

Document downloaded from:

<http://hdl.handle.net/10251/35358>

This paper must be cited as:

Picón, A.; Rodríguez-Vaamonde, S.; Jaén Martínez, FJ.; Mocholí Agües, JB.; García, D.; Cadenas, A. (2012). A statistical recommendation model of mobile services based on contextual evidences. *Expert Systems with Applications*. 39(1):647-653.
doi:10.1016/j.eswa.2011.07.056.



The final publication is available at

<http://dx.doi.org/10.1016/j.eswa.2011.07.056>

Copyright Elsevier

A statistical recommendation model of mobile services based on contextual evidences

Artzai Picón^a, Sergio Rodríguez^a, Javier Jaén^b, Jose A. Mocholi^b, David García^c, Alejandro Cadenas^c^aInfotech Unit, Tecnalía Research and Innovation, Zamudio, Spain^bGrupo ISSI - Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València, Valencia, Spain^cTelefónica I+D, Madrid, Spain

Abstract— Mobile devices are undergoing great advances on the last years allowing users to access an increasing number of services or personalized applications that can help them select the best restaurant, locate certain shops, choose the best way home or rent the best film. However this great amount of services avoids the user find and select those services needed for each specific situation. The classical approaches link some preferences to certain services, include the recommendations given by other users or even include certain fixed rules in order to choose the most appropriate services. However, these methods do assume that user needs can be model by fixed rules or preferences causing this models to fail when model different users or making them difficult to train. In this paper we propose a new algorithm that learns from the user's actions in the different contextual situations to properly infer the most appropriate for that user in a specific contextual situation. This model have been specifically designed to face the inherent lack of learning evidences, computational cost and continuous training requirements for user's behaviour modeling by the use of a double knowledge diffusion approach and it overcomes the performance and convergence rates offered by other learning methodologies.

Index Terms— Context awareness, machine learning, personalization, mobility, user preferences, services.

I. INTRODUCTION

In the last few years, mobile devices are undergoing great advances in terms of capacity, connectivity and adaptability, thereby allowing providers to supply numerous services or personalized applications to the final user and even making it possible for the user to be the provider and to foster these services[1].

This great quantity of available services creates problems for the user when trying to find and select those services needed in a specific situation. A partial solution to this problem is the recommendation of services based on previously established preferences in the user's system. Based on this approach, the catalogued services are linked to specific preferences. An evolution of these methods allows establishing preferential services that have been recommended by other users [2] or providing services associated to a specific location [3] [4].

However, these approaches do not take into account the diversity of user context variables, including climate, location, the day of the week, user's availability, current activity, etc.

An evolution of location based models is known as context based services recommendation [5] [6] [7] which allows recommending services based on

different context variables that could affect the user's decision making (for example, a bicycle rental service would not be offered when raining). In this way, we take into account the variations of the user context in order to provide a more precise recommendation of services.

However, these context-based recommendation models are based on fixed rules and do not take into account possible variations in the user's decision-making when facing similar contexts. Therefore, the users' actions in past situations (user history) must be taken into account in order to model and infer their needs in the present situation.

Despite their great potential, traditional inference techniques, such as neural networks, are not feasible for these types of problems. This is due to the great quantity of necessary evidence required for the system training for each variable combination. Another disadvantage of these approaches is that the models can not be scaled. In this way, adding new evidences would cause the retraining of the entire system, leading to problems of resource management and model validation.

Besides the aforementioned problems, these inference systems are not flexible enough in order to make an adequate services recommendation when the presence of each contextual variable in each

contextual situation is variable (e.g. the geographical location, weather forecast may not be available in certain circumstances).

Another problem associated with classical inference models is that the system must be capable of recommending services in a coherent manner, even in situations where there is a low number of evidences. This can not be achieved with systems based on neural networks.

Therefore, this work aims to provide an enhanced solution to the recommendation of services by resolving some of the earlier problems in order to obtain the most adequate services for the user based in the current contextual situation and its prior history.

Accordingly, we propose the creation of a model of services consumption that allows its estimation in a specific contextual situation, even in the case where there are no prior behaviour evidence of the user in that situation.

