How Knowledge Science is Studied --- a vision from the KSS'2006

Xijin TANG   Zheng WANG

Institute of Systems Science, Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100080 P. R. China
{xjtang, wzheng}@amss.ac.cn

Abstract

As knowledge engineering, knowledge management and even knowledge creation have been widely accepted in academics, business and industry, knowledge science is still regarded as a new area and has been approaching from different disciplines. In this paper, we try to draw a vision how knowledge science is studied at present based on information of all accepted papers to the 7th International Symposium on Knowledge and Systems Sciences (KSS'2006). A keyword network and an author network are constructed and then characteristics of both networks are analyzed, such as cutpoint, community structure, etc. from which to acquire some kind of emerged explanations on knowledge science studies, especially under systemic perspectives. Those information detected by network analysis could be pushed to conference organizers and participants for activities facilitation and to the other interested people to acquire basic concepts or understanding of the concerned disciplines or topics, such as what are the major research topics? Who are the principal investigators? How about the major special interesting groups? etc. Such kind of work could be regarded as conference mining.

Keywords: knowledge science, idea map, human net, conference mining

1 Introduction

As the world declared entering into an era of knowledge economy, the significance of developing knowledge has grown to a level where it is coming to dominate other socio-economic factors. The recent developments challenge people to understand the nature of knowledge and its role in applications, to effectively utilize the knowledge for improving the corporate competitive advantage and national comprehensive power. In reality, knowledge engineering, knowledge management and even knowledge creation have been widely accepted in academics, business and industry. People try to approach knowledge studies from different disciplines and establish a new discipline - knowledge science. The School of Knowledge Science at Japan Advanced Institute of Science and Technology (JAIST) founded in the late 1990s is the world's first research and education institute established under the theme of knowledge. At the web site of the school, some explanations about knowledge science are given, which reflects their explorations of this discipline from social sciences, humanities, engineering and the natural sciences. Those four disciplines could be regarded as the basic constructs among knowledge science education and research. People, especially new comers may prefer more detailed information.

In 2000, the First International Symposium on Knowledge and Systems Sciences was held at JAIST and extended those endeavors of confluence of different ideas and opinions, methods and technologies, etc. to a wider scope of participants from different schools and disciplines, theorists and practitioners, who aim to develop knowledge science from systemic perspective. After 3-year work, the International Society for Knowledge and Systems Sciences (ISKSS) was founded in Guangzhou of China during KSS'2003 by the emerging community dedicated into this goal, and then as a forum for researchers as well as practitioners to exchange innovative ideas and to be aware of each other's efforts and results in the exploration of knowledge science. Then another three year passed, the 7th International Symposium on Knowledge and Systems Sciences was held in Beijing during September 22-25, 2006 to show some new achievements and prospects for continuous thinking and studying under the theme "towards knowledge synthesis and creation".

Each participant may have their own understanding and impressions toward those activities held during KSS'2006. As the local organizer, we try to provide more helpful information about the hot topics of KSS'2006 not only to the participants, but also to those interested people who were not at present. Based on all accepted submissions to KSS'2006, we try some new ways to expose the vision of knowledge science study. A keyword network and an author network are constructed and social network analysis methods are applied to detect the main topics discussed in KSS'2006. Next the network construction and analysis methods are addressed and then applied to the analysis of all KSS'2006 accepted papers. Some further analyses are also taken to all past KSS series symposia. At last, concluding remarks are given.

2 Knowledge Vision by Idea Map and Human Net

If we want to know the meaning of some terms, a search engine such as google is a convenient way. We can also read wikipedia to get more detailed information. Many specialized search engine for the retrieval of journal articles relevant to a dedicated scientific database with narrow focus have been developed for specialists. Even with specialized supporting tools, researchers still need much effort to acquire a rough vision of the concerned fields.

We consider such a problem as an unstructured problem. Fro example, if a student wants to know what knowledge science is, google can provide many urls. He may then browse some web sites of education institutions, such as JAIST, to get the answer. Those urls also include web sites of conference about knowledge research. The scopes and topics, the titles of keynote speeches and final programs may bring more information. All those depend on readers' active search and understanding. More augmented information supports are barely required to help a new comer acquire some basic threads or constructs during his unstructured problem solving process.

