Ontology Supported Personalized Search for Mobile Devices

Daniel Aréchiga, Jesús Vegas and Pablo de la Fuente Redondo

Departamento de Informática, Universidad de Valladolid, 47011 Valladolid, Spain
vdaniel@cusur.udg.mx, {jvegas, pfuente}@infor.uva.es

Abstract. Web search is a frequent activity on Internet connected devices, yet still a nuisance when user is using a mobile device, taking into consideration their limited keypad and screen, and that search results could be mostly irrelevant for the user needs given its mobile context. When in move, the user needs an efficient way to introduce query terms and receive more precise information. We propose a Context Model to represent the context-aware information, which will be used to offer a better word recommendation and an autocompletion system to improve the user experience. This Context Model uses ontologies and a thesaurus to support a word recommender system that enhances the typing task and takes into account the user's context and device characteristics.

Keywords: mobile device, web search, ontology, recommendation system.

1 Introduction

Web pages are virtually the most useful and known Internet service, essentially because it brings the users an enormous volume of any kind of information. An usual way to find that information within this sea of possibilities is through the use of web search services like Google[1], Yahoo[2] and Altavista[3].

On the other hand, modern mobile devices have evolved in their screen size, processing power and connectivity possibilities, transforming them into a powerful Internet access device, consequently bringing the web search capacities to the palm of our hands.

There were near to 4 billion cellular subscribers around the world in early 2009[4] this implies a big potential market for Internet services for mobile users. Recent studies [1] reveal that web search will become a meaningful tool for those mobile users in the next years. Some Nokia's strategic research is related to mobile applications and services, with convergence between cell phones and computers in mind as a main issue for the next 5 to 10 years.

[1] Google, http://www.google.com
[2] Yahoo, http://www.yahoo.com
[3] Altavista, http://www.altavista.com
[4] 3G Americas, http://www.3gamericas.org

Mobile web search represents a different paradigm in front of desktop search. When in a mobile environment, users need a time and cost efficient web search as well as obtaining significant results taking into account the context that surrounds him, all in a limited-interface device with a higher bandwidth cost.

One important challenge to improve the web search for a proper mobile user experience is related to the mobile device interface and input method, because of most of actual devices have a 9 or 12 key keypad, and only a minority of them include a full physical or virtual keypad.

An intelligent interface that helps users to write fewer letters with a recommender or autocompletion system is a desirable component on limited keyboard devices [2]. If these recommendations also consider context information, the recommended words will be more precise and effective. For example, a recommender system will not propose at first place the term "snowboard" when user is at the beach or in summer time.

To solve this issue, we propose to use a Context Model which includes context information by means of an ontologies and a thesaurus use. This model creates and manages the user and device profiles which consist of environmental data like time and place, user preferences, device capabilities description, etc.

We incorporate this Context Model to a working prototype which also includes a user interface designed specifically for mobile devices which considers their limitations of screen and keypad. The use of ontologies and a thesaurus help us to offer better results according to our main objective in our recommendation system. The ontologies let us model the spatial, temporal, user dependent and device capabilities concepts easily.

The rest of the paper is organized as follows. Section 2 describes the background and related work regarding web search on mobile devices, mobile context-aware ontologies, personalization and recommender systems. The detailed description of MYMOSE Project[5] can be found in section 3. Section 4 includes our conclusions and the guidelines for future work.

2 Related work

There are many research and developments related to context-aware sensitive web search, studies from various points of view but the vast majority of the efforts have been based on desktop environments. Therefore, the environment of mobile devices is less explored than its desktop counterpart. However, there is some previous research related to web search enhancement and personalization on mobile devices.

Major search engines as Google and Yahoo offer a mobile version that, in some cases, includes some autocompletion and word recommendations, most of them based on their gathered knowledge of general previous searches.

There are some proposals which include an ontological user profile [3] and even define a context model to determine user's interests [4] or re-rank search results

[5] MyMoSe Project, http://mymobilesearch.morfeo-project.org/mymobilesearch/lng/es

according to the profile created while the user is surfing the web [5] [6], but essentially they consider desktop users only.

Focusing on the mobile environment, there is previous research related to web search. Some proposals already uses ontologies to represent user profile [7], others include some context elements [8] but usually limited to the geographic location. Some others approaches use ontologies to create clusters of results [9]. There are other works which deal with word recommendation and autocompletion [2] on mobile devices, but they rarely include the ontology knowledge.

