http://ansinet.com/itj



ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL



Asian Network for Scientific Information 308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Exploring the Academic Collaboration Chance in Open Research Community

Bin Xu

College of Computer and Information Engineering, Zhejiang Gongshang University, Hangzhou Zhejiang, 310018 China

Abstract: Due to the advance of information technology and the widely-used social network on Internet, research is being done in an open environment. While the researchers may share the published research achievements in paper format, it's essential for the research groups to find the potential academic collaboration chance between them and the large scale academic social network. This article is to present our solution to find the chance of academic collaboration between different research groups in such an open research community. A prototype has been made and the core algorithms and an experiment with main interfaces are introduced.

Key words: Social network, research interesting similarity, collaboration potential exploring, collaborative innovation

INTRODUCTION

Facilitated by the advance of information technology and the widely used internet (Memon et al., 2007), research is being done in an open environment which we name it "research community" in this article. While, blog (Lin et al., 2007), Wikipedia, E-Learning systems (Fiaidhi and Mohammed, 2003; Jayanthi et al., 2007) and professional search engines (Curran, 2004; Bal and Nath, 2010) provide plenty information for could research. researchers share achievements together thev look forward for efficient collaborative innovation between them (Wu and Zeng, 2009).

Collaborative innovation begins when universities and industries come together to solve problems and/or develop customer-centric solutions that are beyond the scope, scale or capabilities of the individual universities or companies. Collaborative innovation has been paid great attention by different domains during these year, people establish the networks and centers for collaborative innovation (Gloor, 2006; Venkatesh, 2006; Mcdaniel, 2003; Klein et al., 2006).

Gloor (2006) introduced "Swarm Creativity", a Collaborative Innovation Network. He recognized Collaborative Innovation Network as a cyber-team of self-motivated people with a collective vision, enabled by technology to collaborate in achieving a common goal innovation by sharing ideas, information and work. He explained the traits that characterize the network members and their behavior. His solutions include creation of self-motivated teams, collective vision, enabled by technology, common goal and sharing of ideas, information and work.

Zhu et al. (2006) introduced role-based collaboration model E-CARGO for collaboration. He suggested to establishing the development/business environment as a role net. Each role provides certain services and applies certain services in the proposed role net. His solutions to build a more efficient collaborative system includes regarding roles as agent dynamics in multi-agent systems, reducing the impact of role transfer in emergency management systems.

The collaborative innovation with business partners, customers. consultants. associations and even competitors have many benefits such as access to markets and customers, higher quality, sharing of risk, financial and intellectual benefits, reduction in technology quickly, leveraging of shared gap infrastructure, significant scale, faster time-to-market and time-to-revenue and increased customer (Venkatesh, 2006). However, the pitfalls are also significant which include:

- "Not invented here" syndrome
- Co-ordination failure
- Risk perception
- Opportunism and monitoring costs

In this study, the authors focus the research on establishing stronger relationship between other research groups with similar interest. The authors view the research community as a social network which made of research groups. Within each research group, the researchers are connected by similar research interest and domain knowledge. The authors will propose a model to value the similarity of research interest between two artifacts and then determine the similarity of research

interest between every two groups by analyzing the similarity of their artifacts. After that, collaboration graph will be generated and used to illustrate the research similarity between different research groups, high value relationship denotes large similarity between the two groups and low value denotes a small similarity between the two groups. Finally, the possible chance of academic collaboration can be found out so as to make the research activities more efficient.

There are some researches related to academic social network. For example, Tang et al. (2007) proposed a unified approach consisting relevant page identification, preprocessing and tagging, to capture the online academic information and generate the social network of the authors with a few academic relationship (Tang et al., 2007, 2008; Wang et al., 2010), ArnetMiner system was established on the basis of their proposed work. Strufe (2010) analyzed the online social networks including the voluntarily maintained and automatically exploitable databases of personal profile, with the analysis on a period of three months, the evaluation indicates that there is a strong relation between both the number of accepted contacts and the diligence of updating contacts versus the frequency of requests for a profile, while the overall activity, gender and participation span of users have no significant impact on the profile's popularity (Strufe, 2010). Considering an academic social network may help students to share commons academic interests, preferences, profiles and their historical studies, Jorge and Porfirio (2010) proposed an idea of building a social network for assisting European students in the personalization of their studies plans especially organizing their mobility issues.

