Context Awareness in Learning Human Habits

Szymon Bobek
AGH University of Science and Technology
al. A. Mickiewicza 30
Krakow, Poland
Email: sbobek@agh.edu.pl

Weronika T. Adrian
AGH University of Science and Technology
al. A. Mickiewicza 30
Krakow, Poland
Email: wta@agh.edu.pl

***Abstract*—Mobile devices have gained steadily increasing popularity over the last few years. They collect and process various kinds of data which can be used as rich inputs to infer user preferences and habits. In this paper, we propose an architecture of a mobile application that will serve as an *intelligent assistant* capable of learning human habits and giving suggestions and recommendations. The system will learn patterns in its user's behavior with respect to his or her location, social preferences and activities that can be measured with a mobile phone sensors. Data gathered by the mobile device will be used to model daily, monthly or annual routines of the user. Based on that model, the mobile assistant will be able to find deviations in an organizational rhythm of the user and perform appropriate actions. The system will uses machine learning algorithms and ontologies to learn and model human behavior.**

***Keywords*-Context-awareness; machine learning; ontologies; mobile applications.**

I. Introduction

Various information and data sources can be nowadays reached from nearly every place in the world. Internet access is now possible through notebooks, tablets and more often – mobile phones, which are omnipresent in human daily routine. The last tend to play a role of *all–in–one* devices that serve as phones, calendars, web browsers, GPS navigations, and social media interfaces.

Modern mobile phones (or more commonly named: smartphones) can be themselves valuable data sources of human habits with respect to:

- one's usual location over time (for instance home, work, cinema, etc.),
- one's social preferences (who, where and how often the one meets),
- one's entertainment preferences (e.g., one more often goes to the opera than cinema),

and combinations of the above.

Monitoring human behavior can lead to development of a rich knowledge base, not only useful for sociologists, but also being a great input into systems that, based on the individual habits, will try to optimize user daily routines. Intelligent houses [1] are one of the most popular way of implementing ambient intelligence [2] solutions in real life. Tools monitoring human social behavior [3] with smartphones show that ubiquitous computing and machine learning techniques can be successfully implemented on mobile platforms as real-time hybrid systems.

One of the biggest advantage of using mobile platforms is that a system deployed on it can accompany its user almost everywhere. Hence, it can become a powerful source of information about the user behavior. This paper presents an idea of using smartphones to monitor and learn human habits from heterogeneous sources: phone sensors (accelerometer, Bluetooth, light sensor), GPS data or data available from functionalities like location sharing via Facebook. Information gathered from these sources can be treated as an input for a system that will learn human habits and act upon this *knowledge* as a intelligent mobile assistant.

The paper is organized as follows: In Section II, we present an overview of selected related work. The original contribution of our approach is discussed in Section III. Section IV describes techniques that can be used for knowledge acquisition about human behavior. The proposed method of modeling in a way that can be used for further inference is presented in Section V. Section VI describes challenges and main problems faced by our approach. The paper is summarized in Section VII.

II. Related work

Ambient intelligence applications and ubiquitous computing are nowadays widely used in intelligent systems. Various sorts of such systems cover different aspects of human living. One of the most popular automation systems are intelligent houses. Projects like CASAS [4], MavHome [5], or Intellidomo [1] build user profiles based one their activity at home during a day. Several techniques are incorporated within these projects that are crucial in discovering human habits. They include:

- temporal reasoning for describing time dependencies between the activities,
- methods of automatically constructing universal models by taking the output of sequential data mining algorithms and sequential prediction algorithms,
- methods of discovering sequences of user actions in a system based on speech recognition, and
- ontologies and production rules for describing model of human behavior.

However, intelligent houses build human profile and can adapt to it only within a limited space. They create human habits models using data from sensors installed within a house. Hence, the behavior profile is limited to describing user activities inside this building.

 ISBN: 978-1-61208-230-1

A different approach is represented in the SocialCircuits platform [3]. The platform uses mobile phones to measure social ties between individuals, and uses long- and short-term surveys to measure the shifts in individual habits, opinions, health, and friendships influenced by these ties.

Sociometric badge [6] has been designed to identify human activity patterns, analyze conversational prosody features and wirelessly communicate with radio base-stations and mobile phones. Sensor data from the badges has been used in various organizational contexts to automatically predict employee's self-assessment of job satisfaction and quality of interactions.

Reality Mining is a term coined by Eagle and Pentland [7]. The authors used mobile phone Bluetooth transceivers, phone communication logs and cellular tower identifiers to identify the social network structure, recognize social patterns in daily user activity, infer relationships, identify socially significant locations and model organizational rhythms.

