# Research on Intelligent Wireless Channel Allocation in HAPS 5G System Based on Reinforcement Learning

Zhou Wu, Ming-xiang Guan(✉), Yingjie Cui, Xuemei Cao, Le Wang, Jianfeng Ye, and Bao Peng

Shenzhen Institute of Information Technology,
Shenzhen 518029, Guangdong, China
gmx2020@126.com

**Abstract.** An intelligent wireless channel allocation algorithm for HAPS 5G systems based on reinforcement learning was proposed. Q-learning reinforcement learning algorithm and the back-propagation neural network were combined, which made HAPS 5G systems autonomous learn according to the environment and allocate channel resources of the system efficiently.

**Keywords:** High altitude platform system · Reinforcement learning · Artificial intelligence

## 1 Introduction

The future 5G network will be an intelligent system with multi-service, multi-access technology and multi-level coverage [1–3]. The purpose of development of HAPS is to provide supplementary wireless services for ground stations and satellites. The application of 5G key technologies in HAPS communication systems has many outstanding advantages.

A wireless dynamic channel allocation algorithm for HAPS communication based on distance decision is proposed, which guarantees the quality of service of all kinds of services and maximizes the resource utilization ratio of high altitude platform communication in reference [4]. Aiming at the problem of horizontal swing caused by stratospheric crosswind on high-altitude platforms, a channel allocation algorithm combining channel reservation with handoff queuing is proposed to solve the problem of handoff between cellular for the ground calling users to continue to obtain reliable services in reference [5]. The algorithm takes full account of the service level requirements of different types of user terminals, differentiates the priority of user terminals, and queues the handoff callers on the basis of channel reservation from the point of view of reducing the handover failure rate.

This paper proposes an intelligent wireless channel allocation algorithm for HAPS 5G communication systems based on reinforcement learning, which adopts Q-learning reinforcement learning algorithm in artificial intelligence algorithm and combines back-propagation neural network to enable HAPS 5G communication system to learn


M. Jia et al. (Eds.): WiSATS 2019, LNICST 281, pp. 648–655, 2019.
https://doi.org/10.1007/978-3-030-19156-6_61

independently according to environment, intelligently according to channel load and blocking condition. The channel resources are allocated effectively in the system.

## 2 Q-Learning Reinforcement Learning Algorithm

Q-learning is one of the most famous algorithms in the field of reinforcement learning. It learns how to choose the next action (a) by perceiving reward and punishment (r). The detailed algorithm steps are as follows:

1. for each state s and action a, the initialization table Q (s, a) is 0.
2. observe the current state s;
3. Repeat it all the time.
   (1) select an action a and execute it.
   (2) receive an immediate return of r;
   (3) observe the new state s′, and update the table item Q according to the following form:

$$Q(s,a) = r(s,a) + \gamma * \max Q(s',a'), s = s' \tag{1}$$

This paper proposes an intelligent wireless channel allocation algorithm based on reinforcement learning for HAPS 5G communication system, which uses Q-learning reinforcement learning algorithm in artificial intelligence algorithm and combines back-propagation neural network to enable the HAPS 5G communication system to learn independently according to the environment and intelligently according to the channel, load and blocking. Agents perceive the state information in the channel environment through continuous interaction with the channel environment, learn from the environment state to the action mapping, and use the back-propagation neural network for learning training, use the neural network instead of the Q value table, and train the network with each Q update as a training example to update the evaluation function, and repeat cycle iteration until convergence condition is satisfied, then stop learning.

## 3 Intelligent Wireless Channel Allocation Algorithm Based on Q-Learning Reinforcement Learning

The channel assignment problem of HAPS 5G communication system is solved based on Q-learning reinforcement learning algorithm. The channel assignment problem is modeled as a Markov process, which generates an instantaneous return value at each step of learning, and the state converges at the end of learning. Therefore, in order to realize the algorithm modeling, the instantaneous return value R, channel state S and channel assignment action A must be determined.

