

Paper Review Assignment based on Human-Knowledge Network

Xijin Tang Zhengwen Zhang

Institute of Systems Science
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
Beijing 100190, PRCINA

Abstract—Along with the fast-growth in both scale and frequency of academic conferences and advances in communication and information technologies, digital revolutions are happening to conferencing affairs. Paper review assignment is a hard and time-consuming job due to many constraints, even emotional factors affect the review assignment. In this paper, the human-knowledge network, aggregated based on the past conference publications, social networks of PC members' or practical reviewers' and their expertise, is adopted into paper review assignment to tackle those constraints so as to generate an integer programming model to produce optimal alternatives for practical assignment.

Keywords—paper review assignment, limited rationality, decision support, iView analysis

I. INTRODUCTION

Science conferences provide opportunities to expand academic views and personal social networks for researchers and other people who are interested to be involved, and are vital in social processes of collaboration, communication, and consensus that bring out transformations in knowledge by P. Thagard's theory on scientific knowledge growth^[1]. One enabler of the expansion in both scale and frequency of conferences is the advances in communication and information technologies which bring digital revolutions to conference management. Currently the organization of scientific conferences often requires a web-based management system to deal with electronic processing of paper submission, review, registration and information publishing. The machine processing of conferencing affairs decreases a lot of labors of conference management, while the paper review assignment, as the hardest and most time-consuming task is still a heavy burden to the program committee (PC) chairs who usually undertake this job manually. Due to many constraints to be satisfied, such a manual task is very tedious and difficult, and sometimes the results are not satisfactory. No response from the potential reviewers may worsen the situation. Moreover, the review assignment of conference submission is somewhat different from that of journal papers, whose potential reviewers can be firstly considered from those authors of citations. It is beyond PC chairs' capabilities to look through all references of all submissions within a very narrow time window. Assignment support is barely required to be explored for most of conference

management systems.

In this paper, the support to the paper review assignment is studied. The paper review assignment is approached from a human-knowledge network aggregated during the previous conferences. The paper is organized as follows. Section II addresses the general process of paper review assignment, where a variety of constraints are indicated, together with some emotional factors which reflect the limited rationality of PC chairs. Section III addresses our approach to paper review assignment based on human-knowledge network, referred as iView network, which denotes a pair of networks, human network and knowledge network. Based on keyword-sharing link between vertices, human net shows a social network where community structure detection may give the clusters of candidate reviewers whose topics may match the submissions. The keyword network based on keywords shared between humans shows a knowledge vision which may depicts the capacity of current PC members and tell PC chairs to invite new members to fulfill review tasks. An integer programming model is then developed to give the assignment plan. In Section IV an example is illustrated to show a paper review assignment happened at the 7th International Workshop on Meta-synthesis and Complex Systems (MCS'2007). The concluding remarks are given at last.

II. PAPER ASSIGNMENT PROCESS AND COMPUTERIZED SUPPORT

Nowadays, the composition of a science conference is of more diverse styles than before. Besides general sessions, plenary sessions and tutorial sessions, workshops or technical tracks for specific topics become popular which may help to decrease the burdens of organizers of the individual events while make use of resources and services of the whole conferences. Whatever it is, all submissions will be under peer-reviewing process so as to select submissions with quality under an acceptance rate. Different acceptance rates apply to conferences among different disciplines. During the paper submission process, the corresponding author need to select topics or workshops, which are then regarded as attributes of that submission.

After the deadline of paper submission, it comes to the tough period for PC chairs to deal with paper review assignment in a short time to leave more time for sending request, dynamically adjusting review plan according to feedback, tracking the state of paper review process, etc. Here a

This work is supported by Natural Science Foundation of China under Grant No. 70571078 & 70221001.

general review assignment process is depicted to indicate as many factors as possible during the problem solving process.

A. General Paper Review Assignment Process and a Mathematical Model

Step 1. Invite each PC member m_i ($i = 1, 2, \dots, M$) to set his preference toward the available list of submissions. Usually he can browse the titles of the submission. 3 preference options are available: favorable, neutral and unfavorable.

However, most PC members always ignore this option.

Step 2. Invite each PC member m_i to set his working interests according to the topics of the conference or the workshop.

Authors usually have already declared their topics during paper submission. More often, many authors select as many topics as possible. Due to different understandings toward the wide topics, sometimes the submissions may not actually belong to the topics.