The proposed model provides solutions to the following problems:

- 1- No available evidences for all contextual situations.
- 2- Not sufficient available evidences for all service typologies.
- 3- Automatic updating of the model without retraining.

In order to solve these problems, the proposed design takes advantage of several concepts:

- The similarity or "contextual proximity" between different contextual situations.
- The similarity among service typologies.
- The knowledge of the preferences of the average user, or average compatible user profile.

The rest of the paper is organized as follows: Section II describes the formal model which is proposed for the estimation of the service consumption model of the user according to his contextual domain. Section III shows the developed implementation of the model and the obtained results and Section IV provides clear conclusions about the performance of this model,

II. SYSTEM DESCRIPTION

The proposed Contextual Services Consumption Model is based on two main elements:

- **The matrix of contextual situations**, which will help establish the proximity of the user when facing different contextual situations.

- **The services consumption model**, which provides knowledge on the consumer's desired services in a contextual situation.

The joining of these two elements defines the Contextual Services Consumption Model (CSCM) that allows knowing, in each contextual situation, the services most likely desired by the user.

A. SYSTEM ARCHITECTURE

Firstly, we define a contextual situation \vec{V} as a vector where each component defines a specific contextual variable, as show in (1). Such contextual components may be of different nature, including location, weather, user's activity, mood, etc. A contextual situation \vec{V} represents a position in the K-dimensional contextual space that reflects a specific situation of a user and means contextual evidence in itself.

$$\vec{V} = (v_1, v_2, \dots, v_K) \quad (1)$$

Additionally, a distance function is established between two contextual situations to express the degree of similarity between them.

$$Dist(\vec{V}_i, \vec{V}_j) = f(\vec{V}_i, \vec{V}_j), i, j \in N \quad (2)$$

As an optional simplifying condition, we require that this distance function not be affected by interrelations of contextual variable of different typology:

$$D(\vec{V}_i, \vec{V}_j) = D(v_{1i}, v_{1j}) + D(v_{2i}, v_{2j}) + \dots + D(v_{ki}, v_{kj}) \quad (3)$$

This previous premise will simplify the behavior of the system when not all the contextual variables are present. However this is not necessary if the distance functions are specifically designed to model the lack of contextual information maintaining the relationships among variables.

A contextual situation matrix (CSM) is defined as a K-dimensional set of situations \vec{V}_p , where each node represents a specific contextual situation that is defined by a maximum number of K variables within the contextual space. The number of elements of each of the dimensions is selected based on the chosen level of discretisation.

$$CSM = \{\vec{V}_1, \vec{V}_2, \dots, \vec{V}_p\} \quad (4)$$

Where P represents the total number of contextual nodes.

The distance between two contextual situations in the CSM is defined as the distance previously defined for contextual situations (2).

Being based in previous definitions, the contextual space is defined by a K-dimensional matrix of contextual nodes (**CSM**), where K is the number of contextual variables. A distance function (*Dist*) is associated to allow obtaining the similarity between different contextual situations \vec{V} .

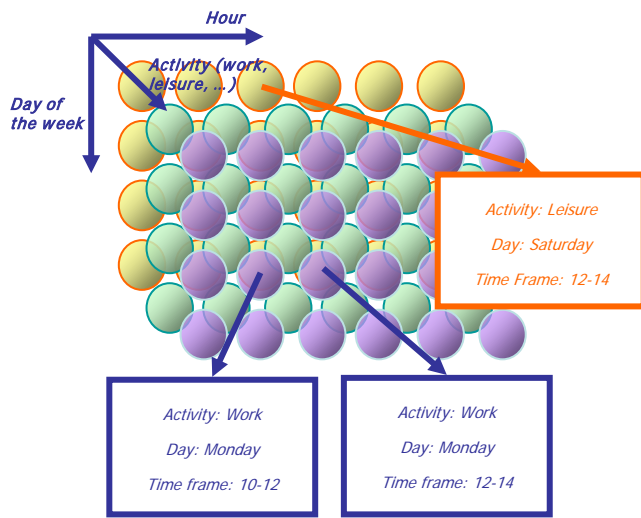


Figure 1: Contextual node matrix.

