Intelligent Semantic Universal Search in Mobile Platform

Jinhyung Kim, Myunggwon Hwang, Hanmin Jung, Doheon Jung
Information Software Research Center,
Korea Institute of Science and Technology Information,
52-11, Eoeun Dong, Yuseong Gu, Daejeon, Republic of Korea

Abstract: Problem statement: As the performance of mobile devices is getting higher, a large quantity of information is stored and managed in mobile devices. Therefore, the precise and intelligent universal search becomes the most important issue in the smart mobile platform. **Conclusion/Reommendation:** In this study, we suggest the semantic universal search in smart mobile platform facilitated by the ontology; ontology modeling, ontology learning and ontology reasoning. Through the suggested system, we can enhance usability and convenience compared to conventional universal searches.

Key words: Data Dependency (DD), Data Function (DF), Data Sources (DS), universal search application, provide snapshots, conventional universal searches

INTRODUCTION

As a specification (ex. CPU speed, storage capacity, memory capacity) of mobile phone is getting higher revolutionarily, current mobile phones manage various kinds of data in the phone such as contacts, calendar, email, schedule, multimedia (ex. image, music, video), application lists. In 20 years, mobile processor was progressed from 1MHz to 1GHz and the storage capacity was developed more than 100 times. However, precise searching of requested information is becoming more difficult since there is a considerable increase in the volume of data in mobile phones. Precise information retrieval of the requested information is more important than just fast search of the data (Kim, 2011; Kim *et al*., 2009; Chaisatien and Tokuda, 2011; Sagayeshi *et al*., 2009; Bellavista *et al*., 2006).

For overcoming above limitation, many platform companies support a universal search in smart mobile platform for ease to search about diverse data in mobile phones. Most of mobile platform suppliers made an effort for effective search. The universal search application of many platforms provides an integrated search view about many kinds of data in the mobile phone. Given a keyword, the universal search application execute search various kinds of data such as contacts, schedule, email, image, video, music and so on. Then, the universal search application provides the integrated view about unified search results.

However, universal search applications mentioned above are based almost purely on the occurrence of words in documents as the traditional information retrieval (Chen and Koskela, 2011; Xuan *et al*., 2011).

The universal search application just allows users to enter some keywords to retrieve data that contain the given keyword. Therefore, this search method cannot retrieve more exact and diverse results. For example, "central park" is given as a keyword, then contacts, image, music data including given keyword are viewed as a search result. However, if we want to search a picture with my friends at the central park and contacts about those friends, the traditional universal search in smart mobile platform cannot support them. Many kinds of data in mobile device are related each other, search function for semantically related term is important feature. Therefore, we believe that the addition of explicit semantics can improve search in smart mobile devices.

For Realizing above improvement, we suggest the semantic universal search in this study. The Semantic universal search is an ontology-based application for supporting search about relationships between a set of terms and concepts. We construct additional ontology for consisting relationship among various kinds of data in OWL format (Cristani and Cuel, 2005; Noy and Hafner, 2007; Garcia-Castro and Gomez-Perez, 2010; Backett and McBride, 2004; Macguinness and Harmelen, 1976; Ren *et al*., 2003). Then, the semantic universal search includes all of related data as search results by using that ontology. In addition, the semantic universal search provides a graph view for search result display. This view is superior to traditional list-based search result view in user convenience and aesthetic (Rich, 1983; Trajkova and Gauch, 2004; Gauch *et al*., 2003; Razmerita *et al*., 2003; Mikroyannidis and Theodoulidis, 2005 Markellou *et al*., 2005).