III. Motivation

Most of the systems described in Section II monitor and build human profile based on data gathered from sources that are highly correlated. They are limited to one domain of user activity, like home daily routine or social activities. Moreover, they are more concerned with monitoring *human environment* than *a human within the environment*. Hence, it is not possible to create a comprehensive human habits profile based on a human daily routine. The project that is the closest to the idea that we want to incorporate in our work is described in [7]. However, the information presented by Eagle and Pentland was not incorporated into any real-life system. In addition to this, data gathered within the project was processed using statistical and machine learning tools, but no formal model of the human behavior was presented.

The original contribution of our approach consists in:

1) extending the idea presented in [7] by additional data sources: GPS location, accelerometer sensors, Wifi/GPRS activity, and RFID sensors,
2) learning and developing a model of human behavior with methods that allow for automated inference (formally grounded ontologies, see Section V), and
3) implementing a system that will work in real-time as a mobile assistant.

A system that has all the above mentioned information at its disposal can act as an intelligent assistant, monitoring regular behaviors of the users and recommending actions or signaling aberrations. Such an assistant can be a helpful tool optimizing daily routines of multi-tasking persons that interact with technology on a daily basis. This *optimization* will mainly consist in suggesting specific actions or decisions inferred from the model of the user habits and external knowledge. However, this is not an only possible area of application.

Equally important is enriching the quality of life for elderly people who wish to stay mobile and independent. Nowadays, technology services are available to and used by more and more adults, and the number of technology-aware raisins will increase over time. Helping them is not only an important goal, but also a trending research field, supported by international funding programs such as Ambient Assisted Living. The program is motivated by the demographic change and aging in Europe, which implies both challenges and opportunities for the citizens, the social and health care systems, the industry and the EU market. Our proposal aligns with the newest call for proposal entitled: "ICT-based Solutions for (Self-) Management of Daily Life Activities of Older Adults at Home". By encompassing other data sources into the profile modeling, we can enhance the solution to assist people also outside their houses.

IV. System proposal

This section is reworked to answer reviewers' comments We propose a mobile assistant system that will learn the human behavior patterns, build a semantic model of them and be able to reason over it. In particular, it will recognize aberrations and react to them or suggest decisions based on observations of human regular behavior. Example use cases may include: deciding on route based on the transportation habits and current traffic, adjusting daily meetings based on the user's location habits, or warning users if some variances are recorded (e.g., being late for work or forgetting shopping).

A. Proposed Architecture

The architecture of the proposed system can be observed in Figure 1. *Data Sources* module is responsible for gathering:

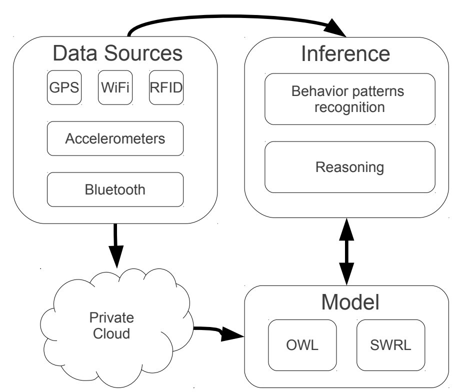

Fig. 1. Proposed Architecture of the Mobile Assistant System.

- GPS location over time,
- the location with respect to other devices (RFID and Bluetooth data),
- accelerometer data determining if the user is walking, running, or not moving at all.

Inference module performs the main learning and reasoning tasks. In the learning phase, it identifies human behavior patterns, by recognizing and clustering heterogenous information (location, time, speed etc.) The other function of the module is inferring relations between activities and suggesting actions to be taken. Some of the suggestions will be automatic, based on rules defined in the *Model* module, e.g., *If it is time that*

 ISBN: 978-1-61208-230-1

regularly the user is driving to work but now the user is not moving, then play a loud alarm to wake them up. It is also planned that the user will be able to define their own rules, e.g., *If I run for 15 minutes at a speed above average, play a tune of... (my favourite song)*.

The *Model* module will store the model of the user profile. The model will comprise of an ontology and rules (see Section V for more details). It initially consists of a top-level ontology and rules defined for abstract concepts. During learning the human behavior patterns, the ontology will be specialized and adapted to the particular user, its common locations (e.g., home, work, park), typical intervals of time (e.g., morning, evening, workday) and objects (persons and devices). The abstract rules are then instantiated to work on specific locations and time intervals characteristic for the user.

B. Learning Human Habits

The system will work in two phases: learning and acting. In the learning phase, the data (in a form of vectors describing different dimensions of information) will be an input for machine learning algorithms [8] that will try discover patterns in human daily routine. The K-Means clustering algorithm will be implemented to identify users behavior patterns in daily routine. After the clusters are identified, they will be semantically annotated by a user.

The system will propose sets of initially classified data and ask to categorize them (identify certain ranges of values). For instance, locations can be flagged as *home, work, cinema* etc. Other devices recognized via Bluetooth or RFID sensors can represent other users or objects. Time intervals can be identified as mornings, evenings, workdays etc. Various speed values can be assigned by user as walking, running, or driving.

