

A hybrid context aware system for tourist guidance based on collaborative filtering

G. Fenza, E. Fischetti, D. Furno, V. Loia
Department of Computer Science &
CORISA (Consorzio Ricerca Sistemi ad Agenti)
University of Salerno
E-mail: {gfenza, efischetti, dfurno, loia}@unisa.it

Abstract— In the area of ambient intelligence there is a need to address user needs according with context features. Recently, the synergy between context aware computing and collaborative filtering is leading to enhance recommender systems with capabilities always nearer to user needs. Specifically, in the domain of tourism it is useful to proactively suggest right sets of attractive locations, events and so on. This work defines a context aware recommender system aimed at suggesting pertinent points of interest (POIs) to tourists. In particular, the approach is strongly based on the synergy between soft computing and data mining techniques. The general framework integrates user profiles, history of social networking and POIs data. Then by defining collaborative filtering approach on the history meaningful POIs are extracted. Indeed, soft computing techniques are mainly applied in order to support activity of unsupervised users and POIs classification. On the other hand, data mining techniques are exploited in order to extract rules able to associate user profile and context features with an eligible set of recommendable POIs. Experimental results show performance in terms of recommendations accuracy.

Keywords—component; context-awareness; fuzzy data analysis; data mining; recommender system; points of interest

I. INTRODUCTION

The tourism sector is one of the biggest in the global market employing 200 million people and serving 700 million tourists worldwide. Furthermore, this trend is intended to grow in the next years. Additionally, Information Technology with its constant progress has decisively affected the way that tourism services are performed and provided. In effect, tourists thanks to advances in Information and Communication Technologies (ICTs) are enabled to access tourist information as well as to reservations and plans in a more reliable, accurate, inexpensive and efficient way. As a result, services related to tourism are increasingly influenced by the Internet and related technologies, so much so that it is possible to find most of them on line without the necessity of relying on travel agency. On the other hand, ICTs and the Internet have increased the number of choices so dramatically that is very difficult for the users to retrieve what they are looking for. The appropriate discovery of tourism services is one of the main drawbacks when people travel [1]. According to [2] research, the reasons of consumers not purchasing travel products online are the lack of personal service, security issues, lack of experience, and time consuming. Furthermore, different users have different preferences and needs. So a remedy for these problems is to develop systems with an ability to adapt their behavior to the

profile, context and other features of individual users and groups of them. An interesting solution for reducing Internet researches complexity and overcoming the aforementioned difficulties is provided by so called recommender systems. Recommender systems have already been applied to several application domains such as finding books, movies, tv-programs, news, music, web-pages, etc. One of the main characteristics of a recommender system is that it can personalize its interaction to each individual user according with context. In that case, personalization involves the design of enabling systems to capture or infer the needs of each person and then to satisfy those needs in a known context [3]. Context-awareness computing and recommender systems can enhance and complement each other; they both help users in finding relevant and interesting objects based on their interests and current context.

In the specific of this paper, our intention is to propose a hybrid context-aware recommender system aimed at adaptively and proactively suggesting points of interest (POIs) to tourists according to their profile, but also to their context. In particular, the work defines an integrated environment applying content based and collaborative filtering on historical data through soft computing and data mining techniques in order to predict and suggest POIs to users. In other words, the main goal of the proposed work is related to define context-aware recommender system whose quality of feedback relies on the acquisition of user context and profile by means of a robust theoretical approach. So, the overall framework is composed of two main workflows: knowledge discovery and recommendation. In particular, Knowledge Discovery Process concerns the fuzzy data analysis of users and POIs. Indeed, fuzzy clustering and classifiers enable us to integrate cluster of users and POIs with historical data. Then association rule mining activity points out a correspondence between user, contexts and cluster of POIs. On the other hand, Recommendation Process foresees activities of user classification and rule matching in order to provide some ranked recommendation about pertinent POIs.

The paper is organized as follows: Section II presents some related works available in literature; Section III proposes a theoretical background for supporting workflows description; Section IV and Section V introduce workflows of the framework by focusing on the Knowledge Discovery Process and Recommendation Process, respectively. Then, Section VI describes the system performance in terms of accuracy. Finally, conclusions and future works close the paper.

II. RELATEDWORKS

The aim of this section is to depict the research trend with reference to context aware recommender systems in ambient intelligence for tourist guidance and to point out the benefits of the proposed approach. Among research works regarding personalized and/or location-based recommender systems based on web service, with specific reference to tourist domain, Kakalettris et al. [4] address the design and implementation issues for delivering location-based tourism-related content services. By illustrating an example of generic tourist guide, Scherp and Boll [5] emphasize on supporting the dynamic generation of personalized multimedia content in respect of the tourist's interests and preferences, current location and environment, and mobile device used. Aiming at effectively coordinating and integrating disparate information and service resources anytime and anywhere, Chiu and Leung [6] propose a ubiquitous tourist assistance system that is built upon multi-agent and semantic web technologies for providing personalized assistance and automation to the tourists with different preferences and often changing requirements during their trips. Five agent clusters specified in the system architecture include tourist assistant, ontology maintenance and search, requirement/preference management, package planning, and local tour planning. Ricci and Nguyen [7] present a critique-based mobile recommender system that lets users expressing session-specific preferences in addition to the long-term collected preferences. The objective is to provide on-tour system support to mobile travelers for selecting products or services based on the integration of both types of preferences. The COMPASS system [8] allows easy creation of context-aware personalized applications and the services that are part of such a platform, including a service registry, a context manager and a recommendation engine. Console et al. [9] created a prototype system called MastroCARonte, which provides personalized services that adapt to the user and his context onboard cars. This research focuses on the effects of having such adaptive systems onboard cars.

