

Semantic Network based Integrated Multi-Level Search System for Smart TV

Myung-Eun Kim¹, Joon-Myun Cho¹, Sang-Ha Kim²

¹Smart TV Research Center, ETRI (Electronics and Telecommunications Research Institute), Korea

{mekim, jmcho}@etri.re.kr

²Department of Computer Science and Engineering, Chungnam National University, Korea

shkim@cnu.ac.kr

Abstract— As the Internet technology has been growing up fast; Internet has been used as a common infrastructure widespread. The broadcast environment has been also changed with transmission network expanded to Internet. TV is no longer than a passive device to receive a signal of terrestrial television broadcasting and it is evolving into a smart device providing Internet-based services. The evolution made Smart TV users get plenty of multimedia contents from broadcasting stations or content providers on Internet. However, it is quite difficult to find multimedia contents using ambiguous user search keywords. In this paper, we have designed the semantic network based integrated multi-level search system to improve search accuracy and expand search coverage. The proposed system can retrieve a target multimedia content what a user exactly wants from various sources using user keywords interpreted by semantic network.

Keywords— Semantic Network, Multi-Level Search, Integrated Search, Smart TV

I. INTRODUCTION

The Internet technology has evolved tremendously over the years, so it has been having a big effect on our daily life such as education, art, game, and broadcast etc. In reality, it is no exaggeration to say Internet is an indispensable infrastructure to lead everyday life these days. Now Internet technology plays a leading role in developing new technologies as well as becomes a fundamental infrastructure to provide various services. As the Internet has been growing up, many technologies have taken another big step forward to a new paradigm shift – the convergence with IT technology. The broadcast technology has been also changed with transmission infrastructure expanded to Internet. TV is no longer than a passive device to receive a signal of terrestrial television broadcasting and it is evolving into a smart device providing Internet based services [20]. Nowadays we can watch plenty of multimedia contents scattered over the Internet and make use of interactive services including SNS, Internet full browsing, and applications on Smart TV.

While Smart TV provides a lot of services in addition to broadcasting, the fact remains the major role of TV is to watch multimedia contents either TV shows or payable videos. These days users can get a great deal of multimedia contents from various sources, broadcasting stations or contents providers on Internet; however, it is quite difficult to find a multimedia content what a user wants to watch using the user

search keywords entered with the limited UI/UX on TV. We think a TV hasn't still been a good interactive multimedia device with easy input devices such as a mouse and a keyboard to get lots of input data easily from users but a fantastic appliance to watch video contents using a large screen. Therefore it is needed to develop a new search system for Smart TV that can retrieve target contents what a user exactly wants using ambiguous keywords from various multimedia sources including broadcast network and Internet. We consider newly emerged broadcast environment and design the semantic network based integrated multi-level search system which searches target contents efficiently from several content repositories by analysing ambiguous keywords. In the rest of this paper, we present our contributions. In Section II, we present the related works and Section III describes the proposed system, the semantic network based integrated multi-level search system. Then we discuss our current works and future works to improve multimedia search efficiency.

II. RELATED WORKS

Unlike textual data, whose content can be searched for using text strings, multimedia information retrieval(MIR) is dependent on processes that have either cumbersome requirements for feature comparison (e.g. color or texture) or rely on associated, more easily processable descriptions, selecting aspects of an image or video and expressing them as text, or as concepts from a predefined vocabulary.[17]

The earliest years of MIR were frequently based on computer vision algorithms which focused on feature based similarity search over images, video, and audio. Influential and popular examples of these systems are QBIC [1] and Virage [2], around mid-90s. Within a few years, the basic concept of the similarity search was transferred to several Internet image search engines including Webseek [3] and Webseer [4]. Significant effort was also placed on the direct integration of the feature-based similarity search into enterprises-level databases such as Informix datablades, IBM DB2 Extenders, or Oracle Cartridges [5] to make MIR more accessible to private industry.

In the area of video retrieval, the main focus in the mid-90s was on robust shot boundary detection; the most common approaches involved thresholding the distance between color histograms corresponding to two consecutive frames in a

video [1]. Hanjalic et al. [6] proposed a method which overcame the problem of subjective user thresholds. Their approach was not dependent on any manual parameters. It gave a set of key frames based on an objective model for the video information flow. Haas et al. [7] described a method of using the motion within the video to determine the shot boundary locations. Their method outperformed the histogram approaches of the period and also performed semantic classification of the video shots into categories such as zoom-in, zoom-out, pan, and so on. A more recent practitioner's guide to video transition detection is given by Lienhart [8].

