

## **Notice of Violation of IEEE Publication Principles**

### **“Ontology-based Semantic Recommendation System in Home Network Environment”**

by Jinhyung Kim, Dongwon Jeong and Doo-kwon Baik

in the IEEE Transactions on Consumer Electronics, Vol 55, No. 3, August 2009, pp. 1178-1184

After careful and considered review of the content and authorship of this paper by a duly constituted expert committee, this paper has been found to be in violation of IEEE's Publication Principles.

This paper contains text and figures that have been copied from the paper cited below. The original text was copied without attribution (including appropriate references to the original author(s) and/or paper title) and without permission.

Due to the nature of this violation, reasonable effort should be made to remove all past references to this paper, and future references should be made to the following article:

### **“Applying User Profile Ontology for Mining Web site Adaptation Recommendations”**

by Tarmo Robal and Ahto Kalja

in the Proceedings of the 11th East-European Conference on Advances in Databases and Information Systems (ADBIS 2007), pp. 126-135

# Ontology-based Semantic Recommendation System in Home Network Environment

Jinhyung Kim, *Member, IEEE*, Dongwon Jeong, *Member, IEEE*, and Doo-Kwon Baik, *Member, IEEE*

**Abstract** — *Home network is one of the predominant applications in ubiquitous computing. In home network environment, we need semantic recommendation system for exact and abundant information recommendation to users. For semantic recommendation system, we construct ontology for user profiles and home network services firstly in this paper. In addition to, we suggest recommendation method for ontology-based semantic recommendation. By using the suggested system, we can support more precise and richer information recommendation.*<sup>1</sup>

**Index Terms** — User Profile Ontology, Service Ontology, Semantic Recommendation System, Home Network

## I. INTRODUCTION

With rapid growth on the Web, the amount of web information is created and propagated as web services. However, exact information retrieval gets more difficult since there is a considerable increase in the amount of information in the web. Therefore, the Semantic Web is suggested for solving these problems [1]-[2]. In addition, regarding services recommendation, the information is mainly represented in the form of web pages that users may browse; and the way they do it is heavily dependent on the information they are searching for. It has been proven that factors, such as visual experience and site attractiveness, logicity of navigation organization (especially on large sites), placement of objects, color schema and page loading time [3]-[5] also affect web surfing. Furthermore, the preferences of users are changing through time. Thus, there exists a constant need for improving and adapting applications and services according to the end-users' needs. For that we need to collect access data and mechanisms for data mining and processing. The rapid expansion of the Internet has provided the possibility to explore users' navigational patterns and interaction with web-based systems. This allows computing of recommendations for users, either by simple pages, or constructing new web site topologies for user groups

Therefore, for exact providing and recommendation of the home network services, we must support user profiles

ontology and services ontology implemented by an ontology language such as RDF, RDF-S [6], and OWL [7] firstly. In addition, we need the effective recommendation model by using user profile ontology and service ontology for more precise recommendation of web services or information.

In this paper, we suggest ontology-based semantic recommendation system in home network environment including service ontology, user profile ontology, and recommendation method. We describe type classification of service ontology and the method for creating user profile ontology. Furthermore, we discuss generation of recommendations for website user profile ontology and service ontology.

This paper is organized as follows. In section 2, we describe preliminaries regarding building of user profiles and recommendation systems. Section 3 represents overall architecture of ontology-based recommendation system in home network environment. In addition, we illustrate ontology for service and user profile. In section 4, we describe the ontology-based recommendation method in detail. Section 5 concludes this paper with future works.

## II. PRELIMINARIES

This section covers related works and researches regarding building user profile and recommendation systems.

### A. Building User Profile

In the last few years the need for software systems to automatically adapt to their users has been recognized in many application areas and the research on user profiling and context has spread into many disciplines which are concerned with the development of computer systems that are to be used by heterogeneous user populations [8]. Context in [9] is categorized in human user context and surroundings context and may also be categorized according to persistence (permanent and temporary) and evolution (static and dynamic). Elaine Rich [10] identifies a three dimensional space of user models: 1) canonical vs. individual user model, 2) explicit vs. implicit user model and 3) long-term vs. short-term user model.

