This is the Pre-Published Version.

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use(https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s10479-018-2919-7.

An Integer Linear Programming Model of Reviewer Assignment with Research Interest Considerations

Jian Jin[a], Baozhuang Niu[b*], Qian Geng[a] and Ping Ji[c]

[a] Department of Information Management, Beijing Normal University,

Beijing 100875, China

[b] School of Business Administration, South China University of Technology,

Guangzhou 510006, China

[c] Department of Industrial and Systems Engineering,

The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

* Corresponding author:

Email: bmniubz@scut.edu.cn

ABSTRACT In the regular work process of peer review, editors have to read and understand the entire set of submissions to choose appropriate reviewers. However, due to a large number of submissions, to select reviewers manually becomes error-prone and time-consuming. In this research, a framework that considers different indispensable aspects such as topical relevance, topical authority and research interest is presented and, an integer linear programming problem is formulated with practical considerations to recommend reviewers for a group of submissions. Specifically, the topical relevance and the topical authority are utilized to recommend relevant and accredited candidates in submission-related topics, while the research interest is to exam the willingness of candidates to review a submission. To evaluate the effectiveness of the proposed approach, categories of comparative experiments were conducted on two large scholarly datasets. Experimental results demonstrate that, compared with benchmark approaches, the proposed approach is capable to capture the research interest of reviewer candidates without a significant loss in different evaluation metrics. Our work can be helpful for editors to invite matching experts in peer review and highlight the necessity to uncover valuable information from big scholarly data for expert selection.

Keywords reviewer assignment problem; expert recommendation; research interest trend; topical authority;

1. Introduction

Consider the editor-in-chief of *Annals of Operations Research* receives a new submission. As illustrated in Figure 1, the regular work process of peer review is as follows.

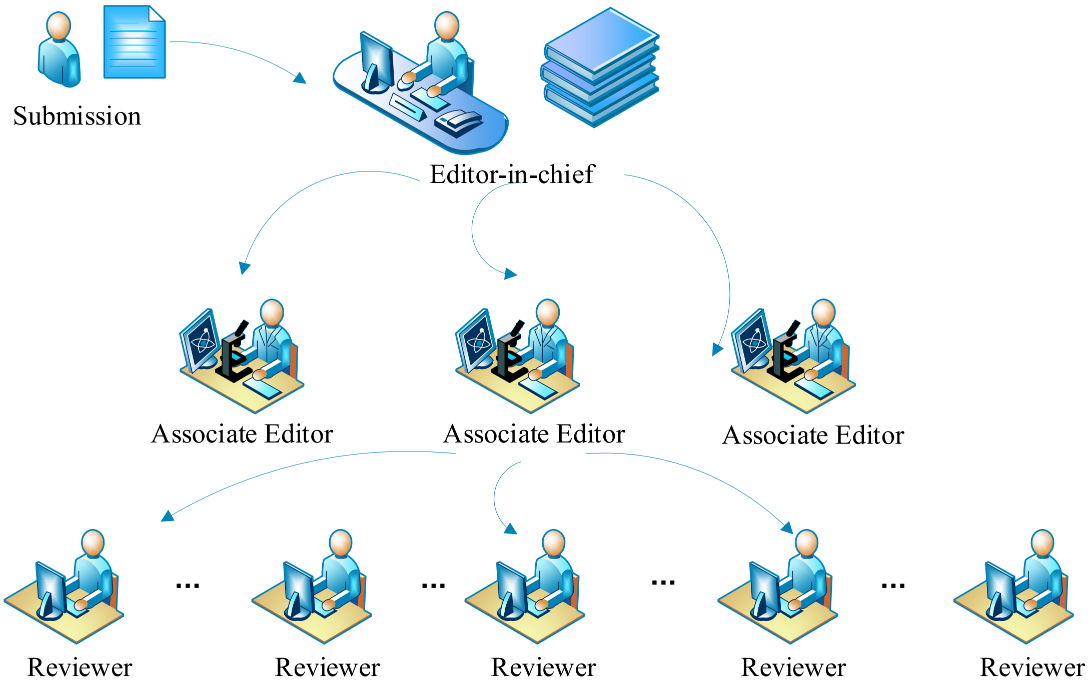


Figure 1. The regular work process of peer review

Step I: He/she browses the entire paper quickly and dispatch the draft to a particular associate editor according to the main research scope.

Step II: The associate goes through the entire paper to comprehend the main research topics of this submission.

Step III: According to the associate editor's understanding about the research fields of each expert in the reviewer repository, he/she tries to match the most appropriate experts with this submission and invite them to review this submission.

Step IV: If invited experts accept the paper review invitation, they begin to read articles for clarity, accuracy, appropriate methodology and theoretical base and send suggestions for revisions and for publication back to the associate editor.

Step V: The associate editor summarizes all these suggestions and sends back to the editor-in-chief.

Step VI: The editor-in-chief decides whether this submission should be accepted, revised or rejected, and sends responses from reviewers to authors.

Note that, in peer review, only high-qualified experts should be recommended to judge the intrinsic value of submissions. However, some famous journals, such as

Annals of Operations Research, receive a large number of submissions. Due to the large volume of submissions, in Step I and Step II, reading all submissions one by one becomes time-consuming. Furthermore, in Step II and Step III, associate editors are expected to be familiar with research interests of many experts and understand the research problem of each submission. The laborious manual reviewer invitation is prone to induce mismatch between submissions and reviewers. Also, in Step III, the manual expert selection is potentially influenced by subjective opinions of the associate editors, which might lead to inaccurate reviews. Note that, in some paper review systems, although invited reviewers are encouraged to pick out a list of labels about their research interests, these predefined subject labels are somewhat coarse-grained to a certain degree and might be not aligned with the detailed research topics exactly. It induces that some submissions might not be assigned well to proper experts.

Generally, the foregoing is referred as a reviewer assignment problem (RAP). It usually implies to choose appropriate experts to review submissions, which appears particularly important in peer review in different fields, such as, research and development (R&D) project selection (Cook, et al., 2005; Sun, et al., 2008; Stephen and Erim, 2015), online knowledge management (Wang, et al., 2013), digital libraries (Gollapalli, Mitra, & Giles, 2011), company recruitment (Balog, Azzopardi, & De Rijke, 2009), scientific evaluation (Fang, Si, & Mathur, 2010; Tayal, et al., 2014), etc.

Due to the obvious limitations of manual expert selection and the importance in the process of peer review, how to develop intelligent approaches for RAP draws an increasing number of researchers. An early work for RAP was conducted (Dumais, & Nielsen, 1992), in which the Latent Semantic Indexing (LSI) was used to measure the topical relevance between reviewer candidates and submissions. Since then, different studies are observed for RAP. Some researchers (Basu, et al., 1999; Biswas & Hasan, 2007; Petkova & Croft, 2008; Fang & Zhai, 2007) treated it as a retrieval problem, focusing on the topic relevance between reviewer candidates and submissions and different approaches were proposed such as LSI, Term Frequent - Inverse Document Frequency (TF-IDF), language model, etc. In addition to the topic relevance, some investigated other aspects for RAP (Karimzadehgan & Zhai, 2012; Tang, et al., 2012; Liu, Suel, & Memon, 2014), such as authority, diversity, etc.

Many studies emphasize the knowledge of reviewer candidates or relations among reviewer candidates and submission authors (Balog, Azzopardi, & De Rijke, 2006;

Zhou, et al., 2007; Tang, Tang, & Tan, 2010). However, most of these studies ignore experts' interest, which is critical in peer review, since it denotes the willingness to review submissions. Also, it is known that the research interest of an expert may change over time. For example, an expert put many efforts on the topic of text mining ten years ago. However, nowadays, this expert turns to study on the social network. It makes that an invitation to this expert for reviewing a submission about text mining is probably not a wise choice. Arguably, an intuitive approach for analyzing the research interests of reviewers is to concern their publications in a recent time window only (Li, et al., 2013). But the length of a time window is somewhat tricky to define. A smaller time window to screen an expert's publication list will potentially lead to the data sparse problem, while a larger time window will fail to capture the trend of his/her research interest. Besides, some models are reported to capture the evolution of topics in corpus, such as DTM (Dynamic topic models) (Blei and Lafferty, 2006), AToT (Author-Topic over Time) (Xu et al., 2014) and SDIM (Supervised Document Influence Model) (Jiang, Liu and Gao, 2015). Nonetheless, although these models made strong assumptions on the dynamic topic evolvement of textual data, few attentions about the strength of an expert's research interests are paid. This induces that they are arguably to be applied to extract the topic evolution of a particular author from big scholarly textual data. Also, topic extraction is focused on in these studies and none of them make further investigations on how to exploit the value of their findings in practical applications. Actually, to design an information system that is able to make effective recommendations of appropriate experts for a large number of submissions and provide a quick response to users is particularly important (Setaputra, Yue and Yao, 2010; Choi, Chan and Yue, 2017). Hence, an effective approach to identify the trend of a reviewer candidate' interest is expected and it helps to guarantee submissions are being allocated to those who have an interest in submission related topics.

Accordingly, in this research, an integrated approach is proposed to recommend experts who are qualified to review submissions. Initially, the Author Topic Model (ATM), which is a famous approach to analyze topics with authors (Rosen-Zvi, et al., 2004), is applied to model the knowledge of reviewer candidates. With the help of the ATM, topic distributions in reviewer candidates' publications and submissions are estimated by the Expectation Maximization (EM) algorithm. Next, according to the estimated topic distributions, three indispensable aspects for RAP are considered, (1) the relevance, which evaluates the topical similarity between a reviewer candidate and a submission, (2) the interest trend of a reviewer candidate, which evaluates the degree of a reviewer candidate's willingness to review a submission, and (3) the authority of a reviewer candidate, which evaluates whether a reviewer candidate have a good recognition in submission-related topics. Finally, to balance these indispensable aspects with practical considerations, the problem of RAP is formulated as an integer linear programming problem.

The contributions of this study are at least threefold. First, the interest trend of a reviewer candidate is captured by analyzing reviewer candidates' publications. It is regarded as one of critical aspects to profile each reviewer candidate for RAP, which is a pioneer study to explore the effect of effective reviewer recommendation. Additionally, different from many approaches that concern about the academic authority globally, this study invites a PageRank based algorithm to estimate the topical authority of a reviewer candidate with respect to topics in each submission. Besides, a framework for RAP is illustrated with practical constraints in peer review and the problem is formulated as an integer linear programming problem. It makes that expected reviewers will be recommended efficiently with algorithms for optimization problems. Categories of comparative experiments were conducted on two large academic datasets with different parameter settings and promising results were obtained. It demonstrates the superiority of the proposed framework for RAP.