Near the turn of the 21st century, researchers noticed that the feature-based similarity search algorithms were not as intuitive or user-friendly as they had expected. One could say that systems built by research scientists were essentially systems which could only be used effectively by scientists. The new direction was geared toward designing systems which would be user-friendly and could bring the vast multimedia knowledge from libraries, databases, and collections to the world. To do this, it was noted that the next evolution of systems would need to understand the semantics of a query, not simply the low-level underlying computational features [18]. This general problem was called "bridging the semantic gap". The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. Techniques for attempting to bridge the semantic gap in multimedia retrieval are followings [19].

A. Automatic Annotation

The current techniques for auto-annotation generally fall into two categories; those that first segment images into regions, or 'blobs' and those that take a more scene-orientated approach, using global information.

Duygulu et al. [9] proposed a method by which a machine translation model was applied to translate between keyword annotations and a discrete vocabulary of clustered 'blobs'. Jeon et al. [10] improved on the results of Duygulu et al. [9] by recasting the problem as cross-lingual information retrieval and applying the Cross-Media Relevance Model (CMRM) to the annotation task. Lavrenko et al. [11] used the Continuous-space Relevance Model (CRM) to build continuous probability density functions to describe the process of generating blob features.

Most of the auto-annotation approaches described above performs annotations in a hard manner; that is, they explicitly apply some number of annotations to an image. A hard auto-annotator can cause problems in retrieval because it may inadvertently annotate with a similar, but wrong label.

B. Semantic Spaces

Instead of applying hard annotations, we have developed an approach in which annotation is performed implicitly in a soft manner. A semantic-space of documents (images) and terms (keywords) is created using a linear algebraic technique. Similar documents and/or terms within this semantic-space share similar positions within the space. Latent Semantic

Indexing is a technique originally developed for textual information retrieval. Berry et al [12] described how Latent Semantic Indexing can be used for cross-language retrieval because it ignores both syntax and explicit semantics in the documents being indexed.

C. Semantic Annotation

Early work on semantically describing images using ontologies as a tool for annotating and searching images more intelligently was described by Schreiber et al [13]. Several authors have described efforts to move the MPEG-7 description of multimedia information closer to ontology languages such as RDF and OWL [14, 15].

The aim of using ontologies to describe multimedia resources is to provide well-structured information to improve the accuracy of retrieval. Semantic web technologies also facilitate the integration of heterogeneous information sources and formats. Well-structured information is crucial for providing advanced browsing and visualisation facilities, as opposed to more traditional query-based systems.

There are several approaches to semantically annotating multimedia. The aceMedia project [16] is developing a knowledge infrastructure for multimedia analysis, which incorporates a visual description ontology and multimedia structure ontology. They have also developed the M-OntoMat-Annotizer tool that allows users to manually annotate multimedia items with semantic information. Won-Ken Yang at al [21] developed an image description and matching scheme using synthesized spatial and statistical features for recommendation service.

III. SEMANTIC NETWORK BASED INTEGRATED MULTI-LEVEL SEARCH SYSTEM

The Smart TV has lots of multimedia content sources such as broadcasting stations and various content repositories scattered on Internet. We classified these multimedia content sources into three categories. The first is the local repository managed by Smart TV businesses similar to IPTV businesses such as QOOK TV, B TV, and U⁺ TV. The second is the EPG (Electronic Programme Guide) repository managing the schedule information of TV program. The third is the global repository such as YouTube having plenty of UCCs.