Among research works regarding more specifically context aware and adaptive recommender systems for tourist guidance, the Intrigue system [10] is an interactive agenda offered by a tourist information server that assists the user in creating a personalized tour along tourist attractions. This research focuses on planning and scheduling a personalized tour taking into account the location of each tourist attraction and the interests of the user. A last example of a tourist guide project is the GUIDE project of the University of Lancaster [11]. This project develops a context sensitive tourist guide for the visitors to the city of Lancaster. Specifically, in this project Users have a portable PC that communicates with a Wireless LAN to retrieve information. The used context information is knowledge of the user and knowledge of the environment (including the physical location of the user). Based on the location and user preferences, visitors can obtain information in two ways. All the required information is broadcasted by cell base-stations to portables either as part of a regular schedule (information relevant for the region the station covers) or in response to user requests. The system determines the location of the user by the cell coverage, or the user can specify in which location he/she is.

The aforementioned works provide the user with tourist recommendations but do not manage uncertainty among user interests. Furthermore, most of them are only content based, omitting collaborative information that is a crucial aspect for user recommendations personalization. In this sense, our approach presents an hybrid system which main aim is related to provide an integrated environment for tourist guidance that combines collaborative and content based filtering. Indeed, the main idea of this work is related to enhanced context awareness by means of data analysis, computational intelligence and rule mining techniques in order to obtain a recommender system able to suggest users in the tourist domain with pertinent POIs. Strictly related approach is described in [12] where a personalization method for tourist Point of Interest (POI) Recommendation is defined by considering both user preferences and collaborative aspects. Indeed, [12] exploits a memory based algorithm to carry out suggestions. In particular, our system define a model based approach that trains itself by performing activities of fuzzy context data analysis in order to support both POIs and user's context identification by means of fuzzy classifiers.

III. THEORETICAL BACKGROUND

The overall system takes into account users data, available POIs and historical data in order to provide proper recommendations (i.e., POIs) according to his/her preferences and context. In particular, algorithms of clustering and classification allow to tackle the huge amount of incoming data (Users and POIs). On the other hand, association rules mining algorithm on historical data allows to extract correspondences with pertinent tourist POI categories.

Since, this work relies on a mathematical model that exploits soft computing and data mining techniques, this section provides some theoretical basis in order to support understanding of the overall framework workflows.

A. Fuzzy clustering

Once data, about users and POIs, are collected and arranged adequately, they are processed by Fuzzy C-Means (briefly FCM) algorithm [15][16]. FCM is an unsupervised clustering algorithm, which considers a prior fixed number of clusters. In order to give more accurate partitions of data, many researchers have studied the cluster validity criteria [17]. In our past studies and experimentations we have achieved good results by the application of Xie and Beni's validity measure [18]; so we have also selected it for the approach proposed in this work. Advanced approaches to (fuzzy) clustering methods could be exploited, according to the nature of the data space and the inherent relationships. As said, we exploit the well-know FCM algorithm, particularly useful for flexible data organizationally. It takes as input a data matrix and it tries to get an "optimal" partitioning of the feature space (composed by the data matrix). FCM aims at maximizing the homogeneity, grouping into the same cluster the patterns which are closer. Herein, each pattern is a row of matrix. FCM recognizes spherical "clouds of points" in a multi dimensional data space (i.e. data matrix) and each cluster is represented by its center point (prototype or centroid). The function minimizes the weighted sum of the

distances between data points x and the centroid v , according to this formula:

$$V(U) = \sum_{l=1}^c \left(\sum_{i=1}^n w_{i,l} \sum_{j=1}^p (x_{i,j} - v_{l,j})^2 + \gamma \sum_{i=1}^n w_{i,l} \sum_{j=p+1}^m \partial(x_{i,j}, v_{l,j}) \right) \quad (1)$$

where $c \geq 2$ is the number of clusters, $w_{i,l} \in [0,1]$ is the membership degree of $x_{i,j}$ in the l -th cluster and m is the features number. In particular, since our feature set is mixed, because of numeric and nominal features, the m features are partitioned in two set, namely a numeric feature set composed by the features from 1 to p and a nominal feature set composed by the features from $p+1$ to m . As a result, the first term in (1) is the squared Euclidean distance measure on the numeric attributes and the second term is a semantic distance on the nominal attributes, where ∂ is equal to a linear combination between a semantic distance taxonomy based [13] and Google similarity distance [14]; finally, γ is used to weight features. After the FCM execution, data partitions are returned, in a prior fixed number c of clusters. Anyway, with the application of validity measures it is possible to execute clustering without a prior number of clusters specification.