While, most of the above research on academic social network focus on the individual expert finding, knowledge sharing and popularity gaining, This study regarding each academic research as the collaboration work of a group of expertise and it has struggled to find out more collaboration chance so as to enhance or extend the achieved research. The main difference of present research work between others' work is that it don't demonstrate the available relationship between researchers but it has explored the possibility of the research collaboration so that the research groups may find out the best partners with our research result.

DEFINITION OF RESEARCH ENVIRONMENT

To make the proposal more easily to be understood, in this study, we consider research environment as self-motivated environment contain private research groups and open research community containing the journals,

conference, blogs, Wikipedia and other published source, as which is demonstrated in Fig. 1. Each research group may have several research interests and may have published a set of articles to the open research community. The research groups know the articles published by other groups but know little about the detail interests the other groups had. It's essential for the groups to find out more collaboration chance to benefit their current research.

In order to find the collaborative chance between different research groups, the similar research interest should be identified and the research interests with largest similarity will be suggested to be a collaboration chance between two groups. Because the research interests may not be opened to others, they must be determined by the analyzing on the published articles. In such way, research interest, academic article, research group should be defined in present research.

Definition 1 research interest: A research interest is a research direction belongs to a research group, it is defined as ri::=<id, dms, goals, key words>, where:

- id is the identification of the research interest
- dms is a set of domains the research group is working
 on
- Goals is a set of goals the research group wants to achieve in the research work
- Key words is a set of key words contained in related papers

Definition 2 academic article: An academic article is the achievement published by a research group. It is defined as aa::=<id, dm, goal, achievement, title, abstract, key words, reference>, where:

- id is the identification of the academic article
- dm is the domain which the article is focused
- Goal is the research goal of the article
- Achievement is the achievements made by the work introduced in the article, which includes methodology, practice, framework, prototype, system and algorithms
- Title is the title of the concrete paper
- Abstract is the abstract of the concrete paper
- Key words is the a set of key words listed in the concrete paper
- References is a serial of references listed in the concrete paper

Definition 3 research group: A research group contains a set of research members and has several research

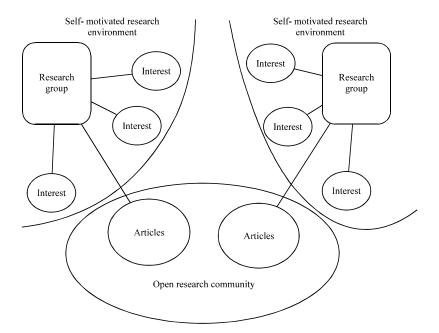


Fig. 1: A simple research environment model

interests and published articles. It is defined as rg::=<id, name, ris, aas, members>, where:

- id is the identification of the research group
- Name is the name of the group
- ris is a set of research interests of the group
- aas is a set of academic articles published by the group
- Members refers to the staff of the group

CORE ALGORITHMS FOR COLLABORATION CHANCE FINDING

To find the collaboration chance, we need to calculate the similarity of the published articles and determine the most suitable research collaboration between different research groups. Here, the related algorithms will be designed to calculate the similarity of the text, the article and determine the most valuable research collaboration for group pairs.

Procedure of text similarity calculation: Since, there are many text similarity algorithms published with different accuracy (Sebastini, 2002; Bedi and Chawla, 2007; Hasany et al., 2010; Chimphlee et al., 2010), the authors only state the procedure of the text similarity calculation without detailing the algorithms. The entire procedure is demonstrated in Fig. 2. The dictionary will be loaded at first; a part of text will be segmented in words according to dictionary. The frequency of the word will then be

calculated. The weights of words are pre-defined in the dictionary so that we could summarize the value of word together with the frequency as the similarity of the part of text

