

Towards combining validation concepts for short and long-term ambient health monitoring

Sebastian Chiriac, Natalie Röhl, Javier Parada, Bruno Rosales

FZI Research Center for Information Technologies

Embedded Systems and Sensors Engineering

Karlsruhe, Germany

{chiriac, roell, parada, rosales}@fzi.de

Abstract— Existing health monitoring systems detect either acute health issues or long-term deterioration of health. This paper presents a validation concept for sensors and algorithms that detect emergencies and monitor long term health development. The reviewed ambient monitoring system combines a short term rule based and a long term scoring approach that are validated in a study with 100 households. The resulting dataset is complemented by diaries and participation of care givers. Care givers monitor the health state through questionnaires and by rating generated alerts and warnings. The validation benefits from this multilateral approach and points towards a further integration of long and short-term monitoring.

Keywords- AAL, health assessment, situation recognition, rule-based monitoring, system validation

I. INTRODUCTION

The use of ambient sensor technology in the area of health care and assisted living [1, 2 and 5] shows the big potential that is seen by the ambient intelligence community. Applications of such sensor technology can be divided in two groups: short and long-term monitoring. While short-term monitoring is targeting especially emergency detection and the recognition of activities of daily living (ADL), long-term situation recognition tries to focus on deviations from usual patterns [8] and health assessment.

Despite the fact that there are studies conducted in the ambient assisted living field [3, 4, 6 and 7], the validation of the benefit of ambient sensors in the area of health monitoring is not yet proven by a significant number of participants involved in a test bed on a long term basis.

The following paper presents two concepts for validating short- and long-term health monitoring systems based on ambient sensor technology. Both concepts include questionnaires to receive information about the actual health status, which are conducted by care givers as a medium to react on warnings and unusual behavior. The aggregated information is used as a reference for long-term sensor data. Both validations will be conducted in the same test bed and therefore rely on the same sensor data.

II. BACKGROUND AND RELATED WORK

Possible occurring problems in everyday life and their characteristics, by which they are distinguished, are described in detail by nursing diagnosis [9]. As a basis for our health monitoring system we used the nursing diagnosis to know

which areas are important to monitor. Therefore we developed a list where nursing diagnoses were connected through their characteristics with a specific sensor and its location (Fig. 1). We clustered the nursing diagnosis in four main areas: health, mobility, cognition and social interaction (TABLE I).

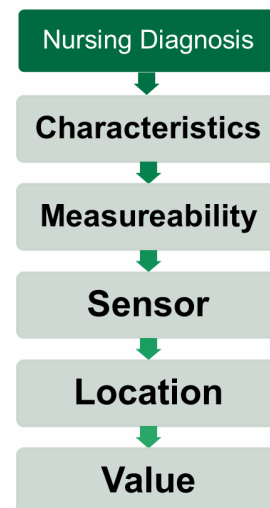


Figure 1. Sensors derived from Nursing Diagnosis

As a next step, we matched the specified sensors with the selected home automation system presented in [10]. The matched sensors with their location are the basis for the deployment of sensors in the home environment.

TABLE I. MONITORING AREAS

Area	Sub-Area
Health	Sleep
	Hygiene
	Dressing
	Nutrition
Mobility	Gait
	Stair Climb
Cognition	Short-term Memory
	Disorientation
Social Interaction	-

Based on this sensor network we build an ambient health monitoring system. First we will present the monitoring system based on rules for situation recognition as well as long-term health assessment. We discuss how assessment tools can be used to aggregate reference data. Related studies in the ambient monitoring environment are shortly described at the end of this section.



