Matching User's Semantics with Data Semantics in Location-Based Services

Shijun Yu
EPFL-IC-LBD
INJ 237 Station 14
CH 1015 Lausanne
41 21 6936706
shijun.yu@epfl.ch

Lina Al-Jadir
EPFL-IC-LBD
INJ 230 Station 14
CH 1015 Lausanne
41 21 6935213
lina.al-jadir@epfl.ch

Stefano Spaccapietra
EPFL-IC-LBD
INJ 236 Station 14
CH 1015 Lausanne
41 21 6935210
stefano.spaccapietra@epfl.ch

ABSTRACT
Mobility is having a tremendous impact on the design and use of information technology. Information services, in particular, are heavily challenged by this new framework, which calls for innovative solutions. One of the major issues for flexible information services is how to be able to correctly understand what is being requested by users, and how to find information that is relevant to the request. This paper focuses on such semantic issues, aiming at outlining the general problem as well as the specificity attached to location-based services, one of the major trends in mobile information systems.

Categories and Subject Descriptors
H.3.5 [**INFORMATION STORAGE AND RETRIEVAL**]: Online Information Services – *data sharing.*

H.2.4 [**DATABASE MANAGEMENT**]: Systems – *query processing.*

H.1.2 [**MODELS AND PRINCIPLES**]: User/Machine Systems – *human factors, human information processing.*

General Terms
Management, Design, Human Factors.

Keywords
Knowledge Management, Ontology, Location-based Services, User Profiles, Query Processing, Databases, Peer-to-Peer

1. INTRODUCTION
Mobile data networks (GPRS and UTMS), positioning systems (GPS), mobile phones and personal digital assistants (PDAs) are just some of the new mobile technologies that participate into composing an ubiquitous computing environment, and make it possible to offer online services to people on the move, wherever they are. Ubiquitous computing denotes a large spectrum of facilities, which includes traditional services such as access to web pages and access to email. A special type of less traditional services, called Location-Based Services (LBS), are becoming increasingly popular as they aim at providing users with "right on the spot" information, i.e., information that belongs to a particular domain of interest to the user and can be of use while the user remains in the area where (s)he currently is. The most frequently quoted example of information provided by an LBS certainly is "find the nearest restaurant". Fortunately, LBS can be designed to provide more elaborated information, in particular by taking into account the user's profile and other contextual data. As stated in [2], there is an urgent need for sophisticated mobile marketing techniques based on detailed knowledge of customer profiles, history, needs, and preferences. This implies techniques able to integrate location information, customer needs, and vendor offerings.

Intuitively, *context* is any information that may influence the result given to users in response to their queries. For LBS, the user's location is definitely the primary component of the context. The environmental context is also important. It includes the time (if it is 11pm, and the restaurant closes at 10pm, then do not return it), the weather (if it is raining, and the restaurant is outdoor, then do not return it), etc. Contexts in mobile environments are dynamic and may change as soon as the user moves or issues a different query. The challenge for LBS is therefore to continuously monitor the semantic relevance of the data in a context.

The *user profile* is the key to personalized services and to obtain tailored information. Two users asking the same request at the same location and at the same time should have different answers according to their profile. The user profile contains personal data (if the user has a dog, and pets are not allowed in the restaurant, then do not return it), and user's preferences or interests (if the user likes vegetarian food, then return restaurants that offer this type of food). Depending on where, when, how, with whom, and why users are navigating in a physical space, their profile and needs will vary (e.g. 'at work' profile, 'at home' profile). Again, profiles in LBS are characterized by their dynamicity, contrasting with the fixed profiles mainly used in web services. User profiles in LBS raise the same semantic relevance issue as context data.

The third semantic player in LBS is the set of data sources the LBS can use to find the answer to a user's query. The set of available data sources changes as the user moves. We say these sources are characterized by a *data profile*. For a restaurant, for example, the data profile includes the location, the opening hours, the type of food, the price range, etc. Relevant data sources and data services are identified by the LBS using a service discovery mechanism based on the matching between the user's request, reformulated according to both the context and the user profile, and the data profile. To help in this matching, the LBS may seek relevant ontological knowledge.