Algorithm for article similarity calculation: Input: Article1, Article2; Factors // Two published articles to be compared. Factors are the adjust factors to the different parts of the article.

```
Output: SV_article //the similarity of the two articles.
Fetch title1, abstract1, key words1 and reference1 from article1.
Fetch title2, abstract2, key words2 and reference2 from article2.
SV_title = text Similarity (title1, title2);
SV_abstract = text Similarity (abstract1, abstract2);
SV key words = text Similarity (key words1, key words 2);
num\_of\_ref1 = num\_of\_ref2 = 0;
FOR(each iref1 in reference1)
num of refl = num of refl+1;
FOR(each iref2 in reference2)
num of ref2 = num_of_ref2+1;
min num of ref = min (num of ref1, num of ref2);
same ref = similar ref = 0;
FOR(each iref1 in reference1)
FOR(each iref2 in reference2)
IF(text Similarity (iref1, iref2)=1)
same_ref = same_ref + 1;
ELSEIF (text Similarity (iref1, iref2)>=lowest Boundary)
similar_ref = similar_ref + 1;
SV reference
                     (same ref+similar ref *
                                                   min factor ref)/
min_num_of_ref;
SV article = SV title * factor title
+ SV abstract * factor_abstract
+ SV_key words* factor_key words
+ SV reference * factor reference;
```

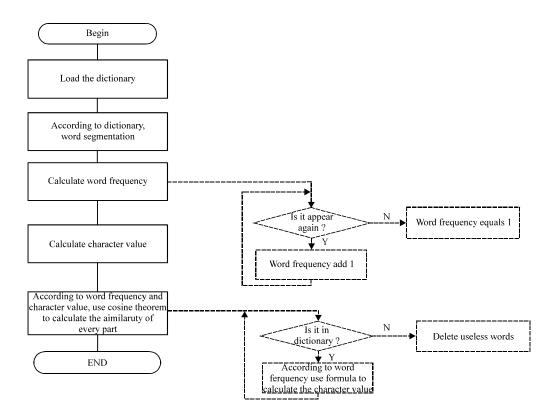


Fig. 2: The procedure of text similarity calculation

This algorithm calculates the similarity of reference different from the similarity calculation of title, abstract and key words. Fetch title, abstract, key words and reference from article can be automatically operated but is manually input in our prototype. The function text Similarity is the algorithm to calculate the similarity of two texts. Regarding the reference similarity calculation, different articles published by the same authors can be considered but is not realized in the prototype.

To be mentioned, the key words similarity can also be calculated as that of reference. The entire procedure of the article similarity calculation is stated as Fig. 3.

After the article similarity has been calculated, the largest similarity article-pair can be calculated for each research group pair. The related research topic and research interests can be considered as the research chance between two research groups. Of course, the lowest boundary should be defined so as to remove the unnecessary research collaboration. Besides, when there are several similar research interests between the research group pair, we should reduce the granularity of research group down to research group and research interests.

EXPERIMENT AND PROTOTYPE

Text and article similarity calculation: The authors have realized the algorithms of text similarity and article similarity calculation and developed a prototype for them.

The following screen snapshots illustrate the article similarity calculations.

Figure 4 demonstrates two articles with low similarity, there is little similar between the title, abstract, key words and references of two papers. While Fig. 5 demonstrates two high relevant papers with high similarity, the similarity value of title, abstract, key words and references shows significant high similarity between two papers.

Academic research groups and papers selection: On the basis of article similarity evaluation, the shared interest between two research groups can be identified with the comparison of the articles published by both groups.

In order to avoid the possible impact to the other research groups, here we only collect some papers from our group and another group with good relationship while the references to their papers has been permitted in this research. The papers may be gotten from the online digital libraries such as IEEE and ACM.