Details about how Fuzzy Clustering has been applied in this framework are given in Section IV and V, where the workflows of the whole proposed system are detailed.

B. Fuzzy Classifier Design

In computational intelligence literature, fuzzy rules have been often exploited in different approaches for decoding the knowledge in a way closer to the human understanding. In neural network [22] approach, if-then fuzzy rules enable the emulation of the decisions or reasoning of a human expert, whereas fuzzy number and compositional rules of inference have been employed for fuzzy decision making in medical diagnosis [23]. The fuzzy if-then rules provide an interpretation of the clusters, closer to the human understanding and easier to evaluate, with respect to the values range of each feature. Outputs of previous phase are prototypes and a partitioning matrix of the data patterns. This matrix contains the memberships of the elements of the given data set in each of the c clusters. Generally, the linear interpolation of memberships is computed; in order to easily assign linguistic label, the membership function is computed by linear interpolation of the projected membership to each axis (according to the n -dimension space), for each cluster. More specifically, exploiting the cylindrical extension technique [15] (a projection-based method of n -dimensional argument vector), the generic i -th fuzzy cluster can be described by a fuzzy rule, where each features is described through a fuzzy set. Each fuzzy cluster K_i with $i = 1, \dots, c$ can be represented through n functions $A_{i1}, A_{i2}, \dots, A_{in}$, obtained by the projected and interpolated memberships. The generated rules are subject to a proper validation by experts of the specific domain and then exploited for suitable classifications. In our work, the fuzzy rules are exploited to classify new tourist POIs and users. Just to provide an example, a fuzzy rule, supposed to be in the form of Takagi-Sugeno model [19] is described as follows:

R_i : **If** x_1 is A_{i1} **AND** x_2 is A_{i2} **AND** ... **AND** x_n is A_{in}
then $y_i = a_{i1} x_1 + a_{i2} x_2 + \dots + a_{in} x_n + b_i$,

with weight $[w_i] i = 1, \dots, c$

where R_i is the i -th rule, $x_1, x_2, \dots, x_n$ are the input variables, $A_{i1}, A_{i2}, \dots, A_{in}$ are the fuzzy sets assigned to corresponding input variables, variable y_i represents the value of the i -th rule output, and $a_i = [a_{i1}, a_{i2}, \dots, a_{in}]$ and b_i are parameters of the consequent function. As said before, the antecedent fuzzy sets are usually achieved by projecting the membership degrees in the fuzzy partitions matrix U onto the axes of individual antecedent variable x_j to obtain a point-wise defined antecedent fuzzy set A_{ij} . Then we approximate it by a normal bell-shaped membership function [20]; indeed the uniform structure of bell-shaped function is suitable for identification, analysis and optimization. Hence, each fuzzy set A_{ij} is calculated from the sampled input data $x_j = [x_{1j}, \dots, x_{nj}]^T$ and the fuzzy partition matrix $U = [u_{ij}]$ as follows:

$$A_{ij}(z) = \exp \left\{ -\frac{1}{2} \left(\frac{z - \alpha_{iq}}{\beta_{iq}} \right)^2 \right\} \quad (2)$$

where

$$\alpha_{iq} = \frac{\sum_{j=1}^n u_{ij} x_{qj}}{\sum_{j=1}^n u_{ij}} \quad \text{and} \quad \beta_{ij} = \sqrt{\frac{\sum_{j=1}^n u_{ij} (x_{qj} - \alpha_{iq})^2}{\sum_{j=1}^n u_{ij}}} \quad (3)$$

represent the mean and standard deviation of the bell-shaped membership function, respectively.

On the other hand, the computation of parameters $a_i = [a_{i1}, a_{i2}, \dots, a_{in}]^T$ and b_i in the consequent part requires a deepened study, in order to evaluate the firing strength w_i of the i -th rule and the value of y_i [21].

Goal is to evaluate the final output of the TS fuzzy model y_i for an arbitrary x_j input sample, which is calculated using the following formula [21]:

$$\hat{y}_k = \sum_{i=1}^c [w_i(x_k)(x_k a_i + b_i)] \quad (4)$$

with $k = 1, 2, 3, \dots, n$ and where $w_i(x_k)$ represent the firing strength of i -th rule for the k -th pattern. Thus, in order to compute the parameters of consequent part of a rule, a regression model of the compact form $\hat{Y} = X' [a_i^T, b_i] + \epsilon$ is evaluated: herein ϵ is the approximation error, X' assumes the form: $X' = [w_1, x_k^T, 1, w_2 x_k^T, 1, \dots, w_c x_k^T, 1]$. Vectors $[a_i^T, b_i]$ are determined using the least-squares method [21].

