Behavior prediction in home telecare systems

Jose Manuel Lopez-Guede\textsuperscript{a,\,d}, Aitor Moreno-Fernandez-de-Leceta\textsuperscript{b}, Manuel Graña \textsuperscript{c,\,d}

May 31, 2015

\textsuperscript{a}Dept. of Systems Engineering and Automatic Control, University College of Engineering of Vitoria, Basque Country University (UPV/EHU), Nieves Cano 12, 01006, Vitoria, Spain, jm.lopez@ehu.eus

\textsuperscript{b}Instituto Ibermatica de Innovacion. Sistemas Inteligentes de Control y Gestion Parque Tecnologico de Alava Leonardo Da Vinci, 9 - 2o - Edificio E5, 01510, Miñano (Alava), Spain, ai.moreno@ibermatica.com

\textsuperscript{c}Dept. of Computer Science and Artificial Intelligence, Faculty of Informatics, Basque Country University (UPV/EHU), Paseo Manuel de Lardizabal, 1, 20018, San Sebastian, Spain, manuel.grana@ehu.eus

\textsuperscript{d}Computational Intelligence Group (UPV/EHU)

Abstract

This paper presents an intelligent system for behavior prediction and personal safety at home environment, principally oriented to elderly people. The system introduces a support decision machine for automatic risk prevention at home that has been tested in real life environments. Accident prediction and odd behaviors prevention of elder people living alone at home has a growing demand. That demand is still not well resolved, or resolved by means of manual monitoring systems in a not effective fashion. The system described in this paper solves automatically this issue, preventing home risks by an advanced analytic method through an expert knowledge system. It is based on the principle of no intrusion, so it uses plug-and-play sensors and machine learning algorithms to learn the elderly usual activity. If the system detects that something unusual happens (in a wide sense), or if something is wrong with respect to its recommendations according to each user’s health habits or medical recommendations, it sends at real-time alarm to the family, care center or medical agents without human intervention. To achieve it, the system uses information from simple sensors at home, knowledge of their physical activities collected by mobile applications and personalized health information based on clinical reports encoded in the system. It is being tested under real-life conditions with respect to their usability and reliability, with an accuracy larger that 81%.

1 Introduction

Ambient Assisted Living (AAL) is defined as the use of Information and Communication Technology (ICT) in intelligent living environments for reacting to the needs of the inhabitants by providing relevant assistance and helping them to live a full and independent life. End users are the stakeholders in the AAL ecosystem: citizens, formal and informal caregivers, service providers, technology providers and policy makers. The beneficiaries will be those people who wish to be able to avoid dependency on nursing homes, preferring to continue to live independently in their own homes. Assistance might be needed in any aspect of daily life, from health safety and security to social integration and mobility support.

The steering board of the European Innovation Partnership on Active and Healthy Ageing (EIP-AHA) asserts [1]: “ICT solutions can prolong independent living of older people and extend the time they remain active and safe in their preferred environment. They also have a huge potential to enhance social inclusion and participation of older people, reduce depression rates, enhance quality of work for carers and make overall care provision economically sustainable (e.g. by avoiding and reducing hospital stays)”. On the other hand, the same document em-
phasizes that “current solutions for telemonitoring, telecare or social interaction are largely proprietary, based on single provider design and cannot be easily adapted to multiple and changing users’ and organisational needs”.

Because the domain of active and independent living is about the daily life of people who might need assistance to be able to avoid dependency on nursing homes and continue to live independently in their own homes, its scope cannot be limited to only certain applications. Assistance might be needed in any possible aspect of daily life, e.g., health, safety and security, daily activities such as personal hygiene, home cleaning, shopping and cooking, comfort and entertainment, social integration, support of mobility, reduction of costs and avoiding waste in consumption (bridge to energy efficiency), etc. This complex spectrum of possible needs and offers is referred to as the domain of AAL. From an investment point of view, the AAL market should allow each individual in danger of losing independency to pick the set of applications and services over time based on the different needs that arise with ageing. However, many stakeholders state their concerns with regard to the maturity of existing platforms so that SMEs do not dare to risk on them before a stable ecosystem has already taken form. Therefore, as one of the nine measures for building such ecosystems, the Lecce Declaration [2] suggests that promising candidate open platforms must be tested under real-life conditions with respect to their usability and reliability in order to help them to mature and provide evidence of reusability.