Another important issue is that a user might be found in various contexts. Thus, a context-aware system has to infer which context the user is in a given moment in time, and consequently adapt the system to that context [10].

An overview of methods for building a user profile is presented in [10]. User modeling issues and guidelines are presented in [8], concentrating on modeling of user knowledge, plans, and preferences in a domain. It focuses on stereotype profiles. The need for a profile that supports reasoning is also stressed out in [10].

<sup>1</sup> This work was supported by the 2nd Brain Korea 21 Projects.

Jinhyung Kim is with the Department of Computer Science and Engineering, Korea University, Seoul, Republic of Korea (e-mail: jinhyung98.kim@gmail.com)

Dongwon Jeong is with the Department of Informatics and Statistics, Kunsan National University, Kunsan, Republic of Korea (e-mail: djeong@kunsan.ac.kr)

Doo-Kwon Baik is with the Department of Computer Science and Engineering, Seoul, Republic of Korea (e-mail: baikdk@korea.ac.kr)

The goals listed above can be achieved through the use of ontologies. Ontologies in the form of hierarchies of user interests have been proposed in [11]. Gauch et al. [12] also proposed a system that adapts information navigation based on a user profile structured as a weighted concept hierarchy. The user may create his/her own concept hierarchy and use it for browsing web sites. A user model can also be built using an ontology schema. Razmerita et al. [13] presented a generic ontology-based user modeling architecture applied in the context of a Knowledge Management System.

In the field of ontology design, efforts have been made by several research groups to facilitate the ontology engineering process, employing both manual and semi-automatic methods. Semi-automatic methods focus on the acquisition of ontologies from domain texts. In [3], for example, a framework is proposed with this objective; it incorporates several information extraction and learning approaches, in order to face the discovery of relevant classes, their organization in a taxonomy and the non-taxonomic relationships between classes. Comprehensive surveys of existing methodologies can be found in [4] and [5]. Throughout the ontology creation process, the designers may take into account a set of ontology design criteria, such as clarity, coherence and extensibility.

### B. Recommendation Methods

Collecting user interaction data during the web sessions and analyzing it in further, enables to develop assistance systems to help users during their web browsing, either directly (e.g. providing related links) or indirectly, i.e., constantly improving and adapting the website to the users' needs. However, usually such analyzing systems, as also noted in [14], suffer from inability to address the semantic aspects of the web. The use of ontology is an emerging trend in web mining, as it is one of the best ways for representing the structure of information regarding to its semantics.

The ability to track down users' actions has made it possible to develop systems which provide users with dynamically discovered recommendations and thus personalize the users' web experience. The aim of any recommender system is to assist users to find needed information in the easiest way via helping them to discover pages or page-sets they otherwise might not find during their site visit. Moreover, additional related information, not contained on that particular web site but which users might also find relevant, can be provided. For employing such systems, usage data has to be collected and processed. As a result, primary recommendation of web sites can be implemented for enhanced and improved user experience.

In general, recommender systems are based on site's overall usage information and therefore transparent for the end-users or being personalized with the aim to adjust the web cognition for a particular user. The latter though needs users to log in to a system or other means of explicit user identification. Recommendation systems are being already successfully used in personal web-based agents such as Letizia, Syskill&Webert, Personal Webwatcher, and OntoSeek. They have also been implemented for personalized e-learning, as discussed in [15].

Data can be gathered using either explicit or implicit methods. The first assumes users to actively participate in data gathering. For

example, we might ask users to rate pages on a sliding scale or give pages, where an evaluation form at the end of web pages is presented and users are asked to rank the pages with the possibility to add comments. Another option is to have surveys for web sites. These surveys though contain far more questions, which usually are not page-specific. Applying explicit data collection methods is a useful technique and provides the users' opinions and ideas; however users are usually not willing to actively participate in such evaluations, moreover fill in forms.