In this paper we focus on semantic issues raised by generic information retrieval location-based services (e.g. get a list of restaurants). We do not discuss the additional issues raised by location-based services aiming at performing some task on behalf of the user (e.g. reserve a table at a restaurant).

The paper is organized as follows. Section 2 describes the main characteristics of location-based services. Sections 3 to 6 discuss the semantic components (contexts, user profiles, user history, and data profiles) that provide the knowledge necessary to perform intelligent location-based services. Section 7 analyzes how this knowledge interoperates in the semantic query evaluation process. Section 8 concludes the paper.

2. LOCATION-BASED SERVICES

Location-based services are defined in [9] as services that integrate a mobile device's location or position with other information so as to provide added value to a user. This loose definition, however, includes traditional database services that would provide location-dependent answers to user requests. We prefer to think of LBS as services mainly using the decentralized knowledge, that is geographically close to the user's location and consequently most likely to hold the information the moving user wants to retrieve, to decide about his/her behavior in the short term (within the next hours or days). Frequently quoted LBS applications include tourist assistance, traffic monitoring, and mobile office. Most proposals consider pull-style services (reacting to user's requests, e.g. find a restaurant), although the mobile environment offers interesting perspectives for the development of push-style services [9] (e.g. a discount offer is advertised to the moving user from a nearby restaurant).

Location is an essential characteristic of LBS that distinguishes them from classical web information retrieval. Thus, a necessary component in the architecture of LBS is the *location engine*, which supports spatio-temporal data processing. Its four basic functions are location capturing, geocoding, reverse geocoding, and geo-data storage. It may have extended functions such as routing, and proximity search. New algorithms for spatial data processing are a current research trend in mobile computing.

Another important research trend focuses on monitoring network connections (namely, dynamic network reconfiguration) and the physical components of the active environment, as the user moves on to next destinations. Communication and operation with a set of local services is the primary concern for LBS, as local services are expected to be more frequently used than remote services.

LBS also inherit a number of issues generic to mobile computing, such as handling authorizations (checking the device parameters, bundling the wireless network connection, validating the user's identification), protecting user's privacy (now including data about user's locations and trajectories), HCI techniques for interacting with users using devices with limited display capabilities, etc.

From a semantic viewpoint, the major characteristic of, and challenge for, LBS is the fact that they serve as mediator between a possibly unknown user and possibly a priori unknown data sources. Moreover, the mediation has to be run on-the-fly, i.e., it cannot be prepared in advance as the partners in the mediation are not necessarily known. We could say that LBS have to perform query evaluation in a hostile environment. To overcome the difficulty, contributions from most advanced techniques are welcome. They include: ontology assistance (to understand what it is all about), peer-to-peer information search (to increase chances of finding relevant information), incomplete information handling (to cope with missing data), and approximation techniques (to determine what could be a reasonable answer when a perfect matching is not possible).

3. CONTEXTS

Everything is context-dependent, in particular the meaning of the terms that appear in user's requests or in service and data descriptions. Yet traditional database technology mostly ignores context, providing data services based on a single, implicit context. Contexts are definitely an area for further research, in particular for the database community.

Contexts in LBS group any information that characterizes the situation of a person, place, or object, as well as the meaning of things at hand, and that can be used to provide more relevant services to the user. LBS contextual information typically includes the user's location, the time, the weather, the traffic conditions, etc. Systems that use contexts are called context-aware systems.

Context can be used as hard or soft criteria in the selection of relevant services [11]. Hard criteria discard the services that do not meet them, while soft criteria order the set of selected services. For example, the user's location can be used to select only the restaurants within a certain distance from the user (hard criterion), or it can be used to sort the selected restaurants according to their distance from the user (soft criterion).

