Integrating the Trend of Research Interest for Reviewer Assignment

Jian Jin, Qian Geng, Qian Zhao, Lixue Zhang
Department of Information Management,
Beijing Normal University
Beijing 100875, China
{jinjian.jay, gengqian}@bnu.edu.cn

ABSTRACT

Reviewer assignment problem in the research field usually refers to invite experts for comments on the quality of papers, projects, etc. Different factors in conventional approaches are reckoned to choose appropriate reviewers, such as the relevance between reviewer candidates and submissions, the diversity of candidates. However, many studies ignore the temporal changes of reviewer interest and the stability of reviewers' interest trend. Accordingly, in this research, three indispensable aspects are analyzed, including the relevance between reviewer candidates and submissions, the interest trend of candidates as well as the authority of candidates. Next, with extracted aspects, the reviewer assignment is formulated as an integer linear programming problem. Finally, categories of comparative experiments are conducted with two large datasets that are built from WANFANG and ArnetMiner, which shows the availability of the proposed approach in modeling the temporal changes of reviewers' research interest. Also, it demonstrates the effectiveness of the proposed approach for the reviewer assignment problem.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*Text analysis*

General Terms

Algorithms, Experimentation, Human Factors

Keywords

expert recommendation, reviewer assignment, research interest trend

1. INTRODUCTION

Peer review is one popular method that is widely used to evaluate research studies and only high-qualified experts should be recommended to judge the intrinsic value of submissions. Generally, reviewer assignment usually implies to invite experts for comments on the quality of papers, projects, etc. One of the major steps in peer review is the selection of reviewers. To avoid some disadvantages about the manual expert recommendation, the automatic reviewer assignment draws an increasing number of researchers. Some approaches are observed to treat the reviewer assignment as a retrieval problem [7, 14], focusing on the topic relevance between reviewer candidates and submissions, such as LSI (Latent Semantic Index), language model, etc. In addition to the topic relevance, other researchers investigated different complementary aspects for reviewer assignment such as authority, diversity, expertise, availability [13, 19, 23], or emphasized mining the knowledge of reviewer candidates or relations between reviewer candidates and authors [2, 10].

However, most studies ignore experts' interest, which is critical in peer review, since it denotes the willingness to review submissions. Generally, it is known that the research interest of an expert may change over time. For example, an expert studied the research topic of text mining ten years ago. However, nowadays, this expert turns to investigate the topic of social network. It makes that to recommend this expert for reviewing a submission about text mining is probably not a wise choice. Arguably, a naive approach for analyzing the research interests of reviewers is to concern their publications in a recent time window only. Nonetheless, the length of the time window is somewhat tricky to define. Specifically, 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. Hence, an effective approach to identify the trend of experts' interest is expected and it will guarantee submissions are being allocated to experts who have an interest in submission related topics.

In this research, an integrated approach is proposed to recommend experts who are qualified to review submissions. Initially, with the help of ATM (Author Topic Model) [21] and the EM (Expectation Maximization) algorithm, topic distributions of reviewer candidates' publications and submissions are estimated. Next, three indispensable aspects 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 candidates' willingness to review a submission, and (3) the authority of reviewer candidates, which illustrates the recognition of a candidate in submission-related topics. Finally, to balance three aspects and to take some practical concerns into considerations, the

WWW 2017 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3053053

problem of reviewer assignment is formulated as an integer linear programming problem.

The contributions of this research are at least three-folds. First, the interest trend of an expert is recognized and utilized to profile a candidate for reviewer assignment. Also, a framework for reviewer assignment is depicted, in which three critical aspects are extracted from submissions and candidates' publications, and an optimization problem is formulated. Finally, with two real datasets, categories of experiments are conducted and benchmark tests on three approaches are compared on different evaluation metrics, which demonstrate the availability of the proposed approach.

The rest of this paper is organized as follows. Related studies are briefly reviewed in Section 2. In Section 3, technical details about how to choose appropriate reviewers automatically for a set of submissions are explained. In Section 4, categories of comparative experiments are presented which shows the availability of the proposed approach. Finally, this research is summarized in Section 5.

2. RELATED WORK

Conventionally, the reviewer assignment is treated as an expert retrieval problem and the topical relevance between reviewer candidates and submissions becomes the main consideration. Specifically, first, publications of reviewer candidates are collected to represent his/her knowledge. Next, submissions are modeled as a query. Finally, reviewers are selected according to the relevance between their knowledge and submissions. Hettich et al. [10] introduced a prototype application to identify prospective experts for proposals. In their research, the reviewer assignment was modeled as a retrieval problem and the TF-IDF weighting was utilized to obtain the match score between reviewer candidates and proposals, which regards each submission as an isolate query. Karimzadehgan et al. [14] regards submissions as a combination of multiple subtopics. Three general strategies were proposed for the reviewer assignment to maximize the subtopic coverage of each submission in a complementary manner. Fang et al. [8] treated the expert recommendation as a classification problem and the logistic model was utilized for expert determination. Kou et al. [15] analyzed topic distributions in submissions and reviewers' publications by the ATM, and, according to the topic weights, a group of experts were recommended. Zheng et al. [30] extracted experts' multiple features by TF, TF-IDF, language model. Next, the approach of learning to rank was applied to sort experts for a particular submission.