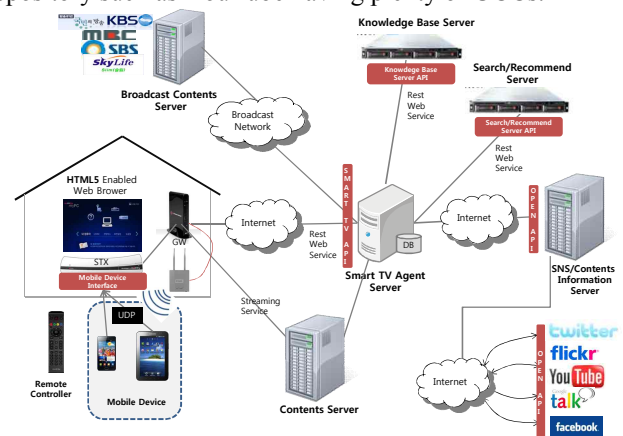


Figure 1 System Overview

So we designed the semantic network based integrated multi-level search system including three search engines for local repository, EPG repository, and global repository. The proposed system overview is illustrated in Figure 1. A user can interact with the Smart TV agent server through a HTML5 enabled web browser and the Smart TV is connected by Internet with back-end search systems including Smart TV agent server, knowledge base server, search/recommend server, contents server, and SNS/contents information server.

The Smart TV agent server performs a two-phase search process. In first phase, it analyses a set of user search keywords using semantic network then searches multimedia contents matching with the analysed keywords from content repositories. The details of the two-phase search process are followings:

Procedure: Phase 1

- The user search keywords entered through the web browser are sent to the Smart TV agent server.
- The Smart TV agent server delivers a set of user keywords to the knowledge base server.
- The knowledge base server identifies the semantic relationships between keywords and expands keywords list using knowledge objects having close relationship with it
- The knowledge base server sends the expanded keywords list to the Smart TV agent server.

Procedures: Phase 2

- The Smart TV agent server sends the expanded keywords list to a local repository search engine.
- The local repository search engine creates a SPQRQL query with the expanded keywords list and executes it.
- The local repository search engine returns the search results, the contents list being grouped, sorted, and classified, to the Smart TV agent server.
- The Smart TV agent server dispatches the expanded keywords list and the contents list to two other search engines, EPG repository search engine and global repository search engine.
- Each search engine generates a query to be appropriate to each repository with the contents list.
 - The EPG repository search engine generates MySQL query with contents list.
 - The global repository search engine generates queries for target web search engine (e.g. YouTube, Daum, Naver, etc.) with contents list.
- Two search engines return the search results being grouped, sorted, and classified to the Smart TV agent server.
- The Smart TV agent Server integrates the search results from each search engine and browses it to a user.

Figure 2 presents the architecture of the proposed system consists of several subsystems. As it shown, each subsystem has several functional software blocks and provides other

subsystems with various interfaces. The functional details of each subsystem are described in the followings.

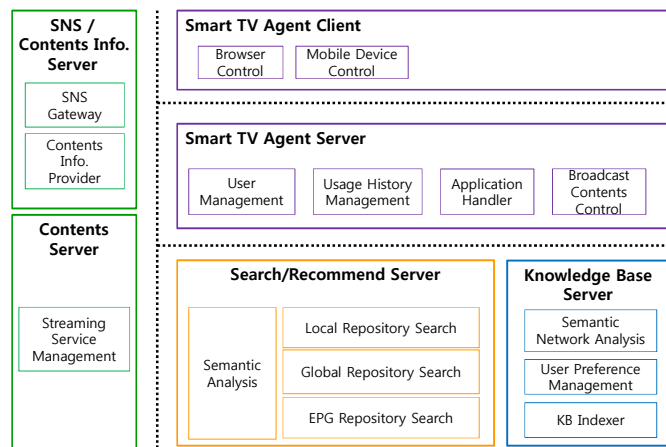


Figure 2 System Architecture

A. Smart TV Agent Server

The Smart TV agent server plays a major role in managing the information of user profile, watching history, and user behaviour pattern. It also controls the execution of Smart TV applications and broadcast contents. The functional blocks in the Smart TV agent server are as follows.

1) User Management

The Smart TV agent server manages user profile information which is created when a new user is added on TV. Some properties in user profile such as the favourite genre of movie, TV show, and music is stored as user preference data in the knowledge base server and used for grouping, sorting, and classifying the search results.

