Paper-to-reviewer assignment, based on expertise degree of reviewers and relevance degree between reviewers and papers

Xinlian Li

Graduate School of Information Science,
Nagoya University,
Furu-cho, Chikusa-ku, Nagoya 464-8603, Japan
E-mail: li@watanabe.ss.is.nagoya-u.ac.jp

Toyohide Watanabe*

Nagoya Industrial Science Research Institute,
c/o Corp. Mizuno,
2-1-16 Seimeiyama, Chikusa-ku,
Nagoya 464-0087, Japan,
E-mail: watanabe@nagoya-u.jp
*Corresponding author

Abstract: Automating the process of paper-to-reviewer assignment is a difficult task to be adequately resolved. Many papers concerning the related topics have been published, but there is still scare of systematic research applications. In most real world conference management, the assignment task is carried out manually by the programme committee, lacking of intelligent assigning rules and efficient matching method. The manual assignment is not only of low efficiency but also does not guarantee to result in the best solution. Given such situation, our paper sets out to analyse the problem of reviewer selection and propose a method for automatically matching papers with reviewers. Our objective is to reduce the loads of both programme committee and reviewers and make the conference-paper assignment task effectual. In this paper, we address this issue of paper-to-reviewer assignment and propose a method to model reviewers, based on the matching degree between reviewers and papers by combining preference-based approach and topic-based approach. We explain the assignment algorithm and show the evaluation results in comparison with Hungarian algorithm.

Keywords: Paper-to-reviewer assignment, Hungarian algorithm, matching degree, expertise degree, relevance degree between reviewer and paper.

Reference to this paper should be made as follows: Li, X. and Watanabe, T. (2014) 'Paper-to-reviewer assignment, based on expertise degree of reviewers and relevance degree between reviewers and papers', *Int. J. Knowledge and Web Intelligence*, Vol. 5, No. 1, pp.1–20.

Biographical notes: Xinlian Li received her BE from Faculty of Engineering, South China Agricultural University in June, 2010. After then, she received her Master degree from Graduate School of Information Science, Nagoya University, Japan in March, 2013. Her research interests include logical design, advanced algorithm development, etc. in computer-support social information systems.

Toyohide Watanabe received his BS, ME and Dr. Eng from Kyoto University in 1972, 1974 and 1983, respectively. In 1987, he was an Associate Professor in Department of Information Engineering, Nagoya University and then was a Professor in 1994. After then, he moved as a Professor to Department of Systems and Social Informatics, Graduate School of Information Science, Nagoya University, in 2003. In March 2013, he was retired from Nagoya University and currently in Nagoya Industrial Science Research Institute. His research interests include knowledge of personal intelligent activity, computer supported collaborative learning, social environment simulations, spatio-temporal model and geographic information systems and so on.

This paper is a revised and expanded version of a paper entitled 'Automatic Paper-to-reviewer Assignment, based on the Matching Degree of the Reviewers' presented at the 17th International Conference on Knowledge-based and Intelligent Information and Engineering Systems (KES2013), Kita-kyushu, Japan, 9–11 September 2013.

1 Introduction

Peer review is an evaluation process for the competence, significance and originality of researches by qualified experts (Sense, 2004). The process is to comment on the validity of research by identifying scientific errors, judge the significance of research by evaluating the importance of the findings, determine the originality of the work, based on how much it advances the field, and recommend the paper to be published or rejected. One of the most important and time-consuming tasks in peer review process is to assign each submitted paper to appropriate reviewers (Goldsmith and Sloan, 2007). The major concern in this task is to take both suitability and efficiency into consideration simultaneously. It is laborious to decide which reviewer has enough knowledge of the research areas related to the papers. While, due to the great amount of reviewers and papers, it is a huge burden for the programme committee to carry out the assignment task.

Due to many constraints to be necessarily fulfilled, automating the process of paper-to-reviewer assignment is still a difficult problem to be adequately resolved. Although many papers concerning the related topics have been published, there is still scarce of systematic research and practical applications. In most real world conference management, the assignment task is carried out manually by the programme committee, lacking of intelligent assigning rules and efficient matching method. The manual assignment is not only of low efficiency but also does not guarantee to result in the best solution. Given such situation, our paper sets out to analyse the problem of reviewer selection and propose a method for automatically matching papers with reviewers, based on decision-making means.

Our objective is to reduce the loads of both programme committee and reviewers and make the conference- paper assignment task effectual. In order to achieve this objective, the following issues must be solved:

- *How to find out proper reviewers?:* In the peer review process, the opinions of reviewers play a significant role in determining whether a paper should be accepted or not. At present, the process of evaluating and selecting reviewers is mainly semi-manual. The programme committee browses and searches the database to find the proper expert. However, the semi-manual selection method adopted currently is not only random but also subjective; and this leads to end up with unfair and inappropriate results. Hereby, selecting the suitable reviewer is the key step to ensure the quality of peer review process.

- *How to assign papers to reviewers?:* Assigning submitted papers to reviewers is another critical part of peer review process, which needs to consider factors such as assigning efficiency and research area similarity. The ideal assignment which takes such factors into consideration is possible if every member reads every paper. However, this is impossible since there are usually several hundred submissions. Moreover, papers are submitted from a wide variety of topics; it is unlikely that every person would have the same ability and interests to review every paper. Therefore, it is necessary to develop an assigning method to balance the load as well as to assign papers to appropriate reviewers.