C. Rule Mining

The proposed approach foresees activities of rule mining on historical data, in order to associate user contexts and tourist POI categories. The problem of mining association rules over was introduced in [24]. An example of an association rule in our case might be that 90% of users arriving in a tourist town in the morning, being scientific researchers and walking, select POI about the nearer science museum. The following is a

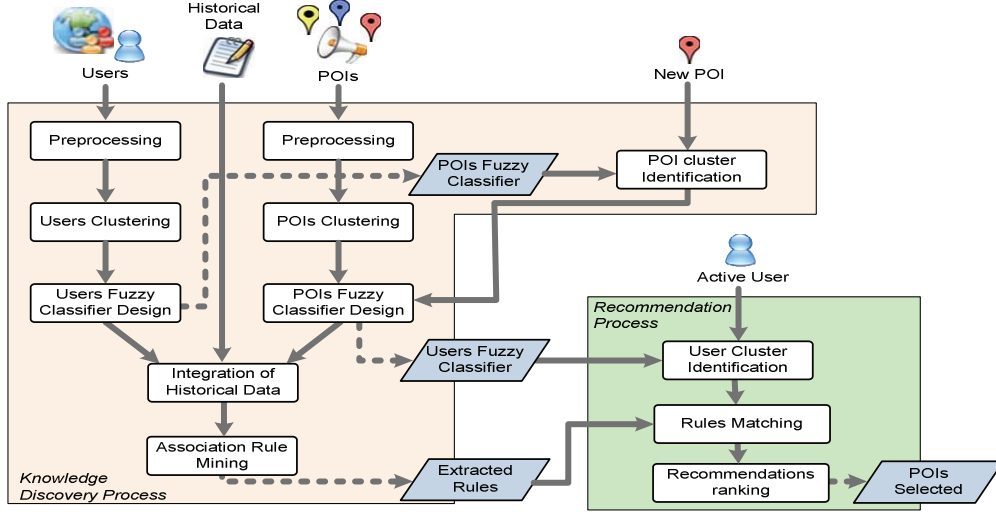


Figure 1. Workflows of the overall framework: Knowledge Discovery Process, Recommendation Process.

formal statement of the problem [24]: given a set of items $I = \{I_1, I_2, \dots, I_m\}$ and a database of transactions $D = \{t_1, t_2, \dots, t_n\}$

where $t_i = \{I_{i1}, I_{i2}, \dots, I_{ik}\}$ and $I_{ij} \in I$, an association rule is an implication of the form $X \Rightarrow Y$ where $X, Y \subset I$ are sets of items called itemsets and $X \cap Y = \emptyset$. X is called the antecedent and Y is called the consequent. We are generally not interested in all implications but only those that are important.

Here, two features called support and confidence are usually used to measure the importance of association rules. The support for an association rule $X \Rightarrow Y$ is the percentage of transactions in the database that contain $X \cup Y$. The confidence for an association rule $X \Rightarrow Y$ is the ratio of the number of transactions that contain $X \cup Y$ to the number of transactions that contain X . Rules that satisfy both the minimum support (MinSup) and confidence (MinConf) thresholds are called strong rules. An algorithm for finding all association rules, henceforth referred to as the AIS algorithm, was presented in [24]. Another algorithm for this task, called the SETM algorithm, has been proposed in [25].

In this work, we use an algorithm, Apriori [26], that differs fundamentally from these algorithms and exhibits better performance. Furthermore, in our future works we would like to experiment a Formal Concept Analysis-based Association Rule Mining Algorithm as proposed in [27].

IV. KNOWLEDGE DISCOVERY PROCESS

As highlighted in Fig.1, the overall framework consists of two main workflows: *Knowledge Discovery Process* and

Recommendation Process. Following subsections detail the first one.

A. Users and POIs Preprocessing

Preprocessing tasks on Users and POIs data allow to start the process of Knowledge Discovery. In particular, gathering and normalization activities take place in order to prepare data for the next phase, that is, *Users and POIs clustering*. Specifically, preprocessing carries out the user and POIs data sets, as shown in Table I (A) and (B). In the example the user data set includes three features (i.e., age, skill and interest), whereas POIs data set is composed of five features (i.e., type, lat, long, when, period). Furthermore, as highlighted in the Table I features may be numerical or nominal. For instance, in the user data set features space, *age* is a numerical attribute whereas *skill* and *interest* are rather nominal.

B. Users and POIs clustering

Analogously with the previous step, clustering execution takes place on user and POIs data sets. We stress that users and POIs are considered in two off-line separate processes and the subsequent extraction of classification rules enable us to properly add new users and POIs to clusters. So, the execution of these two distinct processes should not be time consuming, because there is no need to execute clustering again whenever a new user or POI must be considered. Specifically, the framework exploits FCM algorithm in order to arrange data according to their similarities. In particular, in order to use mixed features, namely, numeric and nominal, a light variant of FCM is performed.