Here we concentrate on how to detect more information about the concerned topics from those accepted submissions from the dedicated academic conferences. A database of all submission is consisted of a set of paper records with the structure as

<topic, authorList, paper title, keywords, time>

Such a record indicates the corresponding *author(s)* submit(s) one *paper* with a set of *keywords* under the *topic* at the point of *time*. The keywords in a paper can denote as a topic, a problem, a method or algorithm, a practical case, etc. The keyword set of one paper could then be understood as the basic ideas toward the problem addressed by the authors. If we count those keywords of all accepted papers of one conference, keywords with highest frequency could be regarded as the popular terms. To acquire a comprehensive understanding of those individual keywords, a network is constructed. Next we define two types of networks, one is keyword network, another is human network.

In a keyword network $G=(K,E)$, the vertex refers to a keyword. If keyword k_i and keyword k_j simultaneously belong to the keyword set of one paper, then an edge exists between two vertexes $e_{ij}=(k_i,k_j), i\neq j$, $e_{ij}\in E$ (E is the edge set). Then each keyword set of one paper constructs a complete keyword graph. The keyword network denotes the aggregation of all keyword graphs. If $G_l=(K_l,E_l)$ indicates the keyword graph of the *l*th paper, $K_l=\{k_1^l,k_2^l,\cdots,k_n^l\}$ is the keyword set of the *l*th record, E_l is the edge set, then $G=(K,E)$, $K=\cup K_l$ = $\cup\{k_1^l,k_2^l,\cdots,k_n^l\}$, $E=\cup E_l=\cup\{e_{ij}\}$, $i,j=1,2,\ldots,m; i\neq j$. This topological map is a weighted undirected network where the weight of edge refers to the frequency of co-occurrence of keywords among all papers and is referred as an idea map contributed by all authors. Given such a network, more senses may be acquired by a variety of network analysis by detecting some features of the idea map, such as cutpoints, centrality of keywords, clustering of keyword, etc. which expose different perspectives of the authors' knowledge scope.

For example, a cutpoint (articulation point) of a graph is a vertex whose removal increases the number of connected component [1]; then the cutpoint keyword in the idea network may reveal the real key ideas (terms). So does the centrality analysis of the keyword vertex. With the keyword clustering by community structure detection, it is then easier for people to understand the major topics from those keyword clusters instead only by individual keywords.

Co-authorship network is a typical network which reflects common interests shared among co-authors, and is usually used to detect interest groups and influential scientists of one discipline. However, general academic conference prefers participation with diverse interest and the majority of participants may not contribute multiple papers. A co-authorship network of a conference may include many isolates and components and then it is difficult to detect practical interest groups. Here keywords-sharing between authors is considered and an author network is constructed. In this network the vertex refers to an author. If two authors share one keyword, then there is link between them. The strength between two authors indicates the frequency of the keywords they share. Obviously, co-authorship network is a sub graph of the keyword-sharing network. From such a human net, social network analysis (SNA) is applied to check the powerful persons by centrality analysis and detect the interest groups by community clustering, etc.

Both idea map and human net constructed based on accepted papers for one conference could be regarded as one kind of structure about the dedicated disciplines. Network analysis aims to detect basic concepts and main topics, principal investigators and the major special interesting groups emerged from the accepted submissions. Those information could be regarded as constructs of the dedicated disciplines and more helpful for the curious students to quickly acquire a rough vision of the interesting discipline.

Next, we construct a rough vision of knowledge science study by those accepted submissions to KSS'2006.

3 Knowledge Science Study by vision of KSS'2006

As the paper review finished, a total of 49 accepted submissions including 179 keywords contributed by 86 authors was collected into proceedings. As a new discipline, knowledge science may somewhat be regarded as an unstructured problem and be accessed from different perspectives. Idea map and human net of KSS'2006 are constructed for perspective detection.

3.1 Idea View of KSS'2006

Before network generation, some preprocessing had to be done. A stop list is applied to deal with synonymous terms. For example, both "complex network" and "complex directed network" refer to "complex network", and then the latter is replaced by the former. Both "scale-free networks" and "scale-free property" are replaced by "scale-free". Finally only 166 keywords remained. Figure 1 is the keyword network of KSS'2006 generated by Ucinet.

Based on Newman-Girvan algorithm for community clustering [2], 26 subgroups are founded (the maximum value of the modularity function Q = 0.901). The largest component (in the middle of Figure 1) including 49 keywords is split into 3 subgroups. The subgroup (vertex in squares) in the central is a group of keywords about the mechanism of knowledge creation, the downside subgroup (vertex in circles) is a cluster of keywords mainly about computerized support, and the upside subgroup (vertex is represented two triangles) is a cluster of keywords about knowledge management and its practice. The major topics of knowledge science concerned by KSS'2006 can be detected by such a way. Actually there are three parallel sessions about those three topics in KSS'2006. The community structure analysis about a keyword network provides some hints about parallel session assignment for conference organizers.