Table 1 User Profile

Property	Type	Meaning
seq_no	Integer	Sequential number
u_name	String	User name
u_id	String	User id
u_passwd	String	User password
u_sex	String	User sex
u_age	String	User age
u_job	String	User affiliation
u_status	String	User family status
f_movie	String[]	Favourite movie genre
f_show	String[]	Favourite TV program genre
f_music	String[]	Favourite music genre
s_id	String[]	User id for SNS
s_passwd	String[]	User password for SNS

2) User History Management

The Smart TV agent server gets the user's watching history from the Smart TV agent client and stores in a user history database. It performs the time-frequency analysis regularly to derive user preference data such as the favourite TV channel, the favourite TV show, and the favourite video from the user

history database. It provides a user with a personalized main page including the user preference data.

3) Broadcast Contents Control

The Smart TV agent server controls a TV tuner card and sets the DVR alarm by user request. It delivers input signal to a TV tuner card and changes the channel of terrestrial TV broadcasting. It also reserves the video recording in specific time by setting an alarm.

4) Application Handler

The Smart TV agent server takes charge in handling applications, especially a multimedia contents search application, so it manages the execution of search process from getting user input data to providing the search results to a user.

B. Search/Recommend Server

The search/recommend server provides semantic analysis of the user search keywords and searches multimedia contents matching with the user search keywords from local repository, EPG repository, and global repository. The functional blocks of the search/recommend server are as follows.

1) Semantic Analysis

The search/recommend server gets the user search keywords from the Smart TV agent server and then searches knowledge objects mapping with the search keywords in semantic network. It also searches the relationships between the user search keywords and expands keywords list using the knowledge objects having close relationship with it to improve search accuracy.

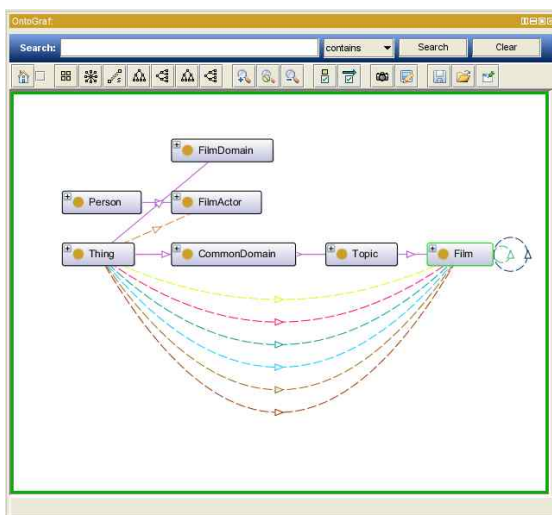


Figure 3 Relationships between Knowledge Objects

2) Local Repository Search

The search/recommend server searches multimedia contents matching with the expanded keywords list in the local repository. First of all it creates SPARQL query with the expanded keywords list and executes it. After searching, it arranges the search results by grouping, sorting, and

classifying according to user preference data stored in the knowledge base server.

3) EPG Repository Search

The search/recommend server searches the schedule information of TV program matching with the contents list as the result of local repository search. It creates MySQL query with the contents list to search target contents in EPG repository and executes it. Before it sends the search results to the Smart TV agent server, it arranges the search results by the broadcast timetable.

4) Global Repository Search

The search/recommend server searches UCC contents matching with the contents list from the local repository search through several web sites such as YouTube, Daum, Naver, etc. It generates queries with contents list and executes it in each web search engine. After searching, it filters the search results with the user preference information and then arranges it by subject.

C. Knowledge Base Server

The knowledge base server manages metadata of multimedia contents stored in a contents server as a form of knowledge objects and provides applications with various interfaces to access to knowledge objects. We chose five domains such as movie, soap opera, entertainment show, sport and music video to implement domain ontologies to be used by Smart TV applications. Figure 4 shows the examples of knowledge base. The functional blocks of the knowledge base server are as follows.