Corresponding Author: Doheon Jung, Information Software Research Center, Korea Institute of Science and Technology Information, 52-11, Eoeun dong, Yuseong gu, Daejeon, Republic of Korea
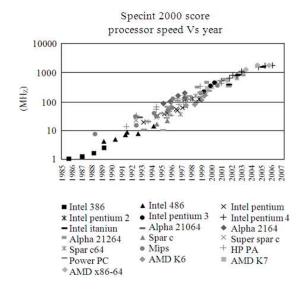
Specint 2000 score
processor speed Vs year

Fig. 1: Mobile CPU progress

Semantic data modeling:

Data modeling: For overcoming limitations of services supported by conventional universal searches and realizing intelligent and semantic services for users, the data modeling based on ontology must be performed as the first priority. In this study, we construct 3 ontologies for more precise and semantic search services: The Data Sources (DS) ontology, the Data Function (DF) ontology and the Data Dependency (DD) ontology.

Data Sources (DS) ontology: The DS ontology includes relationship and hierarchical information among data sources applications such as contacts, SMS, MMS email, music picture, video,. The DS ontology is focusing on the definition about classes in the ontology. The data source class is based on various kinds of applications in the smart mobile platform and can be divided four typical sub data source classes: the key data sources class, the message data sources class, the multimedia data sources class and the web data sources class (Fig. 1). The key data sources class includes indispensable instances such as contacts data and application data. The data sources in the key data source class affect many other data sources. The contacts data affects information stored in the SMS, MMS, Email and the application data affects all of application information in the smart mobile platform. If any of data in the key data sources class is changed, the data affected by those data must be modified. As shown in the Fig. 2, the contacts data and the application data included in the key data sources class have many 'DS: Affected By' relationships with many instances included in other classes.

Data Function (DF) ontology: The DF ontology includes functions information of each data source and inter-operations among several applications. The DF ontology supports the definition about properties in the ontology. The DF ontology is based on functions and operation of several applications. The Fig. 3 shows a part of the DF ontology. Each applications and application data have 'DF:ManageFunc' property and 'DF:ProvideFunc' property. The 'DF:ManageFunc' property can be defined many sub-properties based on characteristics of each application. Characteristics of each application are based on functions and operations supported by applications. For example, the 'DF:ManageFunc' property between the contacts application and contact data has five sub-properties which represents typical function or operation. Generally, the contacts application supports storing, delete, modification, search and move data function. In addition, the search function provides three detail search functions: search by name, phone number and group. The move data property also has additional sub-properties: download and upload functions.

Except for property between the specific application and application data, there are property between different applications or the specific application and the other application data. The SMS and MMS applications have property with the contacts data. Additionally, every application is controlled by the application manager. The red lines in the Fig. 3 represent the specific application and the other application data and the blue line in the Fig. 3 shows properties between different applications

Data Dependency (DD) ontology: The DD ontology represents dependency between applications and environment. For example, there can be following restrictions. (1) The application manager can execute 'DF: delete data' operation about application data restrictively. When some application has no application data, the application manager can delete that application. (2) If some contacts information is deleted, the SMS and the MMS application cannot perform 'DF: search by name' operation by that name. (3) If there is not the SD card or the smart platform does not support the SD card, the application manager cannot execute 'DF: Move to SD' operation. (4) If the smart platform does not support network communication at some point, the contact application cannot perform 'DF: move data' operation. Except for (1), (2), (3) and (4), there can be several another restrictions regarding operation of application in the smart mobile platform. These restrictions are considered when the services are recommended and provided to users by the devices.
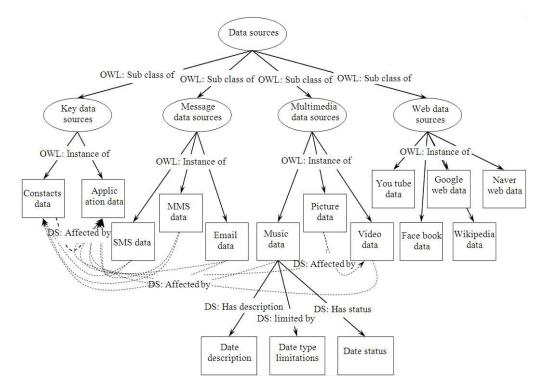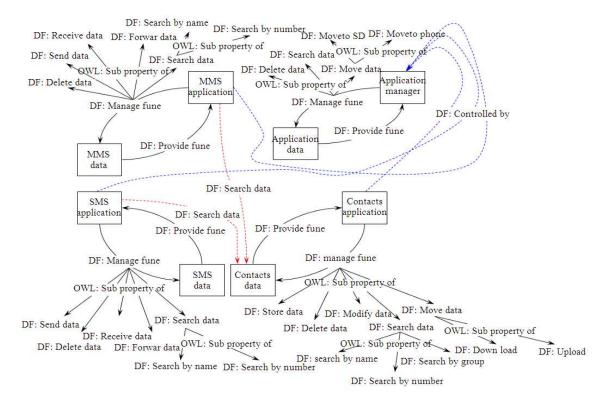
Fig. 2: Data sources ontology
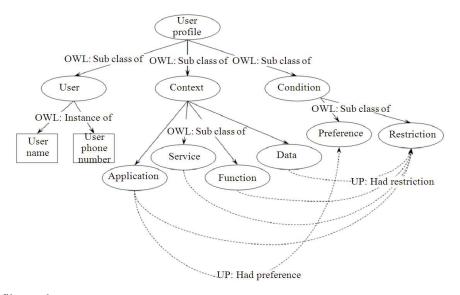