Step 3. PC chairs assign reviewers to each paper p_j ($j = 1, 2, \dots, N$) by satisfying the following constraints. The binary variable x_{ij} denotes whether the PC member m_i review the paper p_j .

i) Each PC member m_i has a maximum number S_i of papers to be reviewed, i.e. $\sum_{j=1}^N x_{ij} \leq S_i$. The number will be different for each PC member.

ii) Each paper p_j should be reviewed by at least N_r reviewers, i.e. $\sum_{i=1}^M x_{ij} \geq N_r$. Usually N_r is assigned to 3 or 5.

iii) PC members cannot review papers written by people from same research group. This is to avoid the conflicts of interests.

Above constraints are common for paper review assignment problem. Then an integer programming model is constructed to be solved.

Step 4. Construct an integer programming model for paper review assignment

$$\begin{aligned} \text{Max } Z &= \sum_{i=1}^M \sum_{j=1}^N \text{obj}(x_{ij}) \\ \text{s.t. } &\begin{cases} \sum_{j=1}^N x_{ij} \leq S_i \\ \sum_{i=1}^M x_{ij} \geq N_r \end{cases} \end{aligned} \quad (1)$$

where $x_{ij} = 0$ if the paper p_j is not assigned to the reviewer m_i , $x_{ij} = 1$ if p_j is assigned to m_i , $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$.

The objective function $\text{obj}(x_{ij})$ measures the match between p_j and m_i . Obviously to maximize the match measure is the objective. We can also suppose it is to minimize the distance between p_j and m_i .

Step 5. Solve the IP model. If the optimal solution can be reached, use that for assignment plan.

Given the above-mentioned process, the paper review assignment is a structured problem. A computerized support based on an integer programming model seems not a quite tough task.

However, it is very difficult to define $\text{obj}(x_{ij})$. Either Step 1 or Step 2 does not produce necessary information to do matching. Computerized support is then to be implemented for this point. GRAPE (global review assignment processing engine), a rule-based expert system for scientific Conference Management System (CMS), was proposed to fulfill the task of match between topics of papers and reviewers preferences successfully [2]. However more factors need to be considered and not easily to be expressed by a rule.

B. Limited Rationality in Paper Assignment

Sometimes there is a group of special papers whose authors are friends of PC chairs. Science conference is for academic exchange. People prefer to get familiar with some ideas directly by face-to-face exchange. Conferences are held to help people report their latest research to avoid loss of reputation for the breaking results, get feedback for sound research and preparation for journal papers. Then PC chairs have to consider how to assign those special papers to the appropriate reviewers who are expected to give both critical comments and considerate ratings to increase the possibility of its acceptance. Sometimes, a reviewer of a friend's paper may give high ratings even the paper's quality is not so good. At this moment, other reviewers are expected to give objective assessment so that the total rating is not too biased. If authors always submit high quality papers, even one biased reviewer give low rating, we still do not need to worry the final assessment. It is very difficult to define a rule for such kind of emotional factors. Moreover, it takes time for PC chair to consider those factors. In summary, PC chairs will take consideration to both emotional and behavioral factors and purport to

- 1) Avoid reject high-quality papers or papers with peculiar ideas which may lead to debate;
- 2) Avoid accept poor quality papers even written by their friends;
- 3) Assign papers of friends' to the honest reviewers for critical comments.

Besides, PC members have their behavioral patterns. Sometimes, formal PC members never review papers. Instead,

they ask their representative(s) to review papers. At this moment, we use technical program (TC) members including those representatives to replace PC members.

In next section, a method is presented to fulfill the match between paper topics and review preferences in consideration of those emotional factors.

III. METHOD OF MATCHING BETWEEN REVIEWERS AND PAPERS BASED ON IVIEW NETWORK

In GRAPE or other systems for paper review assignment support, topics of papers can be acquired by authors' choices or layout structure analysis of the paper. Here, we only consider using simple information of the paper for the match. Actually, each topic can be regarded as one keyword, or a group of keywords defined by the PC chairs. Each paper can be represented by a quintuple structure

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

Such metadata 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 a topic, a problem, a method or an algorithm, a practical case, etc. Then the keyword set of a paper could be understood as the basic ideas toward the problem addressed by the author(s) of that paper^[3]. In fact, the standard keywords representing the topics can also be as members of the *keywords* set of the paper.

A. Similarity between Reviewers and Papers

As keywords can be regarded as attributes of one paper, or one person (author or PC member), we can use keywords to compare the distance between reviewers and papers, and assign papers to the closest reviewers by considering the constraints of maximum review papers. The process is as follows.

Step 1. Combine the keywords of each paper and the keywords sets of potential reviewers to constitute a keyword vector space.