Health knowledge about the elderly’s state is a good starting point to detect behavioral patterns and help to assess his status. With this information, there are some suggestions about daily behaviors, like taking drugs, activity recommendations, social interactions, whose are relevant in order to detect odd situations (early memory loss, disorientation, falls at home, symptoms of weakness, tiredness or fatigue), and there is not systems checking what is happening, or what is the likelihood that the situation is serious or not.

At the end, an intention detection system is desirable. Various experiments, particularly for ambient assisted living environments have demonstrated the possibilities and the complexities of intention detection [3]. User actions measurement and monitorization of user experience in real-time are enablers for a new range of innovative systems and a key enhancement to existing ones.

In this paper we introduce an automatic prediction of elderly behavior system suited for telecare that fulfills the previously exposed requirements. The main objectives of the developed system, already running with dozens of users, are the following:

- Develop a semantic system able to store and understand the clinical status, activity, context awareness and situation about a concrete user, allowing integrate this information in the intelligence system, in order to detect abnormal situations regarding to the health status.

- Create a set of services that enables intelligent monitoring of a particular elderly and his medical issues, so that the system adapts to him, creating automatically rules that determine the usual values for each individual and evolve with the elderly, so that they are always up to date. These rules allow launching fully customized alerts without human intervention.

- Create a telecare third-party system based on an expert system and an inference engine that can automatically detect dangerous situations decreasing false positives, firing events only at abnormal circumstances.

The remainder of the paper is organized as follows. Section 2 introduces the system architecture. Section 3 and 4 give details about the different types of the system data sources, while sections 5, 6 and 7 gives further insight into several main parts of the intelligence system. The obtained results over a real home deployment are discussed in section 8. Finally, section 9 presents our conclusions and future work.

2 System Architecture

The need to understand more about the user and his/her context started with the ubiquitous computing that placed abundant technology in user’s everyday environment [4]. Initial definitions of Context Awareness focused on the situation of an entity, where the user’s context (or the context of the devices carried by the user) was the main focus of the attention. Context was understood to include location, identity, activity and time, but more recently the viewpoint has shifted towards a notion of the user being part of a process or ecology, as exemplified by Ambient Intelligence (AmI) [5]. AmI refers to a vision in which
devices interact to support people in carrying out their everyday tasks and life activities in a natural way using information and intelligence that is hidden in the network of interconnected devices. The system relies on sensor data processing, merging, classification and reasoning. Research in smart environments is often related to ambient assisted living and attempts to derive information about people’s wellbeing (sleeping, awake, daily rhythm, falls, level of general activity) from the various sensors, often including cameras in a smart environment [6].

The present system is designed to fulfill two main requirements. On one hand, the extraction, transformation, and load of sensor information has to be carried out in a simply way: the sensor is plugged on the network, and its raw data automatically are integrated in the platform. On the other hand, the platform has to be able to measure the elderly’s habits in order to track their behavior to find deviations from their daily tasks (e.g. wake up times, sleep habits, diary strolls, etc.), or regarding to their health situation, and provide a detailed summary to the caregivers and the family about their evolution. All services provided by the system are addressed to be a robust, easily-deployable solution and a cost-contained model. The benefits will directly affect the elderly, the family and the caregivers.

As can be seen in Fig. 1, the platform is organized in five main modules to tackle and process the input sensors data:

- Data capturing from a set of sensors of different types
- Inserting relevant clinical record evidences into the teleasistence Subsystem
- Central Ontology and Expert Rules Subsystem
- Intention Detection Subsystem
- Anomaly Detection Subsystem

3 Data capture from a set of sensors of different types

The system contains the three types of sensors: environmental, physiological, embedded service and audiovisual sensors, connected to one or more hubs. In this regards, the main component of this platform at the gateway level is the so-called Home Box Services (HSB), which consists of a structured software system, flexible and portable that can run on different hardware platforms, and is the processing kernel of the system on the remote data received.
An advantage of the system is the ability to locate the user through very cheap hardware sensors in order to be competitive in the market, but with a powerful Context Inference Engine, which minimizes false positives.