The above classification will lead to the development of a personal ontology based on a top-level ontology. The classification will allow to build *context predicates*, e.g., *<user, locatedIn, park>*, *<user, performs, running>*, based on which the *activities* (e.g., shopping, driving to work, meeting friends, relaxing) will be defined (see Section V).

Once the learning phase is done, the data gathered by the *Data Sources* module will be passed to the *Inference* module where on-line pattern identification will be performed and appropriate action taken. Since the on-line classification has to be fast, the artificial neural network implemented on the mobile device will be responsible for this.

The system, based on the real-life data, will predict if the user is realizing daily routine or if there are some derogations that should be reported. To allow this feature to work properly, a special model of human behavior and daily habits should be created. The proposal of this model is sketched in Section V.

V. KNOWLEDGE REPRESENTATION

Data gathered by sensors can be uniformly represented using a graph model. Persons recognized via Bluetooth and objects identified with RFID can be uniquely represented as named nodes in the graph. This approach partially realizes the *Internet of Things* [9] idea in blending the borders between the real world and the virtual model. Accelerometer data will serve to calibrate the ranges for concepts such as running, walking or driving and GPS coordinates will be aggregated and named as locations (home, work, grocery store etc.). Multiple relations will combine into a semantic network representing the user behaviors and habits. The information will be represented with RDF [10], a flexible and universal Semantic Web [11] language that will allow the representation all of the information in a standardized way.

In order to enable automatic reasoning, formalization of the representation is needed. In our approach, we have adopted the top-level ontology for smart environments introduced by Ye et al. in [12]. The authors have identified common semantics for several domains (time, location, speed etc.) and their shared relations, such as: *finer-grained*, *equal to*, *conflicting* or *overlapping*. Consequently, this top-level ontology provides "generic rules to facilitate pervasive tasks including detecting inconsistency of information, generating new knowledge such as new activities for different applications and new relationships between concepts, contexts, and activities" [12].

The chosen ontology provides a framework for several dimensions of information. Therefore, it is suitable for modeling human habits profiles in various locations and time intervals. Each dimension of information contains a set of irreducible *grounded values* (e.g., GPS coordinates, second). Over the ranges of grounded values, *abstract values* are identified. These are mappings to a set of grounded values and are done by semantic tagging by user over preliminary classified data (see Section IV). Abstract values define the ranges of grounded values and allow to create user-specific subclasses (e.g., park, morning) of the top-level classes like *location* or *time*.

Abstract values serve to build the *context predicates* e.g., *<user, locatedIn, park>*, *<user, performs, running>*, which, in turn, are used to define *activities* (e.g., shopping, driving to work, meeting friends, relaxing). Activities are defined by a conjunction of context predicates, e.g. <user, locatedIn, park> $\wedge$ <user, performs, running> $\wedge$ <sensor, time, 7:00>$\Rightarrow$<user, engagesIn, ''jogging''> (see [12] for more details). The rules which govern the logic of the application will operate on the activities and timestamps and will activate certain actions or recommendations.

Using a formally-grounded ontology in the system allows for automated logical inference over the model, which is developed based on data gathered in the application. Intended modeling language is OWL [13]. The initial model has been prototyped in OWL and SWRL [14], because this is the modeling language of the Ontonym ontologies [15].

VI. CHALLENGES AND PROBLEMS

There are several challenges and problems to be considered. One of the main problems is the privacy issue. Users may not want to have information about their location stored on a global server. A separate system called *private cloud* has to be developed, that would allow users to store private data on their computers, and share it only with selected users. The idea of the *private cloud* combines two paradigms of storing data: 1) privacy and 2) distribution of data. The first aspect puts stress on the confidentiality of data: only the owner of the data and authorized users have access to the data. The second aspect touches the problem of accessability of data. If data has to be accessible for many mobile users, the most efficient way is to store it in the distributed environment called *cloud*. That comes with the privacy issues that leads to the conclusion that safety and distribution is very challenging combination that is out of the scope of this work. Another problem is battery cost. The software that would use data from several sensors and process it in a real time will be very power-consuming. Installed on a mobile phone, this may cause the battery to run out very quickly. Another type of challenges is related to users with high entropy of their daily routine. For example, a person who works as a taxi driver probably will not have an usual location that can be tagged as for instance *work*. Those problems and challenges will have to be addressed in later work in order to ensure the software works correctly and efficiently.

