°C.

After centrifugation, four phases are obtained in the tubes (figure 3) :

- The oily phase, containing mainly lipids
- The aqueous phase, containing mainly water
- The oil emulsion phase
- The pellet phase, with insoluble fractions

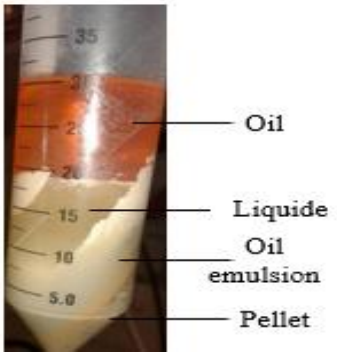


Fig. 3. Phase separation by centrifugation

2.5 Transesterification reaction

The fish oil preheated at 40°C was added with alkoxide mixture (methanol and sodium hydroxide). The methanol/oil ratio is 6:1 and the catalyst amount is about 1% of the total mass of oil. The alkoxide and oil mixture is then stirred throughout the reaction by a mechanical stirrer for 45 min. At the end of the reaction, the contents are poured into a separatory funnel and left to rest for a few hours.

3 Results and Discussion

3.1 Enzymatic hydrolysis

3.1.1 Enzymatic hydrolysis of salmon byproducts

The first test performed was an enzymatic hydrolysis of salmon co-products in double walled glass reactor to determine the optimal reaction conditions and to compare the two processes. Then the extrusion (without enzyme) and reactive extrusion (with enzyme) processes were applied to show the effect of enzyme in the lipid extraction (figure 4).

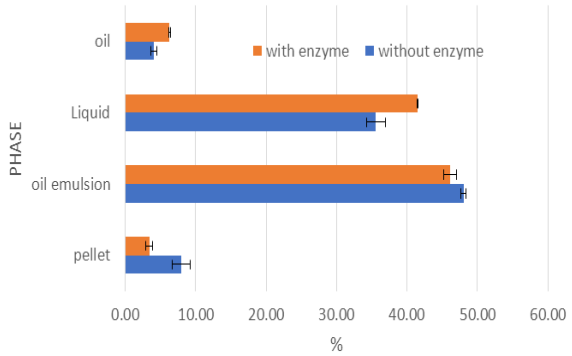


Fig. 4. mass balances of the phases in the extruder with and without enzyme (salmon)

Due to the strength of extruder mixing and the effect of the enzyme in the same time, we got 2.2 % increase of oil recovery than using just the extruder without the presence of an enzyme. The pourcentage of the pellet phase is also reduced by 4.5% by using the reactive extrusion process. This reduction showed the hydrolysis effect of the enzyme on the raw material.

Batch tests were carried out before the reactive extrusion process. Due to high mixture capacity, the enzymatic hydrolysis in the extruder showed an increase of 3.76% of oil phase and 12.07% of oil emulsion phase against to batch process see figure 5.

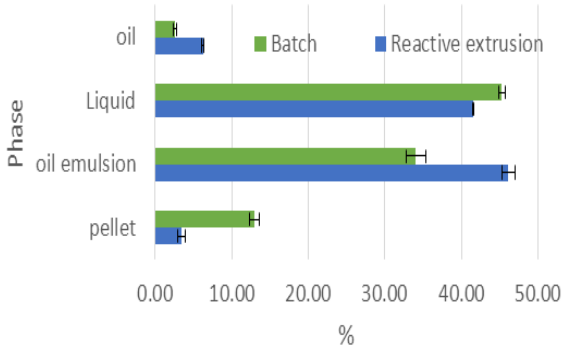


Fig. 5. mass balances of the phases in batch and reactive extrusion (salmon)

3.1.2 Enzymatic hydrolysis of trout byproducts

The percentage of the pellet phase is reduced up to 45% and the oil emulsion phase is increase up to 49.16% with enzyme consumption in the enzymatic hydrolysis in the extruder see figure 6. The tests carried out beforehand in the batch process showed a reduction of 20.21% for the oil emulsion and a small wage decrease of 1% of oil in comparison with the reactive extrusion see figure 7. The combination between enzyme and extrusion is more efficient than the batch process and the extrsuion without enzyme because it provides an important denaturation of the byproducts.

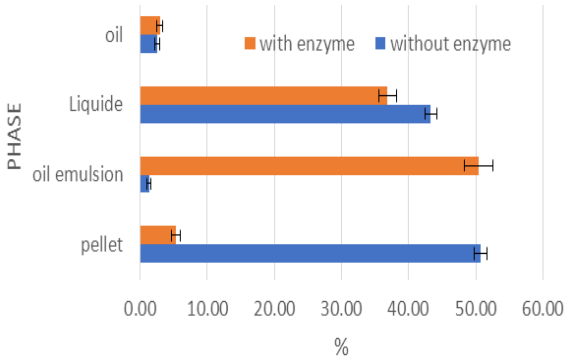


Fig. 6. mass balances of the phases in the extruder with and without enzyme (Trout)

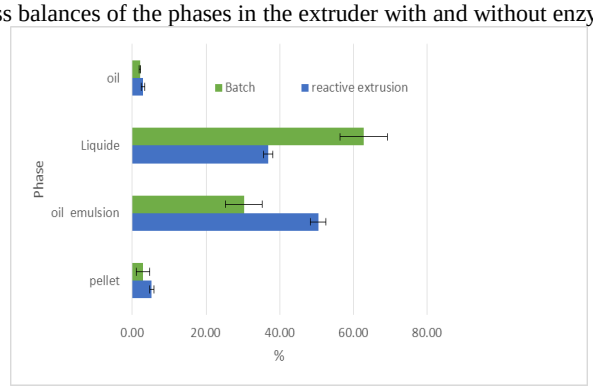


Fig. 7. mass balances of the phases in batch and reactive extrusion (Trout)

3.1.3 Enzymatic hydrolysis of sardine byproducts

There are significant difference between enzymatic hydrolysis in batch and reactive extrusion on oil recovery up to 5%. The percent difference of the oil emulsion phase between the reactive extrusion and the batch is up to 24.05% (figure8). The difference between the batch and reactive extrusion processes is the extruder screws shear force which improves the contact between the raw material and the enzyme.

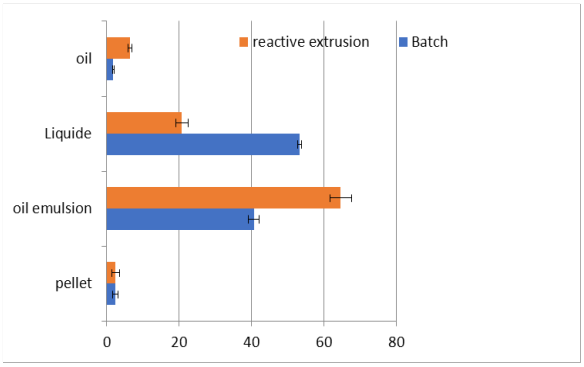


Fig. 8. mass balances of the phases in batch and reactive extrusion (Sardine)

3.2 Determination of total fatty acids

The dosages of total fatty acids were carried out by gas chromatograph (GC) of salmon, trout and sardine oil. These analyses allowed to determine the percentage of total fatty acids and their different classifications.