TABLE I. (A) USER DATA SET AND (B) POIS DATA SETS

(A)				(B)					
ID User	Features			ID POI	Features				
	age	skill	interest		type	lat	long	when	period
U1	28	researcher	Science	P1	zoo	41.24	14.06	all day	always
U2	43	truck driver	restaurant reservation	P2	restaurant	42.23	12.34	all day	ever, except monday
U3	22	student	entertainment	P3	disco	44.55	16.80	from midnight	every weekend
U4	24	student	Science	P4	restaurant	40.20	14.02	all day	ever, except tuesday
U5	28	lawyer	restaurant reservation	P5	museum	42.20	19.46	from 9 a.m.	always

TABLE II. (A) USER PROFILE PARTITION MATRIX AND (B) POI PARTITION MATRIX

ID User	(A) Clusters			ID POI	(B) Clusters			
	UC1	UC2	UC3		PC1	PC2	PC3	PC4
U1	0,68	0,12	0,2	P1	0,6	0,05	0,3	0,05
U2	0,2	0,7	0,1	P2	0,07	0,7	0,03	0,2
U3	0,05	0,35	0,6	P3	0,06	0,25	0,04	0,65
U4	0,7	0,1	0,2	P4	0,07	0,7	0,03	0,2
U5	0,15	0,65	0,2	P5	0,2	0,07	0,7	0,03
...								

The input of this algorithm are matrices similar to Table I (A) and (B), and the output consists of partition matrices that synthesize users and POIs in different clusters (see Section III for more details).

Anyway, since we use a fuzzy algorithm for clustering, we obtain that each user or POI belongs to several clusters with different membership degrees. Table II shows part of the aforementioned partition matrices.

C. Users and POIs fuzzy classifier design

In this step, the previously obtained user and POI clusters are used to design fuzzy classifiers that are useful to relate new incoming data, such as POI or active user, with the existing clusters.

This design process extracts lists of if-then rules characterizing clusters obtained in the previous step. By designing classifiers the system trains itself in order to avoid clustering execution every time to position new POIs or new users. Theoretical details about this phase can be found in Section III. Table III (A) and (B) shows part of rules concerning users and POI classifiers.

D. Integration of historical data

One of the main aim of the proposed approach is to apply collaborative filtering on historical data, in order to recommend POIs to users that meet their context and preferences. Each record in such historical data asserts interest of a certain user that is identified in a specific context for a proper tourist POI. In this step, an integration between historical data and clusters obtained in the previous steps is achieved by exploiting fuzzy classifiers. Specifically, for each row of the historical data, if the user data in a historical transition are:

- age = 20

- skill = student
- interest = science

then these data can be replaced with UC1 (User Cluster 1) according to classifier in Table III (A). Similarly, if the POI attribute values in a transition are the following:

- type = museum
- lat = 42.20
- long = 19.46
- when = from 9 a. m.
- period = always

then, data can be replaced with PC3 (POI Cluster 3) according to classifier in Table III (B). In Fig. 2, are highlighted the substitutions described above, but it is important to point out that a historical transition is also characterized by context data, that is, the context where the user was when he/she selected the specific POI. This further information will be exploit in the next step and it will be of use to provide context aware recommendations to users.

E. Association rules mining

As said before (see Section III) a Apriori algorithm have been selected to carry out association rule on historical data. In particular, in this step, extracted rules synthesize the relation between cluster of users, context and POIs clusters. These rules are useful to recommend eligible POIs to the tourists according with context and user features. Extracted rules are in the following form:

$\{lat=40/42, long=13/15, time=morning, date=saturday, walk/drive=walk, User Cluster=UC1\} \Rightarrow \{Poi Cluster=PC3\}.$

Furthermore, to each rule is associated an importance degree by two features, namely, support and confidence. These features allow to set thresholds (i.e. MinSup, MinConf), in order to trigger only the so called strong rules (see Section III).

V. RECOMMENDATION PROCESS

The results obtained from *Knowledge Discovery Process*, are exploited to recommend POIs to users according to their context and preferences. In particular, the recommendation process, described in this section, consists of three steps detailed below.

TABLE III. (A) USER CLASSIFIER AND (B) POI CLASSIFIER

(A)				
ID User Profile Rule	User Profile Rules			
	IF			THEN
	age	profession	interest	Cluster
UPR1	young	student	science	UC1
UPR2	middle	truck driver	restaurant reservation	UC2
...				

	(B)					
ID POI Rule	POI Rules					
	IF					THEN
	type	lat	long	when	period	Cluster
PR4	restaurant	40/45	13/20	all day	ever, except tuesday	PC2
PR5	museum	40/45	13/20	morning	always	PC3
...						

ID Transaction	Features													
	Context					User Profile				POI				
	lat	long	time	date	walk/drive	age	skill	interest	type	lat	long	when	period	
T1	40.24	13.06	morning	saturday	walk	18	student	science	museum	42.20	19.46	from 9 a.m.	always	
T2	40.24	13.06	morning	sunday	walk	20	student	science	museum	45.36	18.07	from 4 p.m.	every weekend	
T3	42.81	12.30	afternoon	sunday	walk			animals	zoo	42.24			always	
T4	41.35	12.20	morning	saturday	drive			animals	zoo	41.24			always	
T5	41.50	13.88	evening	monday	drive	43	truck driver	restaurant reservation	restaurant	43.20	14.02	all day	ever, except tuesday	
T6	40.50	13.88	evening	tuesday	drive	43	truck driver	restaurant reservation	restaurant	40.20	14.02	all day	ever, except monday	
...														