2 Basis for ordinary approach

Generally speaking, the paper-to-reviewer assignment method can be classified into two categories: preference-based approach and topic-based approach.

2.1 Preference-based approach

Several of the approaches in the paper-to-reviewer assignment problem make use of preference or bidding data from reviewers. In most preference-based approaches, the systems usually require the reviewers to bid the papers to see whether they have their interest for the papers or not. A weakness in this approach is the inadequacy of bidding information; most reviewers return preference only for a small percentage of papers. Rigaux (2004) suggested the use of collaborative filtering techniques to grow the preference by asking users to bid on all or most of the papers in a given topic, instead of a few bids over the entire set of papers. The basic assumption of collaborative filtering techniques is that reviewers who bid similarly on a number of the same papers have likely the similar preference for other papers.

2.2 Topic-based approach

One view of paper-to-reviewer assignments is that papers should be assigned to reviewers with a certain degree of familiarity in the specific field or topic of the paper. This view leads to topic-based approaches that use additional information. By using this information, reviewer assignments can be made so as to ensure a degree of similarity between paper's topic and reviewer's research area. The resultant ranking of each reviewer, based on topical knowledge with respect to a given paper, was called expert-finding or expertise modelling. One problem aroused with this approach is to identify what topics are covered in papers. Early efforts in this field focused mainly on paper abstracts, and topical similarity was determined through common information retrieval means involving keywords. For example, Dumais and Nielsen (1992) matched papers to reviewers by using Latent Semantic Indexing trained on reviewer-supplied abstracts. In Basu et al. (2001), abstracts from papers written by potential reviewers were extracted from the web via search engine, and then a vector space model was constructed for the matching. Yarowsky and Florian (1999) extended this idea by using a similar vector space model with a Naïve Bayes Classifier. More recently, Wei and Croft (2006) proposed a topic-based means by using a language model with Dirichlet smoothing.

Under such approaches, some systems with automatic paper-to-review assignment features (Snodgrass, 1999) are developed practically: Myreview and GRAPE.

2.2.1 Myreview

Myreview proposed by Rigaux is designed to solve the problem of reviewer assignments for scientific conference management. This is based on the preference-based approach. Instead of rating each paper, it asks each user to rate a sample of papers. A collaborative filtering algorithm is then performed to generate predicted preferences of reviewers. The method was implemented in the MyReview web-based system and was used in the ACM/GIS2003 conference.

2.2.2 GRAPE

Another web-based conference management system is GRAPE, based on Di Mauro et al. (2005). GRAPE is notable for considering both reviewers' biddings and topical similarity under two-phases assignment process. The paper's topics from its title, abstract and references, and the reviewers' topics by analysing their previously written papers and web pages are respectively extracted. The system was evaluated on real-world datasets built by using data from a previous European conference.

3 Research view

Although many studies have been published concerning the problem of conference-paper assignment, these studies mainly focused on biddings and did not provide much attention to other input sources. For example, the collaborative filtering approach of MyReview considers only reviewers' preferences. GAPRE considers both reviewers' biddings and topical similarity, but the secondary source is only used when the reviewers fails to provide any preference. Therefore, the current preference-based systems render the following problems:

- consume additional hand-work and time
- place too much emphasis on reviewers' interests: these are questions about the confidentiality of peer review
- require huge amount of calculation for meaning the similarity of research topics.

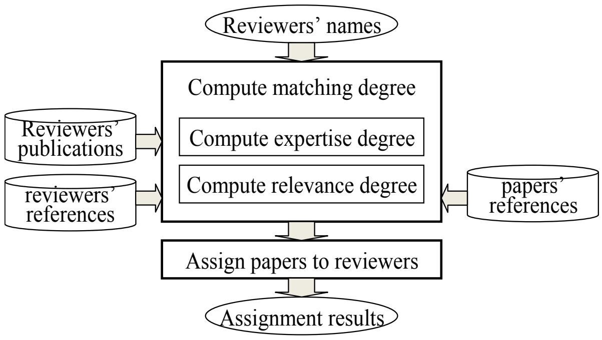
Our approach combines both preference-based approach and topic-based approach in a way that does not require the bidding process of reviewers. We set out to present our approach in solving two viewpoints.

3.1 Evaluation criteria and method

We propose a method to model reviewers, based on the matching degree between reviewers and papers by combining preference-based approach and topic-based approach. The traditional preference-based approaches assume that a person who has high interest for a paper is the suitable reviewer for the paper. However, these approaches suffer from several weaknesses. To solve this problem, we transform the reviewer preference into paper preference. As for the topic-based approach whose objective is to measure the similarity between reviewer's and paper's research area, we employ a new method by using the reference information instead of identifying the topic between reviewer and paper. Our matching degree is divided into two parts:

- Preference of papers: reviewer's expertise;
- Similarity of topics: relevance of references.

Figure 1 Processing flow in our approach



3.2 Assignment criteria and method

After the matching degree is measured, a matrix is constructed for assignment. Several constraints should be fulfilled in order to balance each load, and to ensure each paper is examined by adequate amount of reviewers. Since the existing algorithms cannot be applied to this assignment problem directly, we propose a method which is feasible to solve the assignment problem.

Figure 1 illustrates our system structure. There are two main sub-tasks such as the calculation of matching degree, and the assignment algorithm.