In order to be able to represent a user's **services consumption model**, a taxonomy of services is defined T consisting of a set of categories of services (K) and a set of diffusion coefficients (α). These diffusion coefficients will model the correlation between the different concepts allowing establishing relationships between consumption of different categories in order to enhance the user consumption model estimation. In mathematical terms, a taxonomy T can be seen as a graph where categories would be the nodes and correlations the edges weighted with diffusion coefficients, as shown in Figure 2.

$$T = \{K_1, K_2, \dots, K_M, \alpha_1, \alpha_2, \dots, \alpha_N\} \quad (5)$$

Where M is the number of categories where the different services are categorized and N is the number of edges connecting two different nodes.

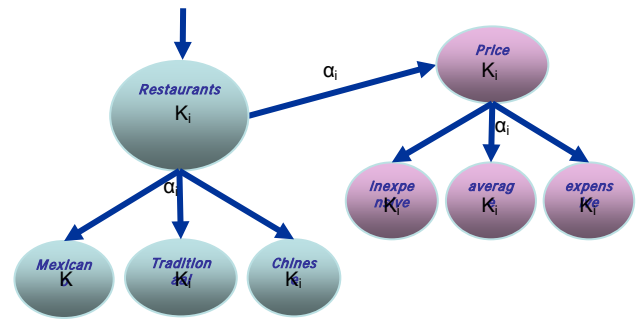


Figure 2: Example of services taxonomy

So, an instance of a service S provided to the user is mathematically described as a set of categories to which it belongs.

$$S = \{K_d, K_e, K_t, \dots\} \quad (6)$$

In this way, a service labeled as "traditional inexpensive restaurant" shall be described as the set of service categories: {restaurant, traditional, inexpensive }.

Each services category K_m is related to other categories through a coefficient of diffusion α . This coefficient will define the degree of diffusion of knowledge between the different service categories, as shall be later expounded.

Starting from this taxonomy, T , we define a *services consumption model* (SCM). Wherein each service category K_m of the taxonomy T has a consumption histogram (H_m) associated to it. This histogram indicates the number of times that a user has *accepted*, *refused* or *ignored* a service that is associated to that category.

$$H_m = (Yes, No, Ignore) \quad (7)$$

In this way, a Services consumption model (SCM) is described by a set of M histograms, being M the number of categories in the taxonomy T.

$$SCM = |H_m|, m \in \{1..M\} \quad (8)$$

Figure 3 shows the visual representation of a SCM.

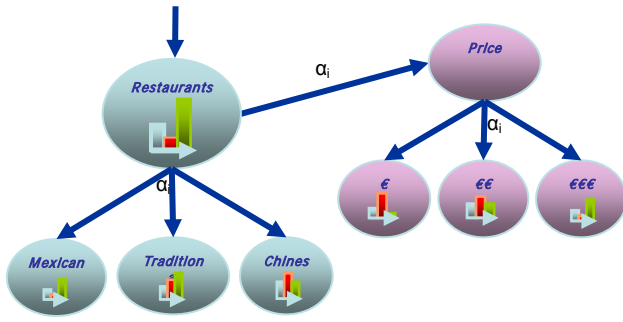


Figure 3: Services Consumption Model (SCM)

In order to relate a services consumption model (SCM) with a contextual situation, each contextual situation \vec{V}_i from the contextual situation matrix (CSM) is associated to a services consumption model (SCM) such as the one previously seen, defining the contextual services consumption model (CSCM) as illustrated in figure 4.

$$CSCM = \{SCM_{C_{v_1}}, SCM_{C_{v_2}}, \dots, SCM_{C_{v_p}}\} \quad (9)$$

Where $p \in \{1..P\}$ and P represents the number of contextual nodes in the contextual node matrix (CNM).