Regarding the technologies involved, UPnP [7] is a highlighted protocol which does not only cover internet protocols such as TCP/IP, HTTP, SOAP, UDP or XML, but also integrates Zigbee, USB, IEEE802.11, BT, BLE Wi-Fi, and security considerations using security techniques like X.509 certificates. Moreover, this protocol is open and it can be extended (e.g. to define a specific type of message adding extra attributes).

4 Relevant clinical record evidences into the teleasistence system

Nowadays, the electronic clinical records written by medical staff about the patient evolution is stored in unstructured plain text (80% of relevant information), while the structured information is minimal (dose, etc.).

However, personal medical information about elder people in a tele-assistance system is vital to customize their habits to a better quality of life, predicting dangerous situations in relation with their pathologies.

All this knowledge, enriched with the sensor information is manually stored in a specialized medical-tele-care aware-context ontology representing information in a hierarchical way, with a dynamic and variable structure, which can be automatically updated online with new clinical information, new kinds of sensors or new measured characteristics. Besides, this ontology is able to implement two kind of rules: on one hand, the rules that allow infer new knowledge based on the raw data, health status, medical or behavior recommendations, and on the other hand, the expert system alerts rules.

Furthermore, the ontology allows the use of a reasoner which can check whether all of the statements and definitions in the ontology are mutually consistent, and can also recognize which concepts fit under which definition. The reasoner can therefore help to maintain the hierarchy correctly. This is particularly useful when dealing with cases where classes can have more than one parent. So, the ontology extracts relevant concepts of queries and questions to annotate the obtained results.

Semantic networks allow concept disambiguation [8], but even using a collection of huge medical resources of relations, it is not sufficient to create semantic networks. These relations are based on co-occurrences and do not contain any systematic description of the ontological concepts. Ontology itself may be used with many relations as: parent-child, related and possibly synonymous, is similar to, has a narrower or broader relationship, has sibling relationship, being the most useful relations for semantic smoothing, among others. To avoid this disambiguation problems, we use a rule-based system that helps to explain inferences for supporting clinicians at care time is still quite relevant. The advent of Semantic Web technologies introduce a framework for cross-linked hypermedia resources that can facilitate the process of cross referencing relevant workflows. Given the large amount of information to be managed, we need to use storage techniques able to manage efficiently large volumes of data (Data Warehouse) and provide easy access and efficient management of the information (Big Data paradigm). So, in this system we use a no-SQL database, storing the information in a triple mode in the Virtuoso system.

5 Central Ontology and Expert Rules System

Expert systems represent formal knowledge to solve human problems. This type of systems are applicable to any domain and are present today in nearly any application that requires high computational cost to automatize processes with some reasoning. Expert systems are suited to specific tasks which require a lot of knowledge derived from a particular domain experience as diagnostics, instructions, predictions or advice to real situations that arise and can also serve as training tools, mimicking the human behavior.

By their nature, the Semantic (Knowledge-based) Management Decision Support Systems (MDSS) work over structured data representation (schema). The knowledge is persistent in data-stores, and the expert knowledge (system rules) are heuristic evidence based rules, with reasoning capacity using an inference engine. This means that the rules are well-know and always are true (there is not possible, in principle, an uncertainty factor). The use of the Resource Description Framework (RDF) standard (and thus its associated representation machinery such as RDF
Schema and OWL) offers the possibility of making inferences when retrieving and querying information, in a way very similar to human natural language, being this the advantage in query-answer systems. Although OWL automated reasoning does not reliably scale for use in large knowledge bases, researchers and practitioners have just begun to explore the problems and technical solutions which must be addressed in order to build a distributed system. On the other hand, there are Non Knowledge-based MDSS, which learn from raw data (semi/un-structured), and are based on probabilistic techniques: patterns are taken as examples or cases in the past and the system has learning and probabilistic prediction capability. Obviously, the direction of the last researches [9] is adding both engine capacities in a hybrid motor platform.