Fig. 3: Data functions ontology

Fig. 4: User profile ontology

Table 1: Axioms in ontology comparison

Axioms	Description
AXIOMS 1	Regarding user U, ∀U is interested to resources R, then is certainly interested to the lower conception resources of R
Axioms 2	Regarding user U, ∀U is interested to resources R, then is interested to the higher position of R by certain proportion
Axioms 3	Regarding user interested in resources R, R's lower position concept is more close to the user's interests compared to higher position
Axioms 4	Regarding resource R, deeper depth is suitable to more users

User profile modeling: The representation method of user profile refers to how to save the user's demand for formation and it's closely related to Web knowledge expression method sometimes. There are many representation methods for user profile nowadays, such as keyword vector, semantic network, XML or RDF and so on. However, conventional methods for representation of user profile cannot satisfy the user's semantic and intellectual demand, so it is imperative to use new methods to express user's information. Ontology has good conceptual hierarchical structure and support logical reasoning well, thus use user profile based on the ontology is kind of new attempt.

User profile ontology includes following 7 classes: user information, application, service, function, data, preference, restriction:

- User Information: Basic user information including name, phone number
- Application: Application name which the user used or related applications

- Service: Service name which the user used
- Function: Function name which the user used for getting the service
- Data: Data list related to the function and service which the user used
- Preference: Application use weight for each user
- Restriction: Specific restriction about application or service which user used

Figure 4 shows ontology structure for user profile. When provides information to the user, compares user profile ontology, DS ontology, DF ontology and DD ontology then provides the similar one. If there is user interested information in the system, provides this resource. Otherwise gives its lower position resources according to the axiom 1 and 3 in the Table 1. If the lower position resource does not exist, provide the higher position resource according to the axiom 2 in the Table 1.

User profile selection:

Classifying and clustering: At present there are mainly two methods to obtain user's characteristic information: the explicit and the implicit methods. The explicit methods is that users provider their own correlative information of interest. This method can reflect the user's need and interest change quite accurately, but require the direct participation of users. The implicit method carries on the analysis of user's information saved by the system, namely the web mining. The realization of this method is complex, but the implicit method does not need user's participation. This point has a vital significance in the practical application.

For realization of the implicit method, we dig user's use pattern and preference, pick up concept subject corresponding to each pattern and analyze these concept subjects to obtain the relativity between the concept and user. This is a kind of classification with completely no information of category. In realization process, may first carry on the cursory cluster to find the basis of classification and reduce the search space of classification, then user supervised clustering method to improve the precision further. The software vector machine and unsupervised clustering method would be taken as an example.