Some researchers claimed that the topic relevance only is not adequate to select the most appropriate group of experts to review submissions. These researchers focused on other complementary aspects. In addition to the topic relevance, several other complementary aspects are investigated, such as expertise, authority, conflicts of interest (COI), diversity. For example, besides the relevance, the authority of reviewer candidates was explored [26], in which a co-PageRank algorithm was proposed to find experts in a specific research filed with the consideration of co-authorship and citations relations. Similarly, both experts' expertise and experts' relevance were considered in [17]. From the expertise perspective, the authority and the freshness were combined to estimate the expertise score. From the relevance perspective, the bibliography and the referring information were combined to estimate the relevance score. Liu et al. [19] claimed that the selection of experts should account for not only the authority and the expertise but also the diverse research background. Then, a convex optimization framework was formulated, which incorporates the expertise, the authority and the diverse research background of experts. In [18], an intelligent decision support approach was described to recommend experts for proposals. In particular, the expertise of reviewer candidates and the COI between reviewer candidates and applicants are investigated. Li et al. proposed another approach for reviewer assignment [16]. Specifically, in addition to experts' knowledge and the COI, the stringent or lenient styles of reviewer candidates were also explored.

The reviewer interest refers to the degree of willingness to review a submission. In previous studies, generally, there are two types of methods on modeling the reviewer interest. The first type regards the reviewer interest as the expertise [4, 17], which describes a reviewer's research achievement on submission related topics in a recent time. Another is to ask reviewers themselves to indicate their willingness on some prepared topics explicitly. Then these preferences are utilized as the prior knowledge for reviewer assignment[5, 20]. Rigaux et al. [20] allowed each expert to express his/her preferences to review different submissions explicitly and, according to their prior preferences, the techniques of collaborative filtering were used to predict their interests on different submissions. Mauro et al. [5] described an expert system, named Global Review Assignment Processing Engine (GRAPE), which considers submission topics and experts' preferences for reviewer assignment. In the GRAPE system, preferences on submissions of all experts were initially collected. Next, the preferences were utilized to attune the prior assignment to experts. Li et al. [17] argued that recent publications have higher capability to represent experts' interest. Then, a time interval was considered and the recently publications were given a higher weight than older ones. These approaches concentrate on the the interest extraction according to expertise labels or reviewers' recent publications. However, the prior knowledge such as preferences is often cumbersome to obtain in increasingly narrow research fields. Comparatively, in this research, the willingness of reviewers is modeled as an interest trend, in which the direction and the smoothness are recognized.

Reviewer authority refers to the impact and the recognition of research achievements, and it is often regarded as an indispensable consideration in reviewer selection. Many studies measured the authority of reviewer candidates according to conventional bibliometrics, such as citations, impact factor of journal [11, 6, 29]. Other studies regard each reviewer as a node in an expert network and rank the node according to the random walk algorithm [9, 19]. Rather than some specific topics, these studies focus on how to measure the global importance of a reviewer, which leads to that some critical facts within different research fields are ignored. For instance, one reviewer's impact is highly possible not to be aligned with all the research areas if he/she has interest on different research fields. It induces that only reviewers who have higher recognition in submission-related topics should be given higher priority.

Different aspects of experts were utilized to obtain the most appropriate reviewers, such as topic relevance, topic coverage for submissions, research impact of publications, diverse background of experts, etc. For many approaches only the semantic features of reviewer candidates and the relations between candidates and authors are exploited. However, the fact that the interest trend of an expert may change over time has been ignored and few studies were reported on mining the research interest trends of reviewer candidates, which is potentially to improve the reviewers' willingness on submissions. Hence, a sound investigation on the preferences of reviewer candidates is expected. Accordingly, in this study, to profile each reviewer candidate, the interest trend, which describes the research tendency over time in a specific topic, is modeled as a substantial aspect for reviewer selection and assignment.

3. METHODOLOGICAL OVERVIEW

3.1 System Architecture

The proposed system architecture for reviewer assignment is presented in Figure 1. As presented, four major steps are involved, including data collection, topic extraction, expert profile construction and an optimization approach.

(1) Data Collection. Two types of data are required to construct reviewer candidates' profile, including candidates' publications and the corresponding citations information of these publications. The publications represent his/her major academic achievements, which mirror the knowledge and research interest, while the citations are utilized to understand the authority and the mutual recognition.

(2) Topic Extraction. The selection of reviewers 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 specific topics. Accordingly, in this step, with the help of techniques about topic modeling, topic distributions of each submission and each expert' publication are estimated.

(3) Expert Profile Construction. Three aspects are utilized to profile each expert: the relevance between an expert and a submission, the research interest trend of each expert and the authority of each expert. The relevance is modeled as the similarity between each reviewer candidate and each submission. For the interest trend, a new model is developed to distinguish the different types of trends, such as the stable upward trend or the fluctuating downward trend. For the authority, the topical PageRank is introduced to estimate the authority in submission-related topics.