“Semantic smoothing for language modeling” emerged recently as an important technique to improve probability estimations using document collections or ontologies, and this was the way followed to design this system. This is the technical way in this project. So, we have developed two joined ontologies, i.e., the home care ontology and the health habits recommendation ontology (see Fig. 2). The sensor and the clinical records system fill them automatically on real time processes, based on a set of semantic rules. So, the Expert Rules System has two main goals:

1. Fill the ontology with the sensor and clinical records raw data by means of “process rules”. Since these rules are easy to adapt, alter and maintain, this feature makes them an attractive solution for non-expert caregivers. The caregiver is able to directly define as well as modify the rules that specify the behavior of a system in a given situation. For example, context-aware behaviors could be specified by a rich set of rules. In addition, the use of rules on top of ontologies can enable adaptive functionality that is both transparent and controllable for users [10].

2. Execute expert rules in order to suggest health recommendations to the elders or send alerts about abnormal situations about the elder’s behaviors, by means of a set of “home care rules”. Context-aware applications should continuously monitor the elder’s environment in order to detect changes and react to them. So, rule-based architectures offer flexibility to tackle the variability of the environment and support the reconfiguration of systems according to changing needs without requiring reprogramming nor human intervention. Rule-based approaches are suitable for highly dynamic context-aware services [11].

Both set of rules are running over the ontologies with a semantic reasoner (Pellet), and the main advantage of the system is that it is open developed: the functions are out of the kernel code, and the program flow or the alarms system can be modified changing the rules, i.e., it is not necessary to change the code of the system kernel. So, this can be done by an expert who does not necessarily have to be a programmer. Defining personalized and adaptive elder interaction/behavior models is a key challenge when considering the issue of analyzing or predicting elder intentions. The Intention Aware elderly care application can predict the activities of the elder based on historic usual activity, profiting automatically the elder usual activity information. So, the system has a hybrid “Inference Engine” that joins heuristics rules (first conditions to check without elder data) mainly with a supervised system, in order to discover new rules, adapting the heuristic rules to the personal behavior changing, with an unsupervised system, with two objectives:

- Detect unknown alerts (not registered in heuristics nor supervised system).
- Join similar users’ behaviors in homogeneous groups (segment usual activity, customizing automatically the user usual activity, state or profile), in order to send new extracted knowledge to the medical and care agents about the general behavior of their patients.

In general, after the home installation, the inference machine runs the following tasks:

- Step 1: Identification of the observable behaviors of the user
- Step 2: Matching the heuristics rules to detect abnormal facts
- Step 2: Start the learning process during the learning period
- Step 3: Put the learnt behavior model into practice by applying the described process, modifying the heuristic general rules by heuristic personal rules
- Finally, back to Step 1
Human actions are influenced by context, knowledge or experience of dependencies between actions, and by expectations of how the situation is going to develop [12]. For example, if it is raining, even though it is summer, the elder probably does not go outdoor unless he likes walking in the rain. These subjective observations should be treated and the inference system must be able to learn these behaviors. In fact, the Inference Engine is an “Activity Tracker” system, and it is intended to track daily activities of the elderly via automatic or manual inputs. The real system uses some cases running over the “Activity Tracker” for example:

- The system infers whether the elderly is eating at a specific time of day. The indoor localization system is aware of the time of day and the amount of time that the elderly stays being staying on the kitchen. Electrical appliance usage is also taken into account to infer if the elderly has been preparing his/her meal. Contact doors on kitchen furniture is also used to track this situation.

- During the day, the system infers whether the elderly is doing the housework. The indoor localization system is aware of the time of day and the amount of time that the elderly spends tracking his/her position. Vacuum cleaning can also be detected.

- Physical exercise like hiking or strolling can be detected through the GPS-enabled smart-phone of the elderly. The position, velocity and path will be analyzed to measure the amount of kilometers of each day.

The “Activity Tracker” system provides three services for the intention aware elderly activities API:

- Current activity: it is the activity that the elderly is performing at the present time.

- Last 14 hour activities: they are the activities performed during the last 24 hours split into 15 minute time slots.