In training stage, assign cluster radius r and find each center of positive example set and counterexample set with unsupervised clustering method. Input positive example set and counterexample set of class α separately and seek their respective center, among:

$$\Omega+ = \{xi \mid (xi, yi) \in E, y1 = 1\}$$
$$\Omega- = \{xi \mid (xi, yi) \in E, y1 = -1\}$$

The center of $\Omega+$ is O1+, O2+,..., On+, the center of $\Omega-$ is O1-, O2-,... Om-.
Then choose some training example and give it to the soft vector machine to study. Its principle is that the training set contains all positive examples and some counterexamples that are close to the center of positive examples. We choose these counterexamples because they are more likely elected for the support vector. The given cutting radius R (R > r), soft vector machine training set may reduce to:

$$\Omega+ \cup \{ x \mid x \in \Omega- \wedge x \in (Oi+) \}$$

Among this, Oi+ is the circle center of BR (Oi+), R is the radius.

In recognition stage, regarding random vector x and decision making threshold value ε, if | dx+ - xx- | > ε. It is greater that x is away from distance difference most close to center of pros and cons examples, then use the unsupervised clustering method to classify class x. Otherwise, making use of the high separating capacity characteristic in two boundaries, use soft vector machine to make decision.

Properly combining classifying and clustering algorithm can enhance the speed and the precision of the classifier and compute interest change of user.

Ontology reasoning: Ontology reasoning is a behavior to reach to the conclusion based on the already known facts (proposition), or a procedure logically leading the process. Reasoning in the ontology field includes description logic based DL reasoning and rule based reasoning. In this study, we perform ontology reasoning by using equivalence relation of ontology such as symmetric, inverse, transitive, hierarchical properties.

Symmetric relation:

{ClassA, PredC, ClassB} ⇔ {ClassB, PredC, ClassA}

The symmetric relation means that the class A is equivalent to the class B. The symmetric relation can be represented as above.

Inverse relation:
{ClassA, PredC, ClassB} ⇔ {ClassB, PredD, ClassA}

The inverse relation means that if the class A is related to the class B with the specific predicate, the class B has relationship with A by the particular predicate. The inverse relation can be illustrated as above.

In the DF ontology, we can reason the typical inverse relation. Each application class has relationship with application data class by 'DF:ManageFunc' predicate. In addition, Each application data class is connected to the application class with the 'DF:ProvideFunc' predicate. As a result, we can reason that if two classes has relationship by 'DF:ManageFunc' predicate, the inverse relationship has 'DF:ProvideFunc' predicate

Transitive relation:

{ClassA, PredC, ClassB} ∧
{ClassB, PredC, ClassD} ⇔ {ClassA, PredC, ClassD}

The transitive relation means that if the class A is related to the class B and the class B is related to the class C, the class A is also related to the class C. The transitive relation can be represented as above.
In the DS ontology, we can extract new ontology information by using the transitive relation. The SMS data class is connected to the contact data class by the 'DS:AffectedBy' predicate and the contact data class and the application data class with 'DS:AffectedBy' predicate. Through above relationships, we can reason that the SMS data class is related to the application data class with 'DS:AffectedBy" predicate.

Hierarchical relation:

{ClassA, PredC, ClassB} ∧ {ClassB, PredE, ClassD} ∧
{PredE, OWL:subPropertyOf, PredC}
⇔ {ClassA, PredE, ClassD}

The hierarchical relation is similar to the transitive relation. The transitive relation is that the three classes has relationship with the same predicate, but the

hierarchical relation represents that the three classes has relationship with a different predicate and two classes have 'OWL: instace of' relationship. The hierarchical instance relation can be illustrated as above.

In the DS ontology, we can perform ontology reasoning by using the hierarchical relation. The MMS data class is related to the contact data class by the 'DF: affected by' predicate and the MMS data class is connected the datatype class with the specific predicate. Finally, we can reason that the datatype class is related to the contact data class with 'DF:AffectedBy' predicate.

