International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:7, No:2, 2013

Research Topic Map Construction

Hei-Chia Wang and Che-Tsung Yang

Abstract**—**While the explosive increase in information published on the Web, researchers have to filter information when searching for conference related information. To make it easier for users to search related information, this paper uses Topic Maps and social information to implement ontology since ontology can provide the formalisms and knowledge structuring for comprehensive and transportable machine understanding that digital information requires. Besides enhancing information in Topic Maps, this paper proposes a method of constructing research Topic Maps considering social information. First, extract conference data from the web. Then extract conference topics and the relationships between them through the proposed method. Finally visualize it for users to search and browse. This paper uses ontology, containing abundant of knowledge hierarchy structure, to facilitate researchers getting useful search results. However, most previous ontology construction methods didn't take "people" into account. So this paper also analyzes the social information which helps researchers find the possibilities of cooperation/combination as well as associations between research topics, and tries to offer better results.

Keywords**—**Ontology, topic maps, social information, co-authorship.

I. Introduction

KNOWLEDGE management has been an important issue recently since the past decade has seen extensive research on knowledge and knowledge management [1]. In 1999, Tim Berners-Lee & Fischetti used the concept of Semantic Web to achieve the goal of solving the problem of users spending a lot of time browsing and searching for required information. Ontology and intelligent agents constitute the basic framework of Semantic Web [2], while ontology serves as a backbone of the Semantic Web by providing vocabularies and formal conceptualization to facilitate information on sharing and exchange [3]. Researcher can obtained may inforamtion through constructed ontologies.

Although there are many resources available for researchers to retrieve on the internet (such as SDOS, ISI, DBworld, and ConfSearch etc.), key contents mainly come from journals and books. Moreover, some websites use the way of keyword search which is designed for focusing on titles or the keywords conference have determined rather than the content of conference articles. There are no detailed or unified classifications for conferences and possible relationships between classifications are not considered either. It leads to incomplete search results. Researchers need to spend time filtering the information they need and even some useful information is uncovered. Therefore, researchers need a suitable platform to assist them.

In existing conference resources such as AllConference.com (http://www.allconferences.com/Search/) and Conference Alerts (http://www.conferencealerts.com/index.htm) only offer simple conference knowledge classification or no classification, even with classification, they do not take conference related domain knowledge into account, not to say the possible relationships between classifications. As a result, if researchers want to know about what topics have been discussed in the conference or what conferences are related to the specific or related topic, they have to spend lots of time browsing related information. And if researchers want to know about the topics cooperation information (for example, what topics has been combined or cooperated with Twitter) or search for the authors of specific topic to cooperate, it is not easy to fine those information in the original search websites.

As mentioned before, simply using keywords to search is not enough for researchers. Conference knowledge (such as keywords, articles and websites etc.) has to be proper analyzed and classified. Through building complete conceptualized architecture so it can help researchers find the information they really need and improve the accuracy of search results. Since ontology plays an more and more important role in text analysis and information exchange between different areas [3], this conceptualized knowledge has been an important tool of search, classification, questions and answers, personalization, and those semantic-oriented technologies and applications [4].

Thus, the purpose of this paper is by incorporating conference information (i.e. title, keywords and abstract) as sources and considering social information to automatically construct research Topic Maps. And by offering relationships between topics, this paper hopes to offer researchers more complete references in conference information.

II. Literaure and Background

A. Ontology and Topic Maps

The word ontology is derived from the field of philosophy, as an important specification, it has great potential to help organizing and managing knowledge [3]. Ontology can be used to present knowledge hiding in a mount of data, it can integrate domain knowledge and demonstrate its consistency. When ontology is constructed, we can understand the relationship between the concepts. In recent years, ontology has been playing an more and more important role in text analysis and information exchange between different areas [3], the topic of ontology construction has been receiving growing attention [5].

Hei Chia Wang is the Professor in Institue of Information Management of National Cheng Kung University, Taiwan (phone: 886-6-2757575; fax: 886-6-2362162; e-mail: hcwang@mail.nckue.du.tw).

Che-Tsung Yang is a Ph.D. student in Institue of Information Management of National Cheng Kung University, Taiwan (phone: 886-6-2757575; fax: 886-6-2362162; e-mail: jason.c.yang@gmail.com).