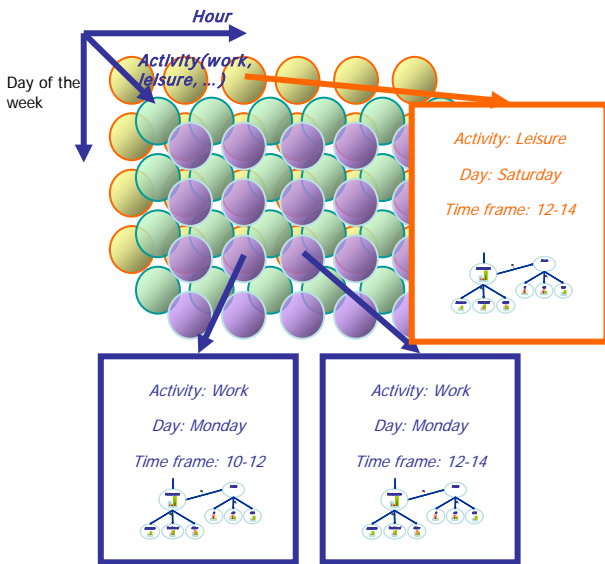


Figure 4: Contextual services consumption model (CSCM).

The following sections explain the interaction between the different elements used for the learning and knowledge inference tasks.

B. LEARNING LOGIC

Given a Contextual Services Consumption Model (CSCM) that is defined by a taxonomy of services T , a

specific discretization for the context variables through a matrix of contextual nodes and a contextual distance function, we proceed to the initialization of the system for a specific user:

We initialize each consumption histogram H_{pm} associated to each category m and to each contextual node p with either the values of the average user or the average user associated to the user profile or with certain predetermined values ($H_{pm} = (1,0,0)$). In this way, the services consumption model starts with previous knowledge that shall be modified through the learning phase.

When a user, in a contextual situation \vec{V}_i , performs a specific action (that includes *acceptance*, *refusal* or *ignorance*) regarding a specific service S_i , each histogram associated with the categories of that service is modified according to the contextual proximity to each contextual node. Therefore, the user's contextual services consumption model (CSCM) is modified.

The updating of the service consumption model associated to each contextual node $SCM_{C_{v_p}}$ is defined in (10) by updating each Histogram associated to each category m and to each contextual situation p in the contextual situation matrix.

$$H_{pm}(t) = \begin{cases} H_{pm}(t-1) + f(Dist(\vec{V}_i, \vec{V}_p)) & \text{if } K_m \in S_i \\ H_{pm}(t) & \text{else} \end{cases} \quad (10)$$

, where $m \in \{1..M\}$ and $p \in \{1..P\}$

Being $f()$, a monotonously decreasing function and $Dist(\vec{V}_i, \vec{V}_p)$ the distance between the present contextual situation and each contextual situation as defined in (4),

As can be seen in equation (10), for a given service category, we update the histograms related with that category of all "nearby" contextual nodes. Such neighbourhood impact is modelled through the $f()$ function. Such function is obtained through optimisation procedures and will depend on the specific context dimension.

In this way, each user action does not only affect the present context, but also has a bearing on the closest contextual nodes, contributing towards a better learning and generalization.

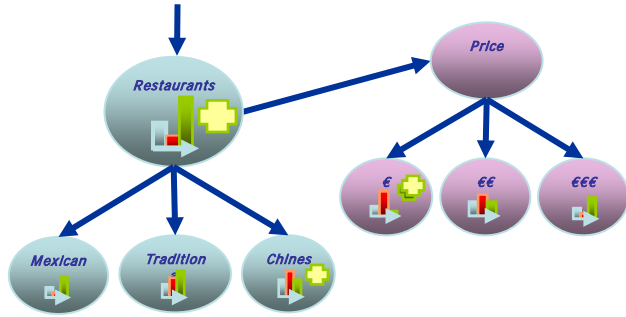


Figure 5: Updating of the consumption histograms when facing the acceptance of a service labeled as (restaurant, Chinese, inexpensive).

C. INFERENCE LOGIC

One of the greatest advantages of the proposed method lies in its capacity to infer knowledge even in situations with scarce evidence.

In order infer the consumption model of the user, the proposed method is based on a double diffusion of knowledge, both in close contextual situations as well as in consumption evidences of compatible services.

There are several types of queries that can be set forth to the Contextual services consumption model.

1) BASIC QUERIES

Given a contextual services consumption model (CSCM) and when facing a contextual situation \vec{V}_j , we can establish the probability $P_{i,j}$ of affinity of a service (S_j) defined by a set of service categories as defined by (6).

