A Dynamic User Profiling Technique in a AmI Environment

Vítor Marques
Ângelo Costa
Paulo Novais
CCTC, Departament of Informatics - Universidade do Minho
pg17203@alunos.uminho.pt
acosta@di.uminho.pt
pjon@di.uminho.pt

***Abstract*—Currently there are many services that can assist the human being in his decision making. Many of these services provide aid as consistent as possible attending the characteristics and preferences of a user, which compiled results in a person's profile. Without profiles, systems could not provide a coherent aid to a user. In addition we can consider the profiles as the basis of recommendation systems. Paramount with cognitive helping systems, that provide decisions and recommendations these actions can more accurate and user driven. However, the profiles need to be updated over time, as a human being changes of preferences or beliefs, the profiles also need to adapt to dynamic environments. It is introduced a project that applies the Bayesian Networks and Case-Based Reasoning techniques to create and modulate user profiles in a coherent and dynamic way, using stochastic models and high-level event relations and characteristics to devise an accurate suggestion of activities the user can perform, being integrated in an Ambient Assisted Living Project.**

I. Introduction

Currently, systems that can help people in relation to decision making are increasingly common. A common example is the recommendation systems that try to suggest a particular decision, based on the characteristics of an individual. These systems are achieved through systems that incorporate artificial intelligence techniques.

The iGenda and TIARAC projects is a Ambient Assisted Living project that has as objective to help elderly persons with cognitive problems [1], [2].

To recommend an activity that the user enjoys making, personal information has to be used to choose the most appropriate one, while taking into account the disabilities and preferences the user has. To be able to recommend a decision is necessary information about the user on which the system can reason, that is, we need a profile. In the simplest version, these profiles are a static set of assumptions that describe the preferences of an individual.

However, as we know, one of the things that distinguishes human beings from the others is its ability to evolve, their ability to acquire knowledge, develop skills, change their behavior, their beliefs, among others. Associated with this human capacity coexists their likes and preferences. Throughout life humans change their likes and preferences, depending on their experience, age and interests among other conditions, rendering the static profiling system inefficient.

Currently there are several studies on profiling techniques that can adapt and change over time [3], among which, those classified as Machine Learning that are based on information collected from user interaction. Various approaches have over time been used for profiling and recommender systems [4], [5], [6], [7], [8].

The main objective of the present study was to develop a dynamic profiling technique, able to adapt to the changing preferences of a user over time. We present a technique for profiling that attempts to incorporate all the capabilities of two very associated techniques with the field of artificial

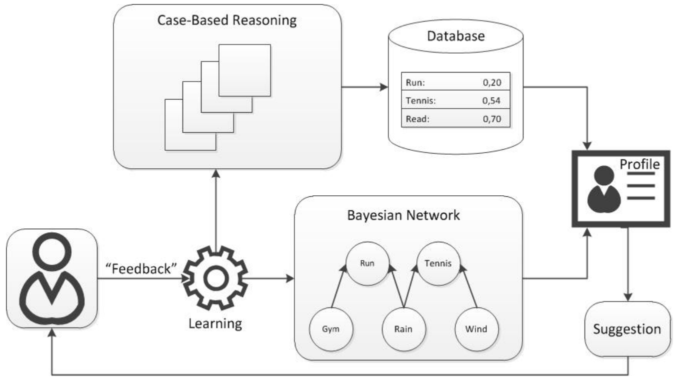

Fig. 1. Modeling a profile

intelligence, which are the Case-Based Reasoning and Bayesian Networks.

This paper is structured in the following form: Section 2 will present the iGenda project and the needs it was in better answer the user requests in terms of personalization, introducing the Bayesian Networks and Case-Based Reasoning to provide and architecture to respond to the proposed problems. The Section 3 will present the suggestion and learning features the project will be using. Finally Section 4 will present the conclusion and future work that this project will generate.

