Workflow Mining Application to Ambient Intelligence Behavior Modeling

Carlos Fernández[1], Juan Pablo Lázaro[1], and Jose Miguel Benedí[2]

[1] Research Group of technologies for Health and Wellbeing (TSB), ITACA Institute, Polytechnic University of Valencia, Spain
{cfllatas,jplazaro}@itaca.upv.es
[2] Department of information Systems and Computation (DSIC), Polytechnic University of Valencia, Spain
jbenedi@dsic.upv.es

Abstract. The handmade human behavior modeling requires too many human resources and for too long a time. In addition, the final result does probably not reflect the current status of the person due to the influence of time. The use on Workflow Mining techniques to infer human behavior models from past executions of actions can be a solution to this problem. In this paper, a Human Behavior modeling methodology based on Workflow Mining Techniques is proposed.

1 Introduction

A workflow is a formal representation of a certain process in order to be managed by a computer-based system. Humans benefit from this formal representation, thanks to the use of authoring tools. Using these tools a non-expert in computer sciences can create complex processes, enabling them to be managed automatically by a system, without the need of writing any line of source code.

The wider purpose of an Ambient Intelligence system is to provide to a certain user a number of services in a proactive way according to his/her current context, as well as hiding the great complexity of the technology, and to provide these services using a communication paradigm that is as easy as possible to understand and to be used.

This is not a trivial problem to solve from the technological point of view, because it requires an important concept to be well implemented in the ambient intelligence system: a good model of the user behaviour and user context.

The implementation of an individualized model of the behaviour of a simple user is a very complex task. In order to implement it, the participation of experts in the field of knowledge of the type of behaviour that is wanted to be modelled, is required. Then, after observing the behaviour patterns of the user in a sufficient period of time (months or even years) the experts in elderly care will be able to define which the behaviour model of that user is, and convert it into a formal workflow for its automatic processing. This methodology has two important disadvantages: firstly, it requires too many human resources and for too long a time, and secondly, the final result probably doesn't reflect the current status of the person due to the influence of time.

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 160–167, 2009.

Using Pattern Recognition techniques enables us to infer the workflows from prior examples. This methodology is commonly known as Workflow Mining. The use of Workflow Mining in the human behaviour modelling field does not only require the use of inference algorithms, but it also requires these algorithms to represent the workflows in a simple structured way, but with the great expressivity needed in real life. An adequate high level of expressivity is considered essential because the utilization of complex systems for representing workflows makes the interpretation of the results of the inference algorithms difficult. Furthermore, if the expressivity is not enough, the real problem might not be well represented by the system. This is the reason why the utilization of finite state models, and in particular the Timed Parallel Automaton (TPA) [10] is probably the best option to define Workflows that have high expressiveness requirements and need to be simple enough to be interpreted by software systems.

Pure workflow mining techniques found in literature are based on the use of mediators for the creation of general purpose workflow models. These techniques can be applied to infer, not only fundamental behaviour models of the user, but also individual models. Therefore, a strategy based on the obtaining of individualized models that overtrain user's parameters and learn the specific behaviour of that person is proposed. By using this trained model it is possible to know, at any given time, whether the behavioural pattern of that user is normal or not. The system could detect changes in individualized behaviour patterns by means of the analysis of the historical behaviour of the user.

In the following paper this methodology is presented as an alternative to discover human behaviour patterns. This technological solution is used to model an individualized user behaviour using automatic learning methods, like Workflow Mining techniques, that allow inferring individual behaviour patterns from the actions performed by the user. This methodology is currently being applied in PERSONA [20] European project where this technology is being used to detect abnormal behavioural patterns in elderly people living at home. The behavioural information is acquired from the information coming from a massive distribution of sensors and detectors and the system is implemented in a computer with limited resources in terms of memory and processing capacity. This fact makes the simplicity of the inferred workflow by the proposed technology very important.

1.1 Ambient Intelligence and Behavior Modeling

The Ambient Intelligence (AmI) model is focused on the creation of physical spaces where the technology is thought to serve the people. The current concept of AmI is the result of the evolution of the ubiquitous computing [19] of Mark Weiser and the vision of ISTAG (Information Society Technology Advisory Group). In the AmI concept, the services are thought to continuously empower the user in terms of time and space. Those services must be as invisible as possible to users' sensors and psichology and be as less intrusive as possible. Therefore, the interfaces among the user and the services must be natural and cannot be an obstacle to the interaction between the user and the environment.