Implicit techniques, being transparent to the end-users, help to overcome these disadvantages. Automated data gathering has enabled monitoring of accessed pages, navigational paths; discovery of usage patterns and user profiles. The techniques usually involve data collection from server, proxy or client level using either web server logs, web browsers modified for data capturing, or special log systems [16].

Web server logs chronicle all the operations and do not produce log data for particular analyses. Most probably we are not interested in every single object accessed on a web server (e.g. dots and lines as elements of graphic design), which is also the case of this study. Moreover, it has been proven in [17] that HTTP traffic logs appear to be flawed and there are some major difficulties due to data incompleteness [16]-[18]. One of the many problems with web server log files is that they do not allow identifying visitor sessions [19].

As shown, standard web server logs occurred to be not suitable, as they did not contain all the needed information, for instance clustering of user actions into sessions and tracking recurrent visits. After analyzing different possibilities of usage data collection and taking into account the aforementioned drawbacks, it was clear that a different kind of approach was in demand to meet the needs of this study. As a result, a special log system was used.

The log system [17] was developed in 2003 and based on a preliminary log system introduced in 2002, which allowed storing only some basic properties of actions users performed and was only aimed on general usage statistics. The major improvement towards the new log system was the ability to capture distinct and recurring user sessions, which is also the basis of users' profiles construction.

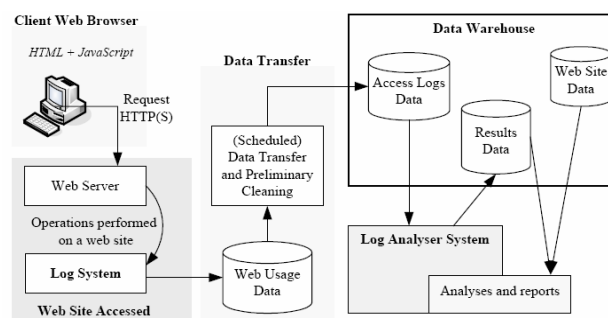


Fig. 1. Data Collection and Processing by the Log system

The web access log contains raw access data, which needs to be cleaned from noise and filtered before it can be used for user profiles extraction. Therefore, crawlers/robots detection rules were added to the log system, and our own experience, as blocking of



hardware and software resources of the device. The Device Status contains volatile information pertaining to the device such as the details of its location, CPU usage, remaining power level, method of power supply, etc.

**2. Function Ontology:** The function ontology identifies function of each device. Each device has several functions. For example, the DVD player can have following functions: storing new title, remove existing title, search a title, playing the title and so on. The indicated electronic devices and functions can be represented as ontology by using OWL:objectproperty relationship. By the function ontology, user can get the information about function lists of each electronic device easily.

**3. Operation Ontology:** The operation ontology describes operation steps corresponding to the function. Each function can be related to more than 2 devices. For example, search function in the DVD player can be controlled by the mobile phone or PDA. As mentioned, many functions in the home network environment are controlled and provided the separated mobile devices. Therefore, the functions must be considered with more than 2 devices. The operation ontology represents the relation between more than 2 devices by functions. In result, the function ontology and operation ontology can be represented as one ontology graph.

**4. Conditional Ontology:** The conditional ontology represents dependency between devices and environment. For example, there can be following restrictions. (1) The printer for printing service must be selected the closest one. (2) The phone and PDA can control only one device in one time. Except for (1) and (2), there can be several another restrictions regarding operation of devices in the home network environment. These restrictions are considered when the services are recommended and provided to users by the devices.