4 Framework of reviewer modelling

The first and essential part of paper-to-reviewer matching task is to model the appropriateness for an expert. Until now, many researchers have carried out on the topic of expertise modelling. An excellent example of expertise modelling is author persona topic (APT), proposed by Mimno and McCallum (2007). APT model contains a number of features designed to capture the better association between a paper and a reviewer. The basic idea of APT model is that even if an author may study and write about several distinct topics, the author's ranking for a given topic would not be decreased by his/her writings on different topics since papers are clustered from these topics into separate author persona. Although APT model is excellent in taking both author's ranking and topic into consideration, it is very complicated for reviewer modelling.

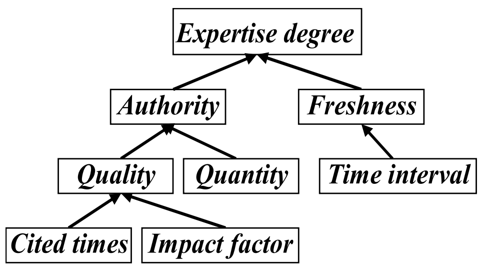
4.1 Reviewer modelling

We present our model for assessing the matching degree. We divide the matching degree between reviewer and paper into two aspects: expertise degree of reviewer, and relevance degree between reviewer and paper.

4.1.1 Expertise degree of reviewer

Publications provide an effective way to evaluate the expertise of reviewer. Many works have been done in the topic for the last few decades. Jauch and Glueck (1988) stated that simple count of publications, modified by the quality index of journals, is the best way to assess the academic contributions of a researcher. Sun et al. (2008) proposed a method to evaluate experts for R&D projects by measuring each expert's performance as publications, projects, historical performance in project selection and other experts' opinions. Although these methods provide us with some illumination, they only offer simple solutions in limited domains.

Figure 2 Expertise degree of reviewer



We propose a model for measuring the expertise degree of reviewer in two main aspects in Figure 2.

4.1.1.1 Authority: quality and quantity of publications

One of the most important attributes in judging the authority of a reviewer is the quality of his/her previous publications. Reviewers who published papers with high quality are more possible to be authoritative and can provide more proper evaluation on submitted papers. The qualities of reviewer's publications are measured by combining two factors together: the impact factor of journal in which papers were published, and the times they were referred by other papers. It can be assumed that the quality of publication relies much on the ranking or grade of journal where it was published. Papers published in different levels of journals have different weights and importance. In general, papers published in journals with high impact factors are regarded to be of high quality. Several methods have been proposed for evaluating the ranking of academic journals (Jauch and Glueck, 1988). One widely accepted approach to assess the levels of journals is by using the impact factor, which was devised by Eugene Garfield in 1995 (Sun et al., 2008). Journals with higher impact factors are deemed to be more important than those with lower ones. In addition to the ranking of journals, the number of paper citations also contributes to the quality. It is important to publish a paper in a high level journal, while it is also more important to be quoted by many other papers. A repeatedly quoted paper is deemed to be a classical literature within its research field. As a result, the number of times that a paper has been cited by other papers indicates the quality of the paper.

Another factor that should be taken into consideration when evaluating the authority of reviewer is the quantity of his/her publications. In general, the more the number of papers that a researcher published in a research field is, the more experience he/she may have in that field. The research experience can certainly enhance the reviewer's authority to evaluate papers. The quantity is measured by the total amount of papers published by a reviewer. In scientific paper written by more than one author, we must look upon the first author as more important person in reviewer assignment because the first author gives the greatest contribution to the paper.

4.1.1.2 Freshness: when was the paper published?

Since science and technology develop at a rapid speed, a paper published decade ago may be not as novel as a paper published recently. The published year reflects the freshness of a paper directly. A researcher who published papers in the recent years may be familiar with the current research trends. On the contrary, a researcher who was inactive in the recent years may be unfamiliar with the diversity of current researches. The freshness of publication is measured by the time interval between the publication date and the current date.

4.1.1.3 Calculating expertise degree

In order to calculate the expertise degree of reviewer, we first measure the quantity, quality and time interval independently. Quantity is calculated by the sum amount of the publications. Quality consists of two factors: number of citations to the paper and the ranking of journal. Here, we denote the ranking of journal to be the ratio of the impact factor of journal for the maximum impact factor within its field. Freshness is measured by using time interval with respect to the ratio for five years as the basis unit. For example, if a paper was published in 2005, then the time interval from 2013 is eight years; and the freshness can be derived as 8/5 = 1.6. Since quality and time contributes independently to the expertise degree, they can be calculated by using multiplicative algorithm. The quality and quantity are the benefit attributes. They positively relate to the reviewer's authority. The greater the values of quality and quantity are, the greater the reviewer's expertise degree is. The time is a cost attribute; it negatively reflects the freshness of a publication.

The quality can be represented as:

$$Quality = \left(c_j + 1\right) \times if_j / if_{\max}. \tag{1}$$

The freshness can be described as time by e-index:

$$Freshenss = \exp\left(-t_j / 5\right). \tag{2}$$

Here, c_j is number of citations for paper j; t_j is time interval between published year of paper j and current year; if_j is impact factor of journal in which paper j was published; and $if_{\max}$ is maximum impact factor of journal within research discipline. Then, the expertise degree (*Expertise*) of each reviewer can be calculated as:

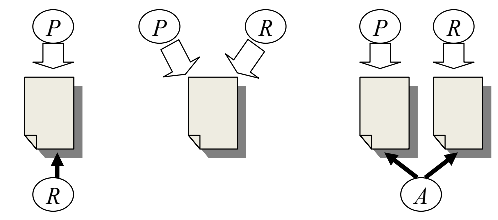
$$\begin{aligned} Expertise &= \sum\nolimits_{Quantity} Quality \times Freshness \\ &= \sum\nolimits_{j=1,n} \left(c_j + 1\right) \times \left(if_j / if_{\max}\right) \times \exp\left(-t_j / 5\right) \end{aligned} \tag{3}$$

Finally, the normalised expertise degree for each reviewer is calculated, so that the maximum of each reviewer's expertise degree is 1.