Figure 2. Sensors in Real-Home Environment

A. The ambient health monitoring system

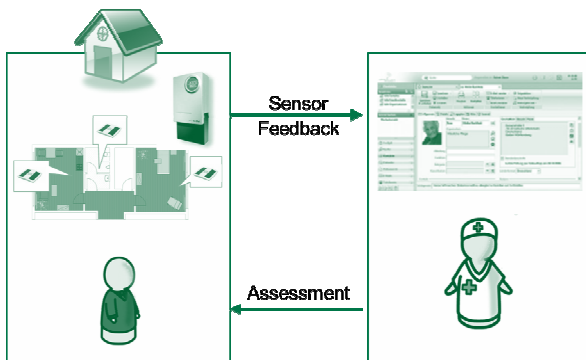


Figure 3. Ambient Health Monitoring System

The ambient health monitoring system is based on the analysis of the nursing diagnosis and the corresponding selection of sensor types. The sensors are deployed in the home environment and are connected wirelessly to a sensor gateway. The data processing is done on the gateway. The abstracted sensor feedback is transmitted to a visualization software for care givers. The care givers on the other hand conduct assessments by interviewing the test persons personally and via phone. Because it is not possible to detect on a large scale ADLs accurately, the system serves two monitoring purposes: short-term situation recognition and long-term health monitoring. Both approaches have the advantage that the reference data is not needed accurately. On a higher level of abstraction, the health of the subjects has to be linked with behavioral data. Assessments will determine the health state of

a subject. This can be used to determine what information and indicators data is hidden in the abstracted sensor data. Next the short-term situation recognition will be described, followed by the long-term monitoring with health assessments.

1) Short-term: Rule-based situation recognition

Considering there are two main short-term recognition techniques, rule-based and learning algorithms [11], we have chosen the former based on the following arguments: (1) the intention was to keep the system simpler; the use of simple sensors provides a simple data set, lacking of noisy and uncertain data. Rule-based systems can handle correctly concrete situations [11], such as standing up after sleeping, and are computationally simpler than learning algorithms, keeping the trade-off between complexity and recognition success. (2) Flexibility: our system is ontology-based containing different levels of information abstraction. This allows us to add or modify rules on different information levels easily, providing flexibility when some rules do not work properly or new functionality needs to be added. (3) No references needed: learning algorithms need normally a training phase where precise annotations are needed, which is usually implemented by additional camera surveillance [12]. Sometimes even the training must be particular for every household, which makes a large field test unfeasible.

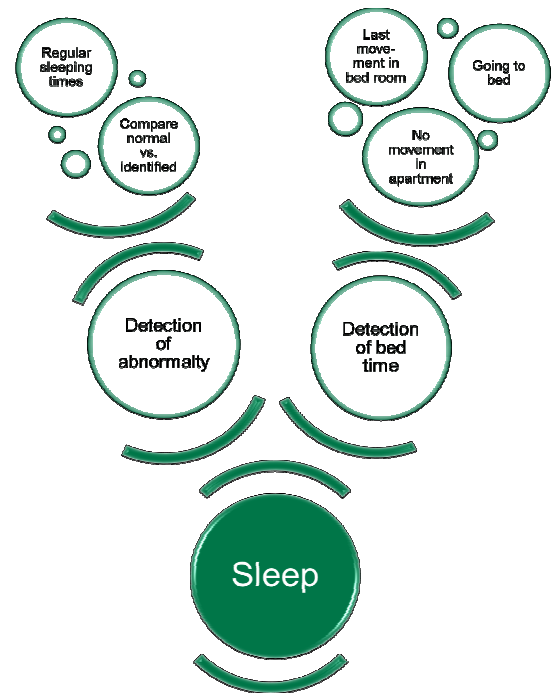


Figure 4. Detection of Sleeping Problems

The ambient middleware openAAL is the basis of the monitoring system [13]. OpenAAL is a framework on top of OSGi which allows easy configuration of AAL-services and communication between those. It especially supports easy uplifting of low-level sensor data to high-level context information. The low-level ontology statements describe the state-based sensor information, whereas the vocabulary for

environment, locations, situations and activities match the generated high-level information.

This way rule-based algorithms can detect complex situations based on the sensor data and context information. The rules for the monitoring system and the required parameters were developed based on the characteristics of the nursing diagnosis and the defined areas of health monitoring, which then were applied in the openAAL middleware.