Into the AmI model there is an essential concept known as Context. The context is the user projection in the system. This projection supposes the mapping of user data in repositories. In this way, the higher the quality of the data, the higher the quality of the services the system can offer. That user information is gathered through sensor systems. The information produced by those sensors is usually processed by intelligent algorithms based on artificial intelligence and pattern recognition techniques. Those algorithms are working as ‘software sensors’ and provide a higher and richer level of information, while the ‘hardware sensors’ provide the basic raw information of the user. For example, hardware sensors can offer information like Heart Rate or number of steps, while ‘software’ sensors can offer information like mood or activity level. This high level information is crucial to know the status of the user.

That data only offer a vision over the static status of the user. Nevertheless, the human behavior is inherently dynamic. Human behavior is the collection of behaviors exhibited by human beings and influenced by the environment. As the environment is continuously changing and the user is more and more experienced, the human behavior is continuously evolving. For that reason, static data is not enough to offer a holistic view of the human behavior evolution. Therefore, we need to study the historical information of the actions of the user to detect anomalous behavioral patterns to improve the diagnosis of the status of the user. For example, some symptoms of the dementia illness are based on anomalous behavioral patterns [18], like excessive flirtatiousness, social withdrawal, or agitation. In addition, the wide variability of humans makes the creation of a general model that explains the behavioral patterns of different kind of human people very difficult. The behavioral patterns are different in each human being and the same aspects can suppose different reactions depending of the person. For example, defining a social withdrawal model is very different for a very shy person than for a very self-confident person. So, the definition of models that allow us to know the behavioral status of the patient and his evolution has to be done individually.

1.2 Workflow Technology

The human behavior has a wide variability and interdependence among the processes involved. This requires models with a high capacity of representation to allow the processes involved to be described. The use of natural language allows professionals to represent the processes with the needed expressivity. Nevertheless, the use of non-formal languages to represent human behavior adds undesirable ambiguity to specify those kinds of processes. In addition, the use of formal languages for representing human behavior allows us to take profit of the big amount of formal frameworks available in literature in order to automate, represent and learn these kinds of processes.

The processes involved in the human behavior at one specific moment are usually based on the previous process and affect the future actions performed by the human being. In literature, there are a research field that fits with this syntactical point of view. This field is known as *Workflow Technology* [1]. The Workflow Technology is intended to provide a framework to represent, automate and mine processes in order to define them in a standardized way. The main objective of this framework is the study of Workflows. According to the Workflow Management Coalition definition, a Workflow *is the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules* [2]. In spite of the help that suppose the use of Workflows for representing human behavior, the design of these kind of systems continue wasting too many computing resources. On one hand, the big amount of variables of an individual to be studied by experts may make the study larger in time. On the other hand the behavior of the individual is continuously changing. So, if the behavior of the individuals study requires too much time, the conclusion of the study will probably arrives too late and the actual status of the individual was not in accordance with the current detected status making the system useless. Therefore, it is crucial to create algorithms and tools that provide real views of the behavior of the individual order to allow the experts to detect anomalies as soon as possible.

The daily actions of the user are gathered by the context of Ambient Intelligence System. Those actions can be used to mine the current status of the individual and providing a graphical view of the current behavior of the user, which can suppose a great help to know the current status. In literature, there is a research framework based on the use of action logs to infer Workflows which explain the flow followed by the processes. This research framework is known as Workflow Mining or Process Mining [3].