II. A Profiling Project

This work is developed upon the requirements the iGenda and TIARAC projects had [1], [9]. The Free Time Manager used a static recommendation system, it could be changed over time but it was an arduous task, consuming precious technician time, and it was almost impossible to update the items in the database each week. The recommendation system of the TIARAC project also is based upon the user preferences, thus requiring that the user file contain his preferences and is always updated, providing in overall a better service by attending to the user characteristics. Also the recommendation system had a narrow group of items that served as modifiers of the decision engine, such as weather conditions, user physical condition, among others.

It was clear that a flexible system has demanded, one that could attend to the user preferences, his capability of proceeding the activities and that changed according to the user current preferences, keeping up-to-date to his likes and dislikes.

When there is free time, choosing an activity to occupy that time may not be an easy task due to the large amount of possibilities and dependencies between these and other attributes. The use of information initially defined that characterize the individual (statistical profiles) may make this choice easier.

However, as already stated, is part of a human being that their preferences change over time, whether as a result of past experiences, whether by the emergence of new points of interest.

A profile is constructed using a Bayesian Network, which models relationships between attributes and activities and data stored in a database, describing the individual in terms of ”satisfaction” in performing certain activities (the probability that an activity is suggested). In both techniques, they adapt over time, depending on the number of feedback that exist, given the suggested activities. The updating of values stored in the database is achieved using the Case-Based Reasoning and its cases repository.

In the Figure 1 is illustrated the process of modeling a profile.

A. *Bayesian Networks*

A major problem that arises when dealing with real problems is the existence of incomplete information on the environment. In these situations it is appropriate to adopt techniques that work on the field of probability. The Bayesian Networks (BN) [10], [11] were introduced in the early eighties in order to solve this problem by adopting this Bayesian Probability Theory.

The inference process is achieved through the definition of evidence (attributes for which we are sure about a value) [10], [11], on which the network updates the Bayesian probabilities for all the relationships that exist. The final result will be a probability distribution in response to evidence.

The Figure 2 illustrates an example of BN for the problem domain.

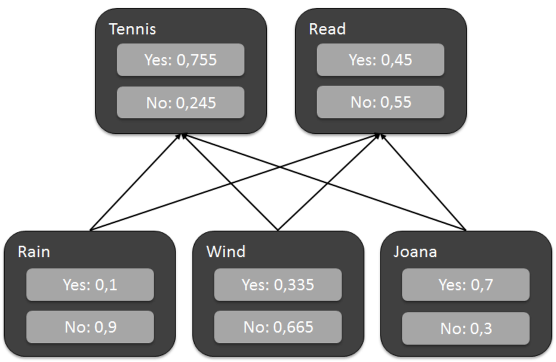

Fig. 2. Bayesian Network Example

The inference process has the main objective to order a set of activities, taking into account the BN probabilistic values, from the placing of new evidence found in the current situation. In the Figure 3 we can check the result of the inference based on the situation of the previous figure.

In the example, *Joana* will do the activity that fits in the conditions rain and wind. It was found that in this situation, it was not raining and there was also no wind. Based on this evidence, it was clear that the activity Tennis is the appropriate activity, followed by Read. If the user does not want these activities, he should tell the system that so a learning process over the network can modulate and adjust according to the user feedback.

In the examples all the information and attributes nodes are known, however there may be situations

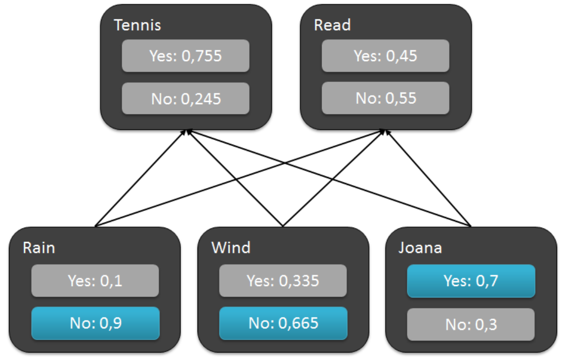

Fig. 3. Inference Process Example

where there are attributes which we have no certainty about the occurrence, in other words, situations of incomplete information. In these situations, the learning process is not possible to be executed by a simple math and specific algorithms had to be adopted[12], [13].