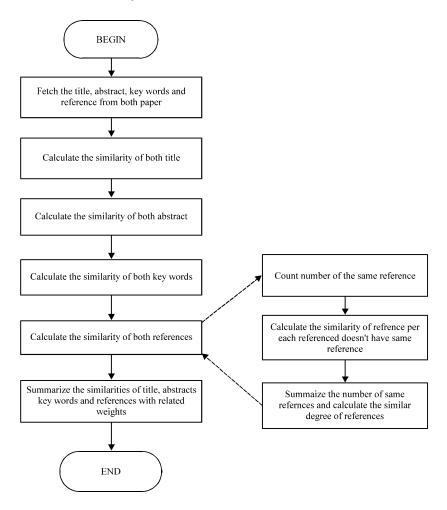


Fig. 3: The entire procedure of article similarity calculation



Fig. 4: Article similarity calculation (example for two low similarity articles)

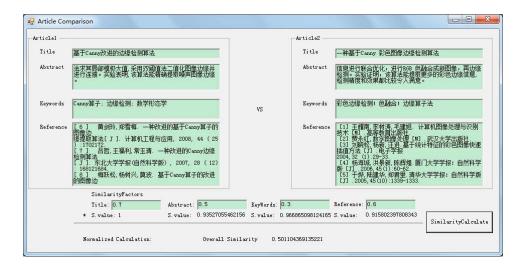


Fig. 5: Article similarity calculation (example for two high similarity articles)

Two categories of papers published by the authors have been collected in the experiment. One category (considered as the first research group) is talking about the collaboration which contains the papers in reference (Xu, 2008a, b, 2009) talking about the collaborative online learning and the paper in reference (Pan and Xu, 2010) talking about the collaborative innovation training to undergraduate and master students, as well as a paper in reference (Xu et al., 2008) talking about collaborative design. Another category (considered as the third research group) contains the papers talking about sensor network, including a paper (Hu and Xu, 2006) talking about on demanding information acquisition from sensor network and a paper (Ling and Xu, 2006) talking about reducing energy usage with a special monitoring protocol, the paper (Chen and Xu, 2006) proposing an efficient protocol for temperature monitoring and a study (Xu and Hu, 2008) presents automatic framework for energy saving in sensor network.

Some papers have been selected to serve as another academic group papers (considered as the second research group), including the paper of Zhu and Zhou (2008a) introduced the problems related to role transfer and the solution, the paper of Zhu and Zhou (2008b) surveyed the roles in information systems, the paper of Zhu et al. (2006) suggested using role in software development and the paper of Zhu and Zhou (2006) established a kernel mechanisms for role-based collaboration.

Shared interest analysis: Currently we didn't adapt an automatic information acquisition method to capture the

Table 1: Similarity evaluation between Group 1 and 2
Group 2

Group 1	Article 1	Article 2	Article 3	Article 4
Article 1	0.7354728	0.5149766	0.6321225	0.6784706
Article 2	0.6663709	0.6180396	0.5605396	0.7517107
Article 3	0.6860579	0.6424781	0.6433232	0.5935819
Article 4	0.6175917	0.7089727	0.5737601	0.5670175
Article 5	0.6853886	0.6073101	0.6519100	0.6131236

Table 2: Similarity evaluation between Group 2 and Group 3

	Group 2				
Group 3	Article 1	Article 2	Article 3	Article 4	
Article 1	0.7053824	0.5567297	0.6386894	0.5401194	
Article 2	0.5025552	0.4507139	0.5725086	0.4352659	
Article 3	0.6214752	0.5911908	0.7680253	0.5646599	
Article 4	0.5961331	0.4717070	0.5193315	0.5237016	

Table 3: Similarity evaluation between Group 1 and Group 3

	Group 3			
Group 1	Article 1	Article 2	Article 3	Article 4
Article 1	0.5212552	0.5961946	0.5686240	0.5353351
Article 2	0.6268912	0.6179277	0.5746366	0.6073374
Article 3	0.6047564	0.7211792	0.6698668	0.5782689
Article 4	0.4209937	0.5717538	0.4198647	0.4306695
Article 5	0.7349350	0.6920872	0.5690916	0.6253412

title, abstract, key words and references from the selected papers. All the information was collected manually from the selected papers and generated into corresponding structured XML files.

We evaluate the similarity of the papers between Group 1 and 2, between Group 1 and 3, between Group 2 and 3 as well. Table 1-3 demonstrate the results of all these evaluations.