Figure 2. Historical Social Networking and clusters association

TABLE IV. ACTIVE USER CLASSIFICATION

ID User	Clusters			
	UC1	UC2	UC3	
U1	0,1	0,2	0,7	
U2	0,65	0,2	0,15	
U _{Active}	0,1	0,7	0,2	
...				

A. User cluster identification

The aim of this step is to identify the cluster of the active user that have “checked-in” to the application system. Thanks to the user classifier, it is not needed to run clustering again in order to determine the belonging of the active user to a specific cluster. When there is a new active user or already registered user changes him preferences the classification take place. In that cases, through the execution of classification rules (see Table III) a new row is added to the user partition matrix as shown in the Table IV. More specifically, Table IV highlights that the considered active user belongs to User Cluster 2 (UC2) with a membership degree higher than the other clusters; therefore, the active user is considered belonging to UC2.

B. Rules matching

In this step, the triggering of a specific rule is achieved according to higher degree of matching between the active user and extracted rules. In particular, a matching degree is calculated by considering a matching matrix M where the rows are represented by the extracted rules ranked according to their support and confidence features, whereas the columns are represented by active user context features and user cluster. Let be M the matching matrix, the maximum matching degree M_{max} is given by the following formula:

$$M_{max} = \max_k \left(\frac{\sum_{j=1}^m \mu_{ij}}{m} \right) \quad i = 1, \dots, n \quad k = 1, \dots, n \quad (5)$$

$$\mu_{ij} = S(u_j, a_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, m \quad (6)$$

$$S(u_j, a_{ij}) = 1 - \partial(u_j, a_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, m \quad (7)$$

where u_i represents the context and profile features of the active user, a_{ij} represents the antecedent clause of each extracted rule in the knowledge discovery process and ∂ is defined as described in Section III. So, Table V shows an example of maximum matching degree between the active user and the extracted rules according to (5), (6) and (7).

TABLE V. MAXIMUM MATCHING BETWEEN ACTIVE USER AND RULES

Matching User/Rule	Context					User Cluster	POI Cluster
	lat	long	time	date	walk/drive		
User _{Active}	41.50	13.88	evening	monday	drive	UC2	???
Rule _{Active}	40/42	13/15	evening	monday	drive	UC2	PC2
$S(\text{User}_{Active}, \text{Rule}_{Active})$	0,75	0,68	1	1	1	0,7	
M_{max}	$(0,75 + 0,68 + 1 + 1 + 1 + 0,7)/6 = 0,855$						

C. Recommendations ranking

In the last step of this process, we determine the ranked list of POIs that may be suggested to the user.

In particular, the previous step determines the maximum matching between a new user that starts a transaction and the extracted rules. Rules matching evaluation, defines the highest match of the active user respect with extracted rules. Since, the consequent of a rule is represented by a POI Cluster, this match points out that POIs belonging to a specific cluster can be recommended as relevant for the active user. For instance, user is in a certain context and with a determinate profile if the active rule is the following:

$$\{lat=40/42, \quad long=13/15, \quad time=evening, \quad date=monday, \quad walk/drive=drive, \quad User \text{ Cluster}=UC2\} \Rightarrow \{Poi \text{ Cluster}=PC2\}$$

then, the ranked list of POIs consists of those belonging to the cluster PC2, as shown in Table V.

TABLE VI. POIS RECOMMENDATION RANKING

ID POI	Features					Cluster	Recommendation Ranking
	type	lat	long	when	period	PC2	$\rho_i = b_{ic} \times M_{\max}$
P8	restaurant	40.33	14.14	all day	ever, except thursday	0,8	0,684
P2	restaurant	42.23	12.34	all day	ever, except Tuesday	0,7	0,6
...	...	...	...	...	...	...	...
P5	restaurant	40.20	14.02	all day	ever, except monday	0,55	0,47
P3	hotel	42.90	15.30	all day	always	0,47	0,4
P1	zoo	41.24	14.06	all day	always	0,34	0,3
...	...	...	...	...	...	...	...

In order to carry out the ranked list of POIs the framework assigns a ranking value to each POI corresponding to b_{ic} , that is the belonging degree of the i -th POI to the cluster c (i.e., PC2) in the consequent of rule with maximum matching in the antecedent and n is the number of POIs. Table VI shows an example of POIs recommendation ranking.

I. EXPERIMENTAL RESULTS

This section presents the performance results of our proposed recommender system. In particular, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used in order to evaluate the accuracy of the system. Specifically, by considering different set of features the system has been evaluated in the following variants:

- LS - System considering Location
- LUC - System considering Location and User Context
- LUP - System considering Location and User Profile
- HRS - Hybrid Recommender System considering Location, User Context and User Profile

The system has been evaluated with 20 users and 200 POIs. Obviously, little by little the history data grow the system retrieve more meaningful results (the history considered contains about 1000 records). As previously said, MAE and RMSE are used for performance evaluation of the proposed recommender system. MAE is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. On the other hand, RMSE measures the average magnitude of the error, that is, the difference between forecast and corresponding observed values, where each of these differences is squared and then averaged over the sample; so, a relatively high weight is given to large errors. MAE and RMSE can be used together to diagnose the variation in the forecasting errors.