Next it is interesting to know who are engaged into the major topics.

3.2 Interest Group of KSS'2006

Figure 2 is keyword-sharing network of KSS'2006 (human net). Excluding four isolates, there are 9 components. Tracing those keywords shared by authors in each component, we find some groups of different research topics as declared during call for papers, such as one component labeled as "cognitive complexity", one component labeled as "scale-free", etc. Labeled as "CORE", the biggest component includes 44 authors who are grouped into 6 communities by Newman-Girvan algorithm as shown in Figure 3.

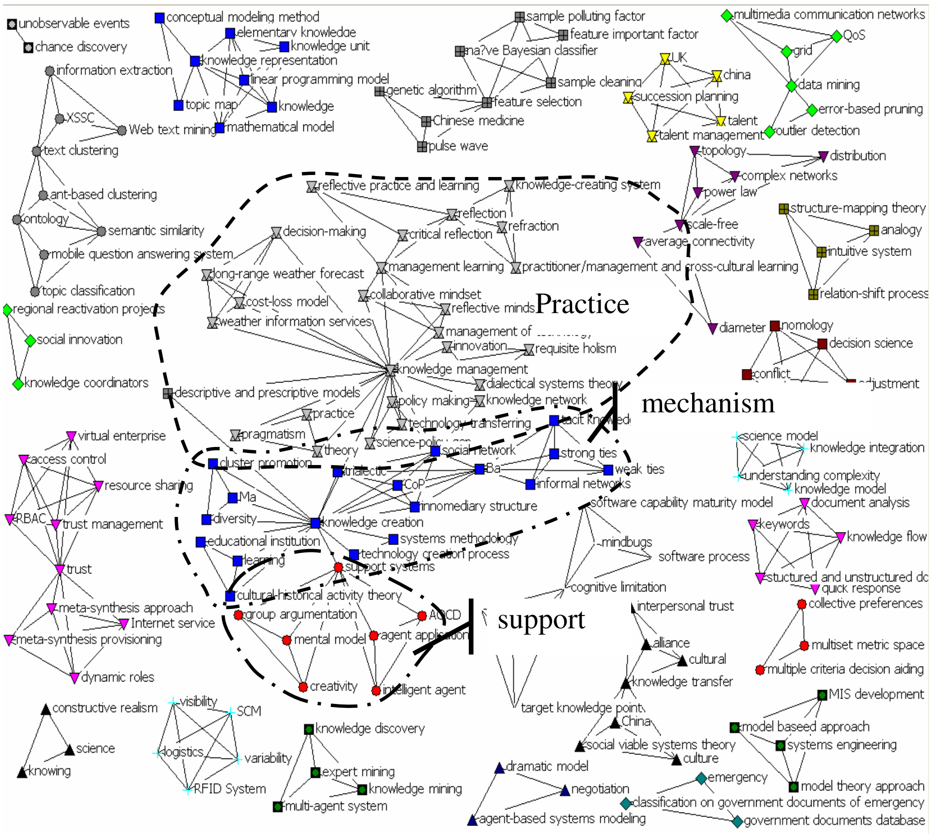

Figure 1. Idea map of KSS'2006 (keyword network)