The contextual information has many alternative representations, which make it difficult to interpret and use. Context providers and context consumers may have different understandings of the same contextual information. Definition of metadata standards is one way to solve the issue, but is unlikely to scale up to all data services that will exist in the future. Another way to go, nowadays very popular, is using ontologies tailored to provide a shared understanding of the concepts used to describe the context and the data services.

In a mobile environment, the contextual information may be continuously changing. For LBS, weather conditions, for example, may change and entail a different strategy in the search for information, or the user can change activity, from work to leisure, for example, entailing different information requirements. Consequently, contexts cannot be retrieved once, but have to be monitored in order to get the up-to-date contextual information.

The semantic challenge in LBS context management is to determine which contextual information is relevant for the current task. It does not make sense to manage very sophisticated contexts if only a few components are actually useful. Conversely, lower quality services are offered if the appropriate contextual information is missing. Determining what is relevant in a context calls for matching the context with the user profile on one hand and with the data profile on the other hand.

Context management has become a hot research issue. Significant contributions include the COMPASS (COntext-aware Mobile Personal ASSistant) mobile tourist application [11], which uses the context to provide relevant business services (restaurants, museums, shops, cinemas, etc.) to the user. The system architecture includes context services (location, time, weather, etc.) that provide contextual information. The context manager retrieves the context by contacting the context services. It notifies the application on changes in the context.

In the COSS (Context-aware, Ontology-based, Semantic Service discovery) system [3], service providers and context providers use domain-specific ontologies to which they commit. These ontologies are: the service type ontology (containing concepts such as shop, restaurant), the product ontology (containing concepts such as DVD, vegetarian food), the payment ontology (containing concepts such as cash, credit card), and the context ontology (containing concepts such as location, time).

The TIP (Tourism Information Provider) system [6] allows end users to get relevant information based on their current location and the current time (besides to their profile and their history).

4. USER PROFILES

User profiles are of major importance to provide intelligent and personalized LBS. They have already attracted much attention, in particular from research in artificial intelligence [1] [4], addressing issues such as acquisition of the user profile, learning from user preferences, use of profiles for better information filtering and decision-making. Different profile-aware filtering algorithms have been implemented in applications such as recommendation systems and web browsing [10] [12]. However, most of the proposals are tailored for some specific computing environment or pre-defined application, such as [8]. In location-based services, due to the inherent mobility framework, the computing environment is continuously changing, as well as the type and function of available data sources. Moreover, the user profile itself often evolves because of diverse factors, such as change in user's location, the social environment, and the user's activity. Hence, LBS rather have to focus on more generic and dynamic techniques, any time capable of adjusting their service to the current environment and user profile.

For user profile acquisition, it is important to avoid lengthy questionnaires. Experiments have shown that users faced to more than 4 prompts for information from a query system tend to give up using the system [5]. LBS can limit the number of interactions with the users by building upon the matching of the user profile with the knowledge they maintain about the data sources. For example, it is obviously useless to ask users which type of cuisine they prefer if the description of restaurants in the data sources does not include the type of cuisine they offer. Thus, selectivity of attributes and frequency of use are relevant criteria in choosing what to look for a user profile.

The same user can have many profiles. At a given point in time, one of these profiles is the one corresponding to the current user's activity and request. Tables 1 and 2 show two possible profiles for the same person, the former relates to his role as a tourist, the latter relates to his role as a computer science professional. User profiles can be organized into two parts, a static part and a dynamic part. The static part stores the information that is seen as inherently related to the user, irrespectively of what are his/her current activity and interest, e.g. age and nationality. The dynamic part contains information closely related to the user's possible activities and requests.