The rest of this paper is organized as follows. A brief review of relevant studies about RAP is given and relevant studies regarding expert interest modeling and expert authority estimation are introduced in Section 2. In Section 3, technical details about how to choose appropriate experts automatically for a set of submissions are explained. In Section 4, categories of experiments are conducted which aim to show the availability of the proposed approach. Finally, this research is summarized and potential future studies are highlighted in Section 5.

2. Related Work

2.1 Reviewer Assignment Problem

Conventionally, the task of reviewer assignment is treated as an expert retrieval problem and the topical relevance between reviewer candidates and submissions is the main consideration. First, publications of reviewer candidates are collected to represent one's knowledge. Next, submissions are modeled as queries. Finally, reviewers are selected according to the relevance between their knowledge and submissions. Hettich & Pazzani (2006) introduced a prototype application to identify prospective experts for proposals. In their research, the task of reviewer assignment was modeled as a retrieval problem. Then, each submission was regarded as an isolated query and the TF-IDF weighting was utilized to estimate the similarity between reviewer candidates and proposals. Similarly, given a query submission, a probabilistic language model was employed for ranking experts (Balog, Azzopardi, & De Rijke, 2006). Besides, with extracted features, some widely utilized machine learning algorithms were reported for RAP. For instance, Fang et al. (2010) treated the expert recommendation as a classification problem and the logistic model was utilized for expert determination. Zheng et al. (2013) extracted experts' multiple features by traditional retrieval methods, such as TF, TF-IDF, and language model. Next, the approach of learning to rank was applied to sort experts for a particular submission. Also, topic distributions of submissions and reviewer candidates' publications were estimated to compare the topical relevance. Karimzadehgan, Zhai, & Belford (2008) regards each submission as a combination of multiple subtopics. Three strategies for RAP were proposed to maximize the subtopic coverage of each submission in a complementary manner. Kou et al. (2015) analyzed the topic distribution in submissions and reviewers' publications by the ATM, and, according to the topic weights, a group of experts is recommended.

However, some researchers claimed that the topic relevance only is not adequate to select the most appropriate group of experts and other complementary aspects are focused on, such as the expertise and the authority of a reviewer, the knowledge diversity of reviewer group, conflicts of interest between reviewer and authors (COI). For example, both experts' expertise and relevance were considered in (Li & Watanabe, 2013), in which the authority and freshness were combined to estimate the expertise score and the bibliography and the referring information were combined to estimate the relevance score. Also, a convex optimization framework was formulated to select the most appropriate experts in Liu et al. (2014), which accounts for not only the authority and the expertise but also the diverse research background. Some researchers also took the COI into considerations for RAP (Liu, et al., 2016). Besides experts' knowledge and the COI, Li et al. (2015) proposed another approach for RAP, in which the stringent or lenient styles of reviewer candidates were explored.

In most of these approaches, semantic features of reviewer candidates and relations between reviewer candidates and authors are exploited. However, the research interest of an expert, which denotes the willingness to review submissions, has been ignored in these studies. But the research interest will significantly affect the quality of reviews in peer review system. Accordingly, in this study, the research interest is modeled as a substantial aspect of RAP.

2.2 Expert Interest Modelling

The expert interest in RAP usually refers to the degree of interest or willingness to review a submission on specific topics. In previous studies, generally, there are two types of methods on the modeling of experts' research interests.

Some studies regard an expert's research interest as a kind of expertise, which describes whether an expert has a certain degree of research experiences that are related to topics of submissions in a recent time (Daud, et al., 2010; Li & Watanabe, 2013). Li et al. (2013) argued that recent publications have a higher capability to represent experts' interest. Then, the time interval was considered and the recent publications were given a higher weight than older ones in the representation of research interest. Such approach is easily understood and implemented, which is welcome by practitioners for developing applications regarding reviewer recommendation.

Another type of modeling research interest is to ask experts themselves to indicate their degree of willingness about some prepared topics explicitly. Then these preferences are utilized as the prior knowledge for RAP (Rigaux, 2004; Di Mauro, Basile, & Ferilli, 2005). Rigaux et al. (2004) motivated each expert to express his/her preferences explicitly when they are invited to review submissions. According to their prior preferences, the techniques of collaborative filtering were used to predict their interests on different submissions. Manually assigned preferences labels are generally assumed to be able to describe experts' research interest more accurately. It induces that models based on manually provided labels are potentially accurate than text analysis based approaches. Di Mauro, Basile, & Ferilli (2005) described an expert system, named Global Review Assignment Processing Engine (GRAPE), which considers submission topics and experts' preferences for RAP. In the GRAPE system, preferences on submissions of all experts were initially collected. Next, these preferences were utilized to attune the prior assignment to experts.

These approaches indeed concentrate on reviewers' interest modeling according to experts' recent publications or expertise labels. As presented, however, it is risky to tune the length of a time window for the screen of an expert's publication list since that a smaller time window may bring the data sparse problem and the ineffectiveness on interest modeling while a larger time window may lose the trend of his/her research interest. Besides, prior knowledge based manually defined or provided explicit labels are often cumbersome to be obtained. Additionally, in increasingly elaborate research fields, it may suffer that different fine-grained topics are discussed in interdisciplinary research studies. Comparatively, in this study, the willingness of experts is modeled as an interest trend, in which the direction and the smoothness of interest trend are captured.

2.3 Reviewer Authority Estimation

The expert authority is often regarded as an indispensable consideration in RAP and it usually refers to the recognition of achievements in related fields. In this subsection, relevant studies on the measurement of expert authority are briefly summarised.

Many studies measured the authority of reviewer candidates according to traditional bibliometric, such as quotations, impact factor, etc. For example, Hirsch (2005) designed the H-index to evaluate the broad impact of a scientist's cumulative contributions, in which academic outputs and citations were combined to calculate the level of expert authority. Egghe (2006) proposed the G-index, which combined the cumulative contributions of previously cited papers and the number of citations of each paper, to compensate that the H-index is insensitive to highly cited papers. Similarly, Zhang (2013) developed the H'-index to improve the performance of the H-index.

Some studies regarded each expert as a node in a network and rank nodes according to the random walk algorithm (Petkova & Croft, 2008; Haveliwala, 2002; Kleinberg, 1999). For instance, Liu et al. (2014) assumed that some experts want to collaborate with those who have a higher authority. Then, an academic network was built based on the co-author relationship and a Random Walk with Restart (RWR) was applied to estimate the authority of an expert. Similarly, another co-author network was constructed to determine the impact of an individual author (Liu, et al., 2005). Also, the aspect of authority was explored in a heterogeneous network (Zhou, et al., 2007). It was constructed according to the social relations among authors and citation network among publications and a random walk was applied to rank authors and documents. Gollapalli, Mitra, & Giles (2011) argued that the venue information should be also considered as an important factor for RAP and a graph-based approach was proposed, in which authors, publications, and venue information were considered as nodes. Davoodi et al. (2013) applied the content-based recommendation algorithms to profile each expert and different social network measures were utilized to determine a particular member's value.

In these studies, some critical facts for RAP are ignored. For instance, the focus is on how to measure the global academic importance of an expert, rather than confined to some specific research topics. However, one expert's impact is highly possible not to be aligned with all fields if he/she has broad interests in different research topics. It can be referred that the topical authority with respect to a submission is conceptually different with global academic expertise. Such phenomenon is increasingly more prevalent on investigations that are required different fields of knowledge. Accordingly, to recommend a list of reviewers, an expert who have a higher recognition in submission-related topics should be given a higher priority. It motivates this study to explore topical authority regarding each submission for accurate reviewer recommendation.

2.4 A Brief Summary

Different aspects of experts were extracted and utilized to seek appropriate experts for submissions, such as topic relevance between experts and submissions, topic coverage for submissions, research impact of publications, diverse background of experts, etc. But few studies were reported on mining the research interest trends of reviewer candidates, which mirrors the willingness for reviewing submissions. Only a few pioneer studies considered expert's willingness. But these studies encourage experts to express their preferences with labels, which were used as the prior knowledge RAP and such expert intervention may result in inefficiency for reviews. Also, many relevant studies invite bibliometric indicators to evaluate the global academic authority. However, given a particular submission, rather than the global academic authority, recommended reviewers should possess a higher authority on submission related topics.

Hence, a detailed study regarding preferences of reviewer candidates is expected, in which the willingness to review submissions is highlighted. In this study, to profile each reviewer candidate, the interest trend is integrated. It describes an expert's research tendency over time about a specific topic, which is assumed to mirror his/her willingness to review a submission. Specifically, in this research, the interest trend regarding the direction and smooth tendency are extensively explored. Furthermore, different from approaches using bibliometric indicators for global academic authority, a PageRank based algorithm is proposed in this study to estimate the topical authority for each reviewer candidate with respect to a submission, which aims to find reviewer candidates who are specialized at topics of submissions. Together with the widely utilized relevance between a reviewer candidate and a submission, the interest trend and topical authority is integrated into a framework to recommend reviewer groups for a list of submissions.

3. Methodological Overview

3.1 Framework

As presented in Figure 1, a conventional work process of peer review requires that a particular associate editor need to handle a large volume of submissions one by one and he/she is familiar with topics in submissions as well as research interests of many reviewer candidates. It potentially leads to that some publications are assigned to invited experts for reviews, which are not fully satisfied with their research interests.

To ease the mismatch dilemma between submissions for and invited experts, a framework for reviewer assignment is proposed in Figure 2. Compared with the conventional workflow in Figure 1, the proposed framework aims to automate the selection of groups of proper experts corresponding to a list of submissions for review without the need to bother associate editors to recommend reviewers according to his/her understanding on submissions and familiarities with reviewer candidates' research interests. As presented in Figure 2, four major steps are involved, data collection, topic extraction, expert profile construction and optimization for reviewer assignment.