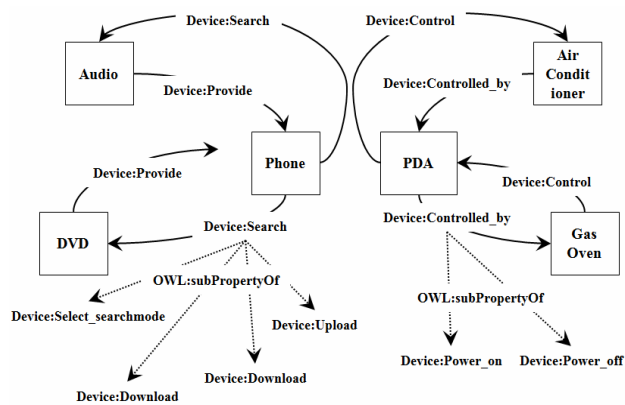


Fig. 5. Function & Operation Ontology

### C. Ontology for user profiles

In general, the ontology used in relation with user profiles is mostly limited to taxonomies of user interests. Bearing in mind that for most applications profiling is not restricted to

user interests but also encompasses other user characteristics (such as education, expertise and computer literacy level), our purpose is to incorporate them in a user profile ontology.

TABLE I.  
USER PROFILE ONTOLOGY

Class Name	Description
Person	Basic user information like name, date of birth
Characteristic	General user characteristics, like eye color, height, weight, etc.
Ability	User abilities and disabilities, both mental and physical
Living Conditions	Information relevant to user's place of residence and house type
Contact	Other persons, with whom the person is related, including relatives, friends, co-workers
Preference	User preferences, for example "loves cats", "likes blue color" or "dislikes classical music"
Interest	User hobby or work-related interests. For example, "interested in sports", "interested in cooking"
Activity	User activities, hobby or work related. For example, "collects stamps" or "investigates the 4th Crusade"
Education	User education issues, including for example university diplomas and languages
Profession	The user's profession
Expertise	Includes all kinds of expertise, like computer expertise
Thing	Living things or Non Living Things the user may possess or otherwise be related to, like a car, a house, a book or a pet

This section presents a brief description of the user profile ontology. The ontology may be extended through inheritance and the addition of more classes, as well as concept instantiation according to the needs of a specific application. As a result, it may be used for the representation of both stereotype profiles (i.e. user profiles that represent a specific user category, like "computer expert" or "woman") and individual ones.

This ontology presents information that is mostly static and permanent. More dynamic characteristics like the current position of the user when moving are currently not included. However, the temporal aspect of some of the ontology classes has been taken into account. The ontology allows the existence of multiple instances of classes that represent characteristics that may change with the passage of time, like living conditions for example. These classes include a period representing the validity period of their instances, for example, "Living Conditions: New York, 12/3/2003 – 18/8/2007".

Table I presents an overview of the proposed ontology upper level classes. The "Person" class is the central one in the ontology, as it contains all the user profile characteristics. These may be of simple type, like the user "name" or "date of birth", or may be instances of other ontology classes, like "physical characteristics", "contacts", etc.

The rest of the classes are used to describe the complex user characteristics. "Living conditions", "Contact", "Education", "Expertise", "Activity" and "Profession" include a set of slots describing the respective aspects of the user's life as well as a time period which represents the duration of that particular



aspect. For example, a user may have had a “Contact” of type “friend” from 1989 to 2004. The slot “person” of the “Contact” class has as type an instance of the class “Person”. This way, relations between different users may be modeled as well.

“Interest”, “Preference”, “Ability”, “Characteristic” and “Thing” contain only three slots: “type”, “name” and “score” (or “value” in the case of “Thing”). “Thing” has two sub-classes, “Living Thing” and “Non Living Thing” as modeled in the WORLDNET ontology [21]. In the case of interests, apart from the “type” slot, which is a String, a slot named “interest type” of type “Interest” has been added to allow the creation of interest hierarchies,

It should be mentioned here that in the case of the “Expertise” class, the aim was to collect from the existing literature user characteristics that may serve as indications or factors during the assessment of the user expertise level. The definition of the expertise levels them and the expertise measures are application-specific and out of the scope of the current work. To sum up, the “Expertise” class has been created as a container for both expertise measures and expertise scores in order to accommodate the particular needs of individual applications that make use of profiling.