Since the lower the MAE, the more accurately the recommendation engine predicts user ratings [28], the obtained results, shown in Fig. 3, highlight a clear superiority of the overall proposed recommender system among the other variants. Furthermore, since the greater difference between MAE and RMSE, the greater the variance in the individual errors in the sample, Table VII points out that all the errors are of the same magnitude.

II. CONCLUSIONS AND FUTURE WORKS

As a result of demands continuously growing in ambient intelligence for systems able to exploit context for satisfaction of user preferences, a context aware recommender system for tourist guidance is proposed in this paper.

The proposed system presents several improvements. Firstly, user context and profile are considered in order to effectively identify his preferences. Uncertainty among user interests is an important issue which is clearly considered in this approach in order to establish an effective recommendation strategy. Additionally, this approach provides dynamic user clustering for active users and supports them with a personalized and adaptive environment. Moreover, any POIs change can be simply notified to all interested users. Anyway, there are some limitations in this research area that we consider as future research directions. FCM alternative clustering algorithms that work better with mixed data can be experimented as well as alternative rule mining algorithms using fuzziness and based on lattice. However, the proposed framework confirms the useful of considering context factor in recommender systems and provides us with reasonable results in terms of recommendation accuracy as evidenced by the experimental results.

TABLE VII. RECOMMENDER SYSTEMS MAE COMPARISON

Recommender System	LS	LUC	LUP	Hybrid Recommender System
MAE	0,932	0,869	0,851	0,811
RMSE	0,932	0,869	0,851	0,811

REFERENCES

- [1] Tak Kee Hui, David Wan, Alvin Ho, Tourists' satisfaction, recommendation and revisiting Singapore, Tourism Management, Volume 28, Issue 4, August 2007, Pages 965-975.
- [2] Wolfe, K., Hsu, C. H. C. & Kang, S. K. (2005). Buyer Characteristics Among Users of Various Travel Intermediaries. Journal of Travel & Tourism Marketing, 17(2), 51-62.
- [3] Doug Riecken. 2000. Introduction: personalized views of personalization. Commun. ACM 43, 8 (August 2000), 26-28.
- [4] G. Kakaletis, D. Varoutas, D. Katsianis, T. Spicopoulos, and G. Kouvas. 2004. Designing and Implementing an Open Infrastructure for Location-Based, Tourism-Related Content Delivery. Wirel. Pers. Commun. 30, 2-4 (September 2004), 153-165.

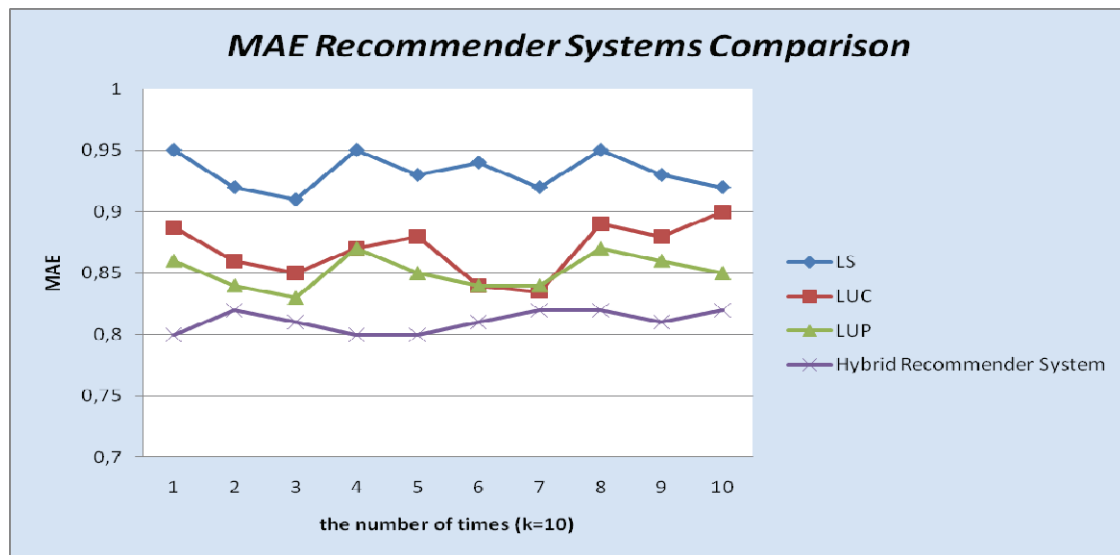


Figure 3. Mean Absolute Error experimental results (k=10)