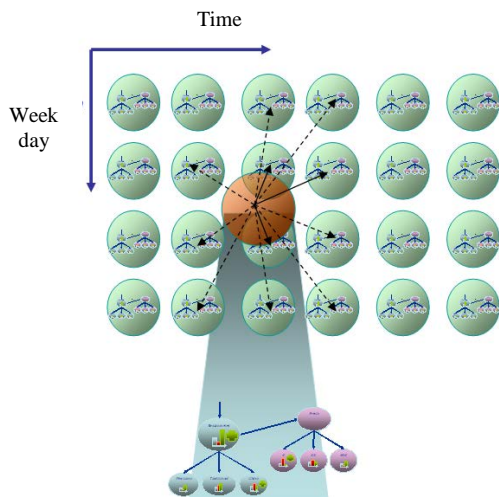


Figure 6: Graphic representation of the calculation of the services consumption model for a contextual situation V_i . Example for two contextual variables.

In order to establish the affinity degree, calculate the services consumption model (SCM) for the contextual situation \vec{V}_j .

In this way, the consumption histograms H_{jm} associated to the contextual situation \vec{V}_j and to a category K_m are defined by:

$$H_{jm} = \sum_{p=1}^P f(Dist(\vec{V}_j, \vec{V}_p)) \cdot H_{pm} \quad (11)$$

, where P represents the number of contextual situations that are present in the contextual situation matrix and H_{pm} represents the histograms associated to category K_m within the contextual node $C_{\vec{V}_p}$.

The graphical interpretation of equation (11) is very simple, being the obtained services consumption model (SCM) a linear combination of each of the services consumption models that are associated to each contextual node weighted in accordance to the similarity of the contextual situation \vec{V}_j with each of the nodes, as can be seen in Figure 6.

Once the services consumption model (SCM) has been obtained for the contextual situation \vec{V}_j in accordance with equation (11), we can proceed to consult, for a given service S_j , the affinity probability of the user with that service for this contextual situation \vec{V}_j .

For this purpose, starting with the services consumption model (SCM) obtained in (11), we can calculate the services consumption histogram for the service S taking into account the diffusion coefficients α , as shown in Figure 7.

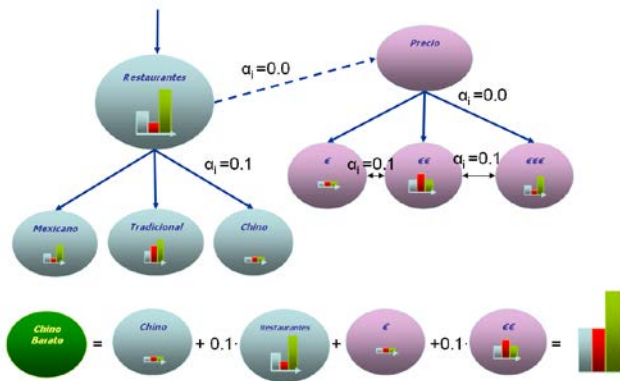


Figure 7: Calculation of the services consumption histogram for a specific service from the diffusion coefficients.

Figures 6 & 7 allow visualizing the learning diffusion concept which permits making very robust inferences even in conditions with scarce evidence. To achieve this, the system uses the learnt knowledge in all contextual situations and for all services. In case there is learnt knowledge in nearby contextual situations, this shall be strengthened by a greater weighting coefficient (9) while in cases where there is lack of knowledge, relevance shall be given to situations with smaller weight. As happens with the diffusion of knowledge within the taxonomy of services, if the user has consumed the service, the system, by means of its design, shall use this information. However, in situations where the user has not consumed a similar service, the system shall disseminate the knowledge from the information contained in the characteristics of services C_i related through the use of diffusion coefficients α_i , as shown in Figure 7, where the system decides to recommend a service based on consumption data that is inferred from other categories of services.

2) ADVANCED QUERIES

The flexibility of this model also allows undertaking other types of queries than those of services recommendations, as a way of supporting other operations.

a) Range queries

The proposed architecture allows queries that span a certain specific range of variables. For example, services consumption model of a user between 9:00 and 14:00 on working days.