Most of the works in literature are based on the use of transactional logs as samples for workflow inferences [6, 8 and 9]. These models use logs from general systems of workflow management as input samples to infer models that explain the whole system. There are two approaches of Workflow Mining methodologies based on the kind of data gathered. One of these methodologies is the Event-based approach. The *Event-based Workflow Mining* [6] approach learns workflows using the available information on transactional logs as input samples. The algorithms that are based on Event-Based data take into account the starting information of processes (i.e., the action name and the starting time) but do not consider the results of the actions. Nevertheless, the behavior modeling problem needs this kind of data to specify the process flow. When an action is made by the user this is probably related with the previous action results made, and the result of current actions affect to future actions that will be made by him. For example, the alarm procedure must be triggered when the detection action finds fire or a gas escape. In this case, the event based approach does not take into account the result of the detection action and, thus, it is not possible to infer this kind of processes. The second approach is the Activity-based Workflow Mining [10]. This model takes into account the results of the actions making possible the inference of behavior modeling processes, which can be based on the results of previous actions. This approach allows to infer flows like the fire example, because it can represent that the Alarm sequence is triggered when the fire detection action returns true.

In Pattern Recognition theory [12], it is usual to select a representation framework bounded with theoretical properties which facilitates the inference. Most of the syntactical pattern recognition works are based on regular languages framework [13, 14 and 15]. Commercial Workflow representation languages are too complex to be covered by this framework making the inference more difficult. In addition, the field of human behavior is very demanding and requires very expressive representation languages to explain the complexity of real processes including parallel patterns and time. In literature, there is a very expressive formal framework able to describe workflows called TPA[10]. This model has theoretically probed its expressivity covering the control flow Workflow Patterns[5], which are the standard way to measure the expressivity in workflow theory. In addition, this model was used to describe formally Life-Style Activity Protocols [16] that represent the human behavior of the user in a specific context. The most important feature of this model is its complexity. This model is equivalent to regular language framework. That means that this model can use the powerful tools available for this framework for inference.

In the literature, there is an Activity-Based Workflow Mining Algorithm called PALIA [11], able to infer TPA from executed actions information. This algorithm can be applied to infer the flow with the labeled transitions representing the result of the actions. In this paper, we use the Activity based approach to infer individualized Workflow models of the human behavior to facilitate the detection of anomalous behavior of people in their own context.

2 Results

In this paper, a new methodology for using Workflow Mining systems to allow experts to infer Workflows that explain the individualized human behavior is presented.

In Figure 1, a general schema of the methodology is presented. The data gathered by the sensors is stored in the AmI context repository. These data are the actions performed by the user and their results ordered in time. This data is used by the Workflow Mining algorithm to infer Workflow models readable by the expert who is able to detect changes in the behavior of the user.

When the user under analysis is introduced in the AmI environment, the system starts to store all the information available into the AmI context in order to create a projection of the user in the system. Using this information, it is possible to create a Workflow Mining corpus selecting the most relevant data associated to the behavior of the user. In this way, the higher the quality of the user data, the clearer the workflow we obtain. The historical data needed is the date and time of the start and end events of each action performed, and the result of the action (i.e. LookAgenda, TakeHeartRate, etc). Using this data, a workflow that represents the user behaviour of a user in a certain environment is inferred by using workflow mining techniques. The first Workflow inferred represents the basic behaviour that will be compared with

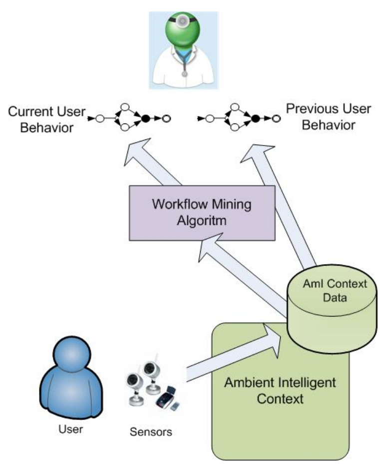

Fig. 1. Workflow Mining process Methodology

future ones. The comparison with new inferred workflows defines the current status of the user indicating if the user behaviour is compliant with the previous model or on the contrary, if the user is following an abnormal (different) behaviour. The abnormal behaviour of the user may have two reasons. The first reason is that the user is having a problem that interferes in his/her normal flow of life, and the second is that the user has evolved and therefore, he/she changed the normal behavioural pattern. The latter will cause a new iteration starting again from the first phase.

The expert user is the mediator in charge to decide if the changes in the behaviour of the user are due to a normal evolution in the user life, or, on the contrary, due to a problem. In this case, experts are able to detect dementia, depression and other problems in early stages.