(4) An Optimization Approach. An integer linear programming problem is formulated to balance the proposed three aspects. In addition, several practical constraints are considered in this optimization problem for reviewers, such as the workload of each reviewer.

3.2 Topic Extraction

Generally, assume that there are T topics and, accordingly, the knowledge of a reviewer candidate r_i, r_i's u-th publication p_u^i ($0 \leq u \leq v_i, v_i$ is the total number of r_i's publications) and a particular submission s_j can be denoted as $\vec{r}_i = (\vec{r}_i[1], \vec{r}_i[2], \ldots, \vec{r}_i[T])$, $\vec{p}_u^i = (\vec{p}_u^i[1], \vec{p}_u^i[2], \ldots, \vec{p}_u^i[T])$ and $\vec{s}_j = (\vec{s}_j[1], \vec{s}_j[2], \ldots, \vec{s}_j[T])$, respectively. For instance, $\vec{r}_i[t]$ refers to the strength of r_i's knowledge about topic t. Similarly, $\vec{p}_u^i[t]$ and $\vec{s}_j[t]$ denote the topic probability in terms of p_u^i and s_j regarding topic t.

Specifically, given reviewer candidates' publications and a list of submissions, topic models such as the LDA (Latent Dirichlet Allocation) and the ATM can be applied to extract the defined topic vectors. Indeed, some other topic models, such as DTM (Dynamic topic models) [3], AToT(Author-Topic over Time) [27] and SDIM (Supervised Document Influence Model) [12], are reported, which aim to capture the evolution of topics in corpus. However, compared with these dynamic models, the ATM made the least assumption on the distributions of underlying topics and it is the most widely utilized model for processing big scholarly textual data. It is also applied in other studies on expert recommendation and reviewer assignment to extract topic vectors for experts, publications and submissions [13, 23, 19]. Hence, in this study, the ATM and the EM algorithm are applied.

Given the entire set about the publications of r_i, the topic vector $\vec{r}_i$ can be estimated by the ATM. For the estimation of $\vec{p}_u^i$ and $\vec{s}_j$, the EM algorithm can be utilized, which can be also found in [28, 15]. Particularly, in the EM algorithm, both publications and submissions are referred as "documents". For a document d, the E-step and M-step are,

E-step:

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

M-step:

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

$p(w_i \mid t)$ refers to the probability of a word w_i is generated by topic t, and $c(w,d)$ is the number of word w in d.

3.3 Reviewer Profile Construction

In this subsection, three aspects of reviewer candidates are introduced to profile reviewer candidates. The first aspect is the topic relevance between a reviewer candidate and a submission, which usually becomes the major consideration for reviewer assignment. Also, intuitively, a high-quality review may be provided if the recommended reviewer is an expert who has a higher degree of research interest in submission-related topics. Accordingly, in the following, the focus is on how to profile a reviewer candidate in three aspects.

3.3.1 *The Reviewer-Submission Relevance*

Intuitively, reviewer candidates who have solid knowledge with the topics that are discussed in submissions should be given a higher priority. In particular, the Kullback-Leibler divergence (KL divergence) is employed utilized to estimate the dissimilarity between the topics of s_j and the knowledge of r_i. 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 r_i and s_j shows a bigger relevance. Hence, the corresponding relevance 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 and the quality of peer review may be affected significantly.

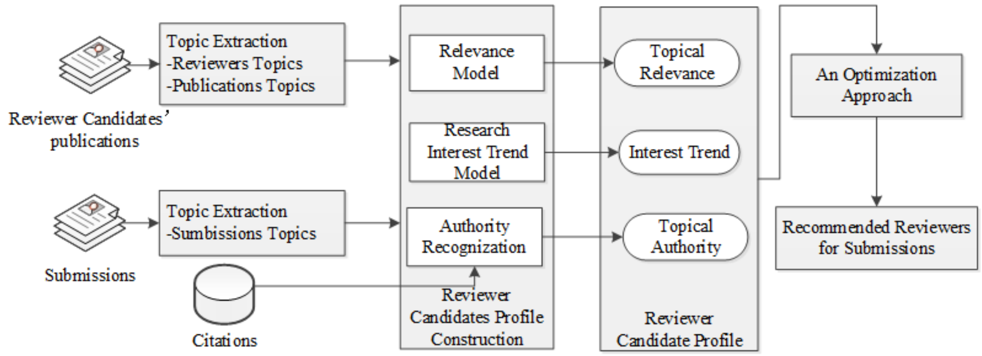
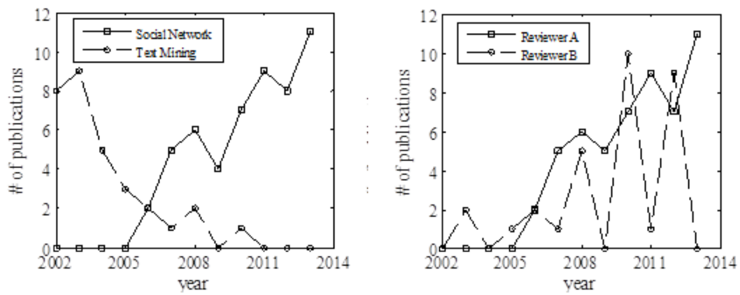