To illustrate how the rules work on the different information levels we consider the example of detecting sleeping problems (Fig. 4): from the very low level, the basic rules abstract the motion sensor information, obtaining the persons location, i.e. person is in bedroom. On the next level, using the location and sensor data, the system recognizes when the person goes to sleep and wakes up. One level above, the system calculates the sleeping times of the person during the day, based on the uplifted information, and with defined context information (sleeping times) detects if punctual sleeping problems arises, i.e. person slept too short last night. Finally at a very high level, the rules infer if the punctual problems recur enough to consider the problem relevant to report the warning to the caregivers, i.e. the person slept too short the last 3 days. The process is a combination of rules, where the output is used concurrently by other rules, abstracting more and more the information until desired information is achieved.

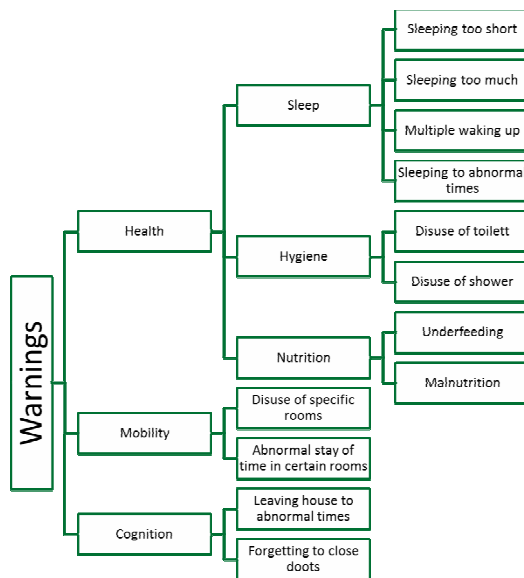


Figure 5. Warnings structured by Monitoring Areas

The rules describe deterioration in one of the clustered areas and indicate a decrease in autonomy. In this context the broad areas were broken down to complex problematic situations resulting from the nursing diagnosis and their characteristics. This offers the possibility to generate specific warnings for each detected problematic situation and appropriate additional information.

The carer receives information about the health status through the rules integrated in the openAAL-middleware

described above. The system simply visualizes the problems which were detected by the rule-based technology. The visualization of the data is implemented in a web based care and case management software (CareCM) as user interface. Case Management is a procedure to plan, coordinate, monitor and evaluate services for different social agencies and staff on behalf of a client. The software allows care givers to efficiently coordinate and keep track of the patients, relatives and other involved stakeholders.

Warnings are sent from the middleware immediately to the CareCM once the system has detected a problem. It informs the nursing staff about new warnings, unusual behavior and emergency situations. They can be viewed in real time to ensure an early response to status changes.

As mentioned before the rule-based system is responsible for the short-term situation recognition. Problematic situations can be detected as they occur. Extending this implementation the following additional health monitoring approach was realized to detect deterioration on a long-term basis.

2) Long-term: Health assessment

In order to offer patients the right treatment health assessments are usually conducted by nursing staff and by doctors. The sensor system should support these assessments and improve the quality of the treatment and reduce workload for the nursing staff. To prove these advantages the recorded sensor data needs a ground truth. This makes it possible to relate long-term deviations with feedback given by test persons and carers. Therefore we analyzed two commonly used assessment instruments, the Resident Assessment Instrument and the Comprehensive Geriatric Assessment. We targeted the question of usability in our study setup [17].

a) Resident Assessment Instrument

The Resident Assessment Instrument (RAI) was designed as an assessment tool for long-term care of patients by providing a standardized summary of their sensibilities. It also serves to create a perfectly tailored patient care plan and measuring the effectiveness of the treatments delivered to the patient [14]. RAI consists of the following parts: Minimum Data Set (MDS), the Trigger System and the Resident Assessment Protocols (RAPs)

The Minimum Data Set (MDS) serves a comprehensive impression of the health of the patient. It collects over 300 items in 17 categories, e.g. cognitive patterns, mood, behavior. Because the RAI was designed as a tool for long-term care, a repetition of the MDS data recording is provided every three months [14].