3.2.1 Different classifications of fatty acids in salmon oil

From the chromatogram of the salmon oil phase, we notice the existence of 11 peaks including 3 with a major representation for oleic acid (51%) , linoleic acid (19.92%) and α-linolénique acid (10.03%) see figure 9 and 10. The salmon oil contains a low level of the erucic acid (2.2%).

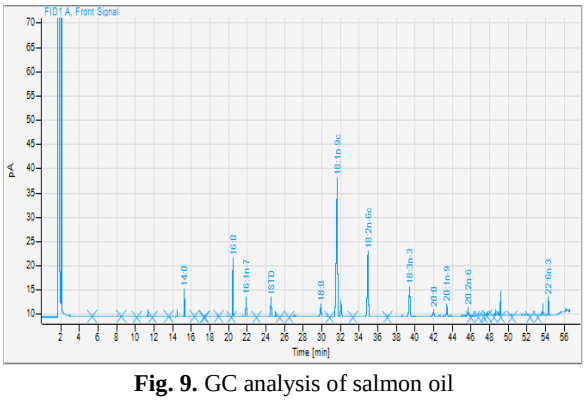


Fig. 9. GC analysis of salmon oil

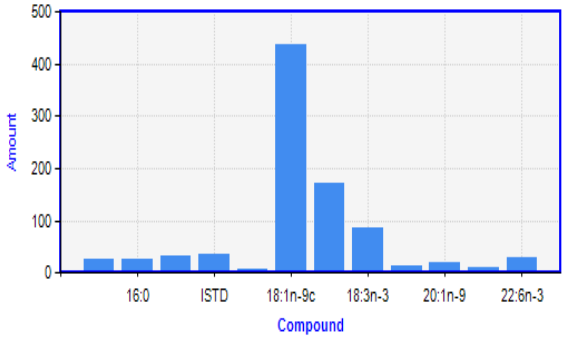


Fig. 10. Amount of different fatty acids: Salmon oil

3.2.2 Different classifications of fatty acids in trout oil

The trout oil is composed of a 15 different fatty acids, principally the oleic acid (49.2%) , linoleic acid (18.44%) and docosahexaénoïque acid (6.78%). It is also composed of erucic acid but with a small amount(2.9%) see figure 11 and 12.

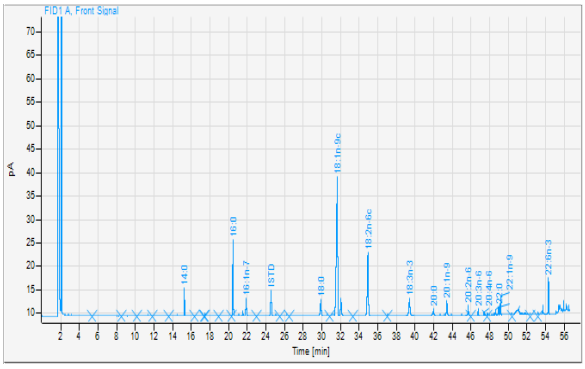


Fig. 11.GC analysis of trout oil

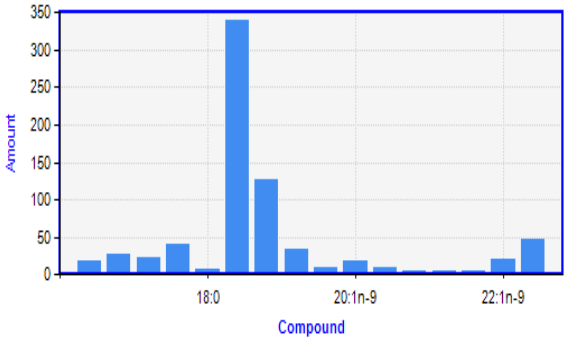


Fig. 12. Amount of different fatty acids: Trout oil

3.2.3 Different classifications of fatty acids in sardine oil

The erucic acid(18.56%), oleic acid (15.87%), docosahexaénoïque acid (13.58%) and Butyric acid (12.52%) are the major fatty acids from 14 different classes of fatty acids in sardine oil.

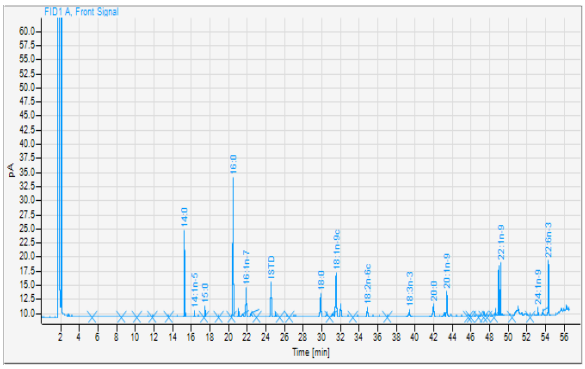


Fig. 13. GC analysis of Sardine oil

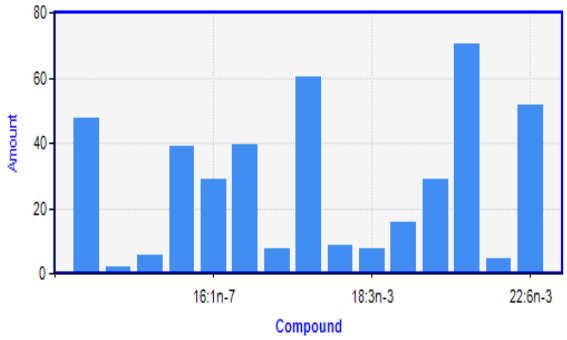


Fig. 14. Amount of different fatty acids: Sardine oil

3.2.4 Total fatty acids in salmon, trout and sardine oil

The salmon by-products contain the highest percentage of total fatty acids (figure15). For this reason, the following steps (transesterifcation and chemical and physicochemical analyses) will be carried out on this type of by-products.

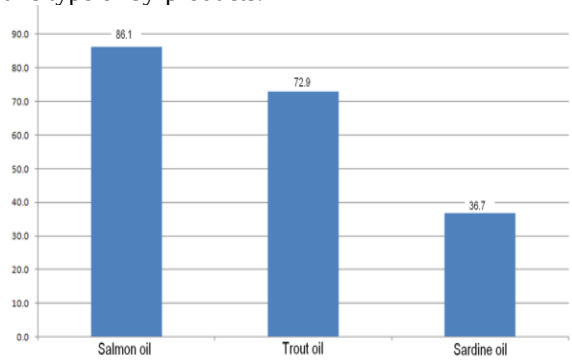


Fig. 15. Percentage of total fatty acids for the 3 fish oil types

3.2 The transesterification reaction

In previous studies, The basic transesterification yield of the high erucic acid content oil is low because of the soap formation so it provides a minimum yield of ester. In addition, the biodiesel from the low erucic acid content oil has the same properties, compared to diesel specifications[6]. On the basis of this study, we chose the salmon oil that has the lowest erucic acid content for the transesterification reaction and biodiesel production.