In this respect, we substitute equation (9) with:

$$\mathbf{H}_m(\text{Range}) = \sum_{\forall p \in \text{Range}} \mathbf{H}_{pm} \quad (12)$$

In this manner, the services consumption model shall be made of the sum of services consumption models that are associated to each of the nodes within a contextual range.

b) Absence of contextual variables

In situations where all contextual variables are not available, the matrix of contextual nodes shall be reduced by one dimension, substituting the reduced dimension by the sum of all associated services consumption models. In case there is more than one variable missing, the previous operation shall be repeated.

III. IMPLEMENTATION AND RESULTS

In order to evaluate the proposed system, we have implemented a user simulator. It allows to simulate the behaviour of a user by executing actions (accept, refuse or omit services) and simulate contextual situations, thus allowing the visualisation of the changes caused in the learning model.

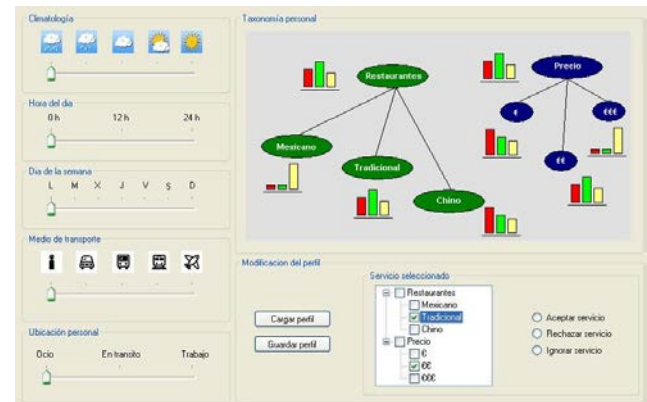


Figure 8: User interface in the fulfilled implementation.

We have added to this simulator all the necessary functions in order to make a correct evaluation of the proposed system.

Five contextual variables have been established: hour, day of the week, means of transportation, activity and climatology. This way, we have discretised the contextual space in the following manner:

TABLE I
DISCRETISATION OF CONTEXTUAL SPACE

VARIABLE	DISCRETISATION	DISTANCE FUNCTION
Hour	3 hour intervals	Lineal
Weekday	7 seven days	Uniform among labor days and lineal among labor days, Saturdays and Sundays
Weather	5 levels: snow, heavy rain, rain, clouds, sun	Lineal
Transport	5 transports	uniforme
Activity	Three activities_ leisure, work and on transit.	uniforme

The implemented simulator allows calculating the services consumption model for a specific contextual situation. It also allows for the continuous learning based on the actions of the simulated user. This functionality will allow simulating the user action when facing a service proposal by the system. It updates the services consumption model in each action and permits the variation of the diffusion parameters. Based on this simulator, the computational cost of the algorithms, the effects of the diffusion mechanisms in learning and the similarity between the user model and the learnt model have been quantified. In this sense, models of 10 users have been tested and analysed where each presenting nearby contextual situations were responded in accordance with the pre-established user model. After each interaction, the system checked the learnt model against the real implemented model until reaching convergence in those contextual situations.

TABLE II

MODEL	Services consumption model	DIFFUSION COEFFICIENTS IN TAXONOMY
Model 1	Only closest contextual node	No
Model 2	Distance Weighting	No
Model 3	Only closest contextual node	Yes
Model 4	Distance Weighting	Yes

As it can be seen in Table II, we have tested four different architectures. The first one uses only the information obtained in the closest contextual node to the contextual situation of the query. The second

architecture takes into account the diffusion from all the nearby contextual nodes (weighting the distances). The third one includes the diffusion coefficients in the services taxonomy but it only uses information from the closest contextual situation in the contextual matrix. The fourth one combines the diffusion among the different contextual situations and the diffusion in the services taxonomy as proposed in this paper.