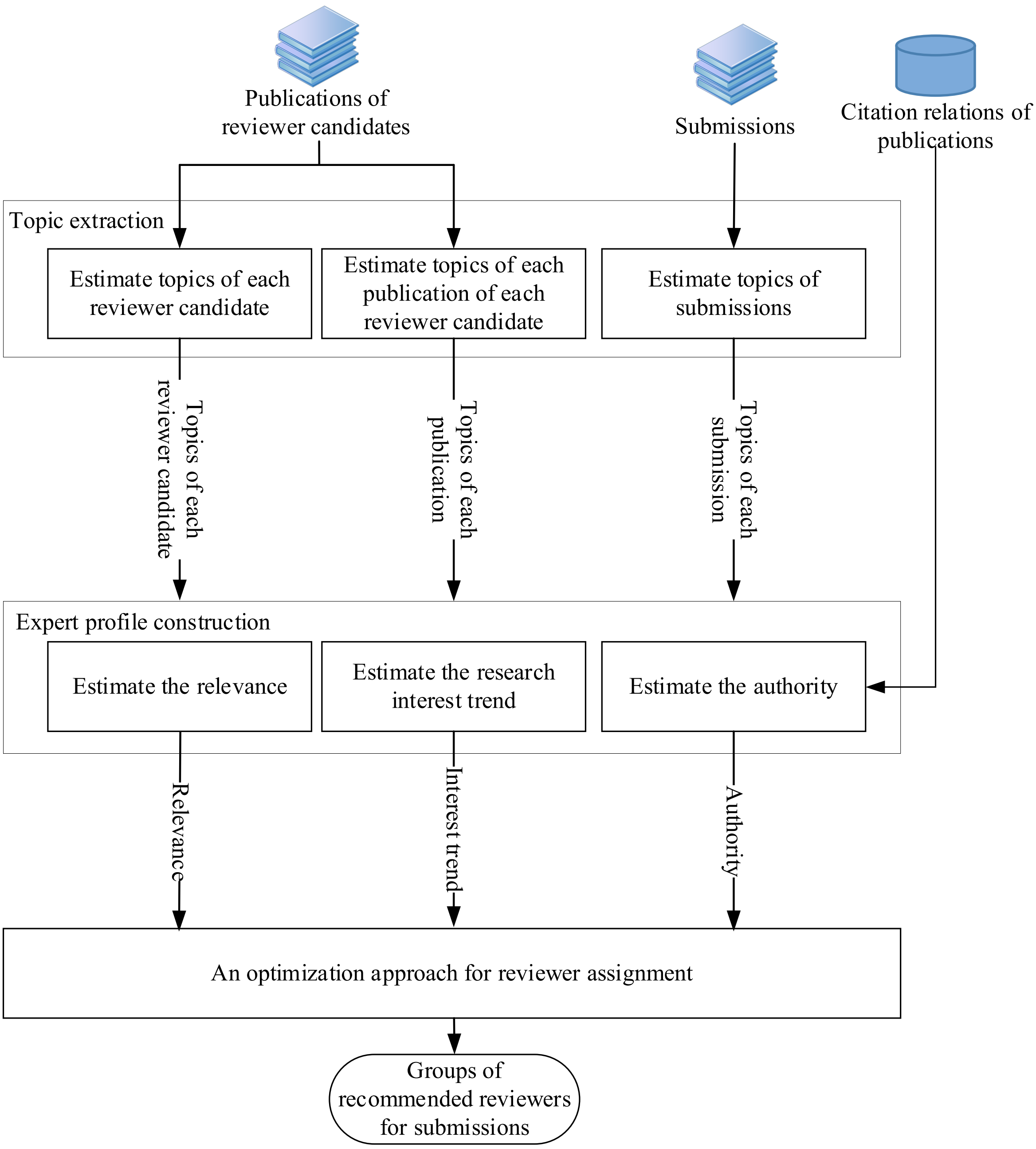


Figure 2. The framework for reviewer assignment

(1) **Data Collection.** In this research, two types of data are required to profile each reviewer candidate, including one's publications and the corresponding citation relations of publications. Indeed, an expert's publications mirror the knowledge and research interest directly, while citations help to understand his/her authority and mutual recognition in the academic area.

(2) **Topic Extraction.** The selection of reviewer candidates in experts' repository with appropriate knowledge often becomes the first concern. For instance, it is expected to ascertain who is a capable specialist to review submissions on some specific research topics. Accordingly, in this step, with the help of techniques on topic extraction, topic

distributions about experts' publications and submissions are extracted to describe the knowledge of each reviewer candidate and central topics of both each reviewer candidate's publication and each submission.

(3) **Expert Profile Construction.** In this step, three aspects are estimated to profile each candidate with respect to each submission, including the relevance of the candidate, the research interest trend of the candidate and the authority of the candidate. The relevance evaluates the similarity between each candidate and a given submission. The interest trend helps to distinguish the different types of temporal changes of each candidate's research interest regarding topics in a given submission, such as the stable upward trend or the fluctuating downward trend. The authority facilitates to estimate the authority of each candidate in submission-related topics with respect to a given submission.

(4) **An Optimization Approach for Reviewer Assignment.** In this step, an integer linear programming problem is formulated to balance the nominated three aspects. In addition, some practical constraints are considered to be embedded in this optimization problem, such as the maximal workload of each reviewer and the required number of reviewers to be selected for each submission.

For the clarity, all symbols in this research are summarised in Table 1.

Table 1 Input parameters and decision variables

Symbol	Description
T	Number of topics
n	Number of reviewers
m	Number of submissions
r_i	The i-th reviewer
v_i	Number of reviewer r_i's publications
r_i	The i-th reviewer
s_j	The j-th submission
p_u^i	The u-th publication of a reviewer r_i
$\overrightarrow{r_i}$	The topic distribution of a reviewer r_i's knowledge
$\overrightarrow{p_u^i}$	The topic distribution of p_u^i
$\overrightarrow{s_j}$	The topic distribution of a submission s_j
R_{ij}	The topic relevance between a reviewer r_i and a submission s_j
I_{ij}	The interest of a reviewer r_i on s_j
M_{ij}	The annual publication number of r_i that relates to s_j
A_{ij}	The authority degree of a reviewer r_i with respect to s_j
N_{ri}	The maximal workload that a reviewer r_i will review
Ns_j	The number of reviewers to be selected for a submission s_j
X_{ij}	A binary variable that indicates whether s_j is assigned to r_i

3.2 Topic Extraction

In the beginning, the techniques of topic modeling are utilized to capture topic distributions in submissions and in publications of each reviewer candidate, which helps to understand the concentration of submissions and the expertise of reviewer candidates.

Generally, assume that there are in total T topics and, accordingly, the knowledge of a reviewer candidate r_i, r_i's u-th publication p_u^i ($u \in [1, v_i]$, v_i is the total number of r_i's publications) and a submissions s_j can be modelled as T-dimensional vectors, which are denoted as $\overrightarrow{r_i} = (\overrightarrow{r_i}[1], \overrightarrow{r_i}[2], ..., \overrightarrow{r_i}[T])$, $\overrightarrow{p_u^i} = (\overrightarrow{p_u^i}[1], \overrightarrow{p_u^i}[2], ..., \overrightarrow{p_u^i}[T])$ and $\overrightarrow{s_j} = (\overrightarrow{s_j}[1], \overrightarrow{s_j}[2], ..., \overrightarrow{s_j}[T])$, respectively. For instance, $\overrightarrow{r_i}[t]$, $\overrightarrow{p_u^i}[t]$ and $\overrightarrow{s_j}[t]$ refers to the weight of topic t in r_i 's knowledge, the u-th publication of r_i and a submission s_j, respectively.

Given reviewer candidates' publications and a list of submissions, the techniques of topic modeling, such as Latent Dirichlet allocation (LDA) and ATM, which are a widely utilized for modelling an author's knowledge, can be applied to extract the defined topic vectors (Rosen-Zvi, et al., 2004; Blei, Ng, & Jordan, 2003). Indeed, many studies applied different algorithms to estimate topic distributions for experts, publications, and submissions (Mimno & McCallum, 2007; Karimzadehgan & Zhai, 2009; Tang, et al., 2012; Liu, Suel, & Memon, 2014). But the ATM is the most widely utilized for processing big scholarly textual data. Accordingly, in this research, given the entire set of publications of r_i , $\overrightarrow{r_i}$ is estimated by the ATM.

Note that, $\overrightarrow{p_u^i}$ and $\overrightarrow{s_j}$ cannot be obtained directly from the result of ATM. However, motivated by studies in (Kou, et al., 2015; Zhai, Velivelli, & Yu, 2004), according to the result of ATM, the EM algorithm can be utilized to estimate $\overrightarrow{p_u^i}$ and $\overrightarrow{s_j}$ without needs to execute other models to analyze topic distributions in p_u^i and s_j. Particularly, in the EM algorithm, both publications and submissions are referred as "documents". For a specific document d, the E-step and the M-step can be derived as follows,

E-step:

$$p^{(n+1)}(z_{wi} = t) = \frac{\bar{p}[t]^{(n)} p^{(n)}(w_i \mid t)}{\sum_{t=1}^{T} \bar{p}[t']^{(n)} p^{(n)}(w_i \mid t')} \tag{1}$$

M-step:

$$\bar{p}[t]^{(n+1)} = \frac{\sum_{w \in W} c(w,d) p^{(n+1)}(z_w = t)}{\sum_{t'=1}^{T} \sum_{w \in W} c(w,d) p^{(n+1)}(z_w = t')} \tag{2}$$

As known, the EM algorithm is an iterative algorithm. In the E-step, the probability that a word w_i is assigned to the t-th topic $p\ (z_{wi}=t)$ is estimated. $p^{(n+1)}\ (z_{wi}=t)$ refers to the estimated probability in the (n + 1)-th iterations. In the M-step, the weight of topic t in a document, $\vec{p}[t]$, is estimated by maximizing the expectation. Similarly, $\vec{p}[t]^{(n+1)}$ refers to the estimated weight of topic t in the (n + 1)-th iterations. $p(w_i|t)$ refers to the probability that a word w_i is generated by topic t. t' denotes the iterative variable regarding each topic. $c(w,d)$ denotes the number of word w in document d.

3.3 Reviewer Profile Construction

In this subsection, three aspects of reviewer candidates are introduced to profile each reviewer candidate. The first aspect is the topic relevance between a reviewer candidate and a submission, which usually becomes the major consideration for RAP. Also, a high-quality review may be provided if the recommended reviewer is an expert in submission-related topics. Finally, he/she is expected to have a higher degree of research interest in that topics. Accordingly, in the following, the focus is on how to model the research interest and the authority of a reviewer candidate.

3.3.1 The Reviewer-Submission Relevance

First, reviewer candidates who have strong knowledge about topics that are discussed in submissions should be given a higher priority. Hence, the distance between the topic distribution of s_j and that of r_i's knowledge should be estimated.