After the transesterification of salmon oil and after mixture settling, two separate phases are obtained. A lower phase containing glycerin which is immiscible with esters settles at the bottom of the funnel. A higher phase that contains esters or biodiesel mixed with excess alcohol see figure 16.

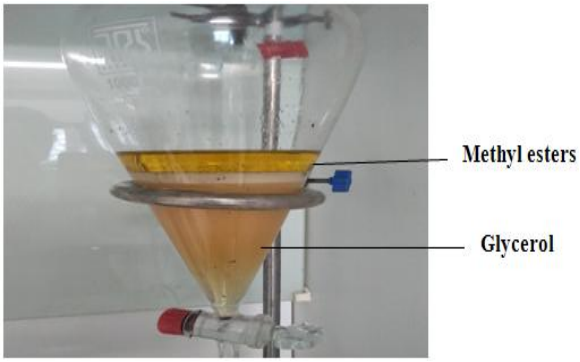


Fig. 16. phase separation after transesterification

A sample of salmon oil and its methyl esters after transesterification were analysed by high-performance thin-layer chromatography (HPTLC). The results of this analysis show the disappearance of triglycerides after transesterification (Figure17 and 18), which validates the transesterification reaction.

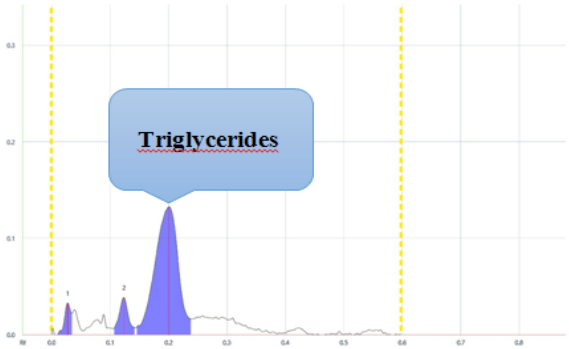


Fig. 17. HPTLC of salmon oil

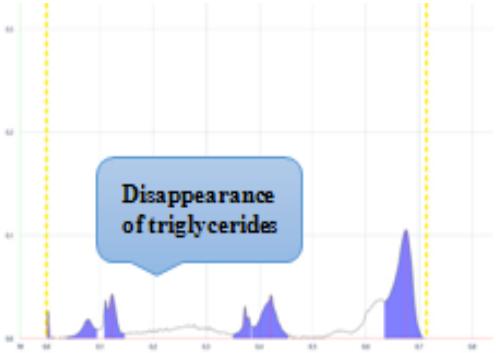


Fig. 18. HPTLC of methyl esters

4 Conclusion and perspectives

The reactive extrusion process is used in order to extract the oil from the fish coproducts. Experimental tests confirm that the used of this process is more efficient compared to the batch process, for many reasons. It offers an important raw material denaturation and a better oil recovery while minimizing the pellet phase. In addition, the used of this process for the fish coproducts hydrolysis compared to batch process optimise the process parameters; the extraction time is reduce from 60 min by using the batch to only 5 min.

It is therefore more advantageous to use salmon byproducts as a raw material for biodiesel production. The salmon recovered oil percentage is more important than the other recovered oil types (trout and sardine). The salmon oil is chosen for its characteristics, it contains the highest percentage of lipids. Futhermore, It has the lowest erucic acid content allowing a viable transesterification so a good ester yield. Besides, the extracted biodiesel from the lowest erucic content oil has a similar properties as the diesel.

The further tests will focus on the determination of the acidity index, chemical composition, viscosity, density of methyl esters. The same analyses will be carried out for diesel in order to compare the parameters. The effect of the different biodiesel and diesel mixture on the engine performances will be subject of future studies.

References

[1] Drvodelic, N.: Analyse de biodiesel selon les normes ASTM D6751 et EN 14214 à l ' aide de l ' ICP-OES Agilent 5100 SVDV. (2014)

[2] Adewale, P., Dumont, M., and Ngadi, M.: Recent trends of biodiesel production from animal fat wastes and associated production techniques. *Renew. Sustain. Energy Rev.* vol. 45, pp. 574–588 (2015)

[3] Sajjadi, B., Aziz, A., Raman, A. and Arandiyan, H.: A comprehensive review on properties of edible and non-edible vegetable oil-based biodiesel : Composition , speci fi cations and prediction models. *Renew. Sustain. Energy Rev.* vol. 63, pp. 62–92 (2016)

[4] Amini, Z., Chyuan, H., Harrison, M. D., Kusumo, F., Mazaheri, H. and Ilham, Z.: Biodiesel production by lipase-catalyzed transesterification of Ocimum basilicum L . (sweet basil) seed oil. *Energy Convers. Manag.* vol. 132, pp. 82–90 (2017)

[5] Nisar, J., Razaq, R., Farooq, M., Iqbal, M., Ali, R., Sayed, M. and Shah, A.: Enhanced biodiesel

production from Jatropha oil using calcined waste animal bones as catalyst. *Renew. Energy*. vol. 101, pp. 111–119 (2017)
[6] Dorado, M. P., Ballesteros, E., Lo, F. J. and Mittelbach, M.: Optimization of Alkali-Catalyzed Transesterification of Brassica Carinata Oil for Biodiesel Production. *Energy & Fuels*. vol. 18, pp. 77–83 (2004)

Estimation of the minimum mean square error channel MMSE And energy efficiency management

[1]Sara Riahi, [2]Azzeddine Riahi
{ [1]riahisaraphd@gmail.com, [2]riahikh@gmail.com }

[1]Department of Computer Science, Chouaib Doukkali University , Faculty of Sciences , El Jadida, Marocco
[2]IMC Laboratory, Department of Physics, Chouaib Doukkali University, Faculty of Sciences, El Jadida, Morocco

Abstract: In order to eliminate distortion due to signal propagation in a transmission channel, a multi-carrier OFDM modulation scheme uses a channel estimation operation. Indeed, an estimate of the frequency response of the transmission channel is often necessary to achieve frequency equalization at the output of the FFT (Fast Fourier Transform). In the literature, most channel estimation techniques have been limited to exploiting the frequency correlation of the channel. Few of them have addressed the interests reflect the temporal correlation. Note that propagation in a mobile radio environment is characterized by frequency selectivity of the transmission channel, but also by mobility of transmitting / receiving equipment translated into time selectivity. This is expressed by a double time-frequency correlation of the frequency response of the transmission channel. In this paper, we propose to design a channel estimation algorithm, half-blind in a mobile-radio propagation environment. The initialization of this algorithm is done through information provided by the pilot OFDM symbols.