4.1.2 Relevance degree between reviewer and paper

The traditional approach for measuring the relevance is to estimate the similarity between reviewer's publication and submitted paper. Methods including vector space model (Salton et al., 1975) and latent semantic indexing techniques (Hofmann, 1999) have been devised for estimating the similarity of text documents. The weakness of the similarity-based methods is obvious: in most case, these methods involve the processes of extracting features from documents and calculating the similarity based on the extracted features. The process is of low efficiency because it is too complex and consumes a lot of time.

Figure 3 Common referring types, (a) direct referring (b) same paper referring (c) same author referring (see online version for colours)



(a) (b) (c)

Table 1 Comparison scales in AHP

Value	*Definition*
1	Equal importance
3	Moderate importance of one over another
5	Strong or essential importance of one over another
7	Very strong or demonstrated importance of one over another
9	Extreme importance of one over another
2, 4, 6, 8	Intermediate values
Reciprocals	Reciprocals for inverse comparison

Our approach is based on the assumption that two papers which refer to the same reference share similar research areas strongly. It is unusual that papers regarding different topics refer to the same references. Therefore, it is true that the more the number of common references two papers have, the more similar research field they are in. In order to compute the similarity of reference between reviewers and papers, we gather all references cited by a reviewer from his/her previous publications, and extract paper's references from its bibliography. Bibliography contains a lot of information, including paper title, author, source and published year. Not only title but also author information should be taken into consideration when computing the similarity of reference. We classify the types of common referring into three categories: direct referring, same paper referring and same author referring. Figure 3 shows three different kinds of common referring. We define three kinds of common referring as follows:

- direct referring: paper *P* quotes one of reviewer *R*'s publication directly
- same paper referring: both paper *P* and reviewer *R* refer to the same reference
- same author referring: both paper *P* and reviewer *R* cite the same author *A*'s publication.

The degree of relevance between reviewer and paper is calculated by combining these three referring information together. We assign different weights to different kinds of referring. The method for determining the relative weight for three kinds of common referring is stressed.

The relevance between paper and reviewer is computed as follows:

$$Relevance = \sum_{i=1,3} (w_i \times r_i) \tag{4}$$

where *Relevance* is the relevance degree between reviewer and paper, *w* is the weight of common referring type, and *r* is the number of common referring type. The relevance degree is normalised so that the range of relevance degree is from 0 to 1.

Figure 4 Matching degree

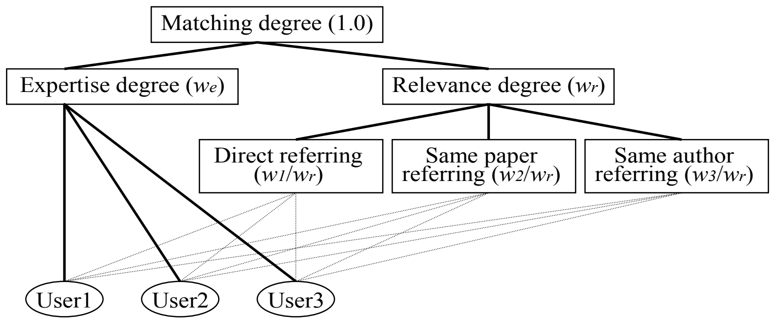


Table 2 Matrix of relevance degree

	w_1	w_2	w_3
w_1	w_{11}	w_{12}	w_{13}
w_2	w_{21}	w_{22}	w_{23}
w_3	w_{31}	w_{32}	w_{33}

Table 3 Matrix of matching degree

	w_e	w_r
w_e	w_{ee}	w_{er}
w_r	w_{re}	w_{rr}

4.2 Matching degree

4.2.1 Determining weights of expertise degree and relevance degree

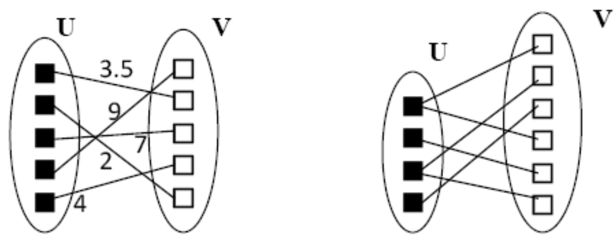
From the above, we can acquire two kinds of criteria for evaluating the matching degree between reviewer and paper. The criterion is integrated in order to form an overall evaluation. Here, we employ analytic hierarchy process (AHP) (Saaty, 1977) to determine the relative importance of expertise degree and relevance degree. As listed in Table 1, the scale '[1, 9]' is used for making the pair-wise comparison judgment in AHP. Since relevance degree is computed by combining three kinds of common referring, we need to determine the weight of different types of common referring at first. Figure 4 illustrates AHP hierarchy structure model for determining matching degree for three users. As described previously, w_1, w_2 and w_3 represent the weights of three kinds of common referring, respectively. w_e stands for the relative weight of expertise degree, while w_r stands for the weight of relevance degree.