#### IV. RECOMMENDATION METHOD IN ONTOLOGY-BASED RECOMMENDATION SYSTEM

The amount of information available over the Internet is enormous - even the quantum of information on an average web site is fairly large for users to maintain easy navigation. This problem can be revealed by providing recommendations based on user profiles mined from web logs. The rationale behind the approach of recommender systems is that users implicitly use a concept model based on their own knowledge of the domain or topic searched, even though mostly they do not know how to represent it [7]. However, having unobtrusively monitored users’ actions and collected them in the log, it is possible to apply detection algorithms to produce the concept models of users, i.e., extract user profiles.

By analyzing the actions of the current user online and comparing it to the user profiles discovered from the web log, it is possible to classify the user as an individual into one of the conceptual user groups and recommend him/her new pages that correlate with that individual (primary recommendation), highlight items, so that they could be easily located or even propose a new topology (analytic recommendation).

The items in extracted user profiles do not contain any semantic information about the content represented on the web-site. Therefore, the extracted user profiles need to be mapped to the concepts of web ontology. For that reason, profiles ontology is constructed. At this point, it was composed manually, however in the future it is planned to implement an automated generation of the OWL-file. In the user profile ontology, several user concepts were defined in

order to give meaning to the mining results and propose recommendations. The user profiles ontology is mapped to the web ontology (concepts “Researcher” and “Student”).

While building the web site ontology, we have to posit ourselves on the users’ side - they behave according to the view they have on a display, and make requests for a web page not a concept. Then and only then we are able to use the concepts of web ontology to characterize the actions of users and map semantics to profiles for producing recommendations.

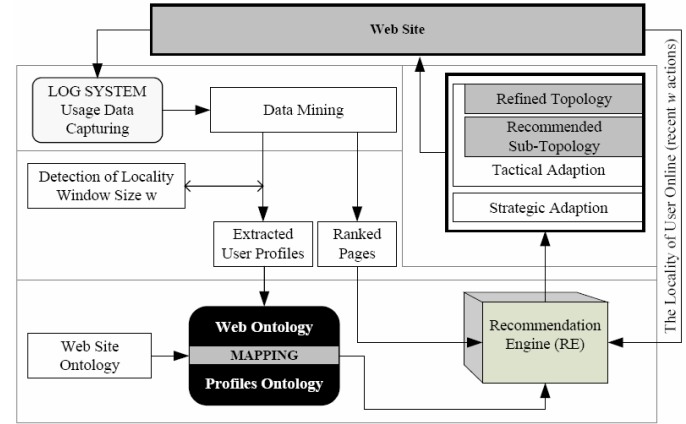


Fig. 6. Recommendations Generation based on Ontology for User Profile and Services

The task of the recommendation engine is to determine the type of the user online and compute recommendations based on the recent actions of that user. The decision is based on the knowledge attained from the ontology and page ranking. No user is attached to a particular outcome, as users are free to move from one locality to another. However, with carefully designed user profile ontology, the recommendation covers the majority of the user session. The pages are ranked using an inverse time weighting algorithm (1).

$$Rank = p \sum_{i=1}^n \frac{Interestvalue(i)}{Age(i)} \quad (1)$$

In the formula,  $Age(i)$  represents number of days into the past,  $Interestvalue(i)$  number of hits for a page during  $Age(i)$ ,  $p$  is a probability value between 0 and 1 and can be predefined. Thus, only the nearest past will play a crucial role in ranking.

##### A. Primary Recommendation

The primary recommendation involves a refined topology by adding items to an existing web topology in a way that they would be more easily accessed by users. In its simplest case, the recommended items in an existing web site topology are raised during the user’s online session. For instance, if the user is classified to be a student, then the set of items to be added into topology or highlighted in the existing one would consist of items belonging to those concepts identified by the recommendation engine and with ranking applied.