(1) Instantaneous return value R

The following principles must be satisfied for intelligent channel allocation:

(a) in the case of existing channel resources, all channels are allocated and the fairness principle is satisfied.
(b) the channel allocation satisfies the outage rate and the minimum principle of GoS (Grade of Service).
(c) channel assignment satisfies the minimum principle of blocking rate.

Therefore, the instantaneous return value of the intelligent channel allocation algorithm is designed to achieve convergence according to the above principles:

- if the a, b and c principles are met, the instantaneous return value of the channel assignment is R = 10.
- if we only meet the a and b principles and do not meet the c principles, then R = 7;
- if we only meet the a and c principles and do not meet the b principles, then R = 5;
- if we only meet the a Principle and do not meet the b and c principles, then R = 3;
- if we do not meet the a principles and satisfy only b and c principles, then R = 0;
- if three principles are not satisfied, then R = −10.

(2) Channel state S

Channel state represents the quality of the channel and the usage of the channel and the idleness of the channel in each period of time before the channel is allocated. The state set of the current allocated channel can be known through the channel state information.

The channel allocated by the intelligent channel assignment algorithm at the same time must meet the following requirements:

(a) Free channel resources;

The number of channels provided by the system is greater than the sum of the peak number of channels needed by each cell in a cluster to satisfy its respective service levels. In this case, the user's demand does not reach the system capacity, so only a certain number of channels need to be allocated to each cell to meet the demand, and the remaining number of channels can be set as a dynamic allocation part, so the enhancements learning algorithm can be used to allocate the free channel.

(b) Scheduling time does not conflict;

In the training phase, a conflicting coefficient is obtained by recording the channel in which a conflict occurs at a certain scheduling time. After the training phase is over, a channel conflict distribution table can be obtained. The conflict coefficient of non conflict scheduling time is 0, and the more conflict time, the greater the conflict coefficient.

(c) Channel quality

According to the channel estimation, the channel quality of the idle channel can be divided into three levels according to the channel quality from high to low:

The best channel quality = 10;
The qualifying channel quality = 5;
The worst channel quality = 0;

(d) GoS (Grade of Service)

According to the requirement of GoS, the priority of channel assignment can be divided into the following 4 categories:

Emergency business level, Level = 100;
High priority business level, Level = 50;
Medium priority business level, Level = 30;
Low priority business level, Level = 10.

(3) Channel assignment action A

Channel allocation action is to select which channel to allocate among the free channel resources, and it also needs to reflect the service level information.

We use a 5-bit binary representation in which the lowest bit denotes whether the channel is allocated, if allocated, is 1, otherwise 0. The 2nd and 3rd bits represent the channel quality, with the best quality is 10, the qualifying quality is 01, and the worst quality is 00, 11 reserved. The 4th and 5th bits are the service level, the emergency service is 11, the high priority is 10, the middle priority is 01, and the low priority is 00.

In this way, we discretize the channel state into the above four variables, the total number of channels is N. So the channel state table of the intelligent channel allocation device contains 4 * N elements, called channel state mode matrix.

The learning process will converge only if each idle channel state and allocation action are used indefinitely and frequently because of the complex and changeable wireless channel environment and the variety of wireless services and the mobility and uncertainty of users. In the 5G era, there will be a large number of traffic connections, and the state-action pair of channel assignment problem will be a huge state space. It is difficult to search such a huge space in practice and it is almost impossible to get all the state-action Q-value tables. So in this case, in order to make the reinforcement learning algorithm achieve the desired effect, we choose to use the back propagation neural network to quickly obtain the estimate of Q value. The neural network is used to replace the Q value table, and every Q update is used as a training example to train the network. When training the intelligent channel assignment BP network, we can quantify the channel state S and take it as the first input of the neural network. Then the neural network finally outputs an estimate of Q value, and compares this Q value with the Q value obtained from the previous learning, and trains the BP network to get the expected Q value.