4.2.2 Calculating matching degree by example

We asked 3 users to give pair-wise comparison matrices for the relevance degree and matching degree as shown in Table 2 and Table 3. The pair-wise comparison matrices on three types of referring are presented as follows:

$$U_1 = \begin{pmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 5/3 \\ 1/5 & 3/5 & 1 \end{pmatrix}, U_2 \begin{pmatrix} 1 & 4 & 4 \\ 1/4 & 1 & 1 \\ 1/4 & 1 & 1 \end{pmatrix}, U_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 3/2 \\ 1/3 & 2/3 & 1 \end{pmatrix}$$

Figure 5 Example of weighted bipartite graph, (a) linear weighted bipartite graph (b) non-linear assignment problem



(a) (b)

And, the pair-wise comparison matrices on expertise degree and relevance degree are:

$$U_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, U_2 - \begin{pmatrix} 1 & 2 \\ 1/2 & 1 \end{pmatrix}, U_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

According to AHP algorithm, the weights for different kinds of common referring are derived as $w_1 = 0.62$, $w_2 = 0.22$ and $w_3 = 0.16$, respectively. The important criterion on expertise degree and relevance degree are $w_e = 0.56$ and $w_r = 0.44$. Then, the matching degree *Matching* is given by:

$$Matching = w_e \times Expertise + w_r \times Relevance \tag{5}$$

5 Assignment problem

How to find out the appropriate and sufficient experts for each paper is a combinatorial optimisation problem to find a maximum weight matching in a weighted bipartite graph. A weighted bipartite graph is a graph whose vertices are divided into two disjoint sets U and V such that every edge connects a vertex in U to one in V with a weight W. Figure 5 gives an example of weighted bipartite graph. Hungarian algorithm (Kuhn and Yaw, 1955) is one of many algorithms which were devised to solve the linear assignment problem. Hungarian algorithm is based on the viewpoint that an optimal assignment for the resulting cost matrix is also an optimal assignment for the original cost matrix if a number is added to or subtracted from all entries in any one row or column of a cost matrix. Unfortunately, the complex process in Hungarian algorithm cannot be fully adapted to our problem for the following reasons:

- To solve the minimum weight matching problem; it cannot be applied to our problem directly.
- To be a linear assignment problem, in which the numbers of reviewers and papers have to be equal. Each paper is only examined by one person and each reviewer inspects only one paper.
- To become extremely complex and time-consuming when the numbers of reviewers and papers are too huge.

Although a lot of methods such as the heuristic algorithm (Hopcroft and Karp, 1973), etc. have been studied on improving these weaknesses, an efficient and effective method has not yet been developed. Our primary aim is to provide an improved algorithm with high efficiency to deal with the non-linear maximum weight assignment problem.

Table 4 Example in conflict assignment

	1	*2*	*3*
a	30	50	*100*
b	70	40	*100*

Table 5 Matching degree matrix

	R1	*R2*	*R3*	*Row average*
P_1	M_{11}	M_{12}	M_{13}	ra_1
P_2	M_{21}	M_{22}	M_{23}	ra_2
P_3	M_{31}	M_{32}	M_{33}	ra_3
P_4	M_{41}	M_{42}	M_{43}	ra_4
Column average	ca_1	ca_2	ca_3	

5.1 Problem formulation

In order to present this problem in all of its complexities, we must consider matching degree as well as load balance. First of all, the assignments of papers to reviewers should be made so that the total matching degree is maximised. Given R for reviewers and P for papers together with a weight function M, the problem is expressed as:

$$\max \sum_{i \in R} \sum_{j \in P} M(i, j) \times a_{ij} \tag{6}$$

where

$a_{ij} = 1$, if reviewer i is assigned to paper j;

0, if reviewer i is not assigned to paper j.

a_{ij} stands for the assignment of reviewer i to paper j, and the weight function M represents the matching degree between each reviewer and each paper. When m is the number of reviewers, the number of papers (n) reviewed by each reviewer is defined as:

$$n = ceil(m \times |R| / |P|) \tag{7}$$

Additionally, since the amount of papers and reviewers is huge, there is a need to balance the assignment to ensure that no single reviewer is overworked. Thus, the following constraints must be fulfilled:

- Each reviewer should be assigned without more than n papers:

$$\sum_{j \in P} a_{ij} \leqq n \quad for \quad i \in R; \tag{8}$$

- Each paper should be reviewed by m reviewers:

$$\sum_{i \in R} a_{ij} \leqq m \quad for \quad j \in P. \tag{9}$$

5.2 Assignment algorithm

5.2.1 Solving assignment problem

Without the above constraints, the best way to obtain the assignment with the maximum weight is to assign papers to the reviewer who has the largest matching degree. However, given the above objective function and constraints, we should find out a way to rearrange the assignment when conflicts occur. We propose our algorithm to solve this maximum weight matching problem by solving the problem of how to deal with the conflict assignments. The fundamental idea is that when the conflict assignments occur it is proper to keep the assignment with larger deviation and remove the assignment with smaller deviation. This idea is based on the theory that if a value's deviation in average is bigger it should be likely that the rest of data are smaller. Consider an example. Although the maximum values of data-sets a and b are both the same (100) as shown in Table 4 and Figure 6, the rest data of a is smaller than that of b since the maximum value of a is more numerically distant from the rest of the data than that of b. If the assignment $b3$ is eliminated, it is still more possible to find a large value in the rest data of b than that of a. Thus, the maximum assignment in Table 4 should be $a3$ and $b1$ where the sum of the assignment is 100 + 70 = 170.