The Trigger System is based on the data obtained from the MDS. If a certain area exceeds a predefined limit then the area is considered an issue. If e.g. incontinence occurs more than once a week, the area of continence is a marked as an problematic area.

The Resident Assessment Protocols (RAPs) serve as a guide to estimate the impact of those areas registered by the trigger system on the care plan of the patient.

b) Comprehensive Geriatric Assessment (CGA)

The Comprehensive Geriatric Assessment (CGA) covers a variety of areas and offers a specialized assessment for each

area. Among others the following assessment tools are used: Screening according to Lachs, Barthel Index, Clock Completion Test, Geriatric Depression Scale, Geriatric Hamilton Depression Scale, Mobility Test according to Tinetti, Mini Mental State Examination (MMSE).

The CGA provides a very comprehensive look at the health of a patient. Despite the nine different methods that are used, the cost of the geriatric assessment, especially when compared to several days of data collection of the RAI, is quite moderate. Each test only takes up some minutes.

c) Conclusion

Both presented tools give a very good insight into the health state of person. They are very well validated in various studies [16]. Two main problems lead us to the decision to derive a questionnaire from both instruments. Firstly, RAI has a very high workload with about four days per patient to aggregate all the necessary data [15] and neglects completely the view of the patient. Secondly, the individual tests of the CGA have to be carried out by different persons due to the different areas.

With the assessment we can understand the impact of deviations of the health state on the long-term sensor data. The long-term validation concept will explain how these two measures are compared to each other.

3) Other studies

Besides our study, in the area of ambient assisted living other studies are undergoing using ambient monitoring systems. CASALA analyses 13 homes occupied by 11 men and 4 women, using a total of 2240 home automation sensors and actuators. In a first step they compared movement data to baseline depression and to mobility data over a 3-month period. [4]. The Emerge project equipped 2 flats over 3 months with multi-sensor nodes to evaluate the ADL recognition and inactivity detection [3]. The eHome project evaluated an engine for rule inferring and ADL recognition with data from 11 apartments over a total of 553 days [19].

III. METHOD

Based on our developed health monitoring system, two validation concepts using the same test bed are pursued in a long-term study. First a concept for the validation of the short-term monitoring is presented. Furthermore we will explain a validation concept for long-term monitoring. These are two different approaches for the recognition of the actual health status and deterioration. The study itself consists of 100 households and is performed in time frame of 18 months. At the moment 60 households are connected to our health monitoring system.

Our target group consists of households where elderly live on their own and are at least 65 years old. The test persons are largely independent in their mobility and mostly live on their own. An hourly care by relatives or nursing services is no exclusion criterion. The participants receive a free home emergency service and may live in an assisted living facility.

The study concept [17] shows that only one dedicated carer is considered for the support of each test person. Due to the constraints we derived an interview guideline with approximately 80 questions. The interview has exactly

verbalized questions and thus offers a good guidance for the interviewer. The areas of the questionnaire are matched with nursing diagnosis showed in section II. At the beginning and at the end of the study, personal interviews are conducted. In monthly telephone interviews abbreviated questionnaires are performed. It is shown that the reliability of assessment tools can be kept at a high level also through questionnaires via phone and mail [18].

The benefit of ambient sensors in health monitoring can be shown through a validation. Below we describe how the system with the rule-based situation recognition and the long-term monitoring are validated.

A. Concept for short-term validation

Despite the fact, that the rules for the short-term monitoring system were derived from the nursing diagnosis, it is important to evaluate the accuracy and validity of the rules.