Setting up a new web site topology as a result of primary recommendation, would be an option as well. However, we

must not forget that there is a risk to confuse the end users by changing the topology too often. Additionally, users might be provided an option to classify themselves into one of the users' concepts by clicking a special link. An option to return to a regular web site based on the full topology is a must. In the long run, reasoning about the recommendations made for primary recommendations may propose strategic ones as well. For either case, the recommendation engine and ontology are needed.

### B. Analytic Recommendation

In terms of analytic recommendation, general site improvement is considered. The result of analytic recommendation is a new web site topology. As analytic recommendations need the approval of webmaster, they should not be applied online and should not be put into effect too often, as it may distress the regular visitors to find a new layout of the site every time they return. Nevertheless, web sites should adjust over time to the preferences of their users and therefore strategic changes are in need as well.

Analyzing the log, sets of pages that are frequently used together, are to be recognized. Running a comparison on the existing web ontology and the results obtained from the actual use of the site, proposals for improvement can be made. The domain of interest covers sets of pages that are related in the ontology but not used together, and sets of pages that are used together but are not described in the web ontology. As can be expected, new relationships not existing in the current web ontology can be discovered. These newly discovered relationships may lead to the need to add links between sets of pages in order to improve content availability for the users. Moreover, over time the sets of pages discovered may impose new web ontology, thus a new structure to be developed.

### C. Recommendation Example

Fig. 7 shows the recommendation result regarding movie titles by the PDA. For movie recommendation, we implement virtual DVD player with storage for storing movie titles because existing DVD players only have playing role of titles without storing and search role about movie titles. Then, we construct ontologies for user profiles and services by OWL language. These built ontologies are stored into the Semantic Home Server and the PDA access the Semantic Home Server to get the information about recommendation of movie title. By using ontologies for user profile and service, and suggested recommendation method, we can acquire more exact and richer recommendation information.



Fig. 7. Movie Recommendation Results by the PDA

## V. CONCLUSION

The enormous amount of information available over the Internet has produced the need for recommender systems which would help users to find the information they are seeking easily. Using various implicit data capturing methods, it is possible to track down users' actions and advise them to visit other related pages or assist in finding information according to their previous actions.

In this paper, we suggested the ontology-based semantic recommendation system in the home network environment. In conventional recommendation system, the method is passive and restrictive. Previous recommendation researches are dependent to keywords or sentences in user profile. However, suggested system can perform semantic and richer recommendation to users through analysis about ontologies for user profiles and home network services.