In order to decide which value is distant from the average of the data, we first need to construct a matrix of matching and calculate the average matching degree of each row and each column. Table 5 is an example of a matching degree matrix with three reviewers and four papers, and Table 6 is an example of assignment matrix.

Figure 6 Example in conflict assignment

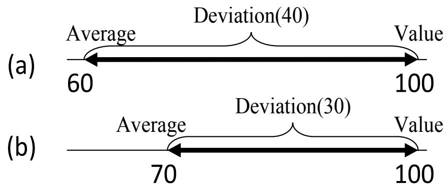


Table 6 Assignment matrix

	R_1	R_2	R_3
P_1	A_{11}	A_{12}	A_{13}
P_2	A_{21}	A_{22}	A_{23}
P_3	A_{31}	A_{32}	A_{33}
P_4	A_{41}	A_{42}	A_{43}

Table 7 Example of matching degree matrix

	R_1	R_2	R_3	*Row average*
P_1	0.75	0.37	0.92	0.69
P_2	0.61	0.87	0.37	0.62
P_3	0.57	0.42	0.75	0.58
P_4	0.18	0.57	0.87	0.54
Column average	0.53	0.56	0.73	

After calculating the average value of each row and each column, the deviation is defined as follows:

$$D_{ij} = M_{ij} - ra_i \tag{10}$$

$$Q_{ij} = M_{ij} - ca_j \tag{11}$$

D denotes the row deviation, which is the difference between the matching degrees of reviewer *j* and paper *i*, and the average matching degree of paper *i*. *Q* denotes the column deviation, which is the difference between the matching degree of reviewer *j* and paper *i*, and the average matching degree of reviewer *j*. Figure 7 gives the pseudo-code for the proposed deviation-based algorithm.

5.2.2 Example

We give an example to illustrate the process of our proposed algorithm. Consider the non-linear assignment problem in Table 7, where the numbers of reviewers and papers are no equal. In this example, we defined that every reviewer is assigned to no more than two papers (n = 2) and every paper is reviewer by exactly one reviewer (m = 1). Find out the maximum matching degree in each row.

5.2.2.1 Round 1: A_{13}, A_{22}, A_{33}, A_{43}

There are totally 3 assignments exist in column R_3. The row deviation of M_{33} (0.17) is the smallest among three assignments. Thus, the assignment A_{33} should be removed. Continue the maximum value in row that does not yet have an assignment.

Figure 7 Assignment algorithm

```
Algorithm: Assigning papers to reviewers
==================================
Input:  M: A matrix of matching degree;
        m: The number of reviewers assigned to each paper;
        n: The number of papers reviewed by each reviewer.
Output: A: A matrix of assignments.
Methods:
SET assignnum to 0
WHILE assignnum ≠ n × number of columns of M
FOR each row ri
WHILE number of assignments in ri < m DO
SET maximum matching degree's columnID into list maxClumnIDs
IF size of maxColumnIDs > 1 THEN
SET j to columnID which maximizes Q in maxColumnIDs
ELSE IF size of maxColumnIDs =1 THEN
SET j to first ID in maxColumnIDs
ENDIF
SET Aij ro 1
SET Mij to 0
ADD 1 to assignnum
ENDWHILE
ENDFOR
FOR each column cj
SET assignRowIDs to list of rowID which be assigned paper in cj
WHILE number of assignments in cj > n DO
SET i to rowID which minimizes D in assignRowIDs
SET Aij to 0
ADD -1 to assignenum
ENDWHILE
ENDFOR
ENDWHILE
```

Table 8 Example of non-linear assignment

	R_1	R_2	R_3		R_1	R_2	R_3		R_1	R_2	R_3
P_1	0	0	1	P_1	0	0	1	P_1	0	0	1
P_2	0	1	0	P_2	0	1	0	P_2	0	1	0
P_3	1	0	0	P_3	0	0	0	P_3	1	0	0
P_4	0	0	1	P_4	0	0	1	P_4	0	0	1
		(a)				(b)				(c)	

5.2.2.2 Round 2: A_{31}

Stop when all papers are assigned to one reviewer. Then, the final assignment of the example is shown in Table 8(c) via their temporary assignments in Table 8(a) and Table 8(b).

5.2.3 Comparison of existing algorithm and proposed algorithm

We discuss the advantages and limitations of the deviation-based method.

5.2.3.1 Advantage

Our method is advantageous in many aspects in comparison with the existing ones. One of the greatest advantages is its simplicity. Another advantage of the deviation-based method is its high efficiency. Besides of its simplicity and efficiency, our method is feasible in the given situations.

Table 9 Comparison in time-consumption

	3 × 3 matrix	*5 × 5 matrix*	*10 × 10 matrix*	*20 × 20 matrix*
Our algorithm	0.005 sec.	0.037 sec.	0.087 sec.	0.174 sec.
Hungarian algorithm	0.01 sec.	0.079 sec.	0.179 sec.	0.313 sec.

5.2.3.2 Limitation

However, it is worth noting that this method has still some kinds of limitation. To some extent, this method is only applicable to situations:

- the matrix is large enough
- the assignment problem is non-linear.

The above conditions must be satisfied in order to obtain a maximum weight matching. In case of conference- paper assignment where the amount of papers and reviewers is huge, it is proper to assume that matching degree matrix is large enough to meet the first condition. Also, since the numbers of reviewers and papers in a conference-paper assignment are not equal in most case, this problem is regarded as a non-linear assignment problem. Considering the above circumstances, our method can be adapted to this conference-paper assignment problem regardless of its limitation.