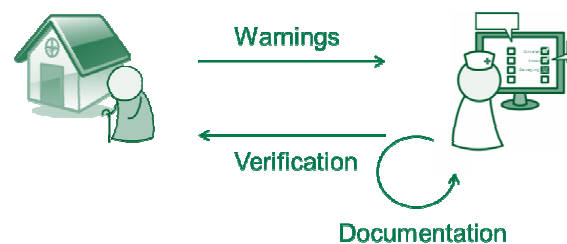


Figure 6. Concept for Short-Term Validation

For the short-term validation, sensor data is aggregated through the ambient monitoring system and are visualized as warnings in the case and care management software presented in section II. Each warning consists of a notice about the problem and additional detailed information that will help the user to classify the severity of the problem and to react appropriate to the detected situation.

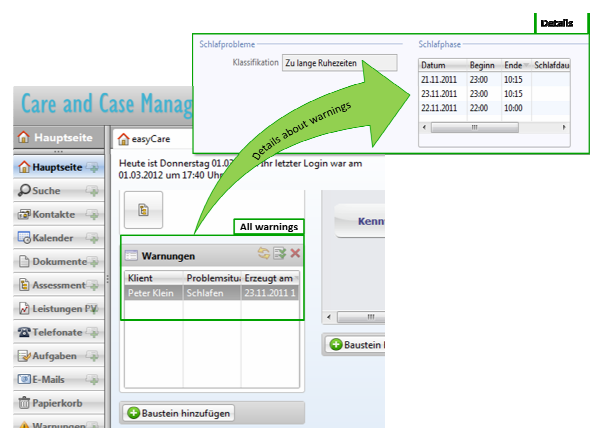


Figure 7. Visualisation of Warnings in CareCM Software

For the validation of the system it is important to identify the validity of the warning. The occurrence of the detected problem has to be checked as well as the validity and the information content of the warnings.

If a certain situation occurs, such as sleeping problems over a longer period, the nursing staff intervenes by supporting the problem as well as verifying the system warning. The software

enables the nursing staff to administer multiple monitored persons and the relating alerts at the same time. Thus, it is easier for the nursing staff to react to the warnings in terms of the validation. The consolidation of the assessment data and the sensor information in the case management software allows to rate the current health status and the implicated problems of the patient.

Twice a day the nurse checks the system and if a warning occurred, gives feedback on the real condition through the care and case management software after checking on the test person personally or via phone.

Particular interesting for us is, if the warning is related to a real condition or if it is a false alarm. This enables the developers of the system to adapt and improve the system online.

With the introduced feedback system we can prove the validity of the short-term health monitoring system.

B. Concept for long-term validation

The validation for the long-term monitoring with health assessments is based on questionnaires conducted by the carers. The assessment consists of personal interviews at the beginning and at the end of the study and telephone interviews every two months. To validate the sensor data, comparable values have to be defined. Therefore in both data areas, a score is calculated for the questionnaire and for the sensor data based on the four main areas derived from the nursing diagnosis.

1) Questionnaire Score

The first score abstracts the answers of an interview in a way that questionnaires of different persons can be compared easily and also changes between questionnaires of the same person can be automatically visualized. In our case, we try to abstract the questionnaire data to compare it with the results of the sensor data. The calculation of the score consists of three parts.

- **Assign a value to an answer:** Each response can receive a value between 1 and 5 (1 being the lowest score and 5 being the highest score) (TABLE II.). For example a response to the question "On how many days have you left the house in the last week?" is rated with 1 for "never" and with 5 for "daily".
- **Assign a weight to each question:** The questions are assigned to specific categories and sub-categories [Sec. II.]. Each question can be assigned to several categories, e.g. to the question "On how many days have you left the house in the last week?" both categories, mobility and social interaction, are assigned. The weight of each question to a category can be set between almost not important (1) to very important (5). In the above example "mobility" is weighted by the value 5, since mobility is needed to be able to leave home very often. In the category "social interaction" the weight is 4, because social interaction works even with limited mobility, for example, if the person is visited at home and they interact via phone / internet.
- **Calculate the score:** The score is calculated first by the value of the response multiplied by the weight.

Then all weighted response values of a category are summed up and divided by the sum of the weights of the category. Hence a score for each category can be calculated. If the answer to the question is refused, the question will not be considered in the rating.