The element of the vector refers the occurrence of keywords instead of the frequency of each keyword.

Step 2. Rerepresent the paper p_j and reviewer m_i with the combined overall keywords set using Boolean model. Then the l th element of p_j or reviewer m_i could be 0 or 1 in condition the keyword exist in the vector.

Step 3. Calculate the similarities between paper p_j and reviewer m_i by cosine of the included angle. That is, let s_{ij} is the similarity between paper p_j and reviewer m_i , then

$$s(i, j) = \frac{p^T \cdot m}{\|p\| \cdot \|m\|} \quad (2)$$

The similarity function is a practical measure of objective function $obj(x_{ij})$ mentioned above.

However, emotional factors could not be reflected by the similarity measure between reviewer' interests and papers topics. The so-called iView analysis is to be exploited.

B. iView Network Analysis

The iView analysis requires constructing the iView network which usually denotes a pair of networks, one is keyword network, the other is human network^[4].

1) Keyword 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 where $K_l = \{k_1^l, k_2^l, \dots, k_n^l\}$ is the keyword set, E_l is the edge set, then $G = (K, E)$ where $K = \cup K_l = \cup \{k_1^l, k_2^l, \dots, k_n^l\}$, $E = \cup E_l = \cup \{e_{ij}\}$, $i, j = 1, 2, \dots, m, i \neq j$. This topological network 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. More senses may be obtained via a variety of network analysis by detecting some of its features, such as cutpoints^[5], centrality of keywords, clustering of keywords, etc. which is superior to general cloud model of text enabled by Web 2.0 technologies. Idea map is a collective vision of all the authors.

For example, a cutpoint (articulation point) of a graph is a vertex whose removal will increase the number of connected components; then the cutpoint keyword may reveal the real key concepts or topics. So does the centrality analysis of the keyword vertex. The clustering of keywords by use of community structure detection methods may help people understand the major topics from those keyword clusters easier instead only by frequencies of individual keywords visualized at a cloud model.

2) Human net via keyword-sharing between humans

In iView analysis, keywords-sharing between people (TC members) is considered and a human network where the vertex denotes a human being is constructed. If two persons share one keyword, a link between them exists. The strength between two persons indicates the number of the different keywords or the total frequencies of all the keywords they share. From such a human net, social network analysis (SNA)^[6] is then applied to find the powerful people by centrality analysis and to detect the interest group by community structure detection, etc. Coauthorship network is a typical social network which reflects common interests shared among co-authors, and is used to study scientific collaboration^[7-9]. However, the interests or topics of the TC members are always more diverse. A co-authorship network of a TC of conference may include many isolates and components; it may then be meaningless to detect major interest groups. Obviously, the co-authorship network is a sub graph of the keyword-sharing network.

3) Finding Appropriate Reviewers by iView Analysis

The closeness of interests or topics between reviewers can be studied via iView network, which is used to locate available TC members for paper review assignment. The process is as follows.

a) Get the available reviewer list.

Find the PC members who share some of those keywords by idea map according to the topics, i.e. the *keywords* set. This step can also be fulfilled by similarity measure between one paper and all available TC members.

b) Conduct power analysis of available PC/TC members.

This is to check if those available PC/TC members are cutpoints or powerful people by betweenness centrality analysis.

c) Find the appropriate TC members.

Here we consider each paper is reviewed at least by 3 reviewers.

Step c_1. Select a TC member with the highest rank of centrality at first. If a tie exists among available options, select the one which is a cutpoint. Or else, randomly select one of them. Assign the paper to that TC member.

Step c_2. Do clustering analysis of human net and find two clusters. Cluster 1 includes the assigned TC member; Cluster 2 is the closest cluster to that TC member.

Step c_3. Select one PC member who is of shortest distance from that TC member with highest centrality in each cluster and assign the paper to them.

The above three steps could be run iteratively considering the constraints of maximum reviewing paper numbers for each reviewer.

Next an example is given to show how the method works.

IV. EXAMPLES OF MCS'2007 PAPER REVIEW ASSIGNMENT

MCS'2007 (the 7th International Workshop on Meta-synthesis and Complex Systems) is a registered workshop of the 7th International Conference on Computational Sciences. The number of submissions to MCS'2007 is beyond 140, which set a record in MCS history. It took MCS'2007 executive PC chair (the 1st author of this paper) one whole week to fulfill paper review assignment by expanding the original PC groups again and again, which drove the study of this paper^[10].