Keywords: Channel estimation, Quadrature Amplitude Modulation, MMSE, Wireless Systems, Energy Efficiency.

1 Introduction:

The OFDM / QAM are a good alternative to OFDM for transmitting high-speed signals in the current conditions of spectrum congestion. Indeed, the use of the cyclic prefix is not required in OFDM / QAM, increasing its spectral efficiency compared with OFDM. In addition, the symbols are not worn by cardinals sinuses, but by functions (called filters prototypes and rated g) greatly reducing out of band interference. Thus, in this paper, we use the frequency response given by **[1]**:

$$G(f) = \sum_{k=-(K-1)}^{K-1} G_k \frac{sin\left(\pi\left(f - \frac{k}{L_f}\right)L_f\right)}{L_f\left(\pi\left(f - \frac{k}{L_f}\right)\right)} \qquad \textbf{(1)}$$

Where K is called the recovery factor (here $K = 4$), $G_0 = 1$ and $G_k = G_{-k}$. L_f is the length of the prototype filter. In Figure 1, it can be observed that the frequency response of the OFDM/QAM using the Bellanger filter has better frequency localization than the OFDM. However, this good spectral property can be obtained at the cost of a relaxation of the requirement of complex orthogonality, since the g prototype filter is only orthogonal in the space of real IR. So instead of transmitting complex symbols at a rate τ_0 as in OFDM, QAM symbols are real and transmitted at a rate $\tau_0 / 2$.

ICCWCS 2019, April 24-25, Kenitra, Morocco
Copyright © 2019 EAI
DOI 10.4108/eai.24-4-2019.2284112

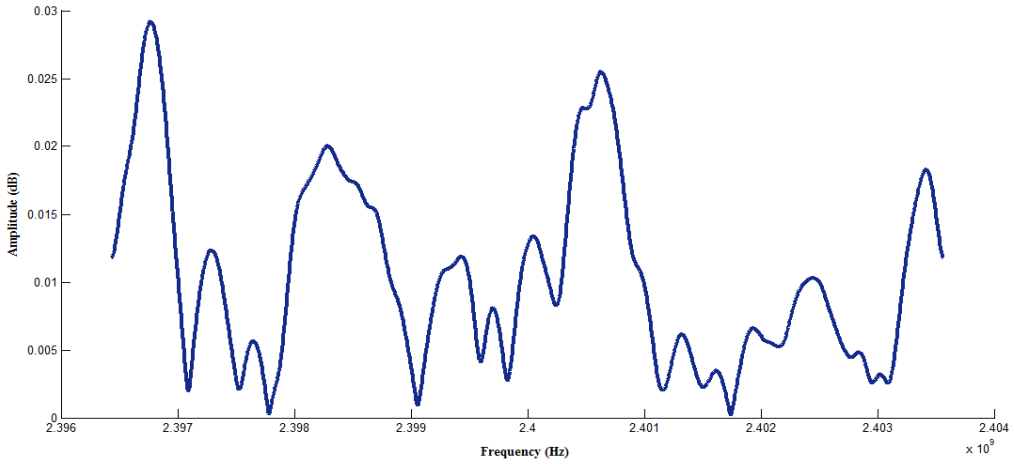


Fig 1: Frequency responses of OFDM

The OFDM/QAM signal received, after the serial /parallel conversion, the filtering in the polyphase network (defined by g) and the discrete Fourier transform (DFT) of size M , can be written **[2]**:

$$y_{m,n} = H_{m,n}x_{m,n} + \underbrace{j\sum_{(p,q)\in\Omega} H_{p,q}x_{p,q} < g >_{m,n}^{p,q} + \omega_{m,n}}_{I_{m,n}} \qquad \textbf{(2)}$$

Where the index (m,n) indicates the position time-frequency, $x_{m,n}$ is the actual transmitted QAM symbol, $H_{m,n}$ is the channel frequency response, and $\omega_{m,n}$ is the additive Gaussian noise with zero mean and variance σ^2. The specificity of the OFDM / QAM modulation is found in the presence of the interference $I_{m,n}$ which is due to the filter g and the channel **[3]**. The term $j < g >_{m,n}^{p,q}$ in (2) is called "intrinsic interference" and is defined as the scalar product $< g_{m,n}, g_{p,q} >$, with $(p,q) \in \Omega$, where Ω is the set of frequency-time (p,q) positions in the vicinity of (m,n). $j < g >_{p,q}^{m,n}$ is a pure imaginary, so that if $H_{m,n} = 1$, then we find the signal transmitted as $\hat{x}_{m,n} = Re(y_{m,n})$. However, the distortion of the complex channel coefficient $H_{m,n}$ induces a complex interference $I_{m,n}$. The channel must be estimated and reversed to recover the actual conditions of orthogonality. Many OFDM / QAM channel estimation methods have been proposed in the literature **[4]**. In this paper, we propose to study the MMSE estimator in OFDM/QAM. This estimate as a filter smoothing the frequency response of the channel, its use may be relevant in modulating considered where the received signal $y_{m,n}$ in (2) is subjected to distortion from noise and interference. In the remainder of the paper, the theoretical expression is derived for the MMSE estimator OFDM/QAM, and it shows that a

simple formulation can be obtained, similar to that known in OFDM. The simulation results compare the MMSE estimator to OFDM/QAM and OFDM.

2 MMSE estimators:

2.1 Theoretical expression of the estimator:

To apply the MMSE estimation to the received signal, it is necessary to rewrite (2) in vector form:

$$y_n = x_n H_n + I_n + w_n \qquad \textbf{(3)}$$

Where y_n, H_n, I_n and w_n are the vectors of size $M \times 1$ containing the samples defined in (2), and x_n is the diagonal matrix whose elements are the transmitted symbols $x_{m,n}$. The developments given in **[5]** are used to give the general expression of the MMSE estimator, which aims at minimizing the cost function J_{MMSE} defined by:

$$J_{MMSE} = E\left\{\left\| H_n - \theta y_n \right\|_F^2\right\} \qquad \textbf{(4)}$$

Where θ is the matrix $M \times M$ whose coefficients are to be optimized. In (4), $E\{.\}$ is the mathematical expectation and $\|.\|_F$ is the Frobenius norm. The cost function J_{MMSE} can be reformulated by substituting (3) for (4) to obtain **[6]**:

$$J_{MMSE} = tr\left(R_{HH} + \theta\left(x_n R_{HH} x_n^H + R_{II} + x_n R_{HI} + R_{IH} x_n^H + \sigma^2 I_d\right)\theta^H - \theta\left(x_n R_{HH} + R_{IH}\right) - \left(R_{HH} x_n^H + R_{HI}\right)\theta^H\right) \qquad \textbf{(5)}$$

Where R_{HH}, R_{II} are respectively the matrix $M \times M$ of covariance of the H_n channel and the interferences I_n, R_{IH} and R_{HI} the intercovariance matrix of H_n and I_n, I_d is the identity matrix.