System implementation: For implementation of the suggested system including the ontology modeling idea and user profile selection idea, we use the smart phone with Android OS. The smart phone for experiment has following specifications. The phone has 1Ghz Cortex-A8 processor, PowerVR SGX530 GPU and TI OMAP 3630 chipset. Additionally, the phone has 512MB RAM and Android 2.2 OS, namely FROYO.

Snapshots: The Fig. 5 shows the search result list of the suggest system. The most different point compared to conventional universal searches is the semantic search. The existing universal searches are based on keyword based search, so those search only search contents including the given keyword in the contents title. As already mentioned, keyword-based universal search has serious limitation. The most important feature of the universal search is that user wants to search all of related contents about a specific keyword. That is, even though contents are related to some person or some topic, if they include same keyword as title, user cannot find those contents simultaneously. However, the suggested system can search all of related contents by using ontology learning and reasoning. As shown the Fig. 5, user input 'Sunryang Kim' as keyword, but the system also provides contents which do not include the given keyword in the title such as video contents, named 'Birthday Party' and photo contents, named 'NewYork Meeting'. For complete realization of the universal search in the smart mobile platform, the system must support not only keyword-based search but also semantic search like the suggested system.

The Fig. 6 shows the graph view of search results about the given keyword. By the graph view, we can enhance understandability about all of related information and relationship among them. Additionally, even though users cannot know the title of some contents, users can search those contents by using related information. The graph view shown in the Fig. 6 is created by ontology modeling, ontology learning and ontology reasoning. Through ontology techniques, we can support not only keyword-based search but also the reasoned secondary and tertiary information.

Fig. 5: Search result list

Fig. 6: Search Result Graph

As shown in the Fig. 6, the SMS and contact information including 'Sunryang' is result by the keyword-based search. The 'Birthday Party' video and the 'New York Meeting' image is the secondary information extracted by ontology reasoning. In addition, we can reason the tertiary information based on the secondary information. The 'Tom Wilson' contacts, the 'Jenifer' contacts, the '20110321' image, 'sub workshop' schedule and 'Jonathan Lee' contact are included in the tertiary information. We can perform ontology reasoning continuously based on the latest reasoned information, but the over-many reasoning may extract information unrelated to the given keyword.

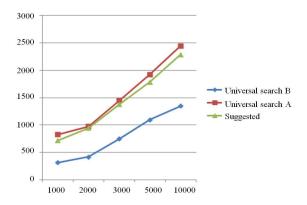Fig. 7: Preference and timeline of search results

Fig. 8: Query processing time

Therefore, in this system, we impose restriction on reasoning level as maximum 4 levels. However, if users want to search the related information of the quaternary information, user can execute re-search based on the quaternary information by clicking that information.

The Fig. 7 shows user preference of search results and timeline information. The x-axis represents timeline and the y-axis describes user preference about each data. User preference is calculated based on search frequency, usage frequency and user's private preference information described in the user profile. By the Fig. 7, user recognize not only related information and relationship about the given keyword but also preference information of each data multilaterally.

Performance test: We execute the comparative performance test by query processing time. The dataset for the experiment is contacts, SMS, MMS, music, image, video, schedule data. We perform experiment from 1,000-10,000 data. As shown the Fig. 8, as the amount of data increases, the query processing time also become longer. The unified search A shows best performance in the query processing time, but the system cannot support semantic search and users cannot retrieve information that they want to find precisely. The

suggested system supports semantic search for effective information retrieval and includes additional process such as the user profile analysis and the ontology reasoning. Nevertheless, the suggested system contains additional operations, the suggested system shows better query processing time than the unified search B.

CONCLUSION

In this study, we proposed the ontology-based intelligent universal search system in the smart mobile platform. For semantic universal search, we execute ontology modeling about data sources, data functions, data dependency and user profiles. Then we adopt classification and clustering method for user profile selection and the ontology reasoning technology for user profile analysis based on user preferences. In addition to, we execute the comparative evaluation with conventional universal searches. The result of experiment represents that in spite of additional processes, the suggested system shows better performance than the other universal search based on same OS.