Here use one paper for example. The title of the paper is “*A hidden pattern discovery and meta-synthesis of preference adjustment in group decision making*”. The keyword set is {“group decision making”, “Markov chain”, “meta-synthesis”, “hidden pattern”}. We use the iView network of MCS'2006 (usually referred as KSS'2006) for analysis.

A. Calculate Similarity between potential Reviewers and the Paper

To calculate the similarity between the case paper and all PC members, and only two PC members, *Tang* and *Gu* whose similarity measures are greater than zero.

$$s(Gu, \text{ case paper}) = 0.4743,$$

$$s(Tang, \text{ case paper}) = 0.1118$$

Since the number of available reviewers is fewer than the minimum requirements, even one is executive TC chair. We need to find other PC members.

Another way to locate the potential TC members is to identify them from via iView network directly. Supported with a visualized tool, such as iView developed by the authors, it is also easier to find whose keyword set includes those keywords of the case paper.

B. Construct the Human Net of MCS'2006 iView Network

Firstly we construct iView network of MCS'2006. Fig.1 is the human net. The keyword-sharing network has 8 components excluding 4 isolates^[4]. Several components are labeled, e.g. one component labeled as “cognitive complexity”, etc. Cutpoints can also be detected.

Table 1 lists some persons' betweenness and the rank. The TC member *Tang* is of highest rank in betweenness centrality. We need to find other TC member to replace *Tang* to review the case paper.

TABLE I. BETWEENNESS ENTRALITY OF PC MEMBERS

Name (TC member)	Centrality of Betweenness		
	Rank	Betweenness	Cutpoint ?
Tang (TC)	1	521.0	True
Kuanjiu Zhou (TC)	2	372.0	True
Yoshiteru Nakamori (TC)	3	165.2	True
Peng Liu	4	156.0	True
Haoxiang Xia (TC)	5	126.0	False
Shuguang Wang	6	126.0	False
Teketoshi Yoshida (TC)	7	126.0	False
Andrzej P. Wierzbicki (TC)	8	113.9	False
Kun Nie (TC)	14	12.9	False
Gu (TC)		0	False
Liu (TC)		0	False

C. Clustering of the Human Net of MCS'2006

Labeled as “CORE”, the biggest component includes 44 authors (above half of all authors of MCS'2006) who are grouped into 4 communities by Newman-Girvan algorithm as shown in Fig. 2. We find two clusters which are linked by *Tang*. One cluster has 10 people including 7 TC members, {Xijin *Tang*, Jifa *Gu*, Minjie Zhang, Wen Zhang, Haoxiang *Xia*, Takatoshi Yoshida, Kuanjiu Zhou} referred as Cluster A; another cluster has 16 people including 4 TC members, {Yoshiteru *Nakamori*, Andrzej P. Wierzbicki, Kun Nie, Yijun *Liu*} referred as Cluster B. Now we acquire available TC members, i.e. those 11 PC members to review the case people. This searching starts to fix one TC member.



Figure 1. Human net of MCS'2006 (keyword-sharing network)^[4]



Figure 2. Community structure of the "CORE" component of MCS'2006 (Modularity measure $Q = 0.712$)^[4]

D. Selecting Appropriate TC Members
 1) Select one member from Cluster A: Among those members in Cluster A, only Gu is of highest similarity measure with the case paper. The similarity between Gu and

Tang is higher than that between other PC members and Tang. Then Gu is selected as the alternative TC member to review that case paper.

2) *Select one member from Cluster B*: Among those TC members in Cluster B, only Liu is of highest similarity measure with Tang.

Till now, we have 3 alternatives to be assigned, *Tang*, *Gu* and *Liu*. As *Tang* (TC chair) is not available, we need to find another TC member. At that time *Ohsawa* is selected, even he is at another component of Fig.1 (close to the bottom). That is to guarantee the fairness of paper review. Another reason is that there is a technical track on Chance Discovery and Meta-synthesis held in ISSS'2007 in Tokyo (August). At this point, we can add meta-synthesis into *Ohsawa*'s keyword list.

The manual assignment plan for the case paper is {*Gu*, *Liu*, *Ohsawa*}. The paper was finally accepted.

In reality, a total of 55 TC members are assigned with 143 submissions to MCS'2007. A total of 55 TC members finished the paper review within a very short period (20 days). Each paper was under at least 3 anonymous reviewers. The least number of assignments for PC members is 1 and the most is up to 22. The average review paper number is around 7 to 8. At last 33 papers were accepted, and 32 papers finally entered into the ICCS'2007 proceedings.

