Scientometrics
https://doi.org/10.1007/s11192-018-2726-6

A novel classification method for paper-reviewer recommendation

Shu Zhao[1] · Dong Zhang[1] · Zhen Duan[1] · Jie Chen[1] · Yan-ping Zhang[1] · Jie Tang[2]

Received: 3 November 2017

Abstract Reviewer recommendation problem in the research field usually refers to invite experts to comment on the quality of papers, proposals, etc. How to effectively and accurately recommend reviewers for the submitted papers and proposals is a meaningful and still tough task. At present, many unsupervised recommendation methods have been researched to solve this task. In this paper, a novel classification method named Word Mover's Distance–Constructive Covering Algorithm (WMD–CCA, for short) is proposed to solve the reviewer recommendation problem as a classification issue. A submission or a reviewer is described by some tags, such as keywords, research interests, and so on. First, the submission or the reviewer is represented as some vectors by a word embedding method. That is to say, each tag describing a submission or a reviewer is represented as a vector. Second, the Word Mover's Distance (WMD, for short) method is used to measure the minimum distances between submissions and reviewers. Actually, the papers usually have research field information, and utilizing them well might improve the reviewer recommendation accuracy. So finally, the reviewer recommendation task is transformed into a classification problem which is solved by a supervised learning method- Constructive

✉ Shu Zhao
zhaoshuzs2002@hotmail.com

Dong Zhang
darkzd_zhang@163.com

Zhen Duan
ycduan@gmail.com

Jie Chen
chenjie200398@163.com

Yan-ping Zhang
zhangyp2@gmail.com

Jie Tang
jery.tang@gmail.com

[1] School of Computer Science and Technology, Anhui University, Hefei 230601, China

[2] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

Published online: 31 March 2018

Covering Algorithm (CCA, for short). Comparative experiments are conducted with 4 public datasets and a synthetic dataset from Baidu Scholar, which show that the proposed method WMD–CCA effectively solves the reviewer recommendation task as a classification issue and improves the recommendation accuracy.

Keywords Reviewer recommendation · Classification · Word embedding · Word Mover's Distance · Constructive Covering Algorithm

Mathematics Subject Classification 68T99

JEL Classification C63 · C89

Introduction

In academic field, lots of papers and proposals are submitted to conferences, journals or administration departments. And generally professional experts are invited to comment on the quality of submissions (Jin et al. 2017). A very important task in this process is the recommendation of reviewers for papers (i.e. paper-reviewer recommendation). As far as we know, the manual reviewer recommendation is time-consuming and sometimes unfair (Yang et al. 2009). Thus, how to effectively and accurately recommend reviewers for the submitted papers or proposals is a meaningful but still tough task.

Significant progress has been made in paper-reviewer recommendation recently. Several approaches have been proposed to improve the quality of recommendation and make it perform more automatically. Most of these approaches mainly fall into two categories: retrieval-based methods and matching-based methods. They mainly treat it as a retrieval problem or a matching problem (Fang and Zhai 2007; Karimzadehgan et al. 2008; Karimzadehgan and Zhai 2009; Shon et al. 2017), applying some unsupervised methods to solve it. First, the retrieval-based approaches mainly focus on the topic relationship between submissions and reviewer candidates, such as language model (Karimzadehgan et al. 2008), Latent Semantic Index (LSI) (Deerwester et al. 1990), topic model (Mimno and Mccallum 2007; Peng et al. 2017), etc. Others investigate different aspects of reviewer recommendation or expert finding, including a number of considerations, such as authority, diversity, expertise, availability (Karimzadehgan and Zhai 2009; Liu et al. 2014; Tang et al. 2008; Liu et al. 2016) and the research interests (Jin et al. 2017) of the reviewer. Second, the matching-based approaches compute a matching based on a bipartite graph between the submissions and the reviewer candidates (Charlin et al. 2012; Conry et al. 2009). They compute the relevance degree and construct a bipartite graph and then the right matching can be obtained according to the maximum weight matching based bipartite graph.

One of the major steps in reviewer recommendation is the selection of reviewers. There were many other approaches developed to promote and improve the selection of reviewers. Optimization approaches were widely used as a search algorithm in various assignment and combinational application where it demonstrated satisfactory performances (Deep and Das 2008). Genetic algorithms (GA) and ant colony optimization (ACO) are combined to quickly find good solutions (Kolasa and Król 2010). Several heuristic algorithms have also been proposed for automatically assigning reviewers to papers that provide effective and good results (Kolasa and Krol 2011). Besides, some intelligent decision support approaches were proposed to recommend experts for proposals (Liu et al. 2016; Xu et al. 2010; Fan et al. 2009; Sun et al. 2008). They used decision models to determine the best solution of reviewer assignment. A content-based recommender system was proposed, which aimed at the selection of reviewers (experts) to evaluate research proposals or articles (Protasiewicz et al. 2016).