- Next 15 minutes forecast activity: the next activity that will perform the elderly based on the historic data of the elderly intention is inferred. This service provides the likelihood of happening the tracked activities, and takes the activity whose probability of happening is higher than the other ones.
6 Intention Detection System

The objective of this system is the development of decision trees that running on central server, and based on information collected by different sensors deployed in the home, make decisions about the status of the patients, depending on their health situation (extracted from clinical records), treatment procedures, level of stress, behaviors, risk of exposure to harsh circumstances, etc. To do this, as exposed in section 5, we have developed an expert system which reflects the knowledge, about which is the best guideline of action as would advised by a specialist in the area.

As each home user could adapt differently to his environment, and could have a recent history of actions that involve a danger due to accumulation of negative elements that can affect him, it is important that the general rules generated by experts adapt automatically to each user inside the platform with implicit knowledge, so the system will personalize the information to each patient, their current parameters, and their last actions, allowing to eliminate the false positives. These supervised models exploit the reasoning power of expert system to derive new knowledge and facts. The Intention Detection System reasons over the base knowledge to infer new facts, resolves context conflicts and maintains the consistence. The situation of a user is derived from his personal context, but the context is derived from the aggregation of all the user’s situations plus the environment situation, too.

Through these techniques of intelligent information processing, we are given special emphasis to the detection and prediction of anomalies (trend analysis, deviations in the data, etc.), such as lifestyle changes, poorly executed exercises, etc...

One easy way to detect changes and behavior anomalies is to compare the actual situation or state of the elderly with one prediction of his state. In this paper, it is proposed a 24 hours sliding window was analyzed with a decision tree (a supervised machine learning algorithm), in order to predict the next user action over the time. For each time window, the decision tree takes into account a multivariate set of values generating a predictive model (decision tree) and extract the next status of the user, with a confidence and probability levels.

Compared to the current technology for evaluating context-aware systems, we focus in particular on the quantitative evaluation of the properties of our rule-based system with a temporal dataset. The challenge of this proposal is a novel and distinctive base technology repository that has been developed in the treatment of time series, and another algorithm repository for rule generation based on probabilistic rules directly from RDF semantic systems, without human knowledge, and automatic insertion of these rules in the central ontology again, assigning weights to the semantic confidence (triples-stores), in order to customize the personal behaviors of each user. The described system is proposed as a new development and rethinking the current model of care taking into account the characteristics and preferences of each person in such a way that a personal behavior is built (see Fig. 3).

7 Anomaly detection system

Usually, interest driven analysis tends to overlook unexpected patterns in data. To avoid this inconvenience, the system contains unsupervised algorithms (clustering and association models). Data Mining deals usually with applications such as anomaly detection to prevent excessive consumption, pollution, escapes, and in general, abnormal patterns of what is an expected profile for each segment.

To detect anomalies, we use the Local Outlier Factor (LOF) algorithm [13]. The algorithm compares the density of data instances around a given instance A with the density around A’s neighbors. If the former is low compared to the latter, it means that A is relatively isolated, i.e., it is an outlier. Such outliers are considered anomalous. With the aim of identify and group those outliers, they can be classified by means of techniques based on statistical models or region density distances [14].

- Statistical Models: they are based on the field of Statistics, given the premise of knowing the distribution of data. Based primarily on measurements of distances between objects, the greater the distance of the object relative to others is considered an anomaly.

- Region Density Distances: Based on the estimation of density of objects, the objects located in regions of low density and relatively distant from their neighbors are considered anomalous. The main feature is that generally it is considered unsupervised learning and a score is assigned to each instance that reflects the degree to which the instance is anomalous.
One of the tools necessary for the Anomaly Detection is the Clustering, which is to group a set of data, without predefined classes based on the similarity of the values of the attributes of different data. This grouping, unlike the classification, is performed in an unsupervised manner, since it is not known beforehand of classes of the training data set.