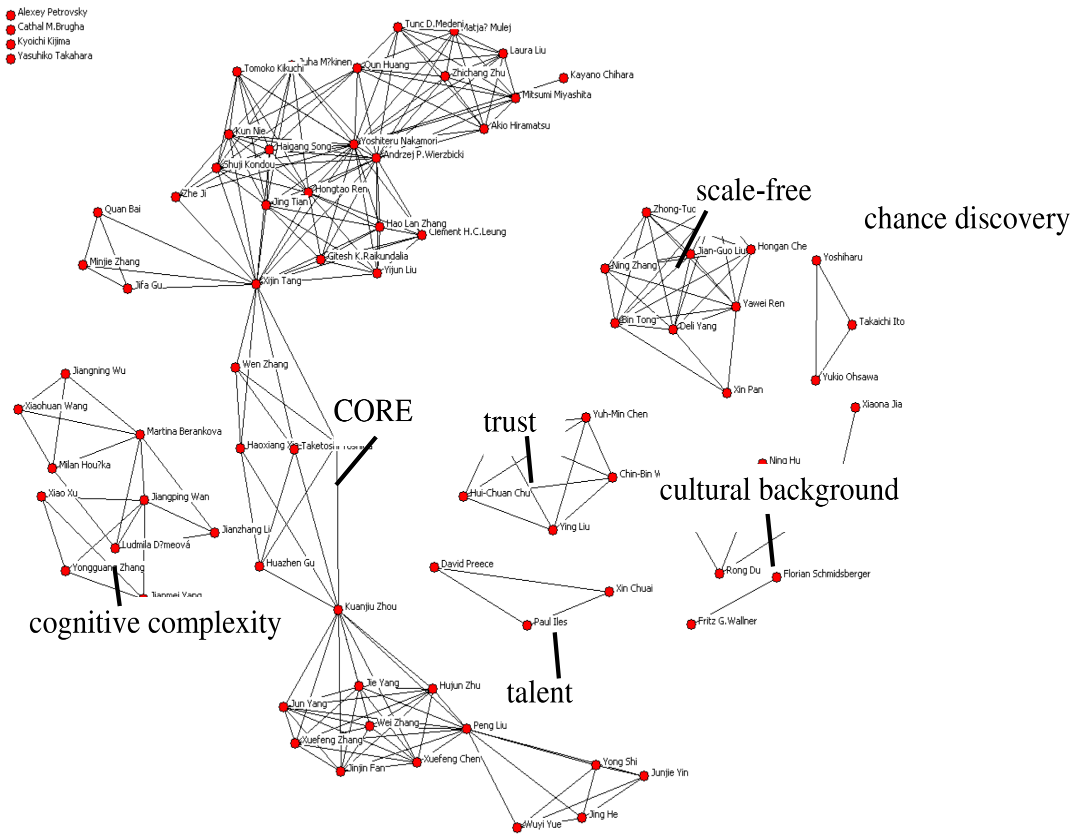

Figure 2. Human net of KSS'2006 (keyword-sharing network)

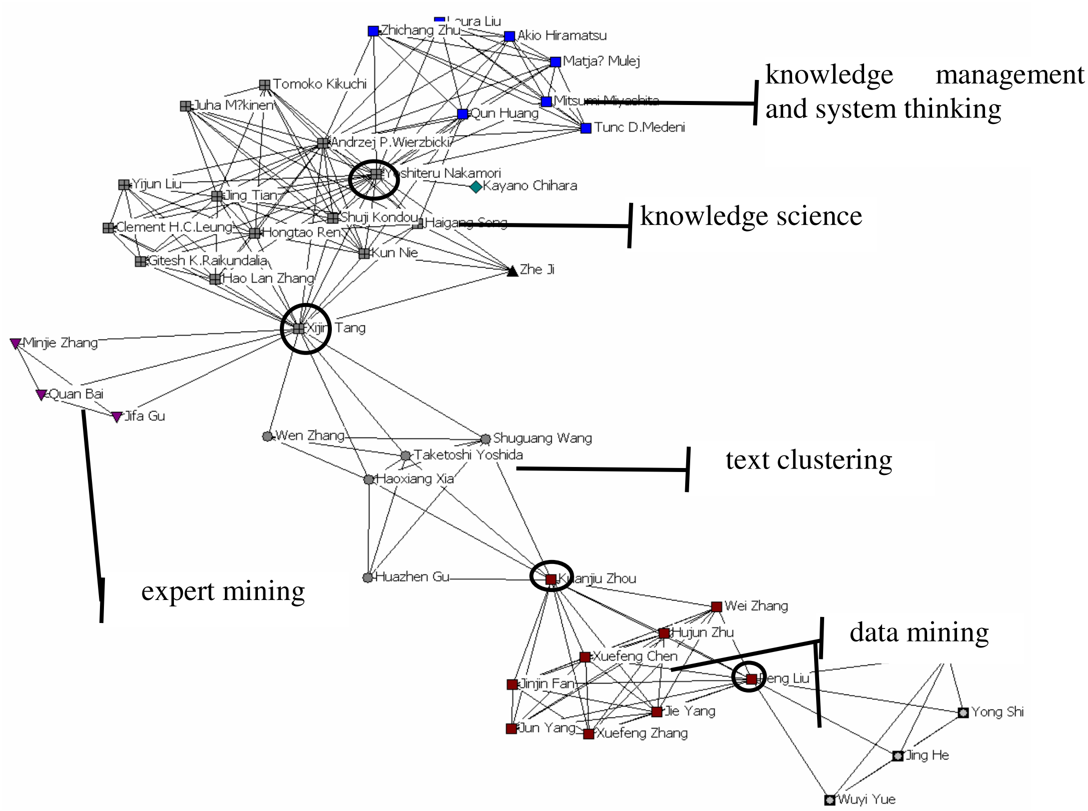

Figure 3. Community structure of the "CORE" component of KSS'2006 (Q = 0.671)