Using the rules of derivation of matrix in C defined in **[7]**, we find the minimum of J_{MMSE} by solving:

$$\frac{\partial J_{MMSE}}{\partial \theta^*} = \theta\left(x_n R_{HH} x_n^H + R_{II} + x_n R_{HI} + R_{IH} x_n^H + \sigma^2 I_d\right) - \left(R_{HH} x_n^H + R_{HI}\right) = 0 \qquad \textbf{(6)}$$

And we deduce **[8]**:

$$\theta_{op} = \left(R_{HH} x_n^H + R_{HI}\right) \times \left(x_n R_{HH} x_n^H + R_{II} + x_n R_{HI} + R_{IH} x_n^H + \sigma^2 I_d\right)^{-1} \qquad \textbf{(7)}$$

The MMSE estimator in OFDM / QAM is expressed from (8), factoring in $x_n(.)x_n^H$, as

$$\hat{H}_n^{LMMSE} = \theta_{op} y_n = \left(R_{HH} + R_{HI}(x_n^H)^{-1}\right)\left(R_{HH} + x_n^{-1}R_{II}(x_n^H)^{-1} + R_{HI}(x_n^H)^{-1} + x_n^{-1}R_{IH} + (x_n x_n^H)^{-1}\sigma^2 I_d\right)^{-1} \times \left(H_n + x_n^{-1}I_n + x_n^{-1}w_n\right) \quad \textbf{(8)}$$

It has been raised in OFDM that the implementation of the MMSE estimator is limited by **[9]**:

- Its complexity (due to matrix inversions and multiplications).
- The absence of a priori knowledge of the covariance matrix of the R_{HH} channel and the noise level σ^2.

This limitation is all the more true in OFDM/QAM since it appears in (8) that matrix R_{II} and R_{HI} are required. Consequently, the MMSE implementation in (8) requires approximations **[10]**. This is discussed below.

2.2 Simplified expression of the estimator:

To simplify the expression of $\hat{H}_n^{LMMSE}$ given in (8), we can assume that **[11]**:

- The source of the interfering symbols with $x_{m,n}$ is limited to Ω_0 defined as the set of positions adjacent to (m,n), $\Omega_0 = \{(p,q) \mid p \in \{m-1, m, m+1\}, q \in \{n-1, nm+1\}, (p,q) \neq (m,n)\}$;
- The channel frequency response is assumed to be constant on the positions $(p,q) \in \Omega_0$, leading to the approximation $H_{p,q} \approx H_{m,n}$. These two assumptions allow rewrite (2) as **[6]**:

$$y_{m,n} \approx H_{m,n}\left(x_{m,n} + x'_{m,n}\right) + \omega_{m,n} \quad \textbf{(9)}$$

$$= H_{m,n} d_{m,n} + \omega_{m,n} \quad \textbf{(10)}$$

Where $$d_{m,n} = x_{m,n} + x'_{m,n} = x_{m,n} + j \sum_{(p,q)\in\Omega_0} x_{p,q} < g >_{p,q}^{m,n} \quad \textbf{(11)}$$

Such that (10) is very similar to the received signal in OFDM.

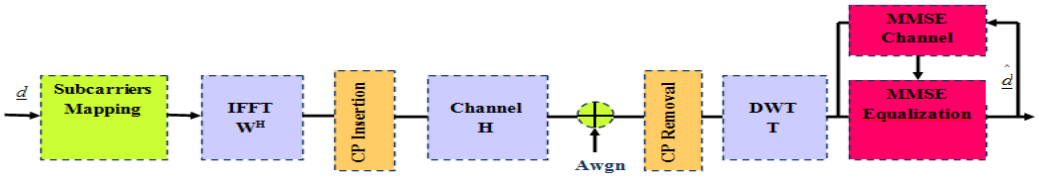


Fig 2: Block diagram for an improved joint MMSE channel estimation.

If $d_{m,n}$ is a pilot value with a known value, then it is possible, by analogy with the MMSE estimator OFDM give a simplified expression (8):

$$\hat{H}_n^{LMMSE} = \underbrace{R_{HH}\left(R_{HH} + \sigma^2\left(d_n d_n^H\right)^{-1}\right)^{-1}}_{K} \hat{H}_n \qquad \textbf{(12)}$$

Where d_n is the diagonal $M \times M$ matrix containing the $d_{m,n}$ drivers, and $\hat{H}_n$ is the estimated channel response by **[10]**:

$$\hat{H}_n = d_n^{-1} y_n \qquad \textbf{(13)}$$

Note that (13) is similar to the OFDM estimator using the least squares criterion, or least square (LS). However, it will be considered later that $x'_{m,n} = 0$ and therefore $d_{m,n} = x_{m,n}$ for all $0 \le m \le M-1$, so as to obtain an estimation in OFDM / QAM very similar to the LS estimator in OFDM. The method used is called ICM (for interference cancellation method) **[11]**. By rewriting the MMSE estimator as in (12), it is possible to calculate $\hat{H}_n^{LMMSE}$ with reduced complexity using simplification methods such as the eigenvalue decomposition of K and the low-order approximation of R_{HH}. Thus, the covariance matrix of the R_{HH} channel is hermitian, so it is diagonalisable in an orthonormal basis such that $R_{HH} = U D_R U^H$, where U is the unit matrix containing the eigenvectors of R_{HH}, and D_R is the diagonal matrix containing the eigenvalues $\lambda_0, \lambda_1, ..., \lambda_m, ..., \lambda_{M-1}$ of R_{HH}. Moreover, assuming that the channel is a finite impulse response filter of length L, then it is also the rank R_{HH} and therefore $\lambda_L = \lambda_{L+1} = ... = \lambda_{M-1} = 0$.

In practice, the non-zero eigenvalues of R_{HH} are rarely known a priori **[12]**. However, R_{HH} can be approximated by a well-known $\tilde{R}_{HH}$ matrix whose spectrum is imposed (corresponding to the power profile of the channel). Without loss of generality, we consider a constant profile, so that the decomposition into eigenvalues (also known) gives $\tilde{R}_{HH} = \tilde{U}\tilde{D}_R\tilde{U}^H$, where the eigenvalues $\tilde{\lambda}_m$ on the diagonal of $\tilde{D}_R$ are equal to $\frac{1}{L}$ if $0 \le m \le \tilde{L}-1$, and 0 otherwise. As noted in [5], the quality of the MMSE estimator using the low rank approximation is guaranteed for $\tilde{L} \ge L$.