- [5] Ansgar Scherp and Susanne Boll. 2004. Generic support for personalized mobile multimedia tourist applications. In Proceedings of the 12th annual ACM international conference on Multimedia (MULTIMEDIA '04). ACM, New York, NY, USA, 178-179.
- [6] Dickson K. W. Chiu and Ho-fung Leung. 2005. Towards ubiquitous tourist service coordination and integration: a multi-agent and semantic web approach. In Proceedings of the 7th international conference on Electronic commerce (ICEC '05). ACM, New York, NY, USA, 574-581.
- [7] Francesco Ricci and Quang Nhat Nguyen. 2007. Acquiring and Revising Preferences in a Critique-Based Mobile Recommender System. IEEE Intelligent Systems 22, 3 (May 2007), 22-29.
- [8] Mark van Setten, Stanislav Pokraev, and Johan Koolwaaij. Context-Aware Recommendations in the Mobile Tourist Application COMPASS. In Adaptive Hypermedia and Adaptive Web-Based Systems, pages 235-244. 2004.
- [9] Console, L., Gioria, S., Lombardi, I., Surano, V., Torre, I.: Adaptation and Personalization on Board Cars: A Framework and Its Application to Tourist Services. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds) Adaptive Hypermedia and Adaptive Web-Based Systems, Proceedings of the second International Conference, AH 2002, Málaga, Spain, May, Springer-Verlag, LNCS 2347 (2002) 112-121.
- [10] Ardissono, L., Goy, A., Giovanna, P., Segnan, S., Torasso, P.: Ubiquitous User Assistance in a Tourist Information Server. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds) Adaptive Hypermedia and Adaptive Web-Based Systems, Proceedings of the second International Conference, AH 2002, Málaga, Spain, May, Springer-Verlag, LNCS 2347 (2002) 14-23.
- [11] Nigel Davies, Keith Mitchell, Keith Cheverst, Gordon Blair: Developing a Context Sensitive Tourist Guide, Proceedings of the First Workshop on Human Computer Interaction for Mobile Devices, Department of Computing Science, University of Glasgow, UK, May 1998, pp. 64-68.
- [12] Eui-young Kang, Hanil Kim, and Jungwon Cho. Personalization method for tourist point of interest (POI) recommendation. In Bogdan Gabrys, Robert Howlett, and Lakhmi Jain, editors, Knowledge-Based Intelligent Information and Engineering Systems, volume 4251 of Lecture Notes in Computer Science, chapter 48, pages 392-400. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2006.
- [13] El Sayed, A.; Hacid, H.; Zighed, D.; , "A New Context-Aware Measure for Semantic Distance Using a Taxonomy and a Text Corpus," *Information Reuse and Integration*, 2007. IRI 2007. IEEE International Conference on , vol., no., pp.279-284, 13-15 Aug. 2007.
- [14] Rudi L. Cilibrasi and Paul M. B. Vitányi. The Google Similarity Distance. Knowledge and Data Engineering, IEEE Transactions on, 19(3):370-383, 2007.
- [15] Hoppner F., Klawonn F., Kruse R., Runkler T. Fuzzy Cluster Analysis - Methods for Image Recognition. J. Wiley, N. York, 1999.
- [16] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York, 1981.
- [17] L.Jegatha Deborah, R.Baskaran and A.Kannan, A Survey on Internal Validity Measure for Cluster Validation, Anna University – Chennai.
- [18] X. Xie and G. Beni, "A Validity Measure for Fuzzy Clustering," IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol. 13, no. 8, pp.841 - 847, 1991.
- [19] Takagi, T., and Sugeno, M.: 'Fuzzy identification of systems and its applications to modeling and control', IEEE Trans. Syst., Man Cybern., 1985, 15, pp. 116-132.
- [20] C.C. Kung and J.Y. Su Affine Takagi-Sugeno fuzzy modelling algorithm by fuzzy c-regression models clustering with a novel cluster validity criterion, IET Control Theory Appl., 2007, 1, (5), pp. 1255-1265.
- [21] Dragan Kukolj Design of adaptive Takagi-Sugeno-Kang fuzzy models, Applied Soft Computing 2 (2002) 89-103.
- [22] S. Mitra,; Y. Hayashi,; Neuro-fuzzy rule generation: survey in soft computing framework IEEE Transactions on Neural Networks, Volume 11, Issue 3, May 2000 Page(s):748 - 768.
- [23] J.F.-F. Yao, J.-S.Yao, Fuzzy decision making for medical diagnosis based on fuzzy number and compositional rule of inference, Fuzzy Sets and Systems, 120 (2001), pp 351-366.
- [24] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In Proc. of the ACM SIGMOD Conference on Management of Data, Washington, D.C., May 1993.
- [25] M. Houtsma and A. Swami. Set-oriented mining of association rules. Research Report RJ 9567, IBM Almaden Research Center, San Jose, California, October 1993.
- [26] Rakesh Agrawal and Ramakrishnan Srikant. Fast Algorithms for Mining Association Rules, Proceedings of 20th International Conference on Very Large Data Bases, September 12-15, 1994, Santiago de Chile, Chile.
- [27] G.Fenza, V.Loia, S.Senatore, "Friendly Web Services Selection exploiting Fuzzy Formal Concept Analysis", Soft Computing, 2008.
- [28] Good, N., Schafer, B., Konstan, J., Borchers, A. Sarwar, B., Herlocker, J., Riedle, J.: Combining Collaborative Filtering with Personal Agents for Better Recommendation. In Proc. of the AAAI conference (1999) 439-446.