Table 1. An example profile (a) for Stefano

Activity: Tourist
Profession: {professor, employee, academic}
Age: senior
Gender: male
Nationality: French
Income: good
Interest: art, culture, hiking, cinema
Languages: French, English, Italian
Food: good, very good
Cuisine: Japanese, Thai, Arabic, Argentinean
Credit cards: VISA, Master Card

Table 2. Another example profile (b) for Stefano

Activity: Professional
Profession: {computer scientist, professor}
Age: senior
Interest: databases, ontologies, semantic web
Languages: French, English, Italian
Memberships: ACM, IEEE, IFIP, SI

In the example profiles for Stefano, the properties 'age', 'gender', 'nationality', 'income' and 'languages' would belong to the static and general part, while 'activity', 'profession', 'interest', 'food', 'cuisine', and 'credit cards' would belong to the dynamic part. Especially, when the user wants to find a restaurant, the 'cuisine' can be referred for selection and the credit card can be one prerequisite for reservation.

It can be noted that even though some attributes are repeated in both profiles, such as 'age' and 'languages', their effect is different in both activities. In addition, some attributes have different values, such as 'profession' and 'interest'. Hence, we propose an activity-based profile modeling approach. In this model, a user has a complete profile, and one or many activity profiles. Each attribute in the *complete profile* has an annotation to indicate what activity it is associated with, when and where it should be used, what other attributes are highly concerned with in the same activity etc. When the user sends a query, one *activity profile* is dynamically created by retrieving the corresponding attributes from the complete profile (using the annotations) as shown in Figure 1.

User profiles are composed of very personalized information on user preferences and activities. The understanding and definitions of properties in user profiles depends on the culture, language, education, etc. For users on the move, it is annoying to adapt their profile according to local languages or habits. It is highly recommended that the LBS could do it for them. Using ontologies could be one solution since they provide general and shared common definitions. In [7] ontologies are consulted to help users to create their own concept hierarchies (including the user profile) for website browsing.

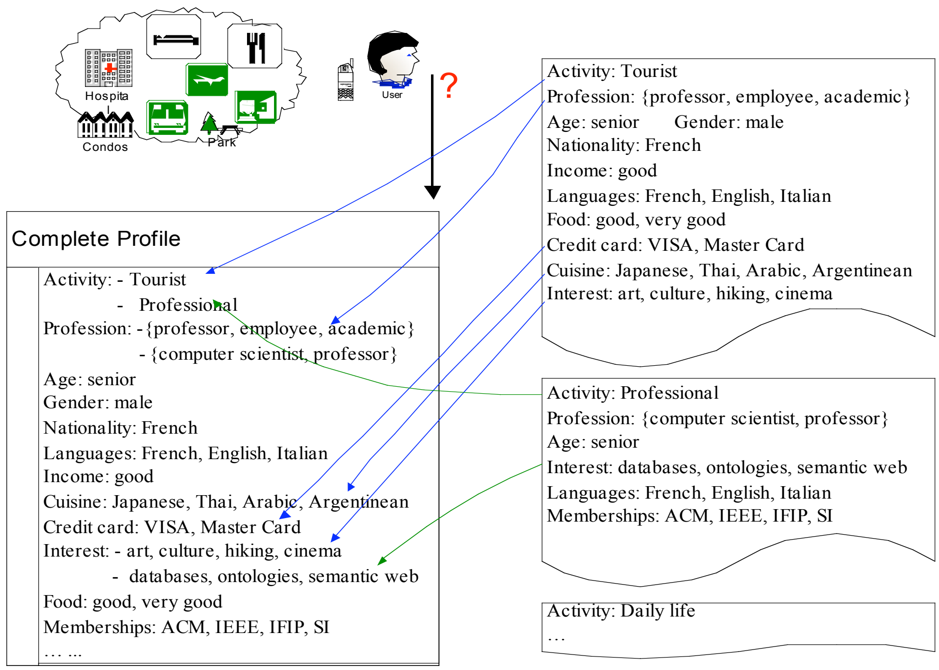

Figure 1. Creation of the user activity profiles in LBS.