TABLE II. STRUCTURE OF QUESTIONNAIRE

Number	Question	Answer	Categories
10	On how many days have you left the house in the last week?	- Daily (5) - Several times a week (4) - Once a week (3) - Less than once a week (2) - Never (1)	Social Interaction (5) Mobility (4)
22	Do you usually walk without help or do you need a cane, crutch or walker?	- Without help (5) - with cane or crutch (3) - with walker (1)	Mobility (5) Gait (5)

The SCORE of the questionnaire is calculated by the following formula:

$$Score_i = \frac{\sum_{j \in \text{Answers of Category } i} (\text{Answer Value} \times \text{Weight})}{\sum_{j \in \text{Answers of Category } i} (\text{Weight})} \quad (1)$$

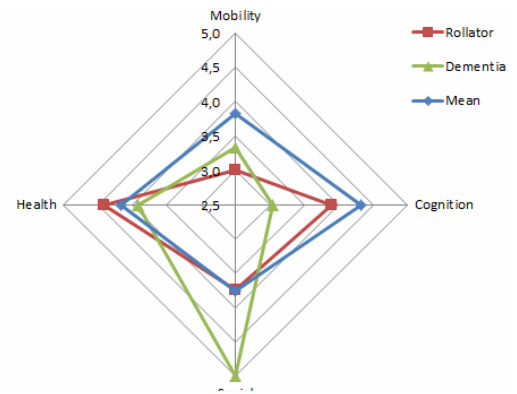


Figure 8. Example for Score of two Participants

In figure 8 the score of two test persons is visualized: one participant with dementia and one participant in need of a rollator. The blue line shows the mean value which is calculated out of the first 15 test persons. The participant with the rollator achieves a value in the area mobility which is much lower than the mean. The reduce perception of the person with dementia is well reflected in the score of the area cognition.

2) Sensor Score

After developing a score for the questionnaire, a sensor specific score will need to be emerged through long-term sensor analysis. The triggering level of each sensor matched with a category adds up to the score. Through the matching of the nursing diagnosis, every single sensor can be traced back to

the area, e.g. the contact sensor at the entrance is assigned to social interaction and to the mobility score.

The trigger level of the different areas is then standardized to a score between 1-5 (1 being the lowest score and 5 being the highest score). In the example shown in figure 9 a participant with cancer is displayed over a duration of 300 days. The participant got better over the first 150 days because of a surgery. The cancer got worse again in the second part of the recorded data. The health issues also affected the social interaction of the person by not going out as much as before. During the validation the sensor score needs to be refined and evaluated with other participants.

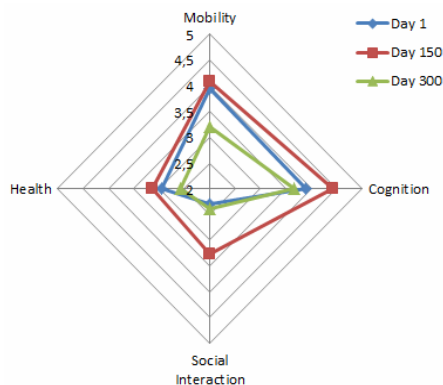


Figure 9. Example of Sensor Score

3) Validation of Sensor Health Assessment

By correlating the scores from the sensors and the questionnaire over the time frame of the study, changes of health and deterioration can be compared. Hence, it can be proven if the ambient monitoring system can supply information about the current health status of a person and support the health assessment on a long-term basis.

IV. RESULTS

First results of the validation show a promising outcome of the study. But in the course of the study some adjustments had to be made.

The validation concept based only on nursing staff was not sufficient. In most cases the warnings are displayed when an accumulation of a problem arises. Only then is it necessary to intervene to prevent health hazards. This approach should help to avoid an information overload.