This work is supported by the National Science Concil of Taiwan under Grant number NSC 101-2410-H-006-011-MY2 & NSC 98-2410-H-006 -023 –MY3 and Ministry of Economic Affairs (MOEA) of Taiwan, under Grant number MOEA 101-EC-17-A-05-S1-192.

Topic Maps is used as a formal syntax for representing and implementing ontologies, it can be used for finding information and suitable for non-scientific areas [6].Topic Maps is a technology for encoding knowledge and connecting this encoded knowledge to relevant information resources [7]. Topic Maps was initially used to represents the structure of indexes in the back of the books so that variety of resources can be integrated. It passed the ISO certification (ISO/IEC 13250) in 1999, after that Topic Maps Organization (TopicMaps.Org) develop Extensible Markup Language (XML) syntax of Topic Maps called XML Topic Maps (XTM). Nowadays, studies about Topic Maps are been developing rapidly.

B. Topic Maps Model

Topic Maps is composed of three elements, Topics, Associations, and Occurrences, TAO for short. Topics are subjects in knowledge domain, associations are relationships between these subjects, while occurrences connect subjects and related information resources [7]. Detailed descriptions of TAO are as follows [8]:

1. Topics: In general, a topic can be any "thing" – a person, an entity, a concept – regardless f whether it exists or has any other specific characteristics. For example, in the context of a dictionary of opera, topics might represent subjects such as "Tosca", "Madame Butterfly", "Rome", "Italy", the composer "Giacomo Puccini", or his birthplace, "Lucca".
2. Associations: A topic association asserts a relationship between two or more topics. Follow the example mentioned before, Tosca was "written by" Puccini, Tosca "takes place" in Rome, Puccini "was born" in Lucca, Lucca "is in" Italy, etc.
3. Occurrences: A topic may be linked to one or more information resources that are deemed to be relevant to the topic in some way. Such resources are called occurrences of the topic. An occurrence could be an article about the topic in an encyclopedia, a picture, a video, commentary or any of a host of other forms in which an information resource might have some relevance to the subject. Occurrences are generally external to the topic map document itself but they may also be inside it.

C. Construction Methods of Topic Maps

The ways of constructing Topic Maps vary considerably, most of the methods use knowledge built in advance to construct Topic Maps. So far, the methods can be roughly classified as follows:

1. Mapping from RDF to Topic Map: RDF is a resource description framework, describing relationships among objects(resources). Such methods can extract knowledge expressed in RDF from different data source (i.e. relational database, websites, enterprise information systems, etc.), and convert to Topic Maps automatically.
2. Generating Topic Maps from XML documents: This kind of method use XSLT (Extensible Stylesheet Language Transformations) algorithm to convert XML documents to Topic Maps. XSLT is a language transferring XML documents while T in XSLT stands for Transformation.
3. Collaborative approach: The approaches are based on construction of Topic Maps with many users, mainly used in E-learning environment.
4. Topic Maps Learning: This kind of methods tries to extract meaningful knowledge and construct Topic Maps from websites.
5. Merging Topic Maps: The idea of "merge" was defined in the label of XTM originally. The thought came from each Topic Maps coming from different sources, when merging two Topic Maps, Topic Maps with the same subject should be merged. So two Topic Maps coming from different sources can be merged, creating bigger Topic Maps.

D. Classification Resource—Open Directory Project(ODP)

Semantic knowledge base is widely used (such as WordNet) to find the meaning, related terms, and other information. But the information is usually too general to extract specific domain associations. For example, we cannot find "Semantic Web" or "Ontology" in WordNet. However, this paper uses corpus in the field of computer science, thus general semantic knowledge base is less suitable. Needing a specific domain knowledge background to help to extract taxonomy associations, this paper selects ODP to assist professional domain associations extraction..

ODP is an open category, constructed and maintained manually. It has 16 main categories and numerous subcategories [9], total containing more than one million categories.

E. The Application of Topic Maps

The basic element of Topic maps, topic, can contain many occurrences and can connect another topics by association. The ability that Topic Maps can encode complicated knowledge structure and link to information source means it plays an important role in knowledge management field.