5. USER HISTORY

Besides the user profile, the user history, i.e. the trace of services already used by the user, can be useful. Users do not wish to get the same service many times unless they explicitly ask for it (e.g. a tourist would not like to go to the restaurant where s(he) ate yesterday). Consequently LBS should store the history or "accumulated knowledge" of the user. A critical issue is the visibility of the history.

In the TIP system [6] the user history is stored on the mobile device for privacy reasons, unlike the user profile, which is stored on the TIP server. Each time users submit a query, their history is shipped to the system.

In [11] an experiment is reported where users look for a restaurant, and the COMPASS system provides them with a list of nearby restaurants that meet their interests and that they did not visit recently. The feedback of these users was then collected. Users believe that using their interests is useful to provide relevant information, but do not like the "last time visited" criterion to be used. They want to be able to decide for themselves which factors are important when selecting a point of interest (e.g. retrieve items based on my interests, location, and prices, but do not include last time visited).

6. DATA PROFILES

Data profiles describe data services. The same way a schema describes a database, a data profile gives information about the data provided by a service. Table 3 gives an example data profile for a restaurant. What characterizes a data profile in LBS is that it includes spatial (e.g. restaurant location, outdoor) and temporal attributes (e.g. opening hours).

Table 3. Data profile for a restaurant

Location: Geneva
Cuisine: {international, Japanese}
Price range: 25-45 CHF
Opening hours: 11:30am-2pm, 6:30pm-10pm
Situation: {indoor, outdoor}
Smoking: {smoking, non-smoking}
Parking: no
Credit cards: VISA, Master Card

In LBS, supporting services are not known a priori, and each of them may have its own description. It is a critical issue for LBS to be able to understand what data is available and where. Traditional integration techniques do not solve the problem as they do not comply with the need for flexibility and fast response. Using ontologies would improve common understanding of services, and also enable semantic reasoning.

In COSS [3], each data service, in a tourism application, is modeled by its type (e.g. restaurants, museums, shops, etc.), its outputs (information delivered by the service, e.g. vegetarian food), its inputs (information needed by the service, e.g. user's location), and attributes (e.g. service location, opening hours). Ontologies are used for the service type, the product, and the attributes.

7. SEMANTIC MATCHING

Most of the existing service discovery mechanisms are keyword-based, i.e. they retrieve services descriptions that contain keywords from the user's query. The retrieved results have low recall (some relevant items are missing) and low precision (some retrieved items are irrelevant) [3]. The reason for the first problem is that keywords might be semantically similar but syntactically different, e.g. 'beautiful' and 'handsome' (synonyms). The reason for the second problem is that keywords might be syntactically the same but semantically different, e.g. 'chair' (to sit) and 'chair' (academic chair) (homonyms). One possible solution is to use ontology-based retrieval.

In location-based services, the matching process should involve not only the data services, but also the context and user profile (and possibly the user history), in order to provide the user with relevant information (see Figure 2).