Specifically, as presented, the topic distribution of s_j and that of r_i's knowledge are represented as two T-dimensional probability vectors $\vec{s_j}$ and $\vec{r_i}$. Then, the famous measure of the difference between two probability distributions, the Kullback-Leibler divergence (KL divergence), can be employed to estimate the dissimilarity between s_j and r_i. In particular, the KL divergence is invited to evaluate the distance between two distributions. Mathematically, it can be denoted as,

$$D(\vec{s_j} \parallel \vec{r_i}) = \sum_{i,j=1}^{T} \vec{s_j} \log \frac{\vec{s_j}}{\vec{r_i}} \tag{3}$$

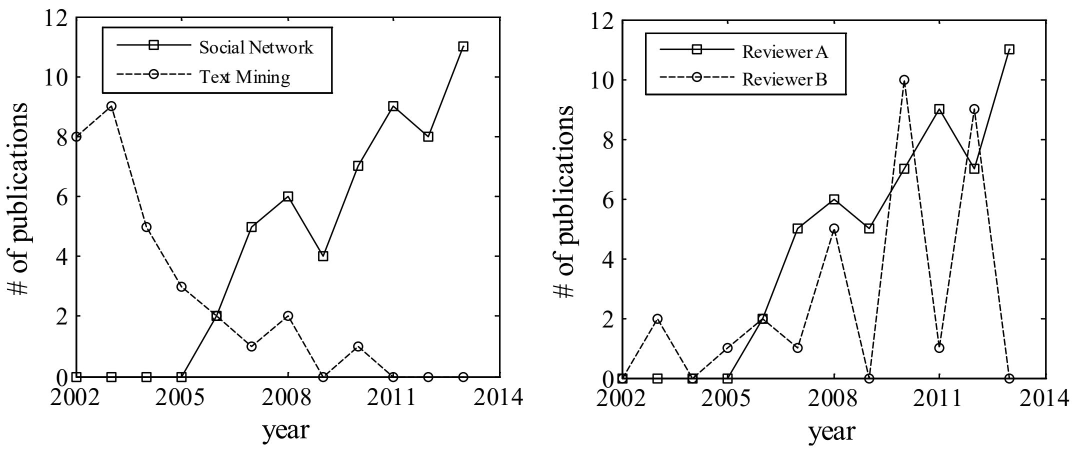
Note that a smaller KL divergence between them shows a bigger relevance. Hence, the relevance between r_i and s_j, R_{ij}, can be defined as,

$$R_{ij} = -D(\vec{s_j} \parallel \vec{r_i}) \tag{4}$$

3.3.2 The Interest Trend of Reviewer Candidates

It is well known that the research interest of a reviewer may change over time, corresponding to the research hotspot or other reasons, which induce that the effectiveness of peer review may be affected.

Assume that, for a specific topic, if a particular reviewer r_i contributes some related papers, he/she might have a certain degree of research interest. Accordingly, the number of papers over time might be a good indicator to reveal the research interest trend of a reviewer. For instance, one exemplary interest trend about a reviewer is shown in Figure 3(a). As presented, an increasing number of papers are observed in the solid line, which shows that this reviewer has greater interest about the social network and he/she is expected to have strong willingness to review submissions on that topics. Similarly, the dashed line appears a contrary appearance. In addition to the direction of a reviewer's interest trend, the stability is another critical factor. Specifically, for a given topic, those reviewer candidates with an upward trend in publication number are preferred to be invited, and additionally, if the trend is smooth and upward, then the candidate is highly preferred. A toy example is shown in Figure 3(b). As seen from this figure, compared with two review candidates A and B, A is preferable since that a smooth and upward trend of research interest is observed.



(a)the direction of interest trend (b)the smoothness of interest trend

Figure 3. Two examples about the differences of reviewers' interest trends

Note that both the cumulative publication number and the annual publication number of an expert in specific topics can be utilized to denote the change of his/her research interest. However, the cumulative publication number stresses the absolute volume of his/her publications, which is better to model the authority in that specific topics. Comparatively, the annual publication number of an expert is a good indicator that focuses on the change of his/her research interest every year. In this study, it is assumed that those reviewers who have more publications in recent years have a higher research interest to review submissions. Hence, the annual publication number of a particular reviewer on specific topics is explored. Another concern to model the research interest is to reckon the number of citations in each year. However, it is argued that the number of citations in each year is perhaps not a good indicator to represent one's research interest trend. For instance, some researcher might have few but classical papers which receive a large number of citations annually in some specific topics. Critical problems are discussed in these classical papers, which help him/her to gain academic reputations but do not help to demonstrate his/her recent research interests. Similarly, suppose that another expert holds a strong publication list in reputable journals with higher impact factors regarding these specific topics. Then, it only shows that he/she has a higher authority in terms of these topics, but it cannot be inferred that he/she holds an obvious recent research interest or a continuous research interest on that topics.

Hence, according to approaches for the analysis of time series data (Wei, 1994), two indicators, including the direction D_{ij} and the smoothness of an interest trend Q_{ij}, are investigated to describe the research interest of a reviewer r_i on topics that are discussed in a submission s_j, quantitatively. However, before two indicators are explained quantitatively, the annual publication number of r_i that relates to s_j, denoted by M_{ij}, which is assumed to be the research interest trend, should be obtained. It can be estimated in the following three steps.

(a) First, top N important topics in $\overrightarrow{p_u^i}$ are extracted to represent r_i's u-th publications. Similarly, top N important topics in $\vec{s_j}$ can be also gained to denoted s_j.

(b) Then, all publications of r_i can be indexed according to T different topics, in which the t-th group is denoted as c_t^i. Thus, publications of r_i that relates to top N important topics in s_j can be selected.

(c) Next, selected publications are sorted by year and, accordingly, the annual publication number that relates to s_j, M_{ij}, can be obtained.

(1) The direction of r_i 's interest trend that relates to a submission s_j, D_{ij}

In this research, a linear relation is utilized to capture the direction of interest trend. Specifically, for a particular submission s_j, the least squares can be applied to estimate the slope and the interception of the linear relation between the annual publication number M_{ij} and the year of publication. Then, the estimated slope of the annual publications is utilized to estimate D_{ij}. If D_{ij} is positive, it means that an increasing trend about the annual publication number is observed, which implies that r_i 's interest over the research topics on s_j is rising.

(2) The smoothness of r_i 's interest trend that relates to a submission s_j, Q_{ij}

Generally, both the standard deviation and the mean of M_{ij} need to be considered for Q_{ij}. Intuitively, if a larger mean of M_{ij} with a smaller standard deviation is observed over time, it implies that the annual changes tend to be stable. In this study, the coefficient of variation (Abdi, 2010), which evaluates the degree of temporal changes, is utilized to define the interest trend, V_{ij}.

$$V_{ij} = \frac{sd(M_{ij})}{mean(M_{ij})} \tag{5}$$

A smaller V_{ij} means the interest trend tends to be stable and, hence, the smoothness Q_{ij} can be defined as,

$$Q_{ij} = e^{-\eta V_{ij}} \tag{6}$$

η is the magnification factor that controls the weights of Q_{ij}.

Intuitively, compared with one reviewer candidate whose publication number declines obviously or fluctuate significantly, a candidate whose publication number exhibits a smooth upward increase over time should be given a higher weight since he/she is expected to have a deeper sustained interest on topics in s_j. Hence, candidates whose interest trend with a smooth upward increase are given higher priorities. Accordingly, the interest trend of a reviewer r_i 's interest trend on s_j, I_{ij}, is modeled as,

$$I_{ij} = D_{ij} \times Q_{ij} \tag{7}$$

In the following, an illustrative example is presented to show how these factors are estimated explicitly. Take three reviewers A1, A2, and A3 in Figure 4 for instance. For a particular topic, interest trends of three reviewers vary substantially. As presented in Figure 4, research interests of both reviewers A1 and A2 have an upward trend, whereas A3’s interest trend is found to be significantly decreasing. Additionally, A1’s interest trend is found to be smoother, while A2's interest trend's degree of fluctuation is higher than that of A1. According to Equation (7), scores of interest trends about these three reviewers can be calculated. Then, A1 has a score of 1.34, A2 has a score of 0.71, and A3 has a score of -1.06. The result, A1 > A2 >A3, does make intuitive sense according to Figure 4.

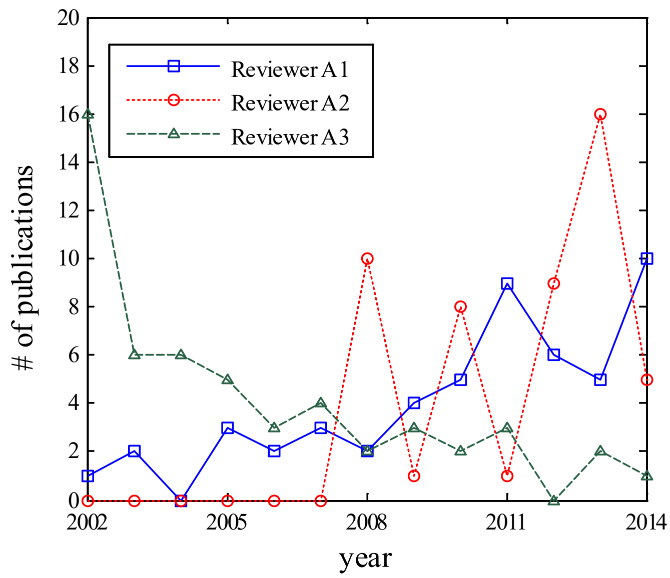


Figure 4. The interest trends of three exemplary reviewer candidates

3.3.3 The Authority Degree of Reviewer Candidates

Many authority-based approaches aim to recommend experts who have a higher recognition in the research field. However, some widely used indexes, such as the H-index, only focus on the global authority. It ignores that experts may investigate different topics at the same time and the level of expert's authority might not be the same on different topics. Accordingly, in this subsection, different levels of authority are distinguished.

Practically, being cited by multiple experts is a good indicator that mirrors the authority of an expert. Now, suppose a reviewer candidate's research topics are highly related to submission topics and this candidate is being cited many times. It means that this candidate is expected to have a higher authority regarding these topics. Furthermore, if this candidate is cited by many authoritative experts on related topics, this candidate is considered to have a higher authority on those topics. Accordingly, motivated by studies in (Wang, et al., 2013; Liu, Suel, & Memon, 2014; Hu, Fang, & Godavarthy, 2013), in this study, a topical PageRank is presented to identify high-authority experts with given topics.

In this subsection, citation activities are formalized in a topical authority graph. In Figure 5, a sample of citation activities among authors and papers are shown. The dashed line denotes the authorship between publications and authors. For instance, r_2 and r_3 co-authored one paper p_3, while r_2 is the only author of p_2. Besides, in the upper part of Figure 5, citations among papers are illustrated, in which each node represents a paper and each edge denotes a citation between two papers. According to the citation relation among papers, citations among authors can be deduced, which is represented in the lower part. In the lower part of Figure 5, each node represents an author and each edge denotes a citation between them.