6 Experiments and evaluation

We conduct experiments and evaluate our approach based on the experimental results. The experiments are divided into two parts:

1. experiments on assignment algorithm
2. overall experiment.

6.1 Data set

We constructed our system based on the proposed approach. The available data sets for reviewer modelling were collected from the existing public data. The impact factor is acquired from the journal citation reports (JCR) (Thomson Reuters, 2005) which offers a systematic means to evaluate the journals based on citation data, the cited time of a paper and its references can be obtained from CiteSeerX (Giles et al., 1998), as a public search engine and digital library for scientific and academic papers. Moreover, the quantity of reviewer's publications and their published year can be gathered from DBLP (Ley, 2002), as a computer science bibliography website. Totally 313,620 papers with 2,084,019 references were stored in the constructed database. All papers are published in the field of Computer Science from 1980 to 2011.

6.2 Experiments on assignment algorithm

The first experiment is to evaluate the efficiency and effectiveness of our algorithm. All experiments were carried out on PC, running Windows 7 with AMD Athlon 64 Processor 3,200 + (2.0 GHz), 2 G RAM. The experimental data sets are random data that range from 0 to 1.

Table 10 Successful rate of matrix weight assignment

	3 × 3 matrix	*5 × 5 matrix*	*10 × 10 matrix*	*20 × 20 matrix*
Successful rate (%)	80	100	100	100

Table 11 Failed example in our approach

	A	*B*	*C*	*Row average*
I	0.99	0.90	0.20	0.70
II	0.20	0.99	0.15	0.45
III	0.99	0.50	0.40	0.63
Column average	0.73	0.80	0.25	

Table 12 Example of assignment

	A	*B*	*C*		*A*	*B*	*C*
I	0	0	1	I	1	0	0
II	0	1	0	II	0	1	0
III	1	0	0	III	0	0	1
		(a)				(b)	

6.2.1 Efficiency: time consumption

We evaluated the efficiency performance of our algorithm by comparing the time used to perform an assignment problem. As shown in Table 9, our algorithm can significantly reduce the time consumption in comparison with Hungarian (or bipartite graph matching) algorithm.

6.2.2 Effectiveness: assignments under constraints

We conducted an experiment to demonstrate the effectiveness of our algorithm. Since Hungarian algorithm is able to fulfil the requirement of maximum weight matching, we estimate the successful rate by comparing the assignment results to Hungarian algorithm. The results demonstrate the rate of assignments which succeed in achieving maximum matching degree, as shown in Table 10. From the result in Table 10, our algorithm is successful in satisfying maximum weight assignment in most cases though it fails in small-scale matrixes.

A failed example is given to explain the limitation of our algorithm. Consider the case of 3-order matrix assignments, shown in Table 11. The result of assignments is displayed in Table 12. The sum of the assignment is sum(1) = 0.20 + 0.99 + 0.99 = 2.18. However, the sum of the maximum matching degree is sum(2) = 0.99 + 0.99 + 0.40 = 2.38 > sum(1). It can be seen from this example that our algorithm may fail to satisfy maximum matching degree assignment in small-scale matrix. The failure is caused by insufficient amount of data for measuring the average value and deviation. When the scale of the matrix becomes larger, our algorithm is able to conform to the maximum assignment requirements.

Table 13 Average results in satisfactory level

Reviewers	*Average satisfactory level (random assignment)*	*Average satisfactory level (our approach)*
A	1	3
B	1.5	3.5
C	2	3.5

Table 14 Composition satisfactory level

	Perfect match	*Good match*	*Fair match*	*Somewhat relevant*	*Poor match*
Our approach	17%	33%	33%	0%	17%
Random assignment	0%	0%	25%	13%	62%

Table 15 Random assignment

Reviewers	*Assigned papers*	*Matching degree*	*Satisfactory level*
A	P_1	0.016	1
	P_2	0.016	1
B	P_3	0.057	2
	P_4	0.057	1
C	P_5	0.017	3
	P_6	0.017	1

Table 16 Our approach

Reviewers	*Assigned papers*	*Matching degree*	*Satisfactory level*
A	P_3	0.457	5
	P_5	0.016	1
B	P_4	0.497	4
	P_6	0.057	3
C	P_1	0.457	4
	P_2	0.457	3

6.3 Overall experiment

In the overall experiment, we estimated the degree of satisfactory and compared it with random assignment result. Since it is difficult to obtain data from real-world conference, this experiment was conducted, based on simulation data. In this experiment, we selected 8 papers, which were published in 2012, from database, and assigned them to three students. Then, each student was given a questionnaire and asked to rate the satisfactory level of assigned paper using 1–5 scales (5: perfect match, 4: good match, 3: fair match, 2: somewhat relevant, 1: poor match), based on the abstracts of the assigned papers.

Table 13 provides the satisfactory level of assigned papers, based on random assignment and satisfactory level, based on our approach. All user's average satisfactory level based on our approach is 3 or above, in average. The results of questionnaire indicate that our approach is able to achieve higher satisfactory, compared to the random assignment. Table 14 illustrates the satisfactory level, respectively. It can be discovered that more than 83% of the assigned papers, based on our approach, are rated higher than fair match. The ratings of low satisfactory level such as poor match are caused by the inevitable low matching degree.