We assume this condition fulfilled in the following, and we can rewrite (12) using $\tilde{R}_{HH}$ instead of R_{HH} and we define $\tilde{K} = \tilde{R}_{HH}\left(\tilde{R}_{HH} + \sigma^2\left(d_n d_n^H\right)^{-1}\right)^{-1}$. Since $\sigma^2\left(d_n d_n^H\right)^{-1}$ is a diagonal matrix, then $\tilde{K}$ is diagonalisable in the same basis as $\tilde{R}_{HH}$, and we can rewrite the MMSE estimator as $\hat{H}_n^{LMMSE} = \tilde{U}\tilde{D}_K\tilde{U}^H\hat{H}_n$, where $\tilde{D}_K$ is the eigenvalue decomposition of $\tilde{K}$ such that its elements are defined by **[10]**:

$$\tilde{\mu}_m = \begin{cases} \dfrac{\tilde{\lambda}_m}{\tilde{\lambda}_m + \dfrac{\sigma^2}{|d_m|^2}} \\ 0 \end{cases} \quad \text{if} \quad \begin{matrix} m = 0,1,....,\tilde{L}-1 \\ m = \tilde{L},\tilde{L}+1,....,M-1 \end{matrix} \qquad \textbf{(14)}$$

Thus, it is possible to use the MMSE estimator for OFDM / QAM modulation without prior knowledge of R_{HH}, and reducing the complexity (given in number of multiplications) of $O(2M^3)$ in (12) to $O(\tilde{L}M^2 + \tilde{L}M)$ using (14).

3 Simulation and results analysis:

In this part, we simulate the validity of the simplified MMSE estimator in OFDM/QAM and compare its performance (Figure 3). For this, we use the parameters of the standard LTE (Long Term Evolution).
The actual QAM symbols are chosen in [-1, +1] (equivalently, the symbols for the OFDM signal come from a QAM), and no channel code is used. The transmission channel is a Gaussian channel, and we consider a simple ZF (zero forcing) equalization on each carrier.
Figure 4 shows the estimation MSE obtained for LS (in OFDM), and simplified MMSE using the OFDM and OFDM/QAM eigenvalue decomposition, as a function of Eb/N0 (dB). The curves are averages obtained on several random realizations of the channel. There is a high gain between the MMSE estimate and the LS and MRC estimators, reflecting the "smoothing" effect of MMSE on the frequency response of the channel.

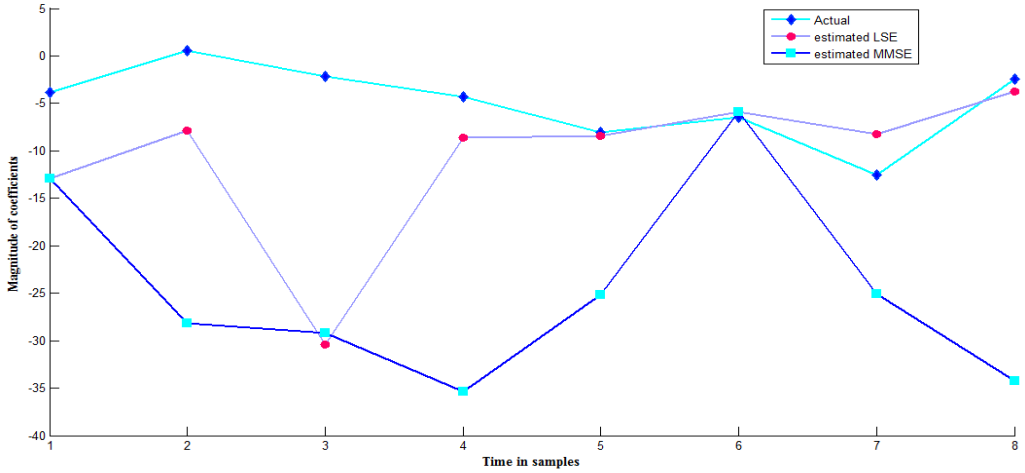


Fig 3: Channel estimation using LS and MMSE estimators

This effect remains valid OFDM / QAM. However, there is a MSE threshold is reached in OFDM/QAM for MRC and MMSE, due to interference inherent to this type of modulation. Figure 5 shows the bit error rate (BER) as a function of Eb / N0 (dB) obtained for the same estimators as before. Moreover, the curves of a transmission in a Gaussian channel (denoted by AWGN, additive white Gaussian noise for) and perfect estimation are displayed as references. A gain obtained for the MMSE estimators compared to the LS and ZF estimators is observed, confirming the good performance of MMSE observed in Figure 5. It should be noted, however,

that the curve corresponding to MRC has a lower slope than LS from Eb / N0 = 15 dB, due to the presence of intrinsic interference in OFDM/QAM.

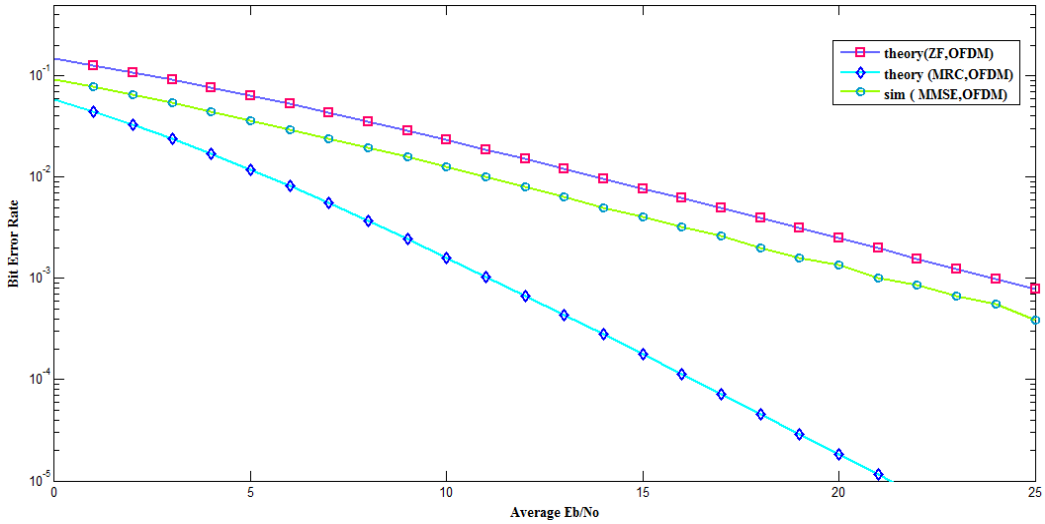


Fig 4: BER for QAM modulation, MMSE equalizer, AWGN channel

These interferences are embedded in noise for low values of Eb/N0, but appear for low noise values. The same phenomenon is observed by comparing the curves of the MMSE estimator in OFDM and OFDM/QAM. However, we note that the loss of OFDM/QAM over OFDM is only 0.2 dB at Eb/N0 = 25 dB. For lower values of Eb/N0, the two curves are combined, and are only 0.1 dB of the perfect estimate.

These results show that despite the complex formulation of the MMSE estimator in OFDM/QAM in (8), it is possible to obtain a simplified version of the estimator while keeping a performance close to the perfect estimate.