Since this implementation does not allow rule verification on a lower base it was necessary to include the patients in the validation. Because the warnings are generated by the monitored patients, they are the ones, who can provide most information about their problems. Based on this knowledge additional diaries were developed. These diaries provided a set of questions about changes in the health status, based on the mentioned assessment tools, as well as questions supporting the rule validation. The patients are asked to complete the questionnaire daily if defined events (e.g. frequent waking up) occur. This provides additional reference data about the situation the system wants to detect.

It was found, that the patients do not want to be monitored and do not want to fill out questionnaires without receiving any feedback. A possible solution could be to transfer the paper-based diary with feedback functionality to an interactive technology such as a tablet pc.

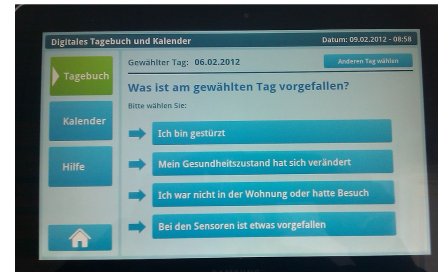


Figure 10. Tablet with Interactive Diary

In the first implementation of the rule-based system many parameters were predefined. For example, it is necessary to define normal sleeping times in order to recognize abnormalities. The response of the nursing staff associated with the occurred false warnings made it clear, that the system needs to be designed more flexible. It needs an interface to parameterize the rules individually for each patient and their habits. Only in this way the system can give correct information about the health status.

We had concerns about the amount of questions to be asked during the interviews. Some reviews with professionals were iterated and the feedback showed that the questionnaire was well developed and the amount of time (around 1h) is matched pretty well with a mean of 54 minutes.

The questionnaire score itself showed variations between the participants. The disabilities remarked by the nurses were very well reflected as shown in Fig 8. The sensor score structure is defined but needs much more refinement as there is not enough experience with more participants over longer time range.

V. DISCUSSION AND OUTLOOK

We defined two validation concepts for short- and long-term health monitoring using ambient sensor technology. The advantages of ambient technologies can be quantified with the validation. Nonetheless some constraints have to be made. The questionnaire is based on validated assessment instruments but is not a validated tool itself. The correlation between the sensors and the questionnaire is not guaranteed. Most questions have to be answered by the participant therefore the response is subjective. We overcome this issue by introducing some questions that are targeted directly to the interviewer. The interviewer expresses his observations e.g. the movement around the home. Is he or she safe or is the participant insecure?

A general issue is the occurrence of health changes during the study. If we record very few cases of emergency or deterioration, than we can only take very few conclusions from the validation. With a time frame of 18 months and 100 participants the probability is high that some interesting cases are evaluated.

For a long-term use of the case and care manager software the information content of the warnings must be ideal so that the nursing staff can intervene adequately. The feedback from the short-term monitoring will show, if the level of detail in the displayed warnings needs to be extended.

In the future we will further improve the system by including also data from smart meters which will aggregate the power consumption of every household. Smart Meters show through the use of different appliances the activity of person throughout the day. This will contribute to a refinement of the sensor score as well as an improvement of the rule accuracy.

The positive outcome of both validations influences the use of the health monitoring system. A wider variety of impairments can be detected. Emergencies as well as long-term health problems like dementia can be handled by the same system. To improve the quality of the results both approaches can be validated by the other one. Are e.g. warnings about sleeping problems also influencing the score of the person? Or is a lowered score reflected in short-term indications?

ACKNOWLEDGMENT

This work was sponsored by the German Federal Ministry of Education and Research (BMBF) in the context of the projects optimAAL and EasyCare.