B. Case-Based Reasoning

The Case-Based Reasoning (CBR) is a problem solving methodology that determines the solution to a given problem, based on the reconstruction of past similar cases, reusing or adapting to the knowledge of such cases [14], [15].

This concept is suitable for the problem domain, since the satisfaction of a user in practice an activity will depend on past experience.

In this work the CBR faces a crucial role in the profiling process, since intervenes in the process of updating the data stored in the database (which describes the individual in terms of satisfaction in practice certain activities).

In the problem domain, a case consists of a set of attributes describing the situations where an activity took place, a field indicating the proposed activity and the user feedback under these conditions.

In the Figure 4 it is represented an example of two cases, corresponding to the suggestion of an activity at the conditions: rain, in the gym and at 4 PM. At first, when the activity Football is suggested , the user refused, opting for the ”Running” activity corresponding to the second case.

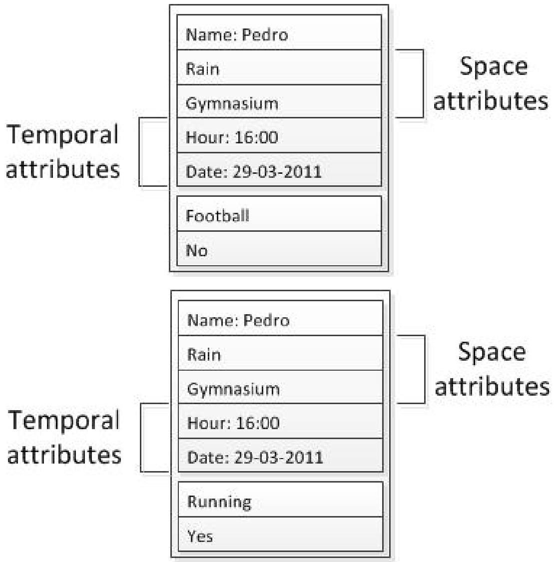

Fig. 4. Cases Example

The main use of CBR is to obtain a set of similar cases with the current conditions. The calculation of this similarity takes into account the similarity existing between the temporal attributes and spatial attributes of past cases with the new case . The integration of temporal attributes aims to overtake the disadvantage of CBR in not be modular with the time factor.

In this case, the use of this factor relates to the possibility of the user changing his willingness to practice an activity after a certain time. In these situations the similarity value should be lower, since there is no certain if the user still wants to practice (a positive feedback will be given if the user accepts the suggestion).

To calculate the similarity it was adopted the following expression:

$$S_F(S_{SA}, S_{TA}) = S_{SA} * S_{TA}$$

Where S_F is the final similarity, S_{SA} the similarity of the spatial attributes and S_{TA} the similarity of the temporal attributes. The lower the similarity of the spatial attributes and the similarity of temporal attributes, the lower the final similarity.

Unlike most uses of CBR, in this work, its use aims to obtain a set of percentages representing the most n similar cases. We do not want to have a complete description of the case, only the value of similarity and the user feedback.

This situation occurs because the target value is compared with a set of probabilistic values and the user feedback, to determine the percentage X of a value V. Theron is added or subtracted (UV) to the value associated with the enjoyment to exercise a certain activity (value stored in the database), according the user preference.

$$UV(X, V) = X * V$$

In short, we want the modifying value to be calculated taking into account the similar past cases and their temporal distances.

The expected result of the use of CBR of $n = 5$ of the activity A, is illustrated by the following table:

Similarity	Feedback
0.92	No
0.90	No
0.80	No
0.75	Yes
0.68	Yes

Analyzing the table, we can see that for past cases where the activity A has been suggested, the user refused the three most recent and accepted the two oldest.

Consider the case where the user refuses to practice the activity A again, the value to subtract of its ”enjoyment” in practice the activity should be relatively large, since in the past and more recent iterations the user refused.

Thus, the value of X resulting from the preceding table and in the case where the user refuses to practice the activity A, should be larger than the next.