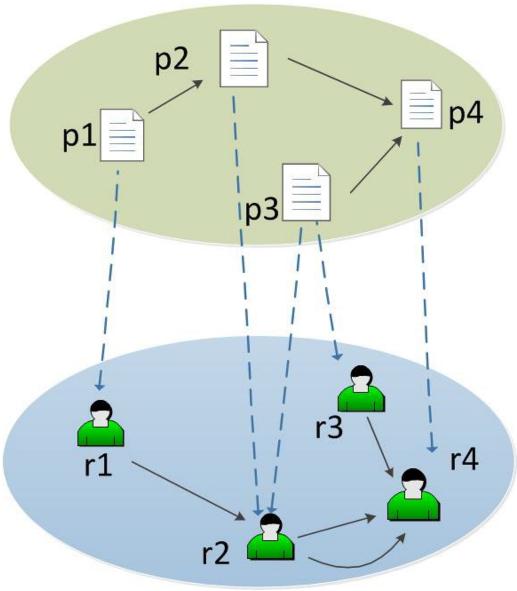


Figure 5. The citation network regarding papers and reviewer candidates

Notably, there might exist more than one edge between expert nodes, which indicates that more citation activities are observed. According to the rationale of PageRank, the authority degree of an expert r_i with respect to a submission s_j, A_{ij}, can be estimated according to the iterative algorithm in Equation (8).

$$A_{ij}^{'} \leftarrow a \sum_{q \in ref(r_i)} w_q^j \times A_{ij} \times B_{qi} + (1-a)\frac{1}{N} \tag{8}$$

$ref(r_i)$ denotes a group of experts who cite r_i 's publications. w_q^j refers to the topical similarity between the q-th expert r_q and s_j. Quantitatively, it is evaluated by the cosine distance between $\vec{r_q}$ and $\vec{s_j}$, which are estimated by the EM algorithm in Section 3.2. B_{qi} denotes the quantity of r_q 's authority that is transferred to r_i, which is estimated according to the ratio between how many times r_q references r_i and the total number of r_q references all reviewer candidates. The parameter a is the damping factor, which compensates for experts with none reference papers in the expert network. N is the total number of expert nodes in the network.

3.4 An Optimisation Approach for Reviewer Assignment

Notice that, in this research, the ultimate goal is to assign appropriate reviewers for a set of submissions. To balance all the nominated three aspects of RAP, an optimization problem is formulated.

Besides, other two practical constraints should be considered. The first one is the maximal workload of a reviewer candidate, which implies that he/she should not be given too many submissions. The other is that each submission should be guaranteed to be assigned to a certain number of reviewers.

Suppose there are a set of reviewer candidates $\{r_1,....,r_n\}$ and a set of submissions $\{s_1,.....,s_m\}$. An $n \times m$ binary matrix, X, where X_{ij} is a binary variable that indicates whether s_j is assigned to r_i, can be utilized to denote the final recommendation result. Then, the reviewer assignment can be accordingly formulated as an integer linear programming problem, which is represented in Equation (9) - Equation (12).

$$\max \quad \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}X_{ij} + \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}X_{ij} + \lambda_3 \sum_{i=1}^{n}\sum_{j=1}^{m} R_{ij}X_{ij} \tag{9}$$

$$s.t. \quad \sum_{j=1}^{m} X_{ij} \le N_{R_i} \qquad \forall i \in [1,\mathrm{n}] \tag{10}$$

$$\sum_{i=1}^{n} X_{ij} = N_{S_j} \qquad \forall j \in [1,\mathrm{m}] \tag{11}$$

$$X_{ij} \in \{0,1\} \qquad \forall i \in [1,\mathrm{n}], \forall j \in [1,\mathrm{m}] \tag{12}$$

Equation (9) describes the objective which aims to maximize the authority A_{ij}, the research interest I_{ij} and the topic relevance R_{ij} between r_i and s_j at the same time. λ_1, λ_2 and λ_3 are the magnification factors that balance different aspects of reviewers. Equation (10) describes that submissions that are assigned to each reviewer r_i for reviews at the same time should be no more than N_{ri}. Equation (11) describes that each submission s_j should be assigned to a predefined number of reviewers, Ns_j, for review comments. Equation (12) indicates the selection of each decision variable X_{ij}, where X_{ij} equals one means that a reviewer r_i should be assigned to a submission s_j. Finally, this optimization problem can be solved efficiently by many software packages for the integer linear programming.

4. Experiments

4.1 Datasets

In this experiment, two datasets are used to validate the effectiveness of the proposed approach. The first dataset was built from the WANFANG Database, which is a famous Chinese scientific library. To build this dataset, 256 scholars with funding support in the subject of information system and management during the period of 2004 to 2013 were obtained from the official website of the National Natural Science Foundation of China (NSFC). Next, 256 scholars were used as seed scholars and their 5,462 collaborators were found in WANFANG Database with 75,023 papers. The second dataset was built from a subset of ArnetMiner (Tang, et al., 2008). It includes 1,712,433 scholars and 2,092,356 abstracts. Next, scholars with more than 40 publications were selected as experts, which makes that 6,173 experts with 427,575 papers were obtained. In each dataset, 500 experts were chosen randomly to build a reviewer candidate repository. Other experts and papers were used to train the proposed model.

4.2 Evaluation Metrics

Generally, for RAP, it is difficult to build a manually labeled dataset accurately, even experienced editors are invited. It makes that some evaluation metrics, such as precision and recall, are not applicable. To tackle this dilemma, different evaluation metrics are proposed according to the problem definition of their own. For instance, in Liu et al. (2014), the relevance, the authority as well as their research background were utilized for the benchmark. Some similar evaluation methods can be also found (Tang, et al., 2012; Tang, Tang, & Tan, 2010).

Similarly, in this research, a group of persuadable evaluation metrics is introduced according to the nominated three aspects. In particular, *Distance@k*, *Interest@k,* and *Authority@k* are utilized as evaluation metrics, which refer to the corresponding score if k submissions are concerned and N_{sj} reviewers are required for each submission.

(1) Distance

First, the distance or relevance between submissions and recommended candidates is widely utilized to evaluate the performance of RAP. In this research, the distance between k submissions and recommended reviewer candidates is expected to be minimized and it is defined as,

$$Distance@k = \sum_{j=1}^{k}\sum_{i=1}^{Nsj} | R_{ij} | \tag{13}$$

R_{ij} is the relevance between the j-th submission and the i-th reviewer, and N_{sj} refers to the number of reviewers required by the j-th submission. *Distance@k* evaluates the sum of the distance between k submissions and recommended reviewer candidates.

(2) Interest Trend

Additionally, it is desirable to invite reviewers with a higher degree of interests so that they are expected to have higher preferences to review those submissions. Given k submissions, the sum of interest trend can be denoted as,

$$Interest@k = \sum_{j=1}^{k}\sum_{i=1}^{Nsj} I_{ij} \tag{14}$$

I_{ij} is the level of interest trend of the i-th reviewer on the j-th submissions according to Equation (7). *Interest@k* calculates the sum of recommended reviewer candidates' interests on all k submissions.

(3) Authority

Last, the total authority of recommended reviewer candidates on submissions is expected to be maximized. In this research, the total authority is estimated as,

$$Authority@k = \sum_{j=1}^{k}\sum_{i=1}^{Nsj} A_{ij} \tag{15}$$

A_{ij} is the authority of the j-th reviewer on the i-th submission that is defined in Equation (8). *Authority@k* evaluates the total authority of recommended reviewer candidates on all k assignments.

4.3 Benchmarking methods

To benchmark the performance in terms of different evaluation metrics, three popular methods are utilized. Admittedly, these approaches are not state-of-the-art algorithms. But they are the most famous approaches for RAP and widely utilized for benchmarks. In the future, other benchmarking algorithms with different considerations will be testified, which helps to polish the performance of the proposed approach.

(1) Vector Space Model

The Vector Space Model (VSM) is a well-known method to measure the text similarity. For RAP, the VSM is utilized to measure the relevance between submissions and reviewer candidates' knowledge. For instance, the VSM based approach was initially utilized for expert recommendation in (Yukawa et al., 2001). In (Yukawa, 2001), the query document and experts' publication are represented as vectors according to VSM and the cosine similarity is utilized as the retrieval function to evaluate the similarity between two vectors.

(2) Language Model

The Language Model is another frequently-used approach in the field of information retrieval. For the RAP, the language model is utilized to rank reviewer candidates for a given set of submissions. In Cao (2005), query words are firstly utilized to find relevant experts' documents set by a simple word matching scheme and a language model based approach is applied to estimate the relevance between query words and each document. Also, according to the language model, an expert centered model, and a document centered model were proposed for expert recommendation in (Balog et al., 2006, 2009). The fundamental language model-based approach can be denoted as,

$$p(q \mid r_i) = \prod_{t \in q} ((1-\lambda) p(t \mid d) + \lambda p(t))^{n(t,q)} \tag{16}$$

q refers to a submission. r_i is a reviewer candidate and d is the combination of r_i authored publications. $n(t, q)$ refers to the number of terms in a submission q.

(3) Author Topic Model

The ATM is an unsupervised learning method to extract authors' research topics from a corpus. In RAP, experts are ranked according to the probability that the word set of a submission is generated. For instance, Kou et al. (Kou et al., 2015) considered the topic distribution in different submissions. First, multiple topics of experts and submissions were extracted by ATM. Next, a group of experts were recommended, according to the different weights topics in submissions. The ATM based approach can be calculated as,

$$p(a_i \mid W) = \prod_{w_q \in W} \sum_{t_j \in T} p(a_i \mid t_j) p(t_j \mid w_q) \tag{17}$$

a_i is one reviewer candidate. W is the word set of a submission, where w_q is the q-th word in W. T is the topic set that is extracted by the ATM. $p(a_i|W)$ is the probability that a_i can be observed given a topic t_j and $p(t_j|w_q)$ is the probability that a topic t_j is observed given a word w_q.