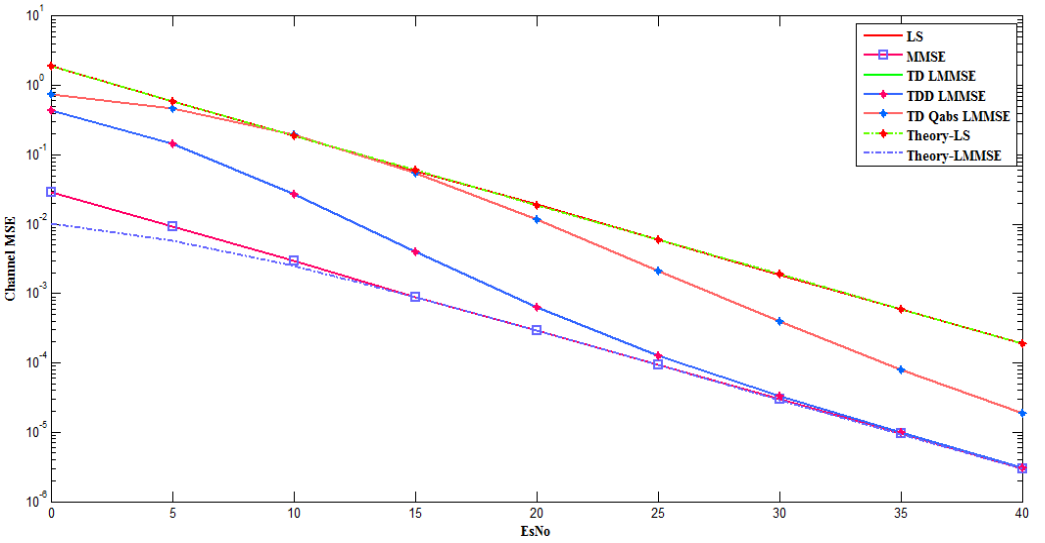


Fig 5: Comparison of different channel estimation techniques with MMSE

In addition, the BER of the MMSE in OFDM/QAM is very similar to that obtained in OFDM. This result is even more interesting that in general, the spectral efficiency of OFDM/QAM (not

using a cyclic prefix) is greater than the OFDM. The analysis of the curves obtained shows that the estimation studied technique introduces improvements in performance of the channel estimation.

4 Conclusion:

In this paper, we introduced the channel estimator based on the criterion of the mean square error (MMSE) in OFDM with the quadrature amplitude modulation. The theoretical expression of the estimator has been proposed, and it shows that it is not applicable, due in particular to its excessive complexity. However, approximations enable a new formulation similar to that known in OFDM. Thus, it is possible to use simplification methods of the MMSE channel estimator, such as the eigenvalue decomposition and the low rank approximation. Simulation results have shown that the MMSE channel estimator in OFDM / QAM is able to achieve the same performance in OFDM, for better spectral efficiency.

Acknowledgment:

We would like to thank the CNRST of Morocco (I 012/004) for support.

References:

[1] M.Tuchler, C.Andrew, Singer and R .Koetter ,"Minimum Mean Squared Error equalization using priori information ,"IEEE Trans, on signal process., vol. 50,no. 3, pp 673-683 mar.2002

[2] Sara RIAHI, Azzeddine RIAHI , '' Applications of game theory to optimization of resource allocation of wireless systems: algorithmic solution '' ,International Journal of Online Engineering (iJOE), ISSN: 2626-8493, PP. 5-25,Vol 14, No 12 (2018).

[3] X.-B. Liang, "Orthogonal designs with maximal rates", *IEEE Trans.Inform.Theory*, vol. 49, no. 10, pp. 2468-2503,. 2003.

[4] M. K. Karakayali, G. J. Foschini, R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems", *IEEE Wireless Commun. Mag.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.

[5] Sara RIAHI, Azzeddine RIAHI, '' Game theory for resource sharing in large distributed systems '', International Journal of Electrical and Computer Engineering , IJECE, Vol. 9, No. 2, April 2019, pp. 1249~ 1257 ISSN: 2088-8708, DOI: 10.11591/ijece.v9i2.pp.1249-1257 , IJECE.

[6] Jittra Jootar, James R. Zeidler, John G. Proakis, "On the Performance of Closed-Loop Transmit Diversity with Noisy Channel Estimates and Finite-Depth Interleaved Convolutional Codes", *Communications IEEE Transactions on*, vol. 56, 2008, ISSN 0090-6778.

[7] K. Cho, D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations", IEEE Trans. Commun., vol. 50, pp. 1074-1080, July 2002.

[8] Y. Yang, Q. Huang, Z. Zhang, "Vehicular Technology Conference 2009", *An Efficient Joint Timing and Frequency Offset Estimation for OFDM Systems*, pp. 1-5.

[9] Mohammed Ali Mohammed MOQBEL1 , Wangdong2 , Al-marhabi Zaid Ali, "MIMO Channel Estimation Using the LS and MMSE Algorithm ",IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN: 2278-2834.Volume 12, Issue 1, Ver. II (Jan.-Feb. 2017), PP 13-22

[10] Ajay Bahadur Singh1 , Vivek Kumar Gupta, "Performance Evaluation of MMSE and LS Channel Estimation in OFDM System ",International Journal of Engineering Trends and Technology (IJETT) – Volume 15 Number 1 – Sep 2014

[11] M.-X. Chang, "A new derivation of least-squares-fitting principle for OFDM channel estimation", *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 726-731, Apr. 2006.

[12] Akash Kumar Shrivas, "A Comparative Analysis of LS and MMSE Channel Estimation Techniques for MIMOOFDM System ",IJIRST –International Journal for Innovative Research in Science & Technology| Volume 1 | Issue 8 | January 2015 ISSN (online): 2349-6010

Proceedings of the Third International Conference on Computing and Wireless Communication Systems (ICCWCS'19)

April 24-25, 2019, Faculty of Sciences, Ibn Tofaïl University -Kénitra- Morocco

ICCWCS 2019

Copyright © 2019 EAI, European Alliance for Innovation

www.eai.eu

https://www.iccwcs.net/

ISBN: 978-1-63190-181-2

ISSN: 2593-7650

Author Index

Abdelmounim, El Hassan 64
Abdelmounim, Elhasan 475
ABDELMOUNIM, Elhassane 186, 485
Abdelmounim, Elhassane 455, 509
Abdelmounim, Hassan 427
Abdlali, Tajmouati 40
Aboulfatah, Mohamed 427, 509
Abounaima, Mohammed Chaouki 256
Abri, Mehadji 57
adadi, amina 274
Addaim, Adnane 465
Adekanle, Oluwaseun 509
Adil, ECHCHELH 215
Ahmed, El Oualkadi 232
Ahmed, Errkik 40
Ait Madi, Abdessalam 465
Alami Massmoudi, Souhail 207
ALAOUI ISMAILI, IMANE 265
Ammor, H. 207
AMMOR, HASSAN 90
aouatif, SAAD 215
Aoulass, Nabila 396, 436
AOURIK, Salaheddine 46
Arhaliass, Abdellah 529
Arioua, Mounir 232
Aytouna, Fouad 160
Aznabet, Mariam 73, 82
Azyat, Abdelilah 265