Tracing back those shared keywords by each community, we recognize the research foci of those 6 communities as "knowledge management and system thinking", "knowledge science", "expert mining", "text clustering" and "data mining". Taking the cluster on knowledge science as a central, then the upper-located cluster in Figure 3 represents group of people on knowledge practice by systemic view, and the communities at the lower side are interested in different technologies intensively applied in knowledge science. Furthermore, 4 cutpoints (*Yoshiteru Nakamori, Xijin Tang, Kuanjiu Zhou* and *Peng Liu*) had been detected. If those cutpoints are deleted, the component will be broken into small groups. Considering the affiliations those authors belong, the collaborations between JAIST, Institute of Systems Sciences, Chinese Academy of Science, Dalian University of Technology, University of Hull, University of Wollongong may be easily understood, which also demonstrated interdisciplinary studies among knowledge science research, especially those keywords referred by those cutpoint authors are considerred. For example, the keyword set referred by Yoshiteru Nakamori is {*knowledge creation (3 times), descriptive and prescriptive models, knowledge management (3 times), technology creation process, system methodology, support system (twice), weather information services, long-range weather forecast, decision-making, cost-loss model, informal networks, tacit knowledge, strong ties, weak ties, Ba, science-policy gap, technology transferring, policy making, regional reactivations projects, knowledge coordinators, social innovation*}; Xijin Tang's keyword set is {*creativity, mental model, support system, group argumentation, knowledge discovery, knowledge mining, expert mining, multi-agent system, information extraction, web text mining, web text summarization, text clustering, XSSC, informal networks, tacit knowledge, strong ties, weak ties, Ba*}. The conjunction of both keyword sets {*informal networks, tacit knowledge, strong ties, weak ties, Ba*} reflects the collaboration between two interest groups on knowledge creation mechanism research.

By both keyword network and author network of KSS'2006, it could be easily understood that those participants expressed their understanding about knowledge science mainly from the mechanism of knowledge creation, how to support knowledge creation, how to apply knowledge in practice (knowledge management) and by what kind of approaches (systemic approaches). Who conducts those research and the principal investigators are also indicated. The network also shows the collaboration between those research groups.

If constructing a keyword network where a pair of keywords connected if they are referred by one author, instead of by one paper, we can also acquire more intensive clusters of keywords. Such a way may be more efficient for analysis if the conference scale is larger.

Above gives network analysis results based on KSS'2006. Due to conference location and difference in organizing, the analysis of one conference is of limitation. Actually, the idea map of human net of one conference only reflects how those participants explain their understanding toward knowledge science study at that year. For more comprehensive scenario of knowledge study, it is necessary to look back all past KSS symposia.

4 Analysis of series KSS Symposia

Table 1 summarizes basic information of all past KSS symposia (2000-2006). It could be seen that KSS is rather a small-scale symposium.

Table 1. Basic Information of KSS (2000-2006)

Year	Paper #	Author #	Keyword #	New Key--word #
2000	37	54	120	120
2001	45	77	142	130
2002	56	104	179	144
2003	57	108	193	149
2004	82	136	254	204
2005	28	64	111	76
2006	49	86	166	111

Due to space limitation, here only part results are given. Centrality of keywords at the individual keyword networks of each year is analyzed and top 10 are listed in Table 2. In 2005, there are only four keywords whose betweenness value is greater than zero.

Table 2. Betweenness of keywords

Year	Keywords whose betweenness are among top 10
2000	complexity (1317.17), system engineering (663.00), culture (517.50), knowledge management (460.00), knowledge (367.80), knowledge creation (261.30), information(235.00), knowledge conversion (183.00), WSR (179.60), chaos (124.00), decision making (124.00), information system (124.00)
2001	knowledge management (364.80), knowledge (210.80), knowledge creation (195.80), soft system methodology (134.20), knowledge representation(75.00), system thinking (75.70), neural network (66.00), genetic algorithm (42.00), fuzzy rules (30.00), agent based simulation (27.00)
2002	evaluation (851.00), knowledge management (820.00), knowledge (704.00), knowledge economy (683.00), methodology (473.00), knowledge value (338.00), meta-synthesis (333.50), knowmetrics (234.00), innovation (104.00), knowledge acquisition (63.00)
2003	complexity (245.00), complex system (161.00), knowledge management (133.00), e-government (69.00), chaos (69.00), strategy (48.00), optimization (28.00), input-output analysis (24.00), organizational learning (17.00), WSR (16.00)
2004	knowledge creation (1885.30), data mining (1586.50), text mining (1032.00), clustering (712.00), knowledge management (615.00), i-system (1525.00), knowledge discovery (485.50), preference (456.00), interaction (389.00), association rule (357.50)