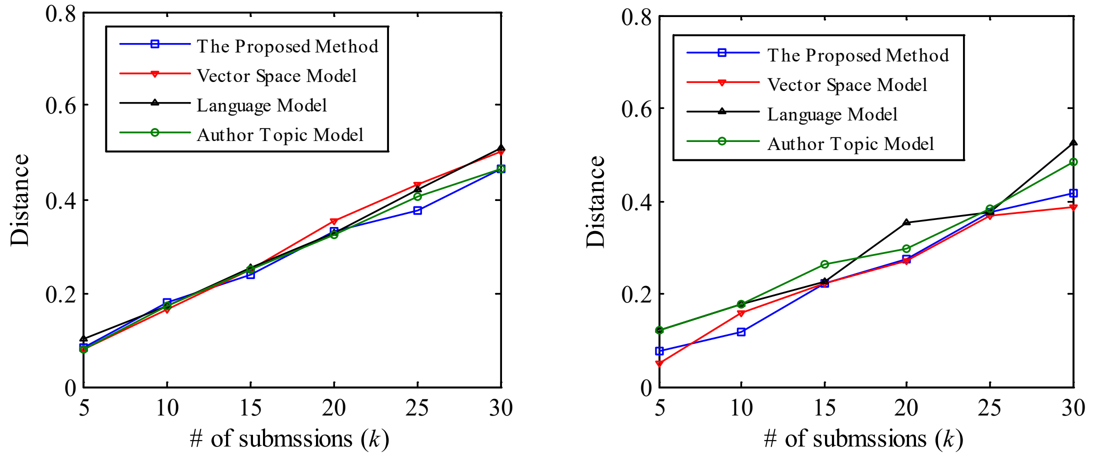
4.4 Experimental Results

In this section, three major parameters k, N_{sj}, and N_{ri} are testified to evaluate the effectiveness of the proposed approach. Meanwhile, different approaches are compared in terms of three evaluation metrics on both datasets. In the following experiments, three aspects of reviewers are regarded to be equally important. Hence, λ_1, λ_2 and λ_3 are set 1. The parameter a in Equation (8) is a damping factor. Like the algorithm of PageRank, it can be set to be 0.85. The parameter η with respect to the interest trend in Equation (6) is a magnification factor that controls the weight of coefficient of variance Q_{ij} about the research interest trend. In this study, η is chosen to one, which means the coefficient of variance Q_{ij} has an equal weight with the estimated direction parameter D_{ij}.

In Figure 6, different approaches are compared in terms of *Distance@k* on the WANFANG Dataset and the ArnetMiner Dataset, where the horizontal axis represents the number of submissions and the vertical axis represents the corresponding topical distance. Note that, in this experiment, for each submission, four reviewers are invited and the maximal workload for each reviewer is also set to four. Actually, similar phenomena can be observed if different number of reviewers and different number of maximal workload are set. As seen from this figure, given a fixed number of reviewers, the topical distance between reviewers and submissions increase gradually if more submissions are considered. Perhaps that more research topics tend to be discussed in more submissions and it induces that some topics potentially fail to be covered by a fixed number of reviewers. Also, no significant differences in these approaches are observed since that, in all four approaches, the relevance between reviewers and submissions is all reckoned as a major concern.

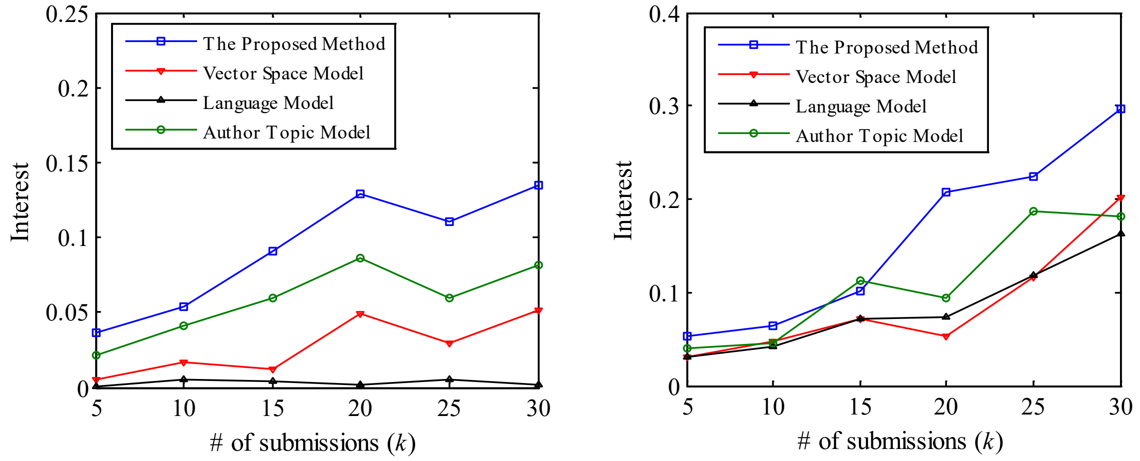
In Figure 7, the proposed approach and three benchmarking approaches are compared in terms of *Interest@k*. As seen from two subgraphs, the proposed approach performs much better than the other three. It shows the obvious strength of the proposed approach in understanding the research interest of reviewer candidates.

In Figure 8, four approaches are compared in terms of *Authority@k*. As presented, the proposed approach also outperforms the other three approaches. It demonstrates that, the topical authority of recommended reviewer candidates is relatively higher, compared with other three benchmarking approaches.



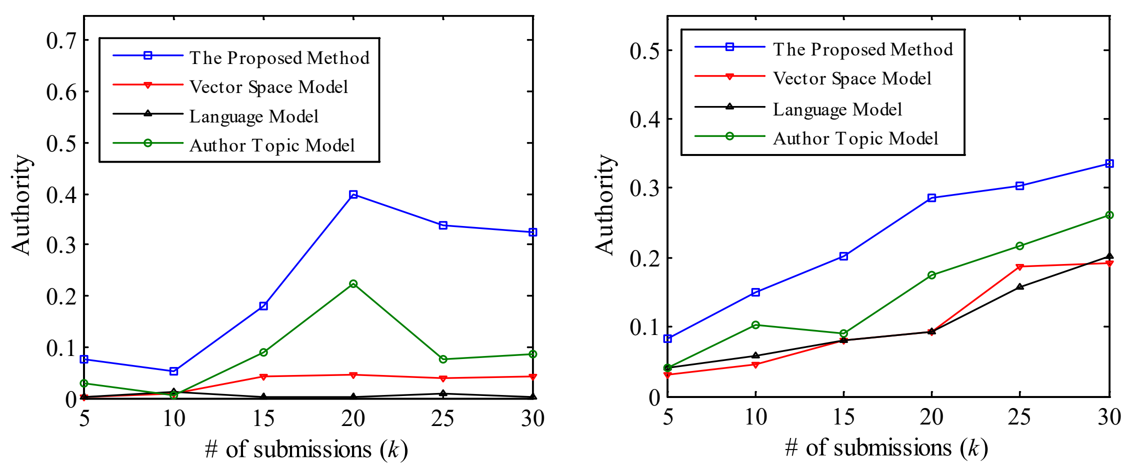
(a)WANFANG Dataset (b) ArnetMiner Dataset

Figure 6. Performance comparisons on *Distance@k*



(a)WANFANG Dataset (b) ArnetMiner Dataset

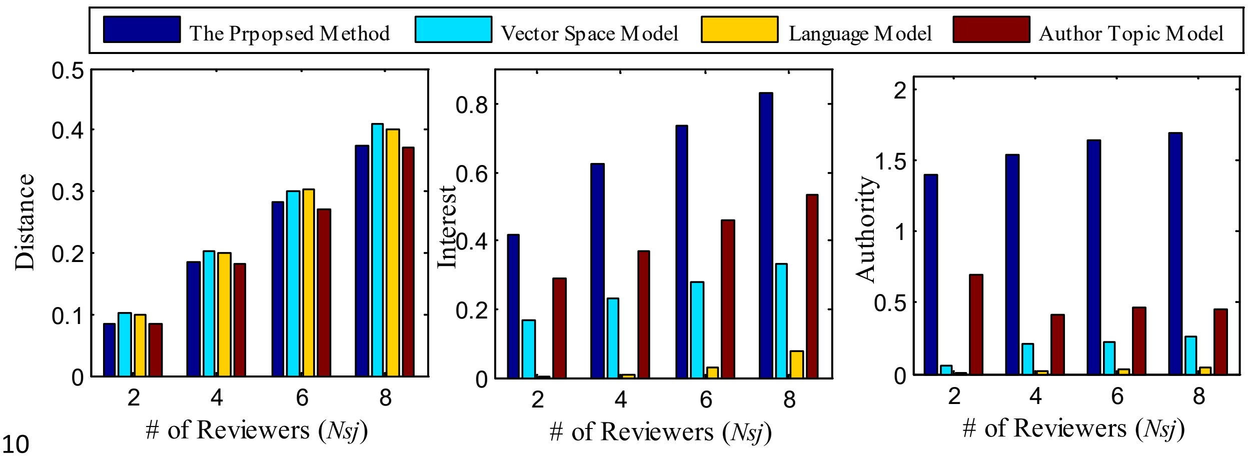
Figure 7. Performance comparisons on *Interest@k*



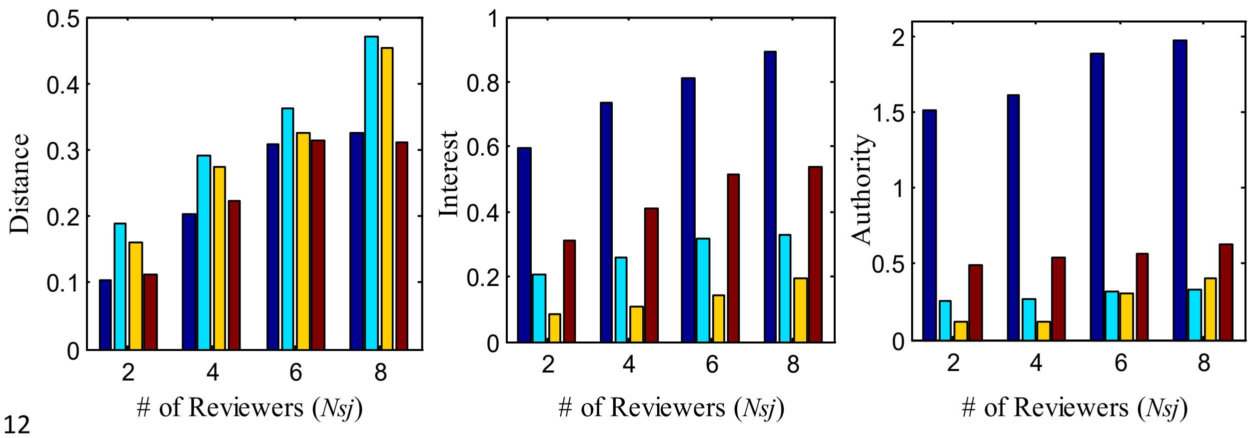
(a)WANFANG Dataset (b) ArnetMiner Dataset

Figure 8. Performance comparisons on *Authority@k*

Figure 9 shows the performance comparison of both datasets with different predefined number of reviewers, when the number of submissions is set to 30 and the maximal workload is set to 4. With an increasing number of reviewers, values of all three metrics of the proposed method start to rise in both datasets. Compared with other three approaches, the degree of interest and authority of the proposed method are improved significantly, though a marginal improvement is reported in terms of the topical distance. It can be deduced that the proposed method presents constantly competitive performance regarding the topical relevance between submissions and recommended reviewer candidates for different predefined number of reviewers.