We further analysed the correlation between matching degree and satisfactory level, based on the returned questionnaires. Table 15 and Table 16 list the assignments and their matching degree using random assignment and our approach. According to the above data, the correlation between matching degree and satisfactory level can be derived as 0.82. This result suggests that satisfactory level is correlated with matching degree. From the above experiments we conclude that our approach succeeds in providing useful help in paper-to-reviewer assignment.

7 Conclusions

We investigated the problem of automatic paper-to-reviewer assignment in academic conference management. Since peer review process has been widely used in academic conferences, supporting the process of paper-to-reviewer assignment has been concerned in many conference management systems. Given such situation, our paper analysed the problem of paper-to-reviewer assignment and proposed a framework of reviewer modelling, based on matching degree by using their previous publications. More importantly, we showed an assignment algorithm that could lead to efficiency improvements in large size assignments. Experimental results demonstrated that our approach is able to provide a solution to this problem and ensures the satisfaction of users. The main contributions of our paper are:

- in the theoretical aspect, our assignment algorithm throws light on the solution of other kinds of assignment problems
- in the applicant aspect, the paper-to-reviewer assignment method can be fielded to provide support for academic conference management.

However, it should be noted that this research has several limitations:

1 the assignment algorithm is restricted to large-scale non-linear assignment problem

2 with regard to the experimental result, the experiment were conducted based on simulation data since it is difficult to obtain data from real-world conference.

The research associated with our paper touches on a number of areas that could be developed further. A key question for future work is to rearrange the assignments dynamically by allowing reviewers to send feedback to the assignment results. Additional analysis and observation of user's feedback in future could significantly help to improve the satisfaction of users.

References

Basu, C., Hirsh, H., Cohen, W. and Manning, C.N. (2001) 'Technical paper recommendation: a study in combining multiple information sources', *Journal of Artificial Intelligence Research*, Vol. 14, pp.231–252, Doi: 10.1613.

Di Mauro, N., Basile, T.M. and Ferilli, S. (2005) 'GRAPE: an expert review assignment component for scientific conference management systems', *Proc. of 18th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Experet Systems*, pp.789–798.

Dumais, S.T. and Nielsen, J. (1992) 'Automating the assignment of submitted manuscripts to reviewers', *Proc. of 15th ACM International Conference on Research and Development in Information Retrieval*, pp.233–244.

Giles, C.L., Bollacker, K.D. and Lawrence, S. (1998) 'CiteSeer: an automatic citation indexing system', *Proc. of 3rd ACM Conference on Digital Libraries*, pp.89–98.

Goldsmith, J. and Sloan, R.H. (2007) 'The AI conference paper assignment problem', *Preference Handling for Artificial Intelligence*, *Proc. of AAAI workshop on Preference Handling in AI*, pp.53–57.

Hofmann, T. (1999) 'Probabilistic latent semantic indexing', *Proc. of 22nd Annual International ACM/SGIR*, pp.50–57.

Hopcroft, J.E. and Karp, R.M. (1973) 'An alrorithm for maximum weighting in bipartite graphs', *SIAM Journal on Computing*, Vol. 15, No. 3, pp.225–231.

Jauch, L.R. and Glueck, W.F. (1988) *Business Policy and Strategic Management*, 5th ed., McGraw-Hill.

Kuhn, H.W. and Yaw, B. (1955) 'The Hungarian method for the assignment problems', *Naval Research Logistics Quart*, Vol. 2, Nos. 1/2, pp.83–97.

Ley, M. (2002) 'The DBLP computer science bibliography: evolution, research issues perspectives', *Proc. of 9th International Symposium on String Processing and Information Retrieval*, pp.1–10.

Mimno, D. and Mccallum, A. (2007) 'Expertise modeling for matching papers with reviewers', *Proc. of 13th ACM/SIGKDD*, pp.500–509.

Rigaux, P. (2004) 'An interactive rating method: application to web-based conference management', *Proc. of ACM Symposium on Applied Computing*, pp.1682–1687.

Saaty, T.L. (1977) 'A scaling method for priorities in hierarchical structures', *Journal of Mathematical Psychology*, Vol. 15, No. 3, pp.248–281.

Salton, G., Wong, A. and Yang, C.S. (1975) 'A vector space model for automatic indexing', *Comm. of ACM*, pp.613–620.

Sense, B.T. (2004) *Peer Review and the Acceptance of New Scientific Ideas*, Sense about Science, London, ISBN 0-9547976-0-x, May.

Snodgrass, R. (1999) *Summary of Conference Management Software* [online] http://www.acm.org/sigs/sgb/summary.html (accessed 6 February 2014).

Sun, Y.H., Ma, Fan, Z.P. and Wang, J. (2008) 'A group decision support approach to evaluate experts for R&D project selection', *IEEE Trans. on Engineering Management*, pp.158–170.

Thomson Reuters (2005) *Journal Citation Reports, The Journal Evaluation Tool*, Thomson Corp., p.20 [online] http://ip-science.thomsonreuters.com/m/pdfs/mgr/jcr-qrg.pdf (accessed 6 February 2014).

Wei, X. and Croft, W.B. (2006) 'LDA-based document models for ad-hoc retrieval', *Proc. of 29th ACM International Conference on Research and Development in Information Retrieval*, pp.178–185.

Yarowsky, D. and Florian, R. (1999) 'Taking the load off the conference chairs: towards a digital paper-routing assistant', *Proc. of Joint SIGDAT Conference on Empirical Methods in NLP and Very-Large Corpora*, pp.220–230.