In order to test that methodology a prototype experiment is made using Workflow Mining technology. Due to the lack of available activity-based corpus in human behaviour modelling research field a lab experiment was made. In the experiment a Workflow Simulator [17] was used to create activity based corpus which can be used to test it. In addition, a modified version of PALIA [11] algorithm, specifically designed to count the number of accessed arcs and states, was made.

Using the Workflow simulator, a patient life trace was simulated. On one hand, using this software, 90 days of a normal life of a specific patient was expressed in a log. The simulated patient was an old widow who lives alone in an AmI environment. Using those 90 samples, PALIA algorithm was used to infer a Workflow which represents the usual actions made by the user. On the other hand, another simulation was made via modifying the standard way of life of the patient adding some dementia indicating behaviour patterns, such us social withdrawal and memory errors. This

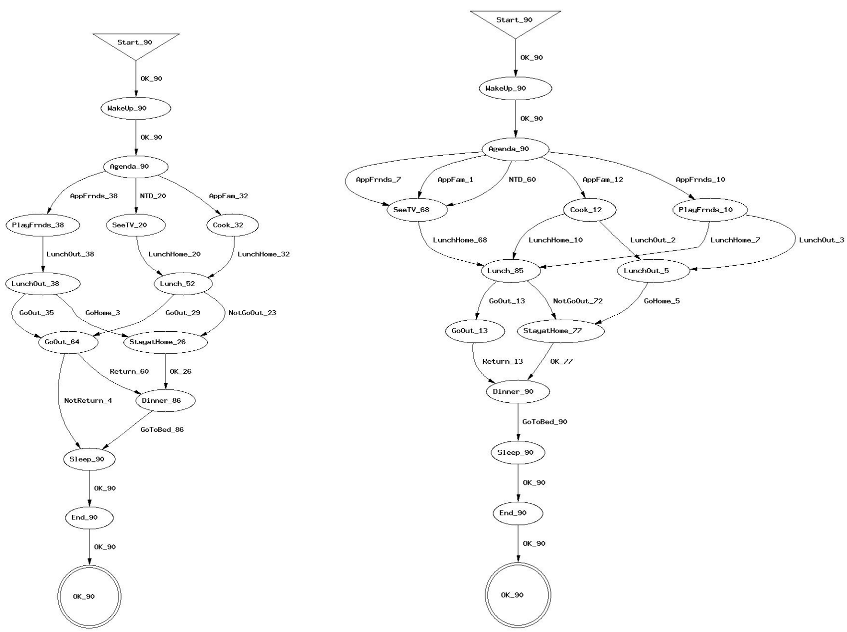

Fig. 2. Example of individual Workflow Mining use for dementia detection

corpus is also inferred by using PALIA Algorithm. Both Workflows inferred are presented in the Figure 2.

The workflow of the left represents the first workflow. In this Workflow each state represents an action, the triangle represents the start action and the double circle represents the final action. The rest of circles represent intermediate actions. The arrows among the circles represent the pass between two actions depending on the results of the actions represent by the labels. The number in the nodes is the number of executions which has achieved the specific node. The number in the arrows is the number of executions which has performed the action destiny after the origin action when the result of the origin action was the indicated on the label of the arrow.

The Workflow on the right represents the second Workflow. This Workflow has the same representing shapes than the previous one. As can be seen in the figure, the simulated workflows show interesting differences:

- In the first Workflow the actions performed before lunch are always dependent on the AmI agenda and the user looks strict with the actions made. In this Workflow, the user goes playing with fiends and cooks with family most of the times. Nevertheless in the second workflow, the user stays at home watching TV avoiding dates with friends and family most of the times.
- In The first workflow, the user always has lunch at a restaurant when he plays with friends. In the second workflow, the user has lunch home most of the times when he is out playing with friends.
- In two occasions, in the second Workflow, the user must go out for lunch after cooking for family. It can be a symptom of errors in cooking that haven't occurred in the first Workflow.
- In the first Workflow, the user sometimes does not have dinner at home probably because he goes out with friends. Nevertheless, the user always has dinner at home in the second Workflow.