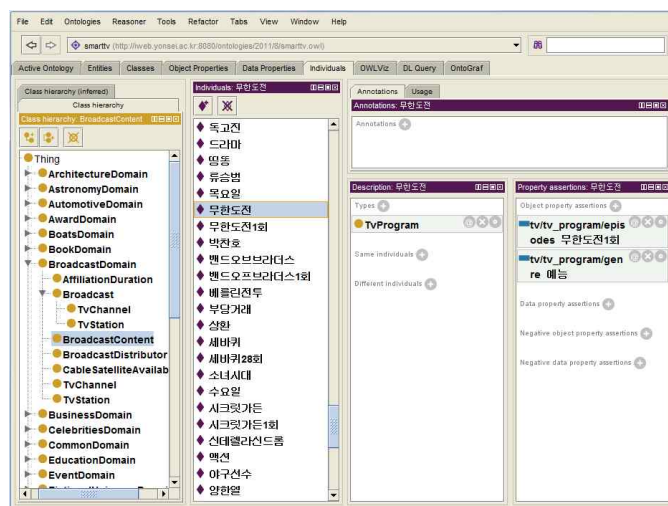


Figure 4 Examples of the Knowledge Base

1) Semantic Network Analysis

The knowledge base server provides Smart TV applications with various interfaces to access to semantic network. It provides interfaces to locate knowledge objects in semantic network matching with the user search keywords and to find a shortest path between two knowledge objects. It also provides interfaces to retrieve properties of given knowledge object and to navigate the semantic network using property information.

2) User Preference Management

The knowledge base server manages user preference data including favourite genre of movie, music, TV show, and favourite actor/actress to be used for grouping, sorting, and classifying the search results. Most of user preference knowledge comes from the user profile information and some of that could be achieved through analysis of user behaviour pattern from the Smart TV agent server.

3) KB indexer

The knowledge base server creates an index of ontology to ensure efficient search and updates it periodically to reflect changes of semantic network.

D. Smart TV Agent Client

The Smart TV agent client provides a user with an interactive browser and asks the Smart TV agent server for a new web page in response to user's input command. It also receives the signal of mobile devices using a UDP server and delivers it to the Smart TV agent server. The functional blocks of the Smart TV agent client are as follows.

1) Browser Control

The Smart TV agent client gets the user input data from a Smart TV browser and sends it to the Smart TV agent server. It gets an alarm message from the Smart TV agent server through WebSocket and browses it.

2) Mobile Device Control

The Smart TV agent client receives the user input data from a mobile device, which is a kind of smart device such as a smart phone or a tablet PC installed a remote control application, using a UDP server installed in a set-top box.

E. SNS/Contents Information Server

The SNS/Contents information server supports interconnection with the social network service, Twitter and Facebook, and manages the detail information of multimedia contents. The functional blocks of the SNS/contents information server are as follows.

1) SNS Gateway

The SNS/Contents information server connects with the social network service using user account information stored in the user profile. Users can share opinions about multimedia contents while watching a Smart TV.

2) Contents Information Provider

The SNS/Contents information server manages the detail information of a multimedia content like poster, synopsis, director, and actor/actress etc. The Smart TV agent server browses the detail information of the content user choose to watch.

F. Contents Server

The contents server manages multimedia contents with unique URL keeping consistency with the knowledge base server. It gets the URL of multimedia content from the Smart

TV agent server and controls the multimedia streaming service - play, stop, fast forward, and rewind.

IV. CONCLUSIONS

As the Internet technology has been growing up fast, the broadcast environment has been expanded to Internet. Therefore, the new broadcast environment needs the search technology to cover the expanded multimedia sources including broadcast network and Internet. We have designed the semantic network based integrated multi-level search system to improve search accuracy and expand search coverage. The proposed system can retrieve target multimedia contents what a user exactly wants from various sources using ambiguous keywords entered with the limited UI/UX on TV. We hope that the proposed system will foster the vision of user centred search system to understand user's intent from limited search keywords for beyond Smart TV.

ACKNOWLEDGMENT

This work was supported by the Broadcasting & Telecommunications R&D Program of KCC Korea [11921-03001, "Development of Beyond Smart TV Technology"].