Also, there is a newer propose for a standardized mobile ontology [10], but it is focused on mobile services instead the web search, so it does not include elements relevant to obtain better search results like the context-aware approach.

This previous efforts attack the problem from different points of view, but they do not include a complete delivery context definition; typically it is used to model the user profile, create a limited context definition or it is used to re-rank results, but it does not include all of this information in one solution. Very few take into account the mobile devices characteristics and the entire environment definition including kind of place, weather or date and time.

3 Context-aware Framework

MyMoSe Project was developed by a collaboration between the GRINBD research group of the University of Valladolid and Telefónica I+D. The main idea of this venture was to develop tools to provide an enhanced web search which bring to mobile users a new experience.

There are two main parts in our approach; the first one is related to the page analysis: crawling, geo-referencing, indexing and ranking. The second component is associated to the mobile environment, offering to the user a web search interface which provides easy to use features like thesaurus-based word recommender, basic context manage and finally a distance and device clustered results (Figure 1). It's on the second part where the ontologies are a fundamental support and it will be the focus of our work.

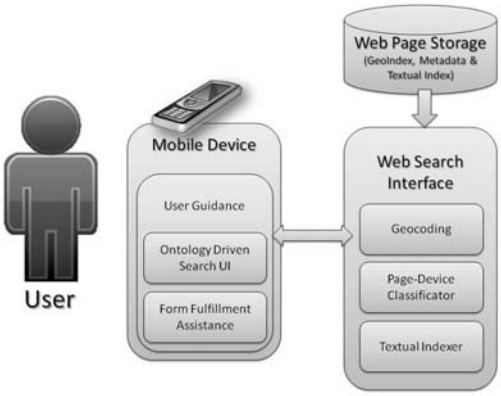

Fig. 1. Search and interface System

We propose a system based in a semantic context-aware framework, which helps the user to build personalized search queries by means of an autocompletion mechanism. Typically, a mobile user will mainly need information related to the entertainment and leisure when in outdoors situation, so we have focused this research effort to the web search of associated terms.

It's on this second stage when the delivery context is defined and where the user can find helpful elements to construct a better search query trying to obtain better results. Moreover, these results will be classified according to different characteristics related to relevance, user needs, distance and page device readiness capability, all within a basic, easy to use and device specially adapted interface.

According to the World Wide Web Consortium[6], the delivery context is defined as *a set of attributes that characterizes the capabilities of the access mechanism, the preferences of the user and other aspects of the context into which a web page is to be delivered*.

To model all the attributes required by our system, we have used ontologies and a thesaurus to create a formal representation of the Context Model. This can be layered into three levels which differences the attributes or properties:

- Directly Fetched Properties. These are properties that can be automatically gathered from context information sources. For example, the location coordinates obtained from a GPS device, the date, the time, etc.
- Derived Properties. These are "implicit" properties that can be inferred or calculated from fetched properties. They constitute an abstraction layer on top of the directly fetched properties layer. For example, the name of the country and region, taking into account the location coordinates provided by a GPS, or the season of the year depending of the date, whether it is day or night from the date and time, etc.
- Application-Specific Properties. Applications might define or redefine additional properties on top of the basic layers by using their own rules, as long as they ensure formal consistency.

Our Context Model is completed with the definition of classes, properties and their relationships; allowing to define the fundamental information required to obtain better web search results in a mobile environment. As a result we define as basic elements of context-aware web searches: the contextual information about the user and his preferences, the device and its capacities; and the environment information like date, time, weather and place (Figure 2).

This is represented through semantic elements, as described below:

- *User Profile*. Contains all the implicit and explicit properties related to the user and his circumstances.
- *Device and Browser*. This is a set of properties that describe the characteristics of the user's device and web browser.
- *Geospatial Context*. Model the user's geographical situation in order to provide better matched results according to the user's location.
- *Environmental Conditions*. These are properties that model the surrounding environment.

[6] W3C's Delivery Context Overview for Device Independence, http://www.w3.org/TR/di-dco

- *Date and Time*. Time and date are fundamental properties when dealing with context-awareness.

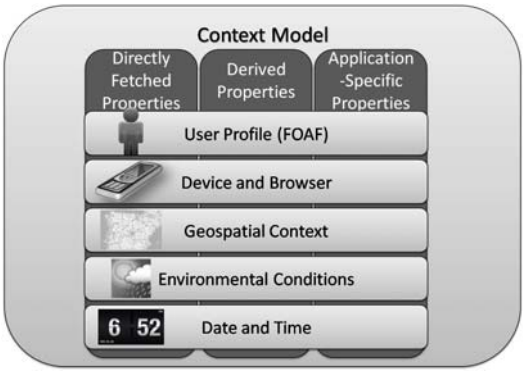

Fig. 2. Context Model

User Profile. This contains basic information about the user in a FOAF[7] extension model and based on existing semantic definitions of a *person* in sources like WordNet ontologies.