Figure 1: The system architecture for reviewer assignment



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

Figure 2: A toy example about reviewers' interest trend

Assume that, for a specific topic, if r_i has a certain degree of research interest, he/she might contribute some papers to discuss it. Accordingly, the number of papers might be a good indicator that reveals the research interest trend of a reviewer. For instance, one exemplary interest trend about a reviewer is shown in Figure 2(a). As presented, an increasing number of papers are observed over the solid line, which shows that this reviewer has greater interest or willingness to review submissions about social network. Similarly, the dash 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 selected, and additionally, if the trend is smooth and upward, then the candidate is highly preferred to be recommended. A toy example is shown in Figure 2(b). As seen from this figure, compared with two candidates A and B, A is preferable since that a smooth and upward trend of research interest is observed.

For instance, an increasing number of papers on social network and a decreasing number of papers on text mining are observed in a reviewer 's publication list over the past ten years, which explains that he/she may has greater interest to review submissions about social network than text mining. Besides 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 selected, and if the trend of a reviewer candidate is smooth and upward, rather than unstable and volatile, then he/she is highly preferred to be recommended. Hence, in this study, for each s_j, two indicators, which are often utilized for the analysis of time series data, are investigated to describe the interest trend of r_i quantitatively. They are the direction D_{ij} and the smoothness Q_{ij}, which quantify r_i's interest trend that is related to s_j.

Before go into the details about two indicators, the annual number of publications about r_i that relate to s_j, denoted by M_{ij}, should be obtained. M_{ij} presents the interest trend of r_i over topics in s_j. It can be estimated in three steps,

(a) Top N important topics in $\vec{p}_u^{\,i}$ are extracted, which are utilized to represent r_i's u-th publications. Similarly, top N important topics in $\vec{s}_j$ are gained to represent s_j.

(b) All v_i publications of r_i are indexed according to T topics, in which the t-th group is denoted as c_t^i. Thus, publications of r_i that are related to the topics of s_j are selected.

(c) Publications that are related to top N important topics in $\vec{s}_j$ are selected. Next, selected publications are sorted by year and the annual publication number that is related to s_j, M_{ij}, can be obtained.

(1) The direction of interest trend D_{ij}

A linear relation is utilized to depict D_{ij}. Accordingly, the approach of the least squares is applied to estimate the slope and the interception about a linear relation between the number of annual publications M_{ij} and the year of publication. Then, the estimated slope is utilized to estimate D_{ij}. If D_{ij} is positive, it means that an increasing trend is observed, which shows the his/her interest over research topics in s_j is rising.

(2) The smoothness of interest trend Q_{ij}

Generally, both the standard deviation and the mean of M_{ij} over time 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. Then, the coefficient of variation, which evaluates the degree of temporal changes [1], 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 the variation coefficient.

Reviewer candidates whose interest trend with a smooth upward increase should be given higher priority. Accordingly, I_{ij}, the interest trend of r_i 's interest trend on s_j, is

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

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, which ignore that the level of expert's authority might not be the same in different topics. Actually, experts may investigate different topics at the same time and the authority degrees in different topics are not same. Accordingly, in this subsection, different levels of authority are distinguished.

Practically, being cited by multiple experts is a good indicator that reflects the authority of an expert. Now, suppose a reviewer candidate's research topics are highly related to the 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 also considered to have a higher authority in those topics. Accordingly, in this study, a topical PageRank is presented to identify high-authority experts within given topics. Some similar approaches can be also found in [9, 25, 19]

In this subsection, citation activities are formalized in a topical authority graph. In Figure 3, a sample of citation activities among authors and papers are shown. In the upper part of Figure 3, the 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 relationship among papers, citations among authors can be deduced, which is represented in the lower part. In the lower part of Figure 3, each node represents one author and each edge denotes a citation between them.

Suppose r_i 's research topics are highly related to s_j 's topics and r_i is being cited many times. It means that r_i is intuitively expected to have a higher degree of authority regarding these topics. Furthermore, if r_i is cited by many authoritative experts on related topics, r_i should be also considered to have a higher degree of authority in those topics. Accordingly, a topical PageRank is utilized to identify high authority experts within given topics.

Notably, as presented in Figure 3, there might exist more than one edges between expert nodes, which indicates that more citations are observed. Accordingly, the authority degree of r_i with respect to s_j, A_{ij}, can be estimated as,

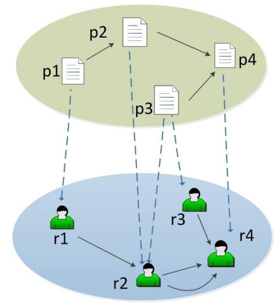


Figure 3: An exemplary citation network

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

$ref(r_i)$ 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, which is evaluated by the cosine distance between $\vec{r}_q$ and $\vec{s}_j$. B_{qi} denotes how much the authority of r_q is assigned 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 others. The parameter a is a damping factor, which compensates for experts with none reference papers in the expert network. N is the node number of the expert network.