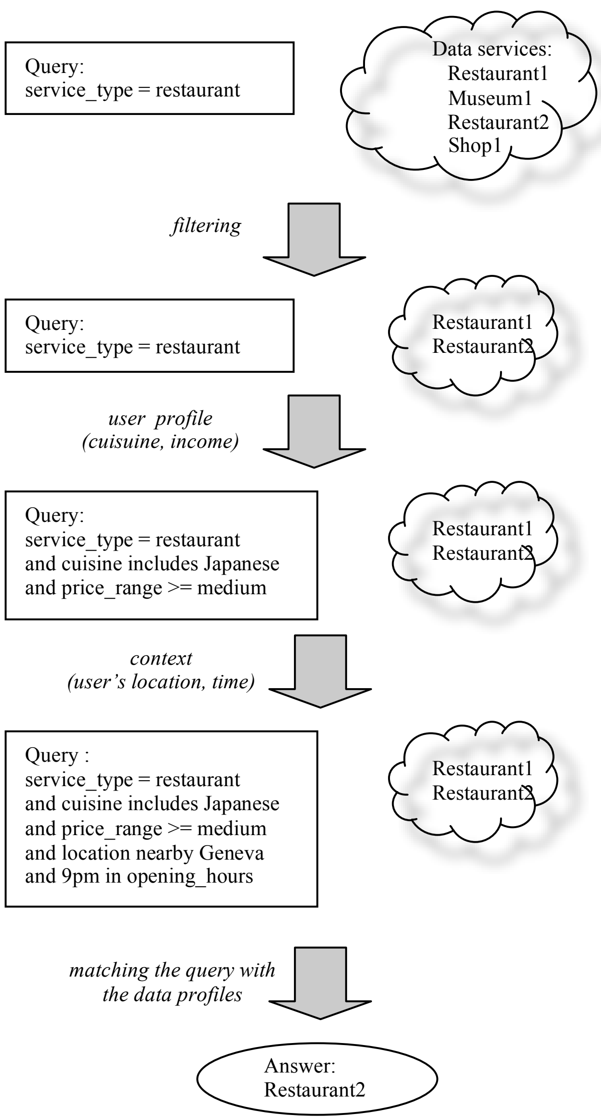

Figure 2: Applying the matching process.

Given a query and currently available data services, a first step would be to filter out those services that do not match the service type (e.g., if the user looks for restaurants, museums should not be considered). Note that if this step applies exact matching between keywords, it may miss some data services (e.g., it may miss cafeteria). Using dictionaries and ontologies can help in the matching. The second step would rewrite the query by making it narrower or more precise, using the user profile [13] [10][10]. The question is which properties of the user profile should be used in this step. One possible solution is to take properties in the user profile that "correspond" to properties in the data profile (e.g. 'income' in the user profile could be linked to 'price range' in the restaurant profile). These correspondences are not easy to find. The third step would rewrite the query by adding conditions to take into account the current environmental context (e.g. the current time is within the opening hours of the restaurant). Again, it is not easy to decide which contexts should be considered. One possible solution is to look at spatio-temporal attributes in the data profile. Finally the query is matched with the data profile of the remaining data services.

In COMPASS [11], the Matchmaker component first uses the service registry to discover the data services that match the request. Services are described in OWL (Ontology Web Language). Once services are discovered, the Matchmaker filters out those that do not match the hard criteria of the context. To do so, the Matchmaker uses the context ontology and domain-specific rules from the application. Then, the obtained list of points of interest (POI) is sent to the recommendation engine, which assigns a score to each POI according to the user's profile and soft criteria of the context.

In COSS [3] the starting point of the matching process (using ontologies) is a set of available data services S. The first step filters out the services that are not of the desired type specified in the user's request. The resulting set is S'. The second step filters out the services that do not have the desired service output. The resulting set is S''. Then this set is queried for the inputs. If the inputs required by a service are provided by the user or the context (e.g. user's location), the match is classified as perfect, otherwise as imperfect. The resulting sets are S_{exact}, and $S_{approximate}$. The final step orders these two sets according to the attributes specified in the user's request or user's preferences in the user profile (e.g., nearby, parking, non-smoking) using concept lattices.

TIP [6] combines location-based services with an event notification system. When a tourist is walking around in a city, and stops at a sight (e.g. a castle), the system provides him/her with general information (e.g. general information about the castle). It also considers the user profile to deliver more specific information (e.g. architectural information about the castle). It looks for relevant information in the scheduled events and spatial databases and notifies the user (e.g. there is a concert tonight in the castle). It compares the situation with the relevant part of the user's history (e.g. you saw a similar castle in London last week). The architecture of TIP is centralized, i.e. the spatial database (including sights and maps), the profile database, and the scheduled events database are on the TIP server. One schema is used for sights, users, and events.

8. CONCLUSION

In this paper, we showed how using semantics can help in finding information that is relevant to the mobile user, and thus improve the quality of location-based services. The context, user profile, user history, and data profile are dynamic semantic components that should be used in the matching process in order to give a tailored and useful information to the user. This is a current research issue.