A significant usage for the User Profile is the possibility to store basic user interests, such as like and dislike features, both implicit and explicit. The user interface provides some options that are stored in his profile that could be selected by the user.

Device and Browser. Device and browser qualification regarding with its display capabilities is solved by separating the different delivery contexts possible in a few classes. Pages are classified based on their structural contents to identify their displaying requirements. These combined elements provide a match between the best adapted content from the results and the device qualification.

The W3C's Mobile Web Initiative Working Group[8] has defined the best practices that should be followed by a web page author to obtain a minimal functional user experience, this is represented with their *mobileOK Basic* scheme; we take this scheme as a baseline and use it as a Default Delivery Context, then we have defined classes above and below for classification purpose.

Previous research and experiments [11] have advised us to divide these classes in the following six different sets of delivery context characteristics:

- Legacy: Legacy voice-intended devices with a WML1 microbrowser and WBMP image support.
- Under DDC: Old feature phones with very basic, entry-level XHTML-MP 1.0 and W-CSS support and no (or very limited) colored image support.
- DDC: W3C's Default Delivery Context, the baseline class. It represents typical feature phones with fair XHTML-MP 1.0, W-CSS and colored image support, no table support and no scripting capabilities.

[7] FOAF Vocabulary Specification, http://xmlns.com/foaf/spec/
[8] W3C's Mobile Web Initiative Working Group, http://www.w3.org/Mobile/

- Over DDC: This class corresponds to smart phones or feature phones that include a mobile web browser with excellent XHTML-MP (1.0 and over), W-CSS, tables and image (in several formats) support.
- Advanced: PDAs or smartphones executing nearly complete, advanced mobile browsers or environments on which a content transformation proxy is available.
- Desktop: Browsers running on non-mobile devices. In our consideration, this does not represent a mobile device environment.

Geospatial Context. Location is defined not only in terms of geographic coordinates but also in terms of the kind of place where the user is located, this allow us to identify if the user is on an airport, beach, countryside, city place, etc. a useful information to improve web search.

The user's location is modeled in order to provide better web search results by taking into account the place where the user is located. Our system classifies these results according to the distance (i.e. walk, bicycle, car, etc.).

Environmental Conditions. This represents the surrounding environment especially weather conditions, which can be obtained from other services over the Internet. This could be relevant on certain conditions, for example, if the user is searching for free time activities and the weather conditions are not recommended for outdoors, therefore the system should propose primary indoor activities.

Date and Time. As simple as this kind of information is, it is also a meaningful component when dealing with context-awareness. This allows to our system to manage temporal concepts like seasons and timetables. For example, it could be less important a "snowboard" concept when in summer or concepts like cinema or theater at seven o'clock in the morning.

3.1 Context Model ontologies and thesaurus

Our Context Model is supported by ontologies and a thesaurus to represent the knowledge required for the system. Ontologies enable us to define abstractions of elements that we consider relevant to the definition of context. This definitions will be used for enhance the user search experience by means of the Context Model we have defined and the recommendation system.

OWL[9] is the language we have decided to use for the ontologies, essentially because of its extended use as an industry standard, besides the wide number of development tools. The SWRL[10] language was used as a complement for the ontologies rule definition. We also use for the thesaurus SKOS[11] as a RDF[12] vocabulary because it is a standard way to represent lattices of concepts.

[9] Web Ontology Language W3 Consortium, http://www.w3.org/TR/owl-features/
[10] SWRL W3 Consortium, http://www.w3.org/Submission/SWRL/
[11] SKOS W3 Consortium, http://www.w3.org/2004/02/skos/
[12] RDF W3 Consortium, http://www.w3.org/RDF/

For our Context Model, we have created four ontology vocabularies, which interact through the rules definitions with the thesaurus and therefore with the recommender system. We also import classes and properties from other public ontologies to complete classes' definition required.

The ontologies definition is divided in two parts, one for the domain classes definition, which is defined by the *domain* ontology and represents concepts closely related with concepts in the thesaurus. The other part is the delivery context definition, which is defined by the *delivery context* ontology, *user profile* and *localization* sub-ontologies.