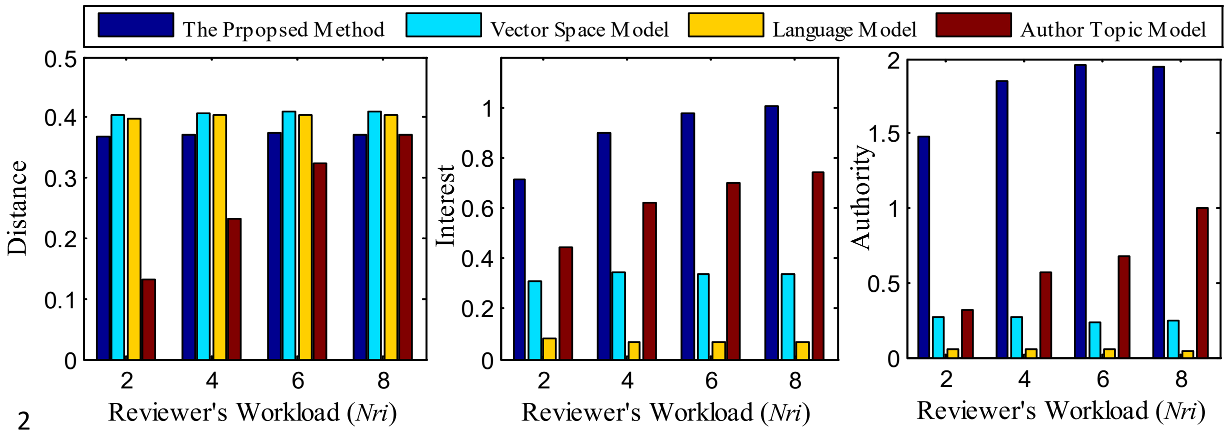
(a)WANFANG Dataset



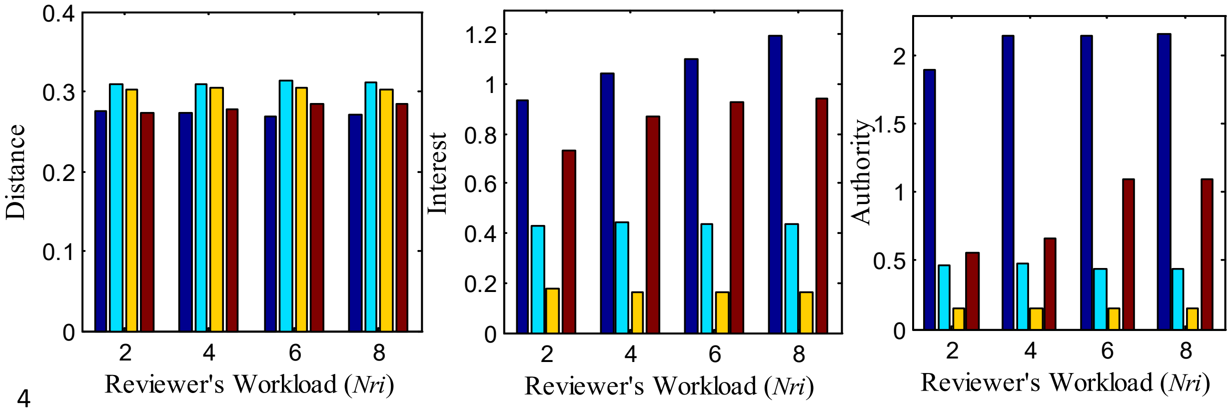
(b) ArnetMiner Dataset

Figure 9. Performance comparisons with different number of reviewers

Meanwhile, effects of the maximal workload of each reviewer are represented in Figure 10. In this experiment, 30 submissions and 4 reviewers are considered. Compared with Figure 9, some similar phenomena are observed, i.e. improved levels of the research interest and the authority with competitive topical relevance.



(a)WANFANG Dataset



(b) ArnetMiner Dataset

Figure 10. Performance comparisons with different number of maximal workload

One of the most significant differences between the proposed method and others lies in the considerations about reviewer’s research interest trend in a fine-grained manner, while the relevance and the authority are paid much attention in many other approaches. As presented in these experiments, all three benchmark methods potentially neglect to capture the reviewer's interest. Comparatively, in the proposed approach, the research interest trend of recommended reviewers is largely improved with not an apparent loss in terms of the relevance between submissions and recommended reviewers.

4.5 A Case Study

To demonstrate how the proposed framework can facilitate the editor-in-chief regarding reviewer assignment, an illustrative example is presented in Figure 11.

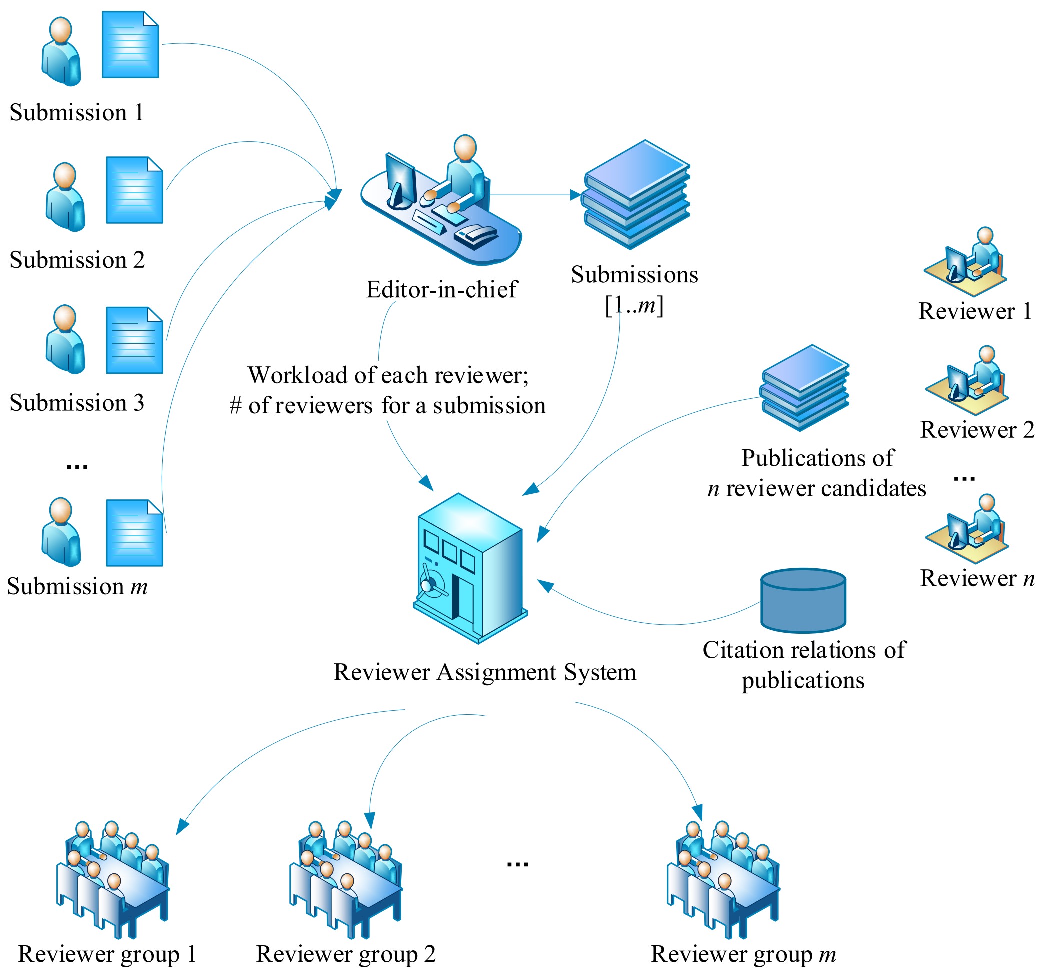


Figure 11. An example of how the proposed framework helps on reviewer assignment

In this example, m manuscripts are submitted to the editor-in-chief at the same time. Next, the editor-in-chief extracts abstracts of m submissions from the submission system and extracts abstracts of publications authored by all n reviewer candidates working for the journal from online academic databases. These abstracts are then sent to the reviewer assignment system which is developed based on the proposed framework. Besides, the editor-in-chief provides the maximal workload of each reviewer as well as the number of reviewers for each submission to the system. Note that, as explained, supporting data of this system all include citation relations between publications which aims to estimate the relative authority of each reviewer with respect to each submission. Then, three modules in the system start to extract topics for m submissions, n reviewer candidates as well as their publications, profile n reviewer candidates from three aspects and execute an algorithm for analyzing the integer linear programming problem. Finally, m groups of reviewers are recommended by the system.

5. Conclusion and Future Work

With an increasing number of submissions and experts, finding proper reviewers to evaluate the quality of submissions becomes obviously cumbersome. It induces that RAP receives much attention in the academic filed and appears increasingly more critical in R&D project selection, scientific evaluation, company recruitment, etc.

As presented, many studies regarding RAP focus on the relevance between recommended reviewers and submissions as well as the authority of reviewer candidates. Most of them ignore experts' interest, a critical aspect in peer review, which indicates the willingness to review submissions. Accordingly, in this study, besides the widely concerned relevance, the relative authority of reviewer candidates and the research interest trend of each reviewer candidate with respect to each submission are taken into considerations. Different from previous studies that employed bibliometric based approaches, according to publication citation relations, a topical PageRank algorithm is introduced to estimate the relative authority of each reviewer candidate with respect to each submission. In addition, the direction of interest trend and its smoothness are embedded to model the research interest trend of each reviewer candidate with respect to each submission. The direction of interest trend reckons whether a reviewer's interest over the research topics of a submission is rising, while the smoothness gauges a candidate whose publication number exhibits a smooth upward increase over time regarding the research topics of a submission. Finally, based on the relevance, the research interest trend, and the relative authority, the reviewer assignment is modeled as an integer linear programming problem with different practical concerns. Categories of experiments were conducted on two real scholarly datasets with a large number of experts as well as their publications, which demonstrate the effectiveness of the proposed approach. Specifically, compared with other approaches for reviewer recommendation, such as VSM, language model, and ATM, the proposed framework enhances the research interest trend and the relative authority of recommended reviewers significantly without an obvious loss in terms of the relevance between submissions and reviewers. Also, the stability of strength over the research interest trend and the relative authority are testified by categories of experiments over different number of reviewers and different number of maximal workload.