3.4 An Optimization Approach for Reviewer Assignment

In this research, the ultimate goal is to assign appropriate reviewers for a set of submissions. Finally, an optimization problem is formulated, in which all the formulated three aspects for reviewer assignment are linearly combined to balance their importance. Notice that, generally, it is difficult to build a manually labeled dataset accurately, even experienced editors are invited. It induces that supervised learning approaches potentially fail to be applied to explore the weights of different factors exactly.

Besides, practically, other two factors should be reckoned. The first is the workload of each reviewer, which implies that a reviewer 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 $\{r_1, \ldots, r_n\}$ and a set of submissions $\{s_1, \ldots, s_m\}$, this reviewer assignment can be formulated as an integer linear programming problem,

$$\begin{aligned} \max \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}X_{ij} + \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}X_{ij} + \lambda_3 \sum_{i=1}^{n}\sum_{j=1}^{m} R_{ij}X_{ij} \\ s.t. \quad \textstyle\sum_{j=1}^{m} X_{ij} \le N_{ri} \quad \forall i \in [1,n] \\ \textstyle\sum_{i=1}^{n} X_{ij} = N_{sj} \quad \forall j \in [1,m] \\ X_{ij} \in \{0,1\} \quad \forall i \in [1,n], \forall j \in [1,m] \end{aligned} \tag{9}$$

X is an $n \times m$ binary matrix, where X_{ij} is a binary variable that indicates whether s_j is assigned to r_i. n and m are the number of reviewers and submissions, respectively. Notice that three aspects, A_{ij}, I_{ij} and R_{ij}, are normalized respectively before they are linearly combined. λ_1, λ_2 and λ_3 are the magnification factor that balance different aspects of reviewers. N_{ri} is the maximal workload that r_i reviews at the same time, and N_{sj} is a predefined number of reviewers that each s_j should be assigned for reviews. This optimization problem can be solved by many software packages for the integer linear programming.

4. EXPERIMENTS

4.1 Datasets

In this experiment, two datasets are used to valid the effectiveness of the proposed approach. The first dataset was built from the WANFANG DATA, which is a Chinese scientific library. To build this dataset, 256 scholars with funding supports 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 [22]. 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 are chosen randomly to build a reviewer candidate repository. Other experts and papers are used to train the proposed model.

4.2 Evaluation Metrics and Benchmarking Methods

As mentioned in the previous sections, a manually labeled dataset is generally difficult to be built. It induces that some evaluation metrics for information retrieval, such as precision and recall, are not applicable. To tackle this dilemma, some researchers proposed different evaluation metrics according to the problem definition of their own. For instance, in Liu et al. [19], a unification model was proposed and the relevance, the authority as well as their research background were utilized for the benchmark. Some similar evaluation methods can be also found in [24, 23].

Similarly, in this research, a group of persuadable evaluation metrics about the relevance, the interest trend and the authority are invited for performance evaluation. Let $Distance@k$, $Interest@k$ and $Authority@k$ denote the evaluation metrics, which refer to the corresponding score if k submissions are concerned.

(1) Relevance

First, the relevance is widely utilized for expert selection and reviewer assignment, which evaluates the degree of relevance between submissions and recommended re-viewers. In this research, the relevance can be defined as,

$$Distance@k = \sum_{j=1}^{k} \sum_{i=1}^{N_{sj}} \mid R_{ij} \mid \tag{10}$$

N_{sj} refers to the number of reviewers required by the j-th submission. $Distance@k$ evaluates the sum of the topical distance between submissions and recommended reviewers.

(2) Interest Trend

Additionally, it is desirable to select reviewers with a higher score of interest trend so that they may have higher preferences to review those submissions. Given k submissions, the sum scores of interest trend can be denoted as,

$$Interest@k = \sum_{j=1}^{k} \sum_{i=1}^{N_{sj}} I_{ij} \tag{11}$$

$Interest@k$ calculates recommended reviewer candidates' total degree of interest trend on all k submissions.

(3) Authority

Last, the authority of recommended reviewer candidates should be maximized. The total score of each assignment can be estimated as,

$$Authority@k = \sum_{j=1}^{k} \sum_{i=1}^{N_{sj}} A_{ij} \tag{12}$$

$Authority@k$ evaluates the total authority degree of recommended reviewer candidates on all k assignments.

Three popular methods are utilized for the benchmark. They are the Vector Space Model (VSM), the Language Model (LM) and the ATM. All these methods are invited to make comparisons in terms of three evaluation metrics on both the WANFANG dataset and the ArnetMiner dataset. Admittedly, three approaches are not state-of-the-art algorithms for the problem of reviewer assignment. In the future, other benchmarking algorithms with different considerations in reviewer assignment will be testified, which helps to polish the performance of the proposed approach.