The semantic elements are included in the ontologies as follows: **User Profile** is included in the *user profile* ontology; **Geospatial Context** is included in the *localization* ontology; **Device and Browser**, **Environmental Conditions** and **Date and Time** are included in the *delivery context* ontology.

Both parts of the ontologies are linked by the rules definition. These rules define the relation between classes, properties and individuals from the ontologies.

Above of this, the thesaurus is related to the ontologies by the rules definitions, both to validate the relationship between thesaurus concepts and ontologies properties, and to identify if certain concept is suitable for actual context. This way, the recommender system will show at first place the context-suitable concepts to the user, if any (Figure 3).

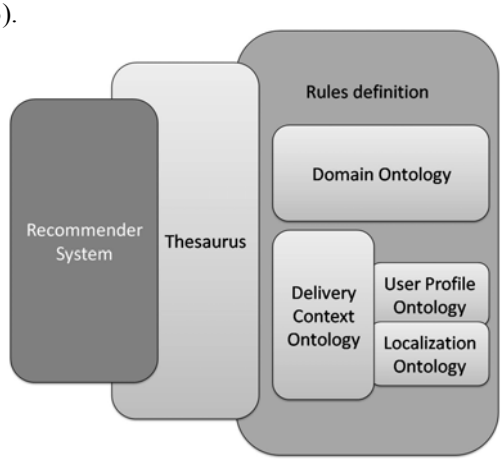

Fig. 3. Semantic Architecture for recommender system

Thesaurus. As previously described, we use the SKOS vocabulary for the RDF definitions. Our thesaurus contains about 100 testing-purpose concepts, and near to 200 relations between concepts. These concepts are introduced in both English and Spanish, and the framework is designed to support the addition of other languages easily.

Our concept definitions allow to clarify ambiguities, like "León" the Spanish city or the animal. There are also different kinds of relations between concepts, like narrowed for a child concept, broader for a father concept and related for equivalent concepts (Figure 4).

```
<rdf:Description rdf:about="http://www.morfeo-project.org/mymw/thes/restaurant">
 <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:prefLabel xml:lang="en">Restaurant</skos:prefLabel
  <skos:prefLabel xml:lang="es">Restaurante</skos:prefLabel>
  <skos:definition   xml:lang="en">
          http://www.tid.es/2007/mymobilesearch/domainOntology.owl#Restaurant
  </skos:definition>
  <skos:related rdf:resource="http://www.morfeo-project.org/mymw/thes/cafeteria"/>
  <skos:broader rdf:resource="http://www.morfeo-project.org/mymw/thes/travel"/>
  <skos:related rdf:resource="http://www.morfeo-project.org/mymw/thes/food"/>
  <skos:broader rdf:resource="http://www.morfeo-project.org/mymw/thes/place"/>
</rdf:Description>
```

Fig. 4. Example of a concept definition, with ontology concept link, broader and related elements

User Profile Ontology. As mentioned previously, the User Profile ontology is based on the FOAF extension model. We add the *Interest* class with more specific properties to allow the management of user preferences, like the favorite pages often consulted by the user (Figure 5).

```
<owl:Class rdf:ID="Interest"/>
 <owl:DatatypeProperty rdf:ID="url">
   <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
   <rdfs:domain rdf:resource="#Interest"/>
 </owl:DatatypeProperty>
</owl:ObjectProperty
```

Fig. 5. Example of the Interest class definition inside the user profile ontology

Location Ontology. We have defined this ontology to manage geospatial information required to establish the delivery context. An important information that could be addressed on this ontology is the description of the spatial situation of the user (Figure 6). For instance, not only the street, city and country, but also if he is inside an airport or train station.

By gathering this information, we can further exploit it to affect the recommendations that will be shown to the user.

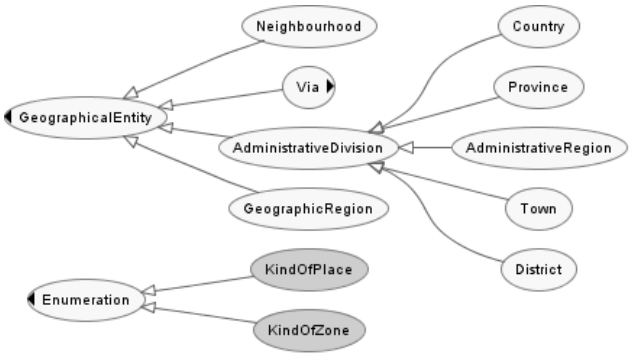

Fig. 6. Some example classes diagram of the *localization* ontology

Delivery Context Ontology. This ontology defines the complete user's environment, and acts as entry point to the whole delivery context. It contains the two sub-ontologies previously described: User Profile and Localization ontologies. As the delivery context ontology, it is also designed to manage some other data which represents the device and user's situation (Figure 7).