As future works, we will execute the usability test for users by a focus group interview or an online survey. Additionally, we execute the additional performance test in ontology construction and ontology reasoning aspect.

REFERENCES

Backett, D. and B. McBride, 2004. RDF/XML syntax specification. W3C Recommendation.

Bellavista, P., A. Corradi, R. Motanari and A. Toninelli, 2006. Context-aware semantic discovery for next generation mobile system. IEEE Commun. Meg., 44: 62-71. DOI: 10.1109/MCOM.2006.1705981

Chaisatien, P. and T. Tokuda, 2011. A description-based approach to mashup of web applications, web services and mobile phone applications. Frontier Artif. Int. Appl., 225: 321-332.

Chen, X. and M. Koskela, 2011. Mobile visual search from dynamic image databases. Proceedings of the 17th Scandinavian Conference on Image Analysis, (SCIA' 11), ACM, Springer-Verlag Berlin, Heidelberg, pp: 196-205.

Cristani, M. and R. Cuel, 2005. A survey on ontology creation methodologies. Int. J. Semantic Web Inform. Syst., 1: 46-69.

Garcia-Castro, R. and A. Gomez-Perez, 2010. Interoperability results for semantic web technologies using owl as the interchange language. Web Semantics: Sci. Services Agents World Wide Web, 8: 278-291.

Gauch, S., J. Chaffee and A. Pretschner, 2003. Ontology-based user profiles for search and browsing. Web Intell. Agent Syst., 1: 219-234.

Kim, H., 2011. Comparison of customers perception of feature and smart phone users mainly in 20s. J. Digital Policy Manage., 9: 115-124.

Kim, J., D. Jeong and D. Baik, 2009. Ontology-based semantic recommendation system in home network environment. IEEE Trans. Consumer Elect., 55: 1178-1184. DOI: 10.1109/TCE.2009.5277973

Macguinness, D. and F. Harmelen, 1976. OWL: Web ontology language. W3C Recommendation.

Markellou, P., I. Mousouroulli, S. Spiros and A. Tsakadilis, 2005. Using semantic web mining technologies for personalized e-learning experiences. ACTA Press.

Mikroyannidis, A. and B. Theodoulidis, 2005. Web usage driven adaptation of the semantic web. University of Manchester, United Kingdom.

Noy, N. and C. Hafner, 2007. The state of the art in otology design, a survey and comparative review. AI Meg., 18: 53-74.

Razmerita, L., A. Angehrn and A. Maedche, 2003. Ontology-based user modeling for knowledge management systems. Lecture Notes Comput. Sci, 2702: 148-148. DOI: 10.1007/3-540-44963-9_29

Ren, Q., M.H. Dunham and V. Kumar, 2003. Semantic caching and query processing. IEEE Trans. Knowl. Data Eng., 15: 192-210. DOI: 10.1109/TKDE.2003.1161590

Rich, E., 1983. Users are individuals:- individualizing user models. Int. J. Man-Machine Stud., 18: 199-214.

Sagayeshi, F.M., M.E. Shiri and A. Faraahi, 2009. Using fuzzy theory to information predicate for improvement of cache operation in mobile databases. Proceedings of the International Conference on Future Networks, Mar. 7-9, IEEE Xplore Press, Bangkok, pp: 104-107. DOI: 10.1109/ICFN.2009.52

Trajkova, J. and S. Gauch, 2004. Improving ontology based user profiles. RIAO, pp: 380-389.

Xuan, K., G. Zhao, D. Taniar, W. Rahayu and M. Safar *et al.*, 2011. Voronoi-based range and continuous range query processing in mobile databases. J. Comput. Syst. Sci., 77: 637-651. DOI: 10.1016/j.jcss.2010.02.005