2005	decision support (24.00), simulation (21.00), ontology(6.00), uncertainty(4.00)
2006	knowledge management (798.20), knowledge creation (699.80), support system (273.00), management learning (264.00), ba (184.00), reflection (141.00), social network (126.70), trust (63.00), knowledge transfer (33.00), scale-free (12.00)

From Table 2, we can see how central keywords shift along the time. It could be seen that the central terms, such as knowledge, knowledge management, knowledge creation, system methodologies are always among the top 10 keyword set of each year, which reflect the lasting theme of series KSS symposia. While the conference location, local hosts and their organizing strategies together with the special interest groups detected from human network of each year could also affect the transition of keywords.

5 Concluding Remarks

In this paper, we try to draw a vision how knowledge science is studied based on organizing work and the accepted papers of the 7th International Symposium on Knowledge and Systems Sciences (KSS'2006). A keyword network and an author network are constructed to depict the main streams or interest groups on knowledge science studies, especially under systemic perspectives.

By P. Thagard's view, scientific knowledge growth consists of the psychological processes of discovery and acceptance, the physical processes involving instruments and experiments, and the social processes of collaboration, communication, and consensus that brought about transformations in knowledge [3], conference mining shown here aims to understand more about both psychological and social process during knowledge growing process.

In-depth analysis is helpful to expose more information such as the vision of knowledge science studies. For example, how knowledge is created (knowledge dynamics), how to support the knowledge dynamics and by what kind of technologies, and how to achieve those theories in practice are the 4 major streams of knowledge science studies acquired in KSS'2006. It is expected to pay attention to those not hot topics in KSS'2006 for a wider understanding of knowledge studies. For more comprehensive vision of knowledge studies, all past KSS paper data have to be analyzed. Lots of work remains further exploration.

Many conference assistant systems have been explored since late 1980s to enrich participants' experiences by advanced information technologies and appliances [4,5]. The in-depth analysis of KSS'2006 papers together with general paper submission and reviewing work is integrated into a so-called on-line conferencing ba (OLCB) and those visualized networks are posted at OLCB discussion area for academic exchanges and idea generation [6].

Our current work is still at a very initial stage at both research and practice. Here only shows very basic analysis. It is also worth adopting or comparing with others' ideas [7].

Acknowledgment

The authors are grateful to all participants and organizing group of the 7th International Symposium on Knowledge and Systems Science for their collaboration this study, which is supported by Natural Sciences Foundation of China under Grant No. 70571078 & 70221001.

References

[1] Harary, F. *Graph Theory*, Addison-Wesley, 1970. Chapter 3.

[2] Newman, M E J, Girvan, M.: Finding and Evaluating Community Structure in Networks. *Physical Review E*, 69:026113, 2004.

[3] Thagard, P.: *How Scientists Explain Disease*, Princeton University Press, Princeton, New Jersey, 1999.

[4] Gitta, B. S.: Designing casual-use hypertext: The CHI'89 InfoBooth. In: Carrasco, J., Whiteside, J. (eds.): Proceedings of the *ACM CHI 90 Human Factors in Computing Systems Conference*, Seattle, USA, 1990, 451-458.

[5] Sumi, Y., Mase, K.: Conference Assistant System for Supporting Knowledge Sharing in Academic Communities. *Interacting with Computers*, 14(6): 713-737, 2002.

[6] Tang, X. J., Zhang, N., Wang, Z. Augmented Support for Knowledge Sharing by Academic Conferences - On-line Conferencing Ba. In Proceedings of the International Symposium on *Information Systems & Management* (ISM'2007, the Management Track of IEEE WiCOM 2007), Shanghai, July 26-28, 2007.

[7] Matsuo, Y., Ohsawa Y. and Ishizuka, M. "KeyWorld: Extracting Keywords from a Document as a Small World". In: Jantke, K.P. and Shinohara, A. eds.: *Discovery Science* (Proc. of 4th International Conference DS 2001), LNAI 2226, Springer-Verlag, Berlin Heidelberg, 2001, pp271-281.