There are classes and properties which describe the different device and browser complexity levels: Legacy, Under DDC, DDC, Over DDC, Advanced, Desktop.

Besides time, date and season information, weather conditions are also considered in our ontology. This information could be valuable for the recommendation system, for example, *museum* will be a better recommendation on rainy days than outdoors activities, in the same manner, the system should not recommend going to the cinema at 8 o'clock in the morning.

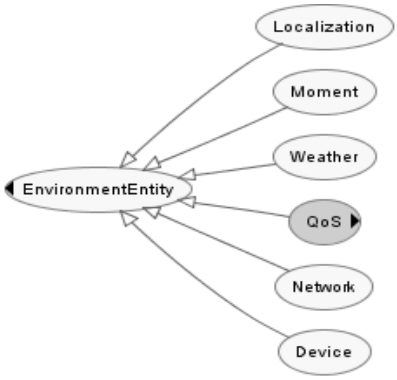

Fig. 7. Some example classes diagram of the *delivery context* ontology

Domain Ontology. This ontology interacts closely with the thesaurus by means of the semantic rules defined. This ontology represents concept's abstractions and relations (Figure 8). These concepts are related with thesaurus concepts, which in combination with rules definition and delivery context determinate the valid text to show on the recommendation interface.

Moreover, the ontology determinates if a concept contains defined properties which are used to enhance the information featured in the screen device by the recommender system. For example, if the user when typing selects the word "restaurant", the system will show options with the kind of restaurant, price ranges and restaurant's category obtained from the ontology.

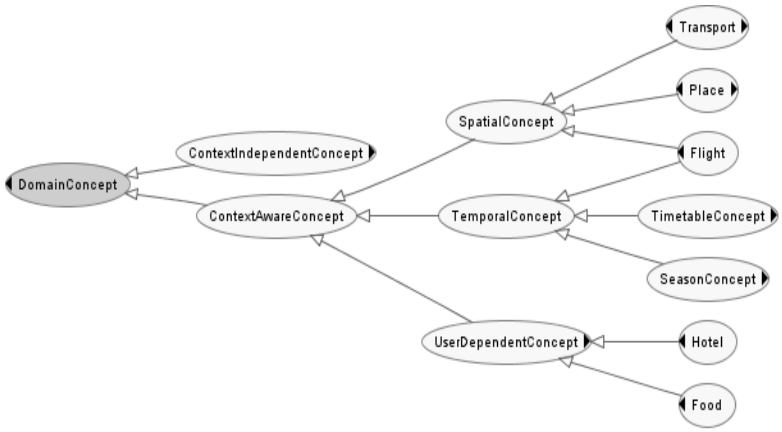

Fig. 8. Main classes' structure diagram of the domain ontology

Rules definition. A fundamental element to obtain the desired behavior in our system is the implemented rules. OWL Lite and OWL DL has expressiveness limitations and SWRL was proposed to enhance the rules on this semantic environment. For that reason, we have decided to use them in our system as the natural extension to OWL.

Rules are defined to integrate the recommender system, the thesaurus and the proposed ontologies. These rules allow us to specify whether a certain concept is suitable or not in the current context, after determining if a relation is valid (if concept is suitable in context) then the recommender system will show it as a recommendation and once selected will show the related properties from the ontology. This relation can be described as following:

$$concept(?c) \wedge context(?t) \wedge \{conditions(?t)\} \rightarrow isSuitableInContext(?c; ?t) .$$

Where ?c is a concept that will be evaluated against the current context and conditions specify which conditions should the context satisfy for that concept to be suitable. The conditions term is a logical equation which represents the conditions that the delivery context must satisfy for that specific concept to be suitable in the context.

3.2 Enhanced User Interface

To complete the semantic information described, in this first approach, we have created an enhanced tab-based user interface. This allows the user to model his search query and modify some parameters. It is divided in three tabs: Search, Results and Context tabs.

The Search tab contains an input box where user can type the search terms. While typing and based upon the words entered, the recommender system will show suggestions based on the thesaurus contents. Once one of the shown concepts is selected, the system provides some more options according to the ontology, to improve the quality of the gathered information.