The algorithm used to carry out the segmentation process takes as an input the data set (sensors data and clinical history) and the cluster model that was generated by a clustering algorithm (K means). It categorizes the clusters into small and large clusters and the anomaly score is then calculated based on the size of the cluster the point belongs to as well as the distance to the nearest large cluster centroid. With this model, we can check automatically which are the anomaly events in the normal life of one person and detect which is the outlier event and why. This information is sent to the Expert Rule System to decide or not to throw one alert to family or medical centers. Fig. 4 shows the abnormal events about one user along one year, each 15 minutes. It can see easily that the anomaly behaviors belong usually to Sundays, with actions that are not usual over the rest of the week.

8 Deployment results

Since the beginning of 2015, 60 homes have installed the system, checking online the users daily living over three different customer: dependent elderly people, elderly whose habits are worsening due to the aging and elderly people who are suffering the first symptoms of dementia. As previously exposed, the expert rules are running since the moment of the installation, but after a month capturing raw data, the system begins to obtain the behavior of the elderly. The steps through the system gets their behavior are the following:

- At a first stage, the sensor raw data are processed by the expert system to determine which is the elder’s context at every time, formatting the data into a structured table with the information about person, date, hour, and the stage in that moment regarding to the user: Sleeping (S), Cooking (C), Eating (E), Doing Housework (D), Outdoor (O), Outdoor Sport (U), Using Tablet (T), Using Mobile (M) or Spare Time (P).

- At the same time, the system checks the physiological status of the users, as temperature, heart pulse, blood pressure, etc. In this way the data will be managed more compactly and only the relevant information to the alerts is managed (for example, "The elderly has a fever unusual at 08:02 more than 15 minutes"). Initially the thresholds that are in the table are defined manually by the doctors (for example, a temperature of 36.5°C), leaving the detection of statistical thresholds for later, where the average temperature threshold is modified by historical and statistical processes and unsupervised algorithms.
These thresholds are able to modify the medical rules directly regarding the user personal historical set.

- The system also checks environmental sensors, such as smoke, temperature and humidity, throwing alerts when activated.

- At every moment, the system is taking external data to integrate them into users data, with a number of values: Haze (C), Fog (N), Low Fog (N), Fog (I), Precipitation (P), Drizzle (L), Rain (U), Torn Rain (V), Tornado Sight (R), Rain Shower (H), Rain (E), Snow (E), Shower Hail (T), Freezing Rain (T).

The first conclusion, as expected, is that the outdoor weather conditions are the most relevant in order to predict which will be the user behavior, in second correlation place (by a Chi Squared statistical test), after the hour of the stage, but before other indicators such as the day week, or even the month of action. With the temporal sliding window method, including the last action in order to predict the new stage of the user, the accuracy trust by the system using cross-validation is about 81.80%, as shown in Fig. 5.

Thus, it is shown that external data on local weather and data from past actions are representative to make predictions about the future status of the elderly immediately. Thus, if the prediction state does not match with the actual state, and this situation has a significant score, the system sends an alert to remote care services to immediately launch protocols. On the other hand, the anomaly clustering method used is a complement to the supervised system, in order to detect not so usual behaviors. The ability to analyze these patterns to improve the system, and even as a starting point of widespread information to improve care systems for the governments and health agents, is a real fact, thanks to this research.

9 Conclusions and future work

In this paper we have introduced a behavior prediction system to be used in an home telecare platform. Such platform is an assisted living system for monitoring elderly people at their homes, with the final aim of providing a robust, easily-deployable and cost-contained solution to ensure the safety of the elderly. It obtains both physiological and environmental data through a multisensor infrastructure, connecting the home with both the carer and the family, being aware of the state of the elderly.

Our next steps in the closed future are two: on one hand, to transform this validated system in one reference at the telecare platforms at home. On the other hand, reinforce the intelligent systems for analy-
ysis of anomalies and behaviors taking advantage of the health information to detect how the clinical records, medical diagnoses and treatments are affecting the usual behavior of different patient profiles in their daily lives, specially, at home, and if these treatments are appropriate or not. Finally, we are researching the challenge of the creation of a new system for automatic construction of medical summaries from clinical records, usually written in natural language and unstructured plain text.

10 Acknowledgments

The research was supported by the REAAL project (CIP ICT PSP – 2012 - 325189) and by the Computational Intelligence Group, funded by the Basque Government with grant IT874-13.

References