In the future, different benchmark approaches will be evaluated, which will help to improve the quality of the proposed approach and promote it to be applied in a real expert recommendation system. Also, the proposed approach is planned to be evaluated in different datasets in other research fields, such as DBLP (Digital Bibliography & Library Project) in computer and information science, APS (American Physical Society) in physics science, etc. In addition, besides the nominated three aspects, other practical factors are welcome to be reckoned in the designed system, such as relations between reviewer candidates and authors, meta information about research topics of publications that are manually labeled in the scientific library, etc.

Acknowledgment

This research was supported by the youth fund project of ministry of education of the humanities and social sciences research (NO. 16YJC870006), the fundamental research funds for the central universities (NO. SKZZB2014037) and the Guangdong Soft Science Project (No. 2016A070705019).

References

Abdi, H. (2010). Coefficient of variation. In N. J. Salkind, D. M. Dougherty, & B. Frey (Eds.), *Encyclopedia of Research Design*, 169-171. Thousand Oaks, CA: SAGE Publications.

Balog, K., Azzopardi, L., & Rijke, M. d. (2006). Formal models for expert finding in enterprise corpora. In *SIGIR* '06, 43-50, Seattle, WA, Aug 06-11

Balog, K., Azzopardi, L., & Rijke, M. d. (2009). A language modeling framework for expert finding. *Information Processing and Management, 45*(1), 1-19.

Basu, C., Hirsh, H., Cohen, W. W., & Nevill-Manning, C. (1999). Recommending papers by mining the web. In *IJCAI* '99, 1-11, Stockholm, Sweden, Jul 31-Aug 06.

Biswas, H. K., & Hasan, M. M. (2007). Using Publications and Domain Knowledge to Build Research Profiles: An Application in Automatic Reviewer Assignment. In *Proceedings of the 2007 International Conference on Information and Communication Technology*, 82-86, Dhaka, Bangladesh, Mar 07-09.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of machine Learning research,* 3(Jan), 993-1022.

Blei, D. M. and Lafferty, J. D. (2006) Dynamic topic models. In *ICML* '06, 113-120, Pittsburgh, PA. Jun 25-29.

Brin, S. and Page, L. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*. 30: 107–117, 1998

Cao, Y., Liu, J., Bao S. and Li, H., *Research on Expert Search at Enterprise Track of TREC* 2005. TREC, 2005.

Choi, T.M., H.K. Chan, X. Yue. Recent Development in Big Data Analytics for Business Operations and Risk Management. *IEEE Transactions on Cybernetics*, 47(1), 81-92, 2017.

Cook, W. D., Golany, B., Kress, M., Penn, M., & Raviv, T. (2005). Optimal allocation of proposals to reviewers to facilitate effective ranking. *Management Science, 51*(4), 655-661.

Daud, A., Li, J., Zhou, L., & Muhammad, F. (2010). Temporal expert finding through generalized time topic modeling. *Knowledge-Based Systems, 23*(6), 615-625.

Davoodi, E., Kianmehr, K., & Afsharchi, M. (2013). A semantic social network-based expert recommender system. *Applied Intelligence, 39*(1), 1-13.

Dumais, S. T., & Nielsen, J. (1992). Automating the assignment of submitted manuscripts to reviewers. In *SIGIR* '92, 233-244, Copenhagen, Denmark, Jun 21 - 24.

Egghe, L. (2006). Theory and practise of the G-index. *Scientometrics, 69*(1), 131-152.

Fang, H., & Zhai, C. (2007). Probabilistic models for expert finding. In *ECIR* '07, 418-430, Rome, Italy, Apr 02-05.

Fang, Y., Si, L., & Mathur, A. P. (2010). Discriminative models of integrating document evidence and document-candidate associations for expert search. In *SIGIR* '10, 683-690, Geneva, Switzerland, Jul 19-23

Gollapalli, S. D., Mitra, P., & Giles, C. L. (2011). Ranking authors in digital libraries. In *JCDL* '11, 251-254, Ottawa, Canada, Jun 13-17.

Haveliwala, T. H. (2002). Topic-sensitive PageRank: A context-sensitive ranking algorithm for Web search. In *WWW* '02, 784-796, Honolulu, HI, May 07-11.

Hettich, S., & Pazzani, M. J. (2006). Mining for proposal reviewers: lessons learned at the national science foundation. In *KDD* '06, 862-871, Philadelphia, PA, Aug 20-23.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. In *PNAS*, 102(46), 16569-16572.

Hu, J., Fang, Y., & Godavarthy, A. (2013). Topical authority propagation on microblogs. In *CIKM* '13, 1901-1904, San Francisco, CA, Oct 27-Nov 01.

Jiang, Z., Liu, X. and Gao, L. (2015) Chronological citation recommendation with information-need shifting. In *CIKM* '15, 1291-1300, Melbourne, Australia, Oct 18-23

Jin. J., Geng, Q., Zhao, Q. and Zhang, L. (2017) Integrating the Trend of Research Interest for Reviewer Assignment. In *WWW* '17, Perth, Australia, April 3-7

Karimzadehgan, M., & Zhai, C. (2009). Constrained multi-aspect expertise matching for committee review assignment. In *CIKM* '09, 1697-1700, Hong Kong, China, Nov 02-06.

Karimzadehgan, M., & Zhai, C. (2012). Integer linear programming for constrained multi-aspect committee review assignment. *Information Processing and Management, 48*(4), 725-740.

Karimzadehgan, M., Zhai, C., & Belford, G. (2008). Multi-aspect expertise matching for review assignment. In *CIKM* '08, 1113-1122, Napa Valley, CA, Oct 26-30.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM, 46*(5), 604-632.

Kou, N. M., U, L. H., Mamoulis, N., & Gong, Z. (2015). Weighted coverage based reviewer assignment. In *SIGMOD* '15, 2031-2046, Melbourne, Australia, May 31-Jun 04.

Li, L., Wang, Y., Liu, G., Wang, M., & Wu, X. (2015). Context-aware reviewer assignment for trust enhanced peer review. *PloS ONE, 10*(6), 1-28

Li, X., & Watanabe, T. (2013). Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. In *Proceedings of the 17th international conference on Knowledge-based and intelligent information & Engineering Systems*, 633-642, Kitakyushu, Japan, Sep 09-11.

Liu, O., Wang, J., Ma, J., & Sun, Y. (2016). An intelligent decision support approach for reviewer assignment in R&D project selection. *Computers in Industry,* 76, 1-10.

Liu, X., Bollen, J., Nelson, M. L., & Sompel, H. V. d. (2005). Co-authorship networks in the digital library research community. *Information Processing & Management, 41*(6), 1462-1480.

Liu, X., Suel, T., & Memon, N. (2014). A robust model for paper reviewer assignment. In *RecSys* '14, 25-32, Foster City, Silicon Valley, CA, Oct 06-10

Mauro, N. D., Basile, T. M. A., & Ferilli, S. (2005). GRAPE: an expert review assignment component for scientific conference management systems. In *Proceedings of the 18th international conference on Industrial and Engineering Applications of Artificial Intelligence and Expert systems,* 789-798, Bari, Italy, Jun 22-24.

Mimno, D., & McCallum, A. (2007). Expertise modeling for matching papers with reviewers. In *KDD* '07, 500-509, San Jose, CA, Aug 12-15.

Petkova, D., & Croft, W. B. (2008). Hierarchical language models for expert finding in enterprise corpora. *International Journal on Artificial Intelligence Tools*, 17(1), 5-18.

Rigaux, P. (2004). An iterative rating method: application to web-based conference management. In *SAC* '04, 1682-1687, Nicosia, Cyprus, Mar 14-17

Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004). The author-topic model for authors and documents. In *UAI* '04, 487-494, Banff, Canada, Jul 07-11

Setaputra, R., Yue X. and Yao, D. (2010) Impact of Information Systems on Quick Response Programs. In: Cheng, T., Choi, T.M. (eds) Innovative Quick Response Programs in Logistics and Supply Chain Management. *International Handbooks on Information Systems*. Springer, Berlin, Heidelberg

Stephen C. H. and Erim K. (2015). Calibration, sharpness and the weighting of experts in a linear opinion pool. *Annals of Operations Research*, 229(1):429–450

Sun, Y. H., Ma, J., Fan, Z.-P., & Wang, J. (2008). A group decision support approach to evaluate experts for R&D project selection. *IEEE Transactions on engineering management, 55*(1), 158-170.

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). Arnetminer: extraction and mining of academic social networks. In *KDD* '08, 990-998, Las Vegas, NV, Aug 24-27.

Tang, W., Tang, J., Lei, T., Tan, C., Gao, B., & Li, T. (2012). On optimization of expertise matching with various constraints. *Neurocomputing, 76*(1), 71-83.

Tang, W., Tang, J., & Tan, C. (2010). Expertise Matching via Constraint-Based Optimization. In *WI-IAT* '10, 34-41, Toronto, Canada, Aug 31-Sep 03.

Tayal, D. K., Saxena, P. C., Sharma, A., Khanna, G., & Gupta, S. (2014). New method for solving reviewer assignment problem using type-2 fuzzy sets and fuzzy functions. *Applied Intelligence, 40*(1), 54-73.

Wang, G. A., Jiao, J., Abrahams, A. S., Fan, W., & Zhang, Z. (2013). ExpertRank: A topic-aware expert finding algorithm for online knowledge communities. *Decision Support Systems, 54*(3), 1442-1451.

Wei, W. W. S. (1994). *Time series analysis: Univariate and Multivariate Methods.* Boston, Massachusetts: Addison-Wesley.

Xu, S., Shi, Q., Qiao, X., Zhu, L., Jung, H., Lee, S., and Choi, S.-P. (2014)

Author-topic over time (AToT): A dynamic users' interest model. In *Mobile, Ubiquitous, and Intelligent Computing, volume 274 of Lecture Notes in Electrical Engineering*, 239-245. Springer Berlin Heidelberg, 2014

Yukawa, T., Kasahara, K., Kato, T. and Kita, T. An expert recommendation system using concept-based relevance discernment. In *Proceedings of the 13th International Conference on Tools with Artificial Intelligence*. 2001. IEEE.

Zhai, C., Velivelli, A., & Yu, B. (2004). A cross-collection mixture model for comparative text mining. In *KDD* '04, 743-748, Seattle, WA, Aug 22-25.

Zhang, C. (2013). The H'-index, effectively improving the H-index based on the citation distribution. *PloS ONE, 8*(4):e59912.

Zheng, H. T., Li, Q., Jiang, Y., Xia, S. T., & Zhang, L. (2013). Exploiting multiple features for learning to rank in expert finding. In *ADMA* '13, 219-230, Hangzhou, China, Dec 14-16.

Zhou, D., Orshanskiy, S. A., Zha, H., & Giles, C. L. (2007). Co-ranking authors and documents in a heterogeneous network. In *ICDM* '07, 739-744, Omaha, NE, Oct 28-31.