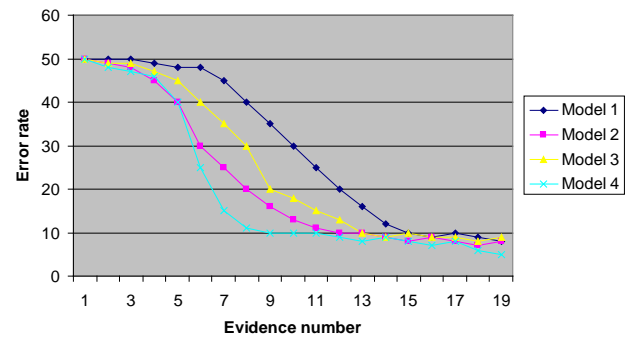


Figure 9. Results obtained with the different models As described in Table II.

The obtained results show that the use of the information of all nearby contextual nodes allows the use of fewer evidences for convergence. In turn, we have checked that double diffusion requires a lesser number of evidence than the other methodologies. This is due to the fact that learnt knowledge in similar situations and the knowledge of similar services consumption allows for a more precise reasoning of the behaviour of the user with a lesser number of evidences.

IV. CONCLUSIONS AND FUTURE WORKS

This article has shown the conceptual design of a learning algorithm for establishing the preferences of the user based on his contextual situation and his contextual history.

The proposed algorithm solves some difficulties that exist in other types of learning algorithms:

- Absence of learning evidence.
- Problems in the model generalisation.
- Need for a continuous retraining.
- Size of the model (memory).

The proposed methodology, specifically designed to model the behaviour and consumption of services by the user, is characterised by being able to be scaled. This allows increasing the size of the system in a dynamic manner and thus representing any service in a generic manner; at the same time, it permits a

progressive adaptation to the user, without risks of overtraining that would exist through the use of traditional learning systems. It also improves the convergence time (required evidences) and diminishes the error more effectively than other methodologies.

Besides this it is able to infer knowledge from unknown situations or for new services based on these diffusion properties and it faces the lack of contextual variables (e.g. forecast) eliminating their influence while keeping the remaining information. Moreover, advanced querying is supported in order to extract enhanced information about user's or group specific behaviour that can be used for marketing purposes.

Future noteworthy tasks will include the enhancement of the model by considering 2nd order interaction among the contextual dimensions and the development of distance functions supporting functions correlated among the different contextual variables, extend the implementation to a real mobile environment based on Android mobile phones and to integrate this learning model with a prediction system of future contextual situations based on ant colony optimisation that will allow not only to predict the most appropriate services for the user in a present contextual situation but also to provide the user with desired services in possible future contextual situations.

Acknowledgments

This work is partially underwritten by the Ministry of Industry, Tourism and Trade of the Government of Spain under the research project CENIT-2008-1019.

References

- [1] « M:Ciudad » Project, 7th Framework Programme, Coordinator Telefonica I+D www.mciudad.org
- [2] Tveit, A. 2001. Peer-to-peer based recommendations for mobile commerce. In *Proceedings of the 1st international Workshop on Mobile Commerce* (Rome, Italy). WMC '01. ACM, New York, NY, 26-29.Kjh
- [3] Kuo, M., Chen, L., and Liang, C. 2009. Building and evaluating a location-based service recommendation system with a preference adjustment mechanism. *Expert Syst. Appl.* 36, 2 (Mar. 2009), 3543-3554.
- [4] Park, M., H. Location-Based Recommendation System Using Bayesian User's Preference Model in Mobile Devices. In Proc. UIC'07 (Hong Kong, China, July 2007). Springer Press:1130--1139

- [5] Luther, M., Fukazawa, Y., Wagner, M., and Kurakake, S. 2008. Situational reasoning for task-oriented mobile service recommendation. *Knowl. Eng. Rev.* 23, 1 (Mar. 2008), 7-19.
- [6] Kwon, O. and Kim, J. 2009. Concept lattices for visualizing and generating user profiles for context-aware service recommendations. *Expert Syst. Appl.* 36, 2 (Mar. 2009), 1893-1902.
- [7] van Setten, M., Pokraev, S., Koolwaaij J.: Context-Aware Recommendations in the Mobile Tourist Application COMPASS. In Nejdl, W. & De Bra, P. (Eds.). AH 2004, LNCS 3137, Springer-Verlag, (2004)