BAAYER, JALAL 90
Badaoui, Hadjira 57
BAHAJ, Mohamed 336
Bahri, Hicham 509
Balogun, Adeola 509
Balouki, Youssef 375
Belaguid, Abdelaziz 455
Ben Achhab, Nizar 265
Benalla, Said 294
Benlafkih, Abdessamad 492
BENNIS, Hamid 186
Bennis, Hamid 1, 64, 242, 327
BENSLIMANE, Mohamed 282
BERRADA, Mohammed 274, 282
BOUCHOUR, Al Mehdi 141
Bouchour, Al Mehdi 131
Bouhouch, Adil 327
Boutyour, El hassan 249
BOUZIDA, Ilham 32
brahmia, hanene 355

Chafik Elidrissi, Mohamed 492
Chahboun, Asaad 265
Chakkor, Otman 396, 436
CHHIBA, MOSTAFA 346

Dahir, Sarah 242
Dherbecourt, Pascal 131
Dherbécourt, Pascal 141

ECHCHAKHAOUI, Khalifa 186
Echchelh, Adil 168, 529
Echchelh, Mehdi 222
EL ABDELLAOUI, Larbi 46, 502
El Abdellaoui, Larbi 177, 197, 520
El Adib, Samir 417
EL AHMAR, latifa 32
EL AKKAD, Nabil 274

El AKKAD, Nabil 407
El Allami, Mhammed 249
El Aoufi, Jamal 294
EL HACHIMI, Younes 113
EL HADRI, Doae 23
El Hafid, Yassine 384
El Hanine, Mustapha 455
El Makhloufi, Assaad 417
El Mazouri, Fatima Zahra 256
EL Mrabet, Otman 82
El Mrabet, Otman 73
EL NABAOUI, Driss 1
EL OUALKADI, Ahmed 446
El Oualkadi, Ahmed 131, 141
EL OUJDI, Abdellatif 215
El Qadi, Abderrahim 242, 327
El Rharras, Abdessamad 384
Elabdellaoui, Larbi 1
ELFAKKOUSSI, SAID 365
Elhabchi, Mourad 122
ELHAROUSSI, Mustapha 485
ELKHALFI, AHMED 365
ELYOUSSFI, Soukayna 215
ERRKIK, Ahmed 32, 46, 502
Errkik, Ahmed 1, 177, 197, 520
ES-Sabry, Mohamed 407
Ez-zazi, Imad 232

FARCHI, Abdelmajid 113
Farchi, Abdelmajid 103
FERFRA, Nawfal 375
Fersiti, Samira 57
Fouad, Aytouna 151

Gadi, Taoufiq 375
Galadi, Abdelghafour 168
Gmih, Yassine 103
González, Antonio 232
Guisser, M'hammed 509

HADDADI, RACHID 455
Hadini, yassine 168
Hamid allah, Abdelhak 520
Hassoine, Mohammed Amine 465
Hefnawi, Mostafa 427
Hicham, BAHRI 475
HILMI, Asmaa 446

Ihamji, Mohamed 64

jarrar, Abdessamad 375
Joubert, Eric 131
Jouha, Wadia 131

KARIM, Karima 407
KHOURDIFI, Youness 336
Khoutar, Fatima 73
Khoutar, Fatima Zohra 82
Krim, Khadidja 57

LAGRAINI, HAJAR 346
Lahlou, Fouad 465
Lahraoua, Mohammed 265
LAKSSIR, Brahim 32
Larbi, El Abdellaoui 40, 151
LAROUI, SARA 314
LATRACH, Mohamed 46, 186, 502
Latrach, Mohamed 1, 64, 177, 197, 520
Latry, Olivier 131, 141
LAZAAR, Mohamed 314
Legrand, Jack 529
Loqman, Chakir 327

MAHBOUB, Oussama 314
Mahieddine, Fatima 57
Makoure, Dounia 529
Mandry, Rachid 197
Masmoudi, Mohamed 131
MERRAS, Mostafa 407
MEZROUI, Soufiane 446
Mohamed, Aboulfatah 475
Mohamed, HATTABI 215
Mohamed, Latrach 40, 160
mohamed, Latrach 151
Mohammed, Bendaoued 151
Mohammed, El Malah 475
Mondir, Anouar 11
MOUHSEN, AHMED 346
Mouhsen, Ahmed 294
MOUSTABCHIR, Hassane 365
Mouzouna, Youssef 294
M'hamed, Guisser 475

Najah, Said 256
Nasraoui, Hanane 294
Nayat-Ali, Oumaima 73
Nissabouri, Salah 249

OMARA, HICHAM 314
Otman, Chakkor 160
OUAZZANI, Kamar 282

OULDZIRA, HICHAM 346
Oumaima, Nayat-Ali 82

rachakh, amine 177
RACHED, Bouchaib 485
Rachid, Mandry 151
Raissouni, Naoufal 265, 417
riahi, azzeddine 303, 541
riahi, sara 303, 541
Ribate, Mohamed 197

Saadane, Rachid 384
SAAIDI, Abderrahim 407
Salhaoui, Marouan 232
Samir, Elkilani 40
SATORI, Khalid 407
Setti, Larbi 11
simon, Adekanle 475
Srifi, Mohamed 122

TABYAOUI, ABDELMOUMEN 346
Tajamouati, Abdelali 520
TAJMOUATI, Abdelali 1, 46, 502
Tajmouati, Abdelali 177
Taouzari, Mohamed 294
TAYBI, Abdellah 502
TMIMI, Mehdi 282
tolba, cherif 355
Touahni, Raja 122

Wahbi, Mohamed 384

Yazid, Yassine 232

ZAKRITI, Alia 23
ZBITOU, Jamal 32, 502
Zbitou, Jamal 1, 427, 520
Zenkouar, Khalid 256
ZUGARI, Asmaa 23

European Alliance for Innovation Community Research Series

Through its groundbreaking Community Research (CORE) Series, European Alliance for Innovation (EAI) provides a publication platform for conferences and workshops organized under its main organizational tenets – global community collaboration and fair recognition. These features are brought to life via free online collaboration and networking tools, giving all EAI members an equal opportunity to actively participate and publish in various peer reviewed research venues.

EAI is a non-profit organization with free membership and the largest open professional society for advancing research careers through community collaboration and fair recognition. Members benefit from finding feedback and mentorship for their work and they are guaranteed to be evaluated fairly, transparently, and objectively through community.

The scope of EAI CORE Series covers the full spectrum of ICT topics:

- E-health
- Networks & Communication
- Intelligent Systems
- IoT
- Arts, Education, Media
- Smart Cities
- Developing Countries
- Systems & Security

ISSN: 2593-7650

ISBN: 978-1-63190-181-2

http://eudl.eu/series/CCER | www.eai.eu