4.3 Experimental Results

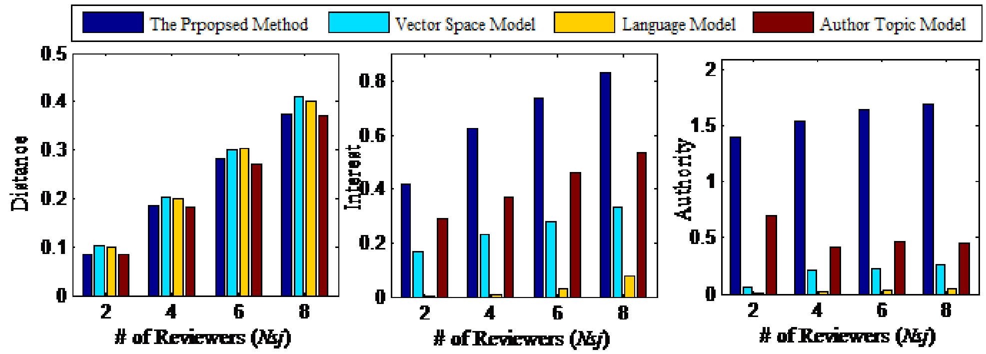
In the following, three aspects of reviewers are regarded to be equally important and, hence, λ_1, λ_2 and λ_3 are 1. Also, a in Equation (8) is 0.85 and the parameter η in Equation (6) is 1 respectively.

In Figure 4, different approaches are compared in terms of the $Distance@k$ on the WANFANG Dataset and the ArnetMiner Dataset, where the horizontal axis represents the number of submissions that are waiting for reviewers 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 be four. Actually, similar phenomena can be observed if different number of reviewers and different maximal workload are predefined. 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 since that more research topics tend to be covered. Also, not significant differences in all of these four approaches are observed since that, in all of these approaches, the relevance between reviewers and submissions are all reckoned as a major concern.

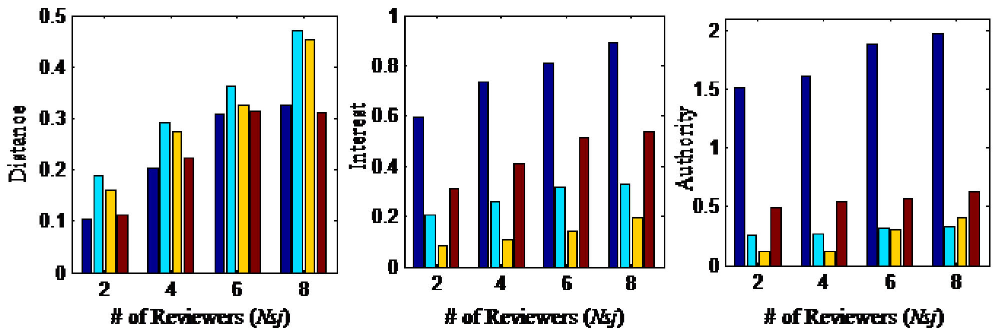
In Figure 5, different approaches are compared in terms of the $Interest@k$. As seen from the two sub graphs, the proposed approach performs much better than the others. It shows the obvious strength of the proposed approach in modeling the trend of reviewers' research interest.

In Figure 6, four approaches are compared in terms of the $Authority@k$. As presented, the proposed approach outperforms all the others. In demonstrates that, the topical authority of recommended reviewers are relatively higher, compared with other three benchmarking approaches.

In Figure 7, the performance is compared with different number of reviewers, if 30 submissions were concerned and the maximal workload of each reviewer candidate was