Intelligent channel allocation BP network is divided into three layers, the number of input layer units is 4 * N channel state. The third layer output layer is only one and the number of hidden layer neural units is chosen 32. All levels of neurons form a fully interconnected connection and hidden layer is S-shaped transfer function and output layer is linear transfer function.

The initial weight matrix of BP network in this paper is:

$$W^2 = 4 * N * 32, W^3 = 32 * 1 \quad b^2 = 1 * 32, b^3 = 1 * 1$$

Initial weights are chosen randomly in (0,1) to avoid possible saddle points without leaving the flat area of the performance surface.

A lot of training data will be generated during our continuous training in the system. Although these data are not the best strategy for dealing with the environment at that time, it is the experience gained by interacting with the environment that is very helpful to our training system. So we set up a replay_buffer to save new interactive data to overwrite the old data, and each time randomly take a batch from the replay_buffer to train our system.

Each record in replay_buffer contains the following:

(a) state; the channel status of the current device;
(b) action: the behavior of our agent in the current state;
(c) reward: the profit from the environment after agent has made the choice behavior;
(d) next_state: the next state that the agent transferred after the agent has made the choice behavior;
(e) done: the flag to indicate if the training is ok.

# 4 Performance Comparison and Analysis

## 4.1 Establishment of Simulation Environment

Next, we will simulate the algorithm in this paper to verify the performance of the algorithm. The simulation model used in this article is shown in Fig. 1. Using a typical 4-platform 32-channel model, the simulation area consists of seven cellular cells, each of which is hexagonal in size, and the seven cells studied are all located in the inner ring area covered by four high-altitude communication platforms. The antenna gain mode meets the ITU standard for high-altitude platform communication. It is assumed that the mobile users distribute uniformly throughout the service area and the antenna of the mobile users points to the high-altitude platforms they access without bias. The transmitting power of each mobile user is the same. The propagation environment obeys the law of free space link loss. We analyze the network performance using the intelligent channel allocation algorithm in a mixed service environment. We choose several business scenarios from the main application scenarios of 5G:

(1) Traffic 1: cloud AR/VR business. VR/AR requires a large number of data transmission. The quality of the channel determines the quality of VR/AR video data transmission. 5G network needs to allocate the corresponding quality level of the channel based on different types of AR/VR services and different environments;
(2) Traffic 2: Vehicle networking business. It includes traditional cars, remote control driving and unmanned autopilot. Vehicle life cycle maintenance, sensor data packets, etc. require secure, reliable, low latency and high broadband connections, which are essential in highways and dense cities. 5G network needs to assign the corresponding quality level and service priority channel according to different types of vehicles in different environments.
(3) Traffic 3: voice business. There are still a large number of voice services in 5G network. 5G network needs to allocate a large number of voice services optimally in the case of channel collision and sudden emergence of high priority services.

**Fig. 1.** HAPS simulation model

## 4.2 Performance Comparison

We choose two classical channel assignment algorithms in HAPS communication system as the comparison algorithm: random channel allocation algorithm [6] and Worst channel acceptable channel allocation algorithm [7].