As can be seen in the Workflow, it is very easy to find these symptoms that clearly show the user is more and more decreasing his social life and probably has some memory problems (forget dates and cooking recipes). As expected, those symptoms point to possible dementia problems.

3 Conclusions and Future Work

The use of Workflow technology allows using a wide spectrum of tools to represent, execute and mine human behavior models. Nevertheless, the representation of Human behavior is a hard task. Due to the wide variability of human beings, the use of generalized human behavior models is not enough to classify the behavior of users. The use of Workflow Mining techniques to model the human behavior facilitates experts to detect anomalous behavior of users in an individualized way, making a picture of the user behavior during a specific period of time. The capability to take a snapshot of the behavior of the user allows the comparison of the evolution of the human being with snapshots taken in previous steps.

In the laboratory experiment, we can see that it is possible to use Workflow Mining techniques to help experts to detect anomalous behavior of the user in order to detect some problems in early stages.

Once this algorithm is successfully tested in laboratory conditions, the next step is to use that algorithm in a real environment to test it with real problems. This algorithm is currently being installed in the PERSONA project platform and it is planned to be used in the second phase of the project (mid 2009) to provide high level knowledge to the Ambient Intelligence context.

References

1. Fischer, L. (ed.): Workflow Handbook 2001, Workflow Management Coalition. Future-Strategies. Lighthouse Point, Florida (2001)
2. WfMC. Workflow Management Coalition Terminology Glossary. WFMCTC-1011, Document Status Issue 3.0 (1999)
3. van der Aalst, W.M.P., van Dongen, B.F., Herbst, J., Maruster, L., Schimm, G., Weijters, A.J.M.M.: Workflow mining: A survey of issues and aproaches. Data and Knowledge Engineering 47, 237–267 (2003)
4. van der Aalst, W., et al.: Workflow Patterns Distributed and Parallel Databases, p. 70 (2003)
5. van der Aalst, W., et al.: Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering 16, 1128–1142 (2004)
6. de Medeiros, A.K.A., Weijters, A.J.M.M.T., van der Aalst, W.M.P.: Genetic process mining: A basic approach and its challenges. In: Bussler, C.J., Haller, A. (eds.) BPM 2005. LNCS, vol. 3812, pp. 203–215. Springer, Heidelberg (2006)
7. Fernandez, C., Benedi, J.M.: Activity-Based Workflow Mining: A Theoretical framework Workshop On Technologies for Healthcare and Healthy Lifestyle (2006) ISBN:978- 84-611-1311-8
8. Fernandez, C., Benedí, J.M.: Timed parallel automaton learning in workflow mining problems. In: Ciencia y Tecnología en la Frontera (2008) ISSN:1665-9775
9. Schalkoff, R.J.: Pattern recognition: statistical, structural and neural approaches. John Wiley & Sons, Inc., Chichester (1991)
10. Oncina, J., Garcia, P., Vidal, E.: Learning subsequential transducers for pattern recognition interpretation tasks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 448–458 (1993)
11. Oncina, J., García, P.: Inferring regular languages in polynomial update time. In: Pattern Recognition and Image Analysis 1991. World Scientific Publishing, Singapore (1991)
12. Coste, F.: State merging inference of innite state classifiers. In: Rapport technique n INRIA/RR-3695, IRISA (September1999)
13. Dominguez, D., Fernandez, C., Meneu, T., Mocholi, J.B., Serafin, R.: Medical guidelines for the patient: Introducing the life assistance protocols. In: Computer-based Medical Guidelines and Protocols: A Primer and Current Trends, vol. 139, p. 282. IOS Press, Amsterdam (2008)
14. Fernandez, C., Sanchez, C., Traver, V., Benedí, J.M.: TPAEngine: Un motor de workflows basado en TPA. In: Ciencia y Tecnología en la Frontera (2008) ISSN:1665-9775
15. Santacruz, K., Swagerty, D.: Early Diagnosis of Dementia American Family Physician 63(4) (February 2001)
16. Weiser, M.: Some Computer Science Issues in Ubiquitous Computing. CACM 36(7), 74–84 (1993)
17. PERSONA Consortium. PERSONA project: PERceptive Spaces promOting iNdependent Aging, http://www.aal-persona.org/