## REFERENCES

- [1] T. Lee, J. Hendler, and O. Lassila, *The Semantic Web*, Scientific American, May. 2001.
- [2] I. Fahmi, J. Ahang, H. Ellermann, and G. Mouman, "System Description: A Case Study in Information Retrieval, Inference and Visualization in the Semantic Web", *LNCIS 4519*, pp. 769-778, Jun. 2007.
- [3] A. Maedche, S. Staab, "Mining Ontologies from Text", *EKAU 2000*, pp. 189-202, Oct. 2000..
- [4] M. Cristani, and R. Cuel, "A Survey on Ontology Creation Methodologies", *International Journal on Semantic Web and Information System*, Vol. 1, No. 2, pp. 46-69, 2005..
- [5] N. Noy, and C. Hafner, "The State of the Art in Ontology Design, A Survey and Comparative Review", *AI Magazine*, Vol. 18, No. 3, pp. 53-74, 1997.
- [6] D. Beckett, and B. McBride, "RDF/XML Syntax Specification", W3C Recommendation, Feb, 2004.
- [7] D. Macguinness, F. Harmelen, "OWL Web Ontology Language Overview", W3C Recommendation, Feb, 2004.
- [8] A. Kobsa, "User Modeling and User-Adapted Interaction", *The Journal of Personalization*, Springer Verlag:Heidelberg, 2008.,
- [9] D. Raz, A. Juhola, J. fernandes, A. Galis, *Fast and Efficient Context-Aware Services*, John Wiley & Sons Ltd: NewYork, 2006..
- [10] E. Rich, "Users are individuals: individualizing user models", *International Journal of Man-machine Studies*, Vol. 18, No. 3, pp. 199-214, 1976
- [11] J. Trajkova, S. Gauch, "Improving Ontology-based User Profiles", *RIAO 2004*, pp. 380-389, Apr. 2004.
- [12] S. Gauch, J. Chaffee, A. Pretschner, "Ontology-based User Profiles for Search and Browsing", *Web Intelligence and Agent Systems*, Vol. 1, No. 3-4, pp. 219-234, 2003.
- [13] L. Razmerita, A. Angehrn, A. Maedche, "Ontology based User Modeling for Knowledge Management Systems, *User Modeling Conference*, pp. 213-217, 2003
- [14] A. Mikroyannidis, B. Theodoulidis, "Web usage Driven Adaptation of the Semantic Web", *UserWeb 2005*, pp. 137-147, May. 2005
- [15] P. Markellou, I. Mousourouli, S. Spiros, A. Tsakadilis, "Using Semantic Web Mining Technologies for Personalized e-Learning Experiences", *WBE 2005*, pp. 461-826, Feb. 2005.
- [16] J. Srivastava, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", *SIDKDD Explorations*, Vol. 1, No. 2, pp. 12-23, 2000.
- [17] T. Robal, A. Kalja, J. Pold, "Aalysing the Web Log to Determine the Efficiency of Web Systems", *DB&IS 2006*, pp. 264-275, Jul. 2006.
- [18] B. Mobasher, R. cooley, J. Srivastava, "Automatic Personalization Based on Web Usage Mining", *Communications of the ACM*, Vol. 43, No. 8, pp. 142-151, 2000.
- [19] R. Kimball, R. Margy, *The data warehouse toolkit: the complete guide to dimensional modeling*, John Wiley & Sons Ltd: New York, 2<sup>nd</sup> ed. 2002.

- [20] P. Tan, V. Kumar, "Discovery of Web Robot Sessions based on their Navigational Patters", *Data Mining and Knowledge Discovery*, Vol. 6, No. 1, pp. 9-35, 2002.
- [21] G. Miller, "WORDNET: An On-Line Lexical Database", *International Journal of Lexicography*, Vol. 3, No. 4, pp. 235-312, 1990.



**Jinhyung Kim** (M'09) This author became a Member of IEEE in 2009. This author received his B.S degree in Computer Science and Engineering from Hongik University, Seoul, Korea, in 2004. He received his M.S. degree in Computer Science from Korea University, Seoul, Korea, in 2006. Currently, He is a Ph. D candidate in department of computer science and engineering, Korea University, Seoul, Korea. At present, he is also a researcher in the computer information and communication center. His research interests include XML data management, Semantic Web, Metadata, XML access control, etc.



**Dongwon Jeong** received his B.S degree in Computer Science from Kunsan National University, Korea, in 1997. He received his M.S. and Ph.D. degree in Computer Science from Choongbuk National University, Korea, in 1999 and from Korea University, Korea, in 2004 respectively. He was a Visiting Research Scholar (PostDoc.), School of Information Sciences & Technology, Pennsylvania State University, US, 2005. Now, he is a Professor in the Dept. of Informatics and Statistics, Kunsan National University, Korea from 2005. His research interests include Data Integration, Semantic Web, Semantic Sensor Network, Semantic GIS, Semantic Grid, and Security.



**Doo-Kwon Baik** received his B.S. degree in Mathematics from Korea University, Seoul, Korea, in 1974. He received his M.S. and Ph.D. degrees in Computer Science from Wayne State University at U.S.A., in 1983 and 1986, respectively. Currently, he is a full professor in the department of computer science and engineering, Korea University, Seoul, Korea. Also, he is a committee chair of ISO/IEC JTC1/SC32- Korea. His research interests include Data Engineering, Software Engineering, and Modeling and Simulation.