All those similarity value greater than or equal to a certain value (here is 0.7 in this study) has been selected in pairs and the common key words shared by the paired

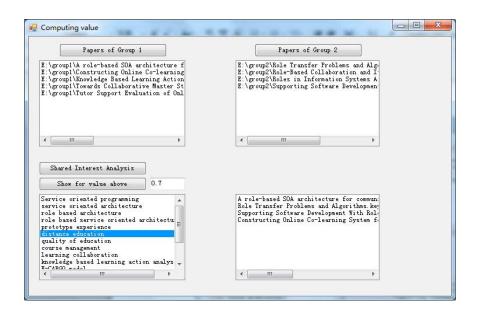


Fig. 6: Shared interests between Group 1 and 2

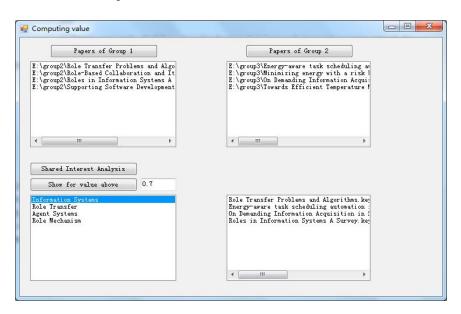


Fig. 7: Shared interests between Group 2 and 3

Table 4: Shared Interests between Group 1 and 2

Group 2

Group 1 Article 1 Article 2 Article 4

Article 1 Service oriented programming

Article 2 Distance education

Article 4 Collaboration

papers are selected as the shared interests. Table 4 lists the shared interests between Group 1 and 2.

Beyond the out expectation, there are also shared research interests between Group 2 and 3 and between Group 1 and 3 as well as what is demonstrated in Table 5 and Table 6.

The shared interests capturing has been realized in the prototype and Fig. 6 shows the founded shared interests and the related papers between Group 1 and 2.

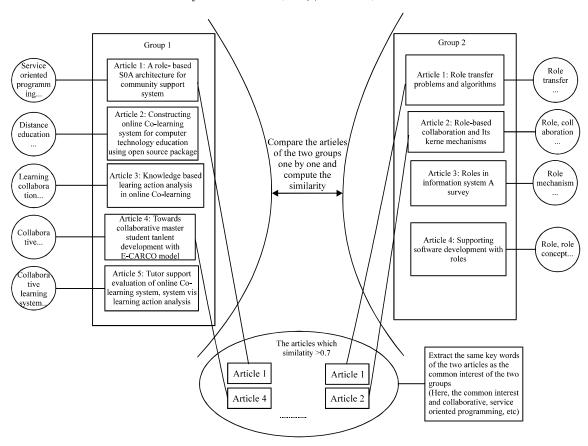


Fig. 8: Collaboration chance indicated in the academic social network

Table 5: Shared Interests between Group 2 and Group 3			
	Group 2		
Group 3	Article 1	Article 3	
Article 1	Information system		
Article 3		Role transfer, agent system	

Table 6: Sha	red Interests between Group 1 and	13
	Group 3	
Group 1	Article 1	Article 2
Article3		Knowledge based learning action analysis
Article5	Collaborative learning system	

Similarly, shared research interest has been founded between Group 1 and 3, as which is shown in Fig. 7 and between Group 2 and 3. As, both interfaces are similar, here we only demonstrate the interface of the Group 2 and 3 as the example.

After all the shared interests have been founded between academic research groups, the research collaboration chance can be indicated with the value of the research interest sharing.

Academic collaboration chance diagram: Figure 8 demonstrates the collaboration chance between Group 1

and 2. With such collaboration chance indication, the academic research groups may easily find the potential research colleague, share their research achievement and extend the available research capability with other research groups with shared research interests.

DISCUSSION AND STATUS OF RESEARCH

While most researchers focus on establishing the personal social network for academic community, this study indicates the potential research collaboration chance based on the analysis of published research papers. The proposal can be adopted by other researchers in the established academic social network so as to enhance the capability of social network evolution.

However, the weakness of this paper is the algorithms ignore the authors who published the research papers. As, the research work should be done by the collaboration of the researcher in the groups, the collaboration will be definitely easier when there are some shared researchers between the research groups. As someone may find from the experiment in this study, the author H. Zhu is both a co-author in the study,

Yu et al. (2009) which belong to research Group 1 and corresponding author in the paper Zhu and Zhou (2006, 2008a, b) and Zhu et al. (2006) which belong to research Group 2. The authors are considering enlarging the collaboration chance value when there are shared researchers in the consequent research. Meanwhile, B. Xu is one of the authors of the papers belong to Group 1 and 3. As a result, there is shared research interest between Group 1 and 3. However, to predict the possible shared research interests indicated by the relationship between co-authors falls outside the scope of this paper.