Results tab shows all the search results classified upon distance and device readiness, leaving the user to choose the desired distance-readiness combination.

Finally, the Context tab lets user to know the context attributes that system detects, let him to modify as required (Figure 9).

Fig. 9. Ontology and Thesaurus supported user Interface screenshots

4 Conclusions and Future Work

We have realized that the increasing number of mobile users requires a new paradigm for web navigation and therefore web search. Mobile user's needs are way different from those in a desktop scenario.

Our proposed solution to this issue includes a semantic definition specifically designed for mobile devices; we have created an ontology and thesaurus-based Context Model to provide a foundation for personalization. Our Context Model in combination with the interface provides an autocompletion recommender and results classification, sensitive to the current context.

With the user experience in mind, we have tested our system with some different kind of mobile devices obtaining good results. This will be a start point for our future work that will be focused to test and improve the Context Model and interface, taking special attention in the context definition and better user profiling.

5 Additional authors

Additional authors are: César Llamas (GRINBD, email: cllamas@infor.uva.es), José Manuel Cantera (Telefónica I+D, email: jmcf@tid.es), Mario Arias (GRINBD, email: marioarias2@gmail.com), Álvaro Zubizarreta (GRINBD, zubisoft@gmail.com),

Jorge Cabrero (GRINBD, email: reybamba@gmail.com) and Guido García (OESIA, email: ggarciab@oesia.com).

Acknowledgments. The present research it's a part of the My Mobile Search project [7] (MyMoSe) developed by the Universidad de Valladolid's investigation group GRINBD with collaboration of Telefónica I+D.

This work has been partially supported by TIN2006-15071-C03-02 project, Ministry of Education and Science, TSI-020301-2008-25 project, Ministry of Industry, Tourism and Commerce, and by the Government of the region of Castilla y León through the Agency for Economic Development, (ADE), Spain.

References

1. Church, K., Smyth, B., Cotter, P., and Bradley, K.: Mobile information access: A study of emerging search behavior on the mobile Internet. In: ACM Transactions on the Web, Vol. 1, No. 1, Article 4 (2007)
2. Kamvar, M. and Baluja, S.: Query suggestions for mobile search: understanding usage patterns. In: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing System, pp. 1013--1016 (2008)
3. Gauch, S., Chaffee, J., and Pretschner, A.: Ontology-based personalized search and browsing. Web Intelligence and Agent Systems Vol. 1, Issue 3-4 pp. 219-234 (2003)
4. Sieg, A., Mobasher, B., Burke, R.: Representing Context in Web Search with Ontological User Profiles. In: Proceedings of the Sixth International and Interdisciplinary Conference on Modeling and Using Context, (2007)
5. Pretschner, A., Gauch, S.: Ontology-based personalized search. In: Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence pp. 391--398 (1999)
6. Sieg, A., Mobasher, B., Burke, R.: Ontological User Profiles for Representing Context in Web Search. In: Proceedings of the 2007 IEEE/WIC/ACM international Conferences on Web intelligence and intelligent Agent Technology, pp. 91--94 IEEE Computer Society (2007)
7. Weiβenberg, N., Voisard, A., and Gartmann, R.: Using ontologies in personalized mobile applications. In: Proceedings of the 12th Annual ACM international Workshop on Geographic information Systems, pp 2--11 (2004)
8. Hattori, S., Tezuka, T., and Tanaka, K.: Query Modification Based on Real-World Contexts for Mobile and Ubiquitous Computing Environments. In: Proceedings of the 7th international Conference on Mobile Data Management, pp. 10--12, (2006)
9. De Luca, E. W. and Nürnberger, A.: Using clustering methods to improve ontology-based query term disambiguation. International Journal of Intelligent Systems, vol. 21, pp. 693--709 (2006)
10. Villalonga, C., Strohbach, M., Snoeck, N., Sutterer, M., Belaunde, M., Kovacs, E., Zhdanova, A., Goix, L.W., Droegehorn, O.: Mobile Ontology: Towards a Standardized Semantic Model for the Mobile Domain. In: Service-Oriented Computing - ICSOC 2007 Workshops, Springer, 248--257 (2009)
11. Cantera Fonseca, J. M., García Bernardo, G., Vegas Hernández, J.: An Automatic Page Classification Method to Improve User Experience on the Mobile Web. In: 2nd International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, pp. 171--177. IEEE Press, New York (2008)