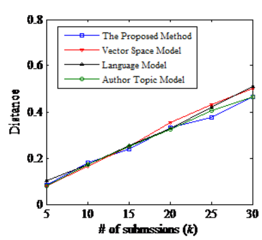


(a) WANFANG Dataset

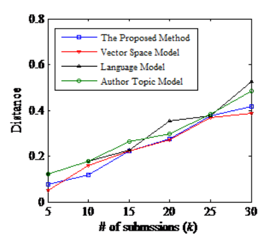


(b) ArnetMiner Dataset

Figure 7: Performance Comparisons with Different Number of Reviewers

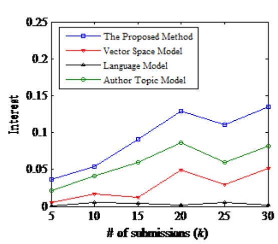


(a) WANFANG Dataset

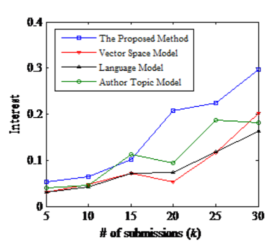


(b) ArnetMiner Dataset

Figure 4: The ***Distance@k***

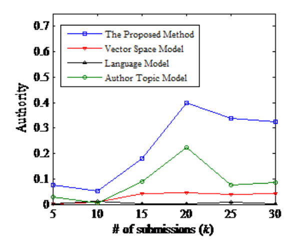


(a) WANFANG Dataset

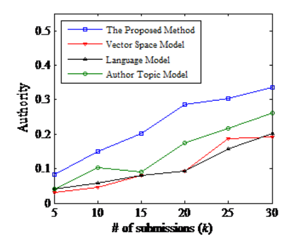


(b) ArnetMiner Dataset

Figure 5: The ***Interest@k***



(a) WANFANG Dataset

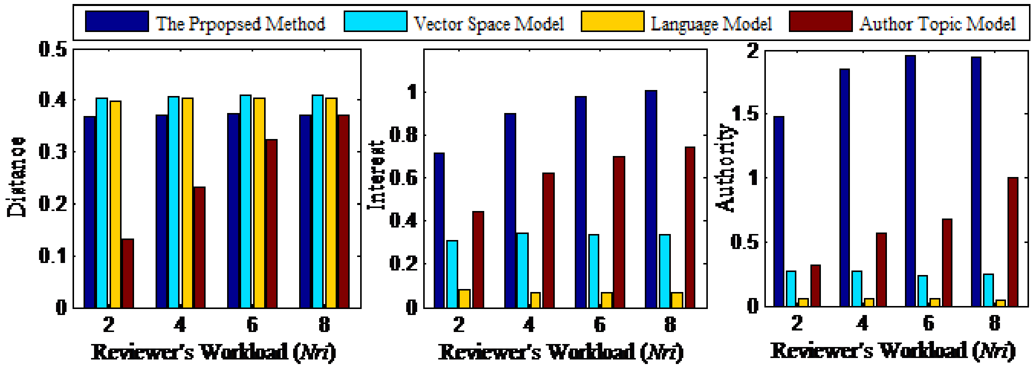


(b) ArnetMiner Dataset

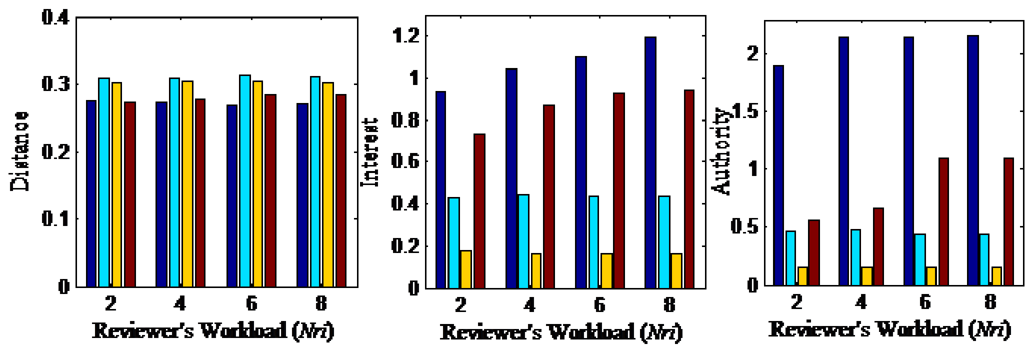
Figure 6: The ***Authority@k***

4. With an increasing number of reviewers, as presented in Figure 7, values of all three evaluation metrics of the proposed method start to rise. Compared with other three approaches, the interest and the authority of the proposed method are greatly improved, though a marginal improvement is seen in terms of the topic distance. It can be deduced that, with enhanced values about the interest and the authority, the proposed method presents constantly competitive performance regarding the topical relevance between submissions and experts for different number of reviewers.

Meanwhile, effects of the maximal workload of each reviewer are represented in Figure 8. In this experiment, 30 submissions and 4 reviewers were considered. Compared with Figure 7, some similar phenomena are observed, i.e. improved values of the interest and the authority with competitive values of the topical relevance.



(a) WANFANG Dataset



(b) ArnetMiner Dataset

Figure 8: Performance Comparisons with Different Maximal Workload

One of the shining differences between the proposed method and others lies in the considerations about reviewer candidates' interest trend in a fined-grained manner. In many approaches, the relevance and the authority are paid much attention, but the interest trend of reviewer candidates is ignored. As presented, all three benchmark methods fail to capture their 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 and the authority.

5. CONCLUSION

With an increasing number of submissions, finding proper reviewers to evaluate the quality of submissions becomes obviously cumbersome. It induces that the reviewer assignment problem receives much attention in the academic filed and appears increasingly more critical in different practical scenarios, such as R&D project selection, online knowledge management, digital libraries, scientific evaluation, company recruitment.

Many research studies mainly focus on the improvement about the matching degree between submissions and recommended reviewers as well as other some complementary aspects, such as the authority of a reviewer, the expertise diversity, the topic coverage, etc. But the research interest of a reviewer is largely neglected, which might affects the quality of submission reviews. In this study, besides some widely concerned aspects, the research interest trend of each reviewer candidate is taken into considerations. Specifically, the direction of interest trend and its smoothness are combined to model the research interest. Then, the reviewer assignment is modeled as an integer liner programming problem with practical concerns. Categories of experiments were conducted on two real datasets with a large number of experts as well as their publications, which demonstrates the effectiveness of the proposed approach.

In the future, other 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 architecture, such as relations between reviewer candidates and authors, meta information about research topics of publications that are manually labeled in scientific library, etc.