REFERENCES

- [1] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Qianhuang Don, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P., "Query by image and video content: The QBIC system", IEEE Comput., (Sept.), 1995, pp. 23-32.
- [2] Bach, J. R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., and Shu, C. F., "Virage image search engine: An open framework for image management", In Proceedings of the SPIE Storage and Retrieval for Still Image and Video Databases, 1996, pp.76-87.
- [3] Smith, J. R. and Chang, S. F., "Visually searching the web for content", IEEE Multimedia 4, 3, 1997, pp.12-20.
- [4] Frankel, C., Swain, M. J., and Athitsos, V., "WebSeer: An image search engine for the World Wide Web", University of Chicago Tech. rep. 96-14, University of Chicago, Chicago, IL, 1996.
- [5] Bliujute, R., Saltinis, S., Slivinskas, G., and Jensen, C. S., "Developing a DataBlade for a new index", In Proceedings of IEEE International Conference on Data Engineering. (March) Sydney, Australia, 1996, pp. 314-323.
- [6] Hanjalic, A., Lagenduk, R. L., and Biemond, J., "A new method for key frame based video content representation", In ImageDatabases and Multimedia Search, A. Smeulders and R. Jain, Eds. World Scientific, 1997, pp.97-107.
- [7] Haas, M., Lew, M. S., and Huijsmans, D. P., "A new method for key frame based video content representation", In Image Databases and Multimedia Search. A. Smeulders and R. Jain, Eds. World Scientific, 1997, pp.191-200.
- [8] Lienhart, R., "Reliable transition detection in videos: A survey and practitioner's guide", Int. J. Image Graph. 1, 3, 2001, pp.469-486
- [9] Duygulu, P., Barnard, K., de Freitas, J.F.G., Forsyth, D.A., "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary", In: ECCV' 02: Proceedings of the 7th European Conference on Computer Vision-Part IV, London, UK, Springer-Verlag, 2002, pp.97-112.
- [10] Jeon, J., Lavrenko, V., Manmatha, R., "Automatic image annotation and retrieval using cross-media relevance models", In: SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, ACM Press, 2003, pp.119-126.
- [11] Lavrenko, V., Manmatha, R., Jeon, J., "A model for learning the semantics of pictures", In Thrun, S., Saul, L., Schölkopf, B., eds.: Advances in Neural Information Processing Systems 16. MIT Press, Cambridge, MA, 2004.

- [12] Berry, M.W., Dumais, S.T., O'Brien, G.W., "*Using linear algebra for intelligent information retrieval*", Technical Report UT-CS-94-270, University of Tennessee, 1994
- [13] Schreiber, A.T.G., Dubbeldam, B., Wielemaker, J., Wielinga, B., "*Ontology-based photo annotation*", IEEE Intelligent Systems 16, 2001, pp.66-74.
- [14] Hunter, J., "*Adding multimedia to the semantic web: Building an mpeg-7 ontology*", In Cruz, I.F., Decker, S., Euzenat, J., McGuinness, D.L., eds.: SWWS, 2001, pp.261-283
- [15] Tsinaraki, C., Polydoros, P., Moumoutzis, N., Christodoulakis, S., "*Coupling owl with mpeg-7 and tv-anytime for domain-specific multimedia information integration and retrieval*", In: Proceedings of RIAO 2004, Avignon, France, 2004
- [16] Kompatsiaris, I., Avrithis, Y., Hobson, P., Strinzis, M., "*Integrating knowledge, semantics and content for user-centred intelligent media services: the acemedia project*", In: Proceedings of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '04), Lisboa, Portugal, 2004
- [17] Richard Arndt, Raphael Troncy, Steffen Staab, Lynda Hardman, and Miroslav Vacura, "COMM: Designing a Well-Founded Multimedia Ontology for the Web", K. Aberer et al. (Eds): ISWC/ASWC 2007, LNCS 4825, 2007, pp. 30-43.
- [18] Michel S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain, "Content-Based Multimedia Information Retrieval: State of Art and Challenges", ACM Transactions on Multimedia Computing, Communications and Applications, Vol. 2, No. 1, 2006, pp. 1-19.
- [19] Jonathon S. Hare, Patrick A. S., Sinclair, Paul H. Lewis, Kirk Martinez, Peter G. B. Enser, and Christine J. Sandom, "Bridging the Semantic Gap in Multimedia Information Retrieval", 3rd European Semantic Web Conference, Budva, Montenegro, 2006.
- [20] Myung-Eun Kim, In-Chul Jeong, and Joon-Myun Cho, "The Technical Trends of Search Service for Smart TV", Electronics and Telecommunications Trends, vol. 26, no. 4, 2011.
- [21] Won-Keun Yang, Ayoung Cho, Weon-Geun Oh, and Dong-Seok Jeong, ETRI Journal, vol. 33, no. 4, Aug, 2011, pp. 589-599