In LBS, data sources are not known a priori. Each of them may have its own way to describe the data services. Using a generic data profile would be helpful in the matching process.

Nowadays, a user can use his mobile phone in Paris or London, without having to know that s(he) is using the telecommunication services of this or that local operator. Location-based services should offer the same flexibility. A tourist, whether s(he) is in Paris or London, should have tourist assistance by local LBS providers, and get relevant information according to his/her profile. This would imply that the user profile is on the mobile device, and that LBS should use a generic user profile.

9. REFERENCES

[1] Basu, C., Hirsh, H., and Cohen, W. Recommendation as Classification: Using Social and Content-Based Information in Recommendation. In *Proceedings of the National Conference on Artificial Intelligence (AAAI '98)* (Madison, Wisconsin, USA, July, 1998), AAAI Press, 1998, 714-720.

[2] Bharat, R., and Minakakis, L. Evolution of Mobile Location-based Services. *Communications of the ACM, 46*, 12 (Dec. 2003), 61-65.

[3] Broens, T., Pokraev, S., van Sinderen, M., Koolwaaij, J., and Costa, P. D. Context-Aware, Ontology-Based Service Discovery. In *Proceedings of the European Symposium on Ambient Intelligence (EUSAI'04)* (EindHoven, The Netherlands, November, 2004). Springer, 2004, 72-83.

[4] Chen, L., and Sycara, K. WebMate: A Personal Agent for Browsing and Searching. In *Proceedings of the International Conference on Autonomous Agents (Agents'98)* (Minneapolis/St. Paul, USA, May, 1998), ACM Press, 1998, 132-139.

[5] Wu, H., Li, M., Zhang, H., and Ma, W. Improving Image Retrieval with Semantic Classification Using Relevance Feedback. In *Proceedings of the IFIP TC2/WG2.6 Working Conference on Visual Database Systems (VDB'02)* (Brisbane, Australia, May, 2002), Kluwer, 2002, 327-339.

[6] Hinze, A., and Voisard, A. Location- and Time-Based Information Delivery in Tourism. In *Proceedings of the International Symposium on Advances in Spatial and Temporal Databases (SSTD'03)* (Santorini Island, Greece, July, 2003). Springer, 2003, 489-507.

[7] Pretschner, A., and Gauch, S. Ontology Based Personalized Search. In *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ICTAI 1999)* (Chicago, Illinois, USA, November, 1999), IEEE Computer Society, 1999, 391-398.

[8] Smyth, B., and Cotter, P. A Personalized Television Listings Services. *Communications of the ACM, 43*, 8 (Aug. 2000), 107-111.

[9] Spiekermann, S. General Aspects of Location-Based Services. In *Location-Based Services*. Schiller, J., and Voisard, A. (eds). Morgan Kaufman, 2004.

[10] Sugiyama, K., Hatano, K., Yoshikawa, M., and Uermura, S. User-Oriented Adaptive Web Information Retrieval Based On Implicit Observation. In *Proceedings of the Asia Pacific Web Conference (APWeb'04)* (Hangzhou, China, April, 2004), Springer, 2004, 636-643.

[11] van Setten, M., Pokraev, S., and Koolwaaij, J. Context-Aware Recommendations in the Mobile Tourist Application COMPASS, In *Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'04)* (EindHoven, The Netherlands, August, 2004). Springer, 2004, 235-244.

[12] Widyantora, D.H., Ioerger, T.R., Yen, J. Learning User Interest Dynamics with a Three-Descriptor Representation. *Journal of the American Society of Information Science and Technology (JASIST), 52*, 3 (Feb. 2001), 212-225.

[13] Yu, S., Spaccapietra, S., Cullot, N., and Aufaure, M.-A. User Profiles in Location-based Services: Make Humans More Nomadic and Personalized. In *Proceedings of the IASTED International Conference on Databases and Applications (DBA'04)* (Innsbruck, Austria, February, 2004). ACTA Press, 2004, 25-30.