REFERENCES

- [1] M.-R. Tazari, R. Wichert and T. Norgall, "Towards a unified ambient assisted living and personal health environment", in: Fourth German AAL Congress, Berlin, Germany, 25 and 26 January 2011.
- [2] M.-R. Tazari, F. Furfari, Á. Fides Valero, S. Hanke, O. Höftberger, Di. Kehagias, M. Mosmondor, R. Wicher and P. Wolf, "The universAAL - Reference Model for AAL", Handbook of Ambient Assisted Living, J.C. Augusto et al. (Eds.), IOS Press, 2012
- [3] T. Kleinberger, A. Jedlitschka, H. Storf, S. Steinbach-Nordmann, "Evaluation of ADL Detection in the EMERGE project.", Third German AAL Congress, Berlin, Germany, 2010.
- [4] B. O'Mullane, B. Bortz, A. O'Hannlon, J. Loane, and B. Knapp, "Comparison of Health Measures to Movement Data in Aware Homes. Response", Aml 2011, Amsterdam, LNCS 7040, pp. 290-294, 2011.
- [5] B. Kröse, T. van Oosterhout and T. van Kasteren, "Activity Monitoring Systems in Health Care", A.A. Salah, T. Gevers (eds.), Computer Analysis of Human Behavior, pp. 325-332, 2011
- [6] P. Mayer, M. Rauhala, P. Panek, "Field test of the eHome system", 4th German AAL-Congress, 2011
- [7] U. Gaden, E. Löhrike, M. Reich, W. Schröer, T. Stevens, T. Vieregge, "SAMDY – Ein sensorbasiertes adaptives Monitoringsystem für die Verhaltensanalyse von Senioren", 4th German AAL-Congress, 2011.
- [8] H. Storf, M. Becker, M. Riedl, "Rule-based activity recognition framework: Challenges, technique and learning," Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009. 3rd International Conference on , vol., no., pp.1-7, 1-3 April 2009
- [9] Th. Herdman, "Nursing diagnoses: definitions and classification 2009 - 2011," Wiley-Blackwell: Singapore. 2009.
- [10] S. Chiriac, B.R. Saurer, G. Stummer; C. Kunze, "Introducing a low-cost ambient monitoring system for activity recognition," 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 , pp.340-345, 23-26 May 2011
- [11] J. Ye, S. Dobson, S. McKeever, "Situation identification techniques in pervasive computing: A review", Pervasive and Mobile Computing, 2011.
- [12] B. Logan, J. Healey, M. Philipose, E. M. Tapia and S. Intille, "A Long-Term Evaluation of Sensing Modalities for Activity Recognition." Proceedings of the 9th international conference on Ubiquitous computing 4717, pp. 483-500, 2007.
- [13] P. Wolf, A. Schmidt, JP. Otte, M. Klein, S. Rollwage, B. König-Ries, "openAAL - the open source middleware for ambient-assisted living (AAL)", 2010 AALANCE conference, pp. 1-5, 2010.
- [14] H. Brandenburg, "Das Resident Assessment Instrument (RAI). Ausgewählte empirische Befunde und Konsequenzen für die pflegewissenschaftliche Diskussion in Deutschland", Pflege & Gesellschaft, 7th edition, 3, pp. 95-102, 2002
- [15] C. Hawes, J. N Morris, C. D. Phillips, B. E. Fries, K. Murphy and V. Mor, "Development of the nursing home Resident Assessment Instrument in the USA", Age Ageing, Vol. 26 (Supplement 2), pp. 19-25, 1997.
- [16] A. Sgadari; J.N. Morris, B.E. Fries, G. Ljunggren, P.V. Jonsson, J.N. Dupaquier and M. Schroll, "Efforts to establish the reliability of the Resident Assessment Instrument", Age and Ageing, Vol. 26 (Supplement 2), pp. 27-30, 1997.
- [17] S. Chiriac, B.R. Saurer, "An Ambient Assisted Living Monitoring System for Activity Recognition – Results from the First Evaluation Stages", Fifth German AAL Congress, Berlin, Germany, 2012
- [18] P.U. Heuschmann et al., "The reliability of the german version of the Barthel-Index and the Development of a Postal and Telephone Version for the Application on Stroke Patients", Fortschritte der Neurologie – Psychiatrie 2005, 73, p.74-82
- [19] P. Mayer and P. Panek, "An AAL Approach to Status and Activity Assessment by Use of Domain Expert Knowledge based on Sparse Nonintrusive Sensors", Fourth German AAL Congress, Berlin, Germany, 25 and 26 January 2011