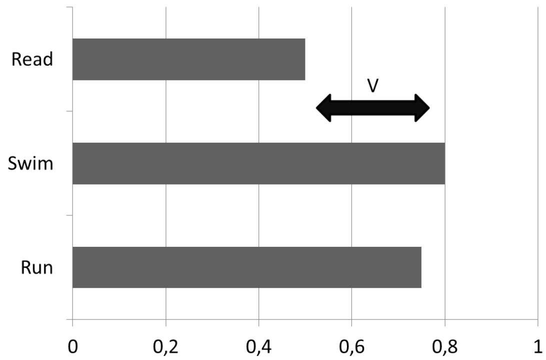

Fig. 5. Example 1 - Calculation of V.

Similarity	Feedback
0.92	No
0.90	Yes
0.80	No
0.75	Yes
0.68	No

Setting the value of V, of which a percentage of X will be added or subtracted to the enjoyment of an activity may be determined in several ways.

One solution may be make V be calculated using the average value of enjoyment of other activities. The goal is to make the activity in question level with the others in terms of suggestion, whether if the feedback is positive or negative.

In the previous example if we select the activity Read and that $V = 0.3$ and $X = 0.7$, in this case, to the value of enjoyment of Read would be added 0.21 percentage value.

III. Suggesting and Learning

The suggestion process is based on game theory, in which among a set of possible activities one will be declared the winner, in this case, the one that is accepted by the user.

After the winning activity was found, the user profile is updated on the BN and on the database (handled by CBR), upping the activity and subtracting to the others. If the user decline all suggestions, no activity should be affected in relation to others.

When asked to suggest activities, the system receives a new case with the attributes but without solution. Based on these attributes, the system infers about the BN, on the probability of each activity to be accepted by the user. Based on these values and the values of enjoyment stored in the database, the system sorts the activities in decreasing order, so that the former are the ones that most fit the profile of the user on the current situation.

The results are presented to the user, through the iGenda, for each activity the user is questioned , when an activity is accepted the profile is updated based on the answer.

In Figure 6 is an example of the process of suggestion and learning.

In this example, it is necessary to sort the activities of Run and Tennis, according to the probability of being accepted by the user (based on the profile). In this case, after the user accept the activity Tennis, two new cases are created corresponding to each activity and the feedback. Based on these cases, the Bayesian Network and the CBR will try to model and update the user profile.

IV. Conclusions and future work

We presented a technique for creating and modulating user profiles, capable of serving for recommendation systems or suggestion in various fields.

With Bayesian Networks it was intended take advantage of its ability to modulate conditional relationships between attributes and its ability to deal with incomplete information.

With CBR we took advantage of their ability to infer about past cases. We proposed a mechanism to deal with one of its limitations, the time factor. With the interconnection of these two technical we judges to have been able to develop a system capable of dealing with profiles in dynamic environments.

Emotions and use of sensors are some domains to be studied [16], [17] to incorporate in this project in the future. These last two are the most appropriate to clarify the value to be modify in the user's preferences.

Acknowledgments: The work described in this paper is included in TIARAC - Telematics and Artificial Intelligence in Alternative Conflict Resolution Project (PTDC/JUR/71354/2006), which is a research project supported by FCT (Science & Technology Foundation), Portugal.