V. CONCLUDING REMARKS

Even electronic conferencing systems have brought out a lot of labor decrease in conference organization, there are still tough tasks left for manual operations. Paper review assignment, the central step of paper review process is the most tedious and heavy burden to PC chairs, where automatic assignment is barely expected for long. Such a problem is of few studies.

In this paper, we at first review the basic process of paper review assignment which seems to be a structured problem approached by an integer programming model. However, more factors have to be considered, such as the emotional and behavioral factors which may be attributed to limited rationality of the decision makers during paper review assignment. We apply iView analysis to consider those factors and acquire the available reviewer list. An example is given to show the operation and effectiveness of our approach. A lot of work is to be done to implement the method into practical application. It is more important to find available TC members for each paper so as to build the integer programming model for optimal solution of paper review plan.

Actually the iView network for appropriate paper reviewers should be different from that of a past conference/workshop. The keywords of those papers that a TC member had reviewed can be regarded as his interested topics and then be included into his keyword set. Improvements wait to be done.

Scientific conference is the principal way for social process of communication, collaboration and consensus along the scientific knowledge growth process. Advanced conference assistant services had been exploited via ubiquitous computing technologies to facilitate participants' communications and interactions, awareness of interested topics and people during the conferences since the late 1980s^[11, 12]. Those extensive activities are helpful to enable on-line conferencing *ba* (OLCB) emerge^[13]. For the emergence of OLCB, the essential requisite is to bring high quality or new peculiar papers which can lead to divergent idea generation and active exchanges. At this point, the quality of paper review assignment plays an important role.

As a matter of fact, automatic paper review assignment only gives the options while it is PC chairs to make the final assignment plan.

REFERENCES

- [1] P Thagard, *How Scientists Explain Disease*. Princeton, New Jersey: Princeton University Press, 1999
- [2] N. Di Mauro, T.M.A. Basile, and S. Ferilli, "GRAPE: An expert review assignment component for scientific conference management systems," in proceedings of IEA/AIE 2005, LNAI 3533, M. Ali and F. Esposito, Eds. Berlin Heidelberg: Springer-Verlag, pp. 789-798, 2005.
- [3] X. J. Tang, "Towards Meta-synthetic Support to Unstructured Problem Solving," *International Journal of Information Technology & Decision Making*, vol. 6, no. 3, pp. 491-508, September 2007
- [4] X. J. Tang and Z. W. Zhang, "How knowledge science is studied - a vision from conference mining of the relevant knowledge science symposia," *International Journal of Knowledge and Systems Sciences*, vol. 4, no.4, pp.51-60, December 2007
- [5] F. Harary, *Graph Theory*, Addison-Wesley, 1970. Chapter 3.
- [6] R. A. Hanneman, and M. Riddle, *Introduction to Social Network Methods*, University of California, Riverside, 2005 (URL: <http://faculty.ucr.edu/hanneman/nettext/>)
- [7] M. E. J. Newman, "The Structure of Scientific Collaboration Networks," *Proc. Natl. Acad. Sci. USA* vol.98, pp. 404-409, 2001.
- [8] M. E. J. Newman, "Scientific Collaboration Networks: I. Network Sonstruction and fundamental Results," *Phys. Rev. E* vol.64, 016131, 2001.
- [9] M. E. J. Newman, "Scientific Collaboration Networks: II. shortest Paths, weighted Networks, and Centrality," *Phys. Rev. E* vol.64, 016132, 2001.
- [10] X. J. Tang, ed. *Meta-synthesis and Complex Systems (2006-2007)*. Research report, no.MSKS-2007-05, pp. 125-127, December 2007. (in Chinese)
- [11] B. S. Gitta, "Designing casual-use hypertext: The CHI'89 InfoBooth," in: *Proceedings of the ACM CHI 90 Human Factors in Computing Systems Conference*, J. Carrasco and J. Whiteside, Eds. Seattle, USA, pp.451-458, 1990.
- [12] Y. Sumi and K. Mase, "Conference Assistant System for Supporting Knowledge Sharing in Academic Communities," *Interacting with Computers*, vol. 14, no.6, pp.713-737, 2002.
- [13] X. J. Tang, N. Zhang, and Z. Wang, "Augmented Support for Knowledge Sharing by Academic Conferences --- On-line Conferencing Ba," in the proceedings of IEEE WiCOM'2007 (the Management Track of IEEE International Conference on *Wireless Communications, Networking and Mobile Computing*), pp.6400 - 6403, 2007.