Currently, we are going to develop a framework with the involved algorithms and would like to try it in finding the possible research collaboration between several groups in computer science. The issues related to share researcher between groups will be studied in further research.

CONCLUSION

The advantage and issues of collaborative innovation has been introduced and academic research chance has been studied in this study. This study presents a method to find out the shared research interests between different research groups to make it easy for the research groups to find out the potential collaboration. Research environment has been modeled and the core concepts have been defined, several core algorithms have been introduced or detail stated. A prototype has been developed to introduce the features of this study. Some interfaces of an experiment have been presented to make the research more readable.

ACKNOWLEDGMENT

This research is supported by Zhejiang Gongshang University, China.

This research was also financial supported by National Natural Science Foundation of China with No. 60873022 to Dr. Hua Hu, The Science and Technology Department of Zhejiang Province, China with No. 2008C11009 to Dr. Bin Xu and No. 2008C13082 to Dr. Xiaojun Li and Education Department of Zhejiang Province with No. yb07035 to Prof. Yun Ling and No. 20061085 to Dr. Bin Xu.

REFERENCES

- Bal, S. and R. Nath, 2010. Filtering the web pages that are not modified at remote site without downloading using mobile crawlers. Inform. Technol. J., 9: 376-380.
- Bedi, P. and S. Chawla, 2007. Improving information retrieval precision using query log mining and information scent. Inform. Technol. J., 6: 584-588.

- Chen, L. and B. Xu, 2006. Towards efficient temperature monitoring and controlling in large grain depot. Proceedings of the 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks, September 28-28, 2006, Reston, VA., pp: 881-885.
- Chimphlee, S., N. Salim, M.S.B. Ngadiman and W. Chimphlee, 2010. Hybrid web page prediction model for predicting a users next access. Inform. Technol. J., 9: 774-781.
- Curran, K., 2004. Tips for achieving high positioning in the results pages of the major search engines. Inform. Technol. J., 3: 202-205.
- Fiaidhi, J.A.W. and S.M.A. Mohammed, 2003. Towards developing watermarking standards for collaborative e-learning systems. Inform. Technol. J., 2: 30-34.
- Gloor, P.A., 2006. Swarni Creativity: Competitive advantage through collaborative innovation networks. Oxford University Press, Oxford, New York, ISBN: 9780195304121, Pages: 212.
- Hasany, N., A.B. Jantan, M.H.B. Selamat and M.I. Saripan, 2010. Querying ontology using key words and quantitative restriction phrases. Inform. Technol. J., 9: 67-78.
- Hu, H. and B. Xu, 2006. On demanding information acquisition in sensor network. Proceedings of the IEEE International Conference on Information Acquisition, August 20-23, 2006, Weihai, pp. 835-839.
- Jayanthi, M.K., S.K. Srivatsa and T. Ramesh, 2007. Learning objects and e-learning system: A research review. Inform. Technol. J., 6: 1114-1119.
- Jorge, A.L. and F.P. Porfirio, 2010. Building an academic social network: For Bologna mobility. Proceedings of the 3rd Workshop on Social Network Systems, April 13, 2010, New York, USA.
- Klein, M., H. Sayama, P. Faratin and Y. Bar-Yam, 2006. The Dynamics of Collaborative Design: Insights from complex systems and negotiation research. In: complex engineered systems, Braha, D., A. Minai, Y. Bar-Yam (Eds.). Springer, Berlin.
- Lin, H.T., T.H. Kuo and S.M. Yuan, 2007. A web-based learning portfolio framework built on blog services. Inform. Technol. J., 6: 858-864.
- Ling, Y. and B. Xu., 2006. Minimizing energy with a risk based temperature monitoring protocol for wireless digital sensor network. Proceedings of the 2006 IET International Conference on Wireless, Mobile and Multimedia Networks, November 6-9, 2006, Hangzhou, pp. 1-5.
- Mcdaniel, E.A., 2003. Facilitating cross-boundary leadership in emerging e-government leaders. Inform. Sci.,