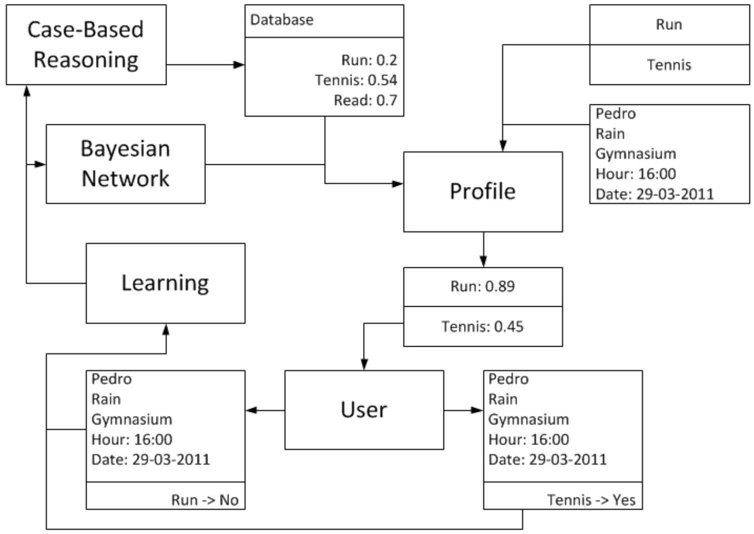

Fig. 6. Example of the process of suggestion and learning

REFERENCES

[1] Â. Costa, P. Novais, R. Costa, J. Corchado, and J. Neves, "Multi-agent personal memory assistant," in *Trends in Practical Applications of Agents and Multiagent Systems*, ser. Advances in Intelligent and Soft Computing, Springer Berlin / Heidelberg, 2010, vol. 71, pp. 97–104.

[2] D. Carneiro, Â. Costa, P. Novais, F. Andrade, and J. Neves, "Providing relevant knowledge in disputes: Umcourt project." in *ODR*, ser. CEUR Workshop Proceedings, vol. 684, 2010, pp. 63–78.

[3] S. Soltysiak, B. Crabtree, M. Heath, and I. Re, "Knowing me, knowing you: Practical issues in the personalisation of agent technology," in *Proceedings of the 3rd International Conference on the Practical Application of Agents and Multi-Agent Technology*, 1998, pp. 467–484.

[4] S. Radde and B. Freitag, "Using bayesian networks to infer product rankings from user needs," in *Proceedings of the UMAP 2010 Workshop on Intelligent Techniques for Web Personalization and Recommender Systems*, 2010.

[5] M. desJardins, E. Eaton, and K. L. Wagstaff, "Learning user preferences for sets of objects," in *Proceedings of the 23rd international conference on Machine learning*, ser. ICML '06. ACM, 2006, pp. 273–280.

[6] M. Dastani, N. Jacobs, C. M. Jonker, and J. Treur, "Modelling user preferences and mediating agents in electronic commerce," *Knowledge-Based Systems*, vol. 18, pp. 335–352, November 2005.

[7] H. H. Syed and P. Andritsos, "A lightweight tree structure to model user preferences," in *10th DELOS Thematic Workshop on Personalized Access, Profile Management, and Context Awareness in Digital Libraries*, 2007.

[8] A. Bellogín, I. Cantador, P. Castells, and . Ortigosa, *Discovering Relevant Preferences in a Personalised Recommender System using Machine Learning Techniques*, 2008, pp. 82–96.

[9] Â. Costa and P. Novais, "An intelligent multi-agent memory assistant," in *Handbook of Digital Homecare*, ser. Communications in Medical and Care Compunetics, Springer Berlin Heidelberg, 2011, vol. 3, pp. 197–221.

[10] R. L. Marques and I. Dutra, "Redes bayesianas: o que são, para que servem, algoritmos e exemplos de aplicações," Universidade Federal do Rio de Janeiro, Rio de Janeiro, Tech. Rep.

[11] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, "Learning bayesian networks from data: An information-theory based approach," *Artificial Intelligence*, vol. 137, no. 1-2, pp. 43 – 90, 2002.

[12] J. Bilmes, "A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models," Tech. Rep., 1998.

[13] F. Dellaert, "The expectation maximization algorithm," Tech. Rep., 2002.

[14] A. Aamodt and E. Plaza, "Case-based reasoning: foundational issues, methodological variations, and system approaches," *Ai Communications*, vol. 7, no. 1, pp. 39–59, 1994.

[15] I. Watson, "An introduction to case-based reasoning," in *Progress in Case-Based Reasoning*, ser. Lecture Notes in Computer Science, I. Watson, Ed. Springer Berlin / Heidelberg, 1995, vol. 1020, pp. 1–16.

[16] Z. Liu, *Computational Emotion Model for Virtual Characters*. InTech Education and Publishing, may 2008.

[17] L. Pomponio, M. L. Goc, P. Eric, and A. Alain, "Discovering models of human's behavior from sensor's data," in *Proceedings of the 6th Workshop on Artificial Intelligence Techniques for Ambient Intelligence, AITAmI'2011*, 2011.