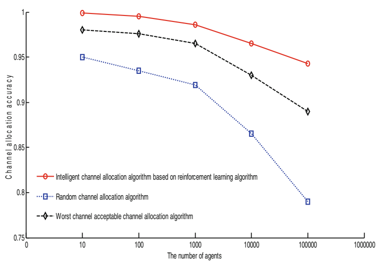


**Fig. 2.** Channel allocation accuracy of traffic 1

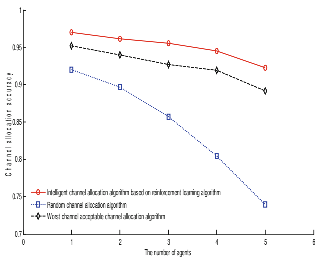


**Fig. 3.** Channel allocation accuracy of traffic 2

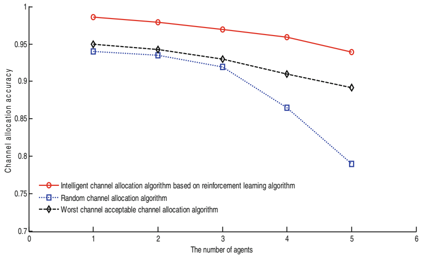


**Fig. 4.** Channel allocation accuracy of traffic 3

Figures 2, 3 and 4 show the channel allocation accuracy of traffic 1, traffic 2 and traffic 3 under three channel resource allocation algorithms respectively. As it can be seen from Figs. 2, 3 and 4, the channel allocation accuracy of the three channel resource allocation algorithms decreases with the increase of the number of agents in the network, especially for the random channel allocation algorithm, when the number of agents increases to a certain number, the decrease is especially obvious. The channel allocation accuracy of the intelligent channel allocation algorithm based on reinforcement learning algorithm is higher than the other two algorithms for different traffic. Even if the number of agents in the network is very large, the channel allocation accuracy is still very high. This is because the algorithm can dynamically adjust the channel allocation according to the channel quality and priority of different services and select the best channel for the current service channel quality and service level.

## 5 Conclusions

In this paper, an intelligent wireless channel allocation algorithm based on the reinforcement learning algorithm for HAPS 5G communication system is proposed, which uses Q-learning reinforcement learning algorithm and combines back-propagation neural network. Finally the network performance of the proposed algorithm is compared with the random channel allocation algorithm and the worst acceptable channel allocation algorithm. The channel allocation accuracy of the proposed algorithm is higher than the other two algorithms for different traffic. Even if the number of agents in the network is very large, the channel allocation accuracy is still very high. It effectively improves the overall performance of the system.

**Acknowledgement.** This paper is supported by the national natural science foundation of China (61401288), the Guangdong Province higher vocational colleges & schools Pearl River scholar funded scheme (2016), the project of Shenzhen science and technology innovation committee (JCYJ20170817114522834,JCYJ20160608151239996), the science and technology development center of Ministry of Education of China (2017A15009) and Engineering Applications of Artificial intelligence Technology Laboratory (PT201701). The author would like to thank the editor and the anonymous reviewers for their contributions that enriched the final paper.

## References


1. Cao, Z., Zhao, X., Soares, F.M.: 38-GHz millimeter wave beam steered fiber wireless systems for 5G indoor coverage: architectures, devices, and links. IEEE J. Quantum Electron. **53**(1), 1–9 (2017)
2. Yuan, Y., Zhao, X.: 5G: vision, scenarios and enabling technologies. ZTE Commun. **13**(1), 69–79 (2015)
3. Cid, E.L., Táboas, M.P., Sánchez, M.G.: Microcellular radio channel characterization at 60GHz for 5G communications. IEEE Antennas Wirel. Propag. Lett. **16**(99), 1–4 (2017)
4. Guan, M., Guo, Q., Gu, X.: A dynamic wireless channel allocation algorithm in HAPS communication based on distance verdict. Electron. J. **41**(1), 18–23 (2013)
5. Jiang, J., Zhang, B., Guo, D., Ye, Z.: A channel assigning algorithm for platform displacement model in HAPS communication system. Telecommun. Eng. **55**(8), 906–912 (2015)
6. Grace, D., Spillard, C., Thornton, J., Tozer, T.C.: Channel assignment strategies for high altitude platform spot-beam architecture. In: The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, vol. 9, pp. 1586–1590 (2002)
7. Pace, P., Aloi, G., De Rango, F.: An integrated satellite-HAP-terrestrial system architecture: resources allocation and traffic management issues. In: IEEE 59th Vehicular Technology Conference, vol. 5, pp. 2872–2875 (2004)