6. ACKNOWLEDGEMENT

This research was supported by the fundamental research funds for the central universities (NO. SKZZB2014037), the youth fund project of ministry of education of the humanities and social sciences research (NO. 16YJC870006) and a research project funded by the ISTIC-EBSCO joint laboratory.

7. REFERENCES

[1] H. Abdi. Coefficient of variation. *Encyclopedia of research design*, pages 169–171, 2010.

[2] K. Balog, L. Azzopardi, and M. de Rijke. Formal models for expert finding in enterprise corpora. In *SIGIR '06*, pages 43–50, 2006.

[3] D. M. Blei and J. D. Lafferty. Dynamic topic models. In *ICML '06*, pages 113–120, 2006.

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

[5] N. Di Mauro, T. M. A. Basile, and S. Ferilli. Grape: An expert review assignment component for scientific conference management systems. In *18th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, pages 789–798, 2005.

[6] L. Egghe. Theory and practise of the g-index. *Scientometrics*, 69(1):131–152, 2006.

[7] H. Fang and C. Zhai. Probabilistic models for expert finding. In *ECIR '07*, pages 418–430, 2007.

[8] Y. Fang, L. Si, and A. P. Mathur. Discriminative models of integrating document evidence and document-candidate associations for expert search. In *SIGIR '10*, pages 683–690, 2010.

[9] S. D. Gollapalli, P. Mitra, and C. L. Giles. Ranking authors in digital libraries. In *JCDL '11*, pages 251–254, 2011.

[10] S. Hettich and M. J. Pazzani. Mining for proposal reviewers: Lessons learned at the national science foundation. In *KDD '06*, pages 862–871, 2006.

[11] J. E. Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46):16569–16572, 2005.

[12] Z. Jiang, X. Liu, and L. Gao. Chronological citation recommendation with information-need shifting. In *CIKM '15*, pages 1291–1300, 2015.

[13] M. Karimzadehgan and C. Zhai. Constrained multi-aspect expertise matching for committee review assignment. In *CIKM '09*, pages 1697–1700, 2009.

[14] M. Karimzadehgan, C. Zhai, and G. Belford. Multi-aspect expertise matching for review assignment. In *CIKM '08*, pages 1113–1122, 2008.

[15] N. M. Kou, L. H. U., N. Mamoulis, and Z. Gong. Weighted coverage based reviewer assignment. In *SIGMOD '15*, pages 2031–2046, 2015.

[16] L. Li, Y. Wang, G. Liu, M. Wang, and X. Wu. Context-aware reviewer assignment for trust enhanced peer review. *PLOS ONE*, 10(6):1–28, 2015.

[17] X. Li and T. Watanabe. Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. *Procedia Computer Science*, 22:633 – 642, 2013.

[18] O. Liu, J. Wang, J. Ma, and Y. Sun. An intelligent decision support approach for reviewer assignment in R&D project selection. *Computers in Industry*, 76:1 – 10, 2016.

[19] X. Liu, T. Suel, and N. Memon. A robust model for paper reviewer assignment. In *RecSys '14*, pages 25–32, 2014.

[20] P. Rigaux. An iterative rating method: Application to web-based conference management. In *SAC '04*, pages 1682–1687, 2004.

[21] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth. The author-topic model for authors and documents. In *UAI '04*, pages 487–494, 2004.

[22] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: Extraction and mining of academic social networks. In *KDD '08*, pages 990–998, 2008.

[23] W. Tang, J. Tang, T. Lei, C. Tan, B. Gao, and T. Li. On optimization of expertise matching with various constraints. *Neurocomputing*, 76(1):71 – 83, 2012.

[24] W. Tang, J. Tang, and C. Tan. Expertise matching via constraint-based optimization. In *WI-IAT'10*, volume 1, pages 34–41, 2010.

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

[26] H. Wu, H. Li, X. Zhang, and S. Yao. Topic-sensitive link-ranking approach for academic expert recruiting. In *2008 International Multi-symposiums on Computer and Computational Sciences*, pages 150–157, October 2008.

[27] S. Xu, Q. Shi, X. Qiao, L. Zhu, H. Jung, S. Lee, and S.-P. Choi. Author-topic over time (AToT): A dynamic users' interest model. In J. J. J. H. Park, H. Adeli, N. Park, and I. Woungang, editors, *Mobile, Ubiquitous, and Intelligent Computing*, volume 274 of *Lecture Notes in Electrical Engineering*, pages 239–245. Springer Berlin Heidelberg, 2014.

[28] C. Zhai, A. Velivelli, and B. Yu. A cross-collection mixture model for comparative text mining. In *KDD '04*, pages 743–748, 2004.

[29] C.-T. Zhang. The h' - index, effectively improving the h-index based on the citation distribution. *PLOS ONE*, 8(4):1–8, 04 2013.

[30] H.-T. Zheng, Q. Li, Y. Jiang, S.-T. Xia, and L. Zhang. Exploiting multiple features for learning to rank in expert finding. In *ADMA '13*, pages 219–230, 2013.