- Memon, S., K. Khoumbati and S.R. Hussain, 2007. Internet based multimedia services and technologies in the context of e-government: A conceptual framework. Inform. Technol. J., 6: 903-908.
- Pan, L. and B. Xu, 2010. Towards collaborative master student talent development with E-CARGO model. Inform. Technol. J., 9: 1031-1037.
- Sebastini, F., 2002. Machine learning in automated text categorization. ACM Comput. Surveys, 34: 1-47.
- Strufe, T., 2010. Profile popularity in a business-oriented online social network. Proceedings of the 3rd Workshop on Social Network Systems, April 13, 2010, Paris pp. 1-6.
- Tang, J., D. Zhang and L. Yao, 2007. Social network extraction of academic researchers. Data Proceedings of the 7th IEEE International Conference on Mining, October 28-31, 2007, Omaha, NE., pp. 292-301.
- Tang, J., J. Zhang, L. Yao and J. Li, 2008. Extraction and mining of an academic social network. Proceedings of the 17th International World Wide Web Conference, April 21-25, 2008, Beijing, pp. 1193-1194.
- Venkatesh, G., 2006. Collaborative innovation. Proceedings of the International Conference on Managing Technological Innovation in IT, June 21-23, 2006, Bangalore, India,.
- Wang, C., J. Han, Y. Jia, J. Tang, D. Zhang, Y. Yu and J. Guo, 2010. Mining advisor-advisee relationships from research publication networks. Proceedings of the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, July 25-28, 2010, Washington, DC, pp. 203-212.
- Wu, C. and D. Zeng, 2009. Knowledge transfer optimization simulation for innovation networks Inform. Technol. J., 8: 589-594.
- Xu, B. and H. Hu, 2008. Energy-aware task scheduling automation framework in large scale temperature sensor network. Proceedings of the 7th World Congress on Intelligent Control and Automation, June 25-27, 2008, Chongqing, pp. 2402-2406.

- Xu, B., 2008a. Constructing online co-learning system for computer technology education using open source package Proceedings of the 2008 International Symposium on Information Science and Engineering, December 20-22, 2008, Shanghai, pp. 219-223.
- Xu, B., 2008b. Knowledge based learning action analysis in online co-learning. Proceedings of the 2nd International Symposium on Intelligent Information Technology Application, December 20-22, 2008, Shanghai, pp. 162-166.
- Xu, B., X.H. Yang, Y.H. Shen, S.P. Li and A. Ma, 2008. A role-based SOA architecture for community support systems. Proceedings of the International Symposium on Collaborative Technologies and System s, May 19-23, 2008, Irvine, CA, USA., pp: 408-415.
- Xu, B., 2009. Tutor support evaluation of online co-learning system via learning action analysis. Proceedings of the International Symposium on Web Information Systems and Applications, May 22-24, 2009, Nanchang, pp. 10-14.
- Yu, B., B. Xu, Y. Ling and H. Zhu, 2009. Equipment conflict checking and removal for innovation talent development in IT undergraduate education via E-CARGO collaborative model. Proceedings of the 4th International Conference on Computer Science and Education, July 25-28, 2009, Nanning, China, pp. 1669-1672.
- Zhu, H. and M.C. Zhou, 2006. Role-based collaboration and its kernel mechanisms. IEEE Trans. Syst. Man. Cybern. C., 36: 578-589.
- Zhu, H., M. Zhou and P. Seguin, 2006. Supporting software development with roles. IEEE Trans. Syst. Man Cybernet. Part A, 36: 1110-1123.
- Zhu, H. and M.C. Zhou, 2008a. Role transfer problems and algorithms. IEEE Trans. Syst. Man. Cybern. A., 38: 1442-1450.
- Zhu, H. and M.C. Zhou, 2008b. Roles in information systems: A survey. IEEE Trans. Syst. Man. Cybern. C. Applied Rev., 38: 377-396.