VII. Conclusion and Future Work

Mobile devices have gained steadily increasing popularity over the last few years. They collect and process various kinds of data from registering calls and storing messages to logging GPS locations and receiving data from other devices via wireless protocols. These collections of data can be used as rich inputs to infer user preferences and habits. In this paper, we presented a proposal of an architecture of a mobile application that will serve as an *intelligent assistant* capable of learning human habits and giving suggestions and recommendations. Several techniques are planned to be implemented in the project, including machine learning algorithms for patterns recognition in human daily routine and ontologies for modeling human behavior. The system can not only be dedicated to people who would like to optimize their daily routine, but also for those who need care and attention during a day. The system can work as an artificial assistant for older people, or as health care tool that will help monitoring and reporting patients behavior.

For future work, we plan to adapt the semantic knowledge-based wiki environment Loki to allow the user to monitor and configure their semantic profile [16]. Using visual methods for designing rules [17], the user will be able to define their own actions to be taken in case of a variance in their behavior. Enriching the functionality by allowing collaboration among users [18] will open up possibilities of e.g., organizing meetings and adjusting behavior profiles of several people.

Acknowledgment

The paper is supported by the AGH University of Science and Technology Grant 11.11.120.859.

References

[1] A. Lozano-Tello and V. Botn-Fernndez, "Analysis of sequential events for the recognition of human behavior patterns in home automation systems," in *Distributed Computing and Artificial Intelligence*, ser. Advances in Intelligent and Soft Computing, S. Omatu, J. F. De Paz Santana, S. R. Gonzlez, J. M. Molina, A. M. Bernardos, and J. M. C. Rodrguez, Eds. Springer Berlin / Heidelberg, 2012, vol. 151, pp. 511–518.

[2] B. Epstein, "Ambient intelligence sources," http://www.epstein.org/brian/ambient_intelligence.htm, [retrieved: September, 2012].

[3] I. Chronis, A. Madan, and A. S. Pentland, "Socialcircuits: the art of using mobile phones for modeling personal interactions," in *Proceedings of the ICMI-MLMI '09 Workshop on Multimodal Sensor-Based Systems and Mobile Phones for Social Computing*, ser. ICMI-MLMI '09. New York, NY, USA: ACM, 2009, pp. 1:1–1:4.

[4] P. A. Valiente-Rocha and A. L. Tello, "Ontology and swrl-based learning model for home automation controlling." in *ISAmI*, ser. Advances in Soft Computing, J. C. Augusto, J. M. Corchado, P. Novais, and C. Analide, Eds., vol. 72. Springer, 2010, pp. 79–86.

[5] G. M. Youngblood, D. J. Cook, and L. B. Holder, "Managing adaptive versatile environments," *Pervasive and Mobile Computing*, vol. 1, no. 4, pp. 373–403, 2005.

[6] D. O. Olgun, B. N. Waber, T. Kim, A. Mohan, K. Ara, and A. Pentland, "Sensible organizations: Technology and methodology for automatically measuring organizational behavior," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART B: CYBERNETICS*, pp. 43–55, 2009.

[7] N. Eagle and A. (Sandy) Pentland, "Reality mining: sensing complex social systems," *Personal Ubiquitous Comput.*, vol. 10, no. 4, pp. 255–268, Mar. 2006.

[8] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[9] K. Ashton, "That 'Internet of Things' Thing," http://www.rfidjournal.com/article/view/4986, [retrieved: September, 2012].

[10] E. Miller and F. Manola, "RDF primer," W3C, Tech. Rep., 2004.

[11] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, May 2001.

[12] J. Ye, G. Stevenson, and S. Dobson, "A top-level ontology for smart environments," *Pervasive and Mobile Computing*, vol. 7, no. 3, pp. 359 – 378, 2011, knowledge-Driven Activity Recognition in Intelligent Environments.

[13] F. van Harmelen and D. L. McGuinness, "OWL Web Ontology Language overview," W3C, Tech. Rep., 2004.

[14] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean, "SWRL: A semantic web rule language combining OWL and RuleML, W3C member submission 21 may 2004," W3C, Tech. Rep., 2004.

[15] G. Stevenson, S. Knox, S. Dobson, and P. Nixon, "Ontonym: a collection of upper ontologies for developing pervasive systems," in *Proceedings of the 1st Workshop on Context, Information and Ontologies*, ser. CIAO '09. New York, NY, USA: ACM, 2009, pp. 9:1–9:8.

[16] W. T. Adrian, S. Bobek, G. J. Nalepa, K. Kaczor, and K. Kluza, "How to reason by HeaRT in a semantic knowledge-based wiki," in *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2011*, Boca Raton, Florida, USA, November 2011, pp. 438–441.

[17] A. Ligeza and G. J. Nalepa, "A study of methodological issues in design and development of rule-based systems: proposal of a new approach," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 2, pp. 117–137, 2011. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/widm.11/pdf

[18] G. J. Nalepa, "Collective knowledge engineering with semantic wikis," *Journal of Universal Computer Science*, vol. 16, no. 7, pp. 1006–1023, 2010. [Online]. Available: http://www.jucs.org/jucs_16_7/collective_knowledge_engineering_with