So far there are many studies applying Topic Maps, [10] used social classification to construct enterprise knowledge map. Topic Maps can be used to render the contents of an enterprise knowledge, such as roles, products, procedures, and link to corresponding document. All in all, the Topic Maps standard offer a flexible infrastructure for variety of information and knowledge management applications.

III. Implementation and Evaluation

The implementation of this paper is built and operated in the operating system of Ubuntu10.04 and Win7 64 bit, using Perl5.10.1 as programming language and MySQL5.1.6 as the database. The hierarchy classification source, ODP, this paper uses contains variety and numerous classes, so this paper selects 16 main categories and its subcategories which are more

related to information science. They are business, computers, reference, and science separately.

This paper uses the proposed method to extract the elements of Topic Maps and visualize. In the system, each node is a topic which may represent conference name, conference topic, ODP class and any information this paper extracted. Different categories will be shown in different colors on the graph. The top left block is the description, containing topic description and other relevant information. The bottom left block is represented in table form. Users can search in the block according to different topic types such as conference name or topic. Users can also search for any association like RT or is the author of. Double clicking then Topic Maps will be shown in the right block. The tools above the Topic Maps can zoom in or zoom out as well as the changing the spacing between topics.

IV. Conclusion and Future Work

At present, many countries have begun to attach importance to study the effectiveness and the promotion of multidisciplinary and Knowledge Base Integration, interdisciplinary research cooperation in science and social sciences are becoming more common, while the rapid growth of the amount of information on the network allowed the researchers in the search conference information is often subject to screening in order to get the information you want, and the general seminar search sites are less easy to get the research topics can be combined or combination of over and find the theme of cooperation information.

The topic map based on the knowledge index, allows users to identify the knowledge needed by the relationship between topics and guidelines. It combines the advantages of traditional indexing, search engines and artificial intelligence and other fields, can be an effective organization of knowledge in order to facilitate exploration, reasoning, and then solve the problems brought about by the large number of disordered information.

References

[1] Ma Z. Z., & Yu, K. H. “Research paradigms of contemporary knowledge management studies: 1998-2007,” *Journal of Knowledge Management,* (14:2), 2010, pp. 175-189.

[2] Weng, S. S., Tsai, H. J., Liu, S. C., & Hsu, C. H. “Ontology construction for information classification,” *Expert Systems with Applications*(31:1), 2006, pp. 1-12.

[3] Jiang, S. Q., Du, J., Huang, Q. M., Huang, T. J., & Gao, W. “Visual ontology construction for digitized art image retrieval,” *Journal of Computer Science and Technology*(20:6), 2005, pp. 855-860.

[4] Jung, Y., Ryu, J., Kim, K. M., & Myaeng, S. H. “Automatic construction of a large-scale situation ontology by mining how-to instructions from the web,” *Web Semantics: Science, Services and Agents on the World Wide Web*(8:2-3), 2010, pp. 110-124.

[5] Santoso, H. A., Haw, S. C., & Abdul-Mehdi, Z. T. “Ontology extraction from relational database: Concept hierarchy as background knowledge,” Knowledge-Based Systems(24:3), 2011, pp. 457-464.

[6] Yi, M. “Information organization and retrieval using a Topic Maps-based ontology: Results of a task-based evaluation,” Journal of the American Society for Information Science and Technology(59:12), 2008, pp. 1898-1911.

[7] Kim, J. M., Shin, H., & Kim, H. J. “Schema and constraints-based matching and merging of Topic Maps,” *Information Processing & Management*(43:4), 2007, pp. 930-945.

[8] Pepper, S. “The TAO of Topic Maps,” 2002 (available online at http://www.ontopia.net/topicmaps/materials/tao.html#d0e632)

[9] Jiang, X., & Tan, A. H. “Learning and inferencing in user ontology for personalized Semantic Web search,” *Information Sciences*(179:16), 2009, pp. 2794-2808.

[10] Liu, L., Li, J., & Lv, C. G. “A method for enterprise knowledge map construction based on social classification,” *Systems Research and Behavioral Science,* (26:2), 2009, pp. 143-153.