To make reviewer recommendation more effective and accurate is an important and arduous task, which still faces some key challenges. First, there is a typically complex lexical gap between submissions and reviewers which are usually text. So an accurate text representation is exactly important. Second, the deeper relationship between submissions and reviewer candidates can improve the recommendation quality. Besides, some methods seldom make full use of the existing key information. They often only utilize the text information of submissions in the reviewer recommendation task.

We aim to tackle these challenges via a proposed classification method called Word Mover's Distance–Constructive Covering Algorithm (WMD–CCA, for short), which transforms the reviewer recommendation problem into a classification problem integrating the research field information of submissions. In other words, given some reviewer candidates and a new submission, we predict the research field labels for reviewers and assign one to the submission owing to the same label.

The main contribution of this paper is to treat the paper-reviewer recommendation problem as a classification issue. We accomplish it with the following steps. First, we represent the submission and the reviewer with their tags, which are the keywords of the submission and the research interests of the reviewer. To find the complex semantic relationship, inspired by the success of word embedding, the tags are transformed into vectors to overcome the lexical gap and the word-to-word relation can be regarded as the similarity of word vector. Then based on the word embedding, the Word Mover's Distance (Kusner et al. 2015) (WMD, for short) is used. And it learns the deep correlation between the submission and the reviewer using an optimization. Finally, the research field labels are considered as supervised information in our method. So the reviewer recommendation problem is transformed into a classification issue with a supervised method—Constructive Covering Algorithm (Zhang et al. 2013) (CCA, for short), and we predict the potential research field information of reviewer. Experimental study on 4 public datasets shows the effectiveness of WMD–CCA method. Furthermore, we utilize the proposed method to recommend reviewers for papers on a synthetic dataset, and it presents the potential of our method.

The remainder of the paper is organized as follows. "Related works" section presents the related works in solving the reviewer recommendation problem. "Problem formulation" section formally formulates the introduced problem. "WMD–CCA method for reviewer recommendation" section mainly introduces the proposed method WMD–CCA. In "Experiments" section, experiments and evaluations present the potential of the proposed method on 4 public datasets and a synthetic dataset. And "Conclusions" section summarizes our contributions and gives future work.

Related works

The research on paper-reviewer recommendation is similar to the researches on peer review and research and development (R&D) project selection. Quality of peer review greatly depends on the degree of matching between reviewers and assigned research proposals (Xu et al. 2010). Peer review and experts play important roles in research project selection, because their opinions affect the potential value of a project (Liu et al. 2016).

How to assign the most appropriate experts to review project proposals might greatly affect the quality of project selection, which in turn could affect the return on investment of the funding organization (Liu et al. 2016).

Conventionally, the research on reviewer recommendation can be mainly categorized into two branches. One is the retrieval-based methods which treat each submission as a query to retrieve the relevant reviewers. Commonly, first, publications of reviewer candidates are collected to represent his/her research field knowledge. Then, submissions are modelled as a query. Finally, reviewers are recommended according to the relevance of their knowledge and submissions. Fang and Zhai (2007) proposed a general probabilistic framework for studying expert finding problem and proposed estimation strategies which are all effective to improve retrieval accuracy. Karimzadehgan et al. (2008) proposed three general strategies for reviewer retrieval and studied how to model multiple aspects of reviewer expertise. Liu et al. (2014) studied how to rank reviewer candidates while balancing three objectives: authority, expertise and diversity. They also proposed a graph constructed on reviewer candidates and the query paper, and then an optimization framework with sparsity principle was applied.

And the other is matching-based methods which commonly have two steps. First, it constructs a bipartite graph and computes the weights of the edges. Then an appropriate matching is formulated based on the graph. Charlin et al. (2012) proposed a framework to optimize paper-to-reviewer assignments and studied several matching objectives. Conry et al. (2009) proposed to learn the weights and computed the maximum weight matching for a recommendation. Li et al. (2017) modelled papers and reviewers based on matching degree considering an academic social network to find a confident, fair and balanced assignment.

With the main two categories methods proposed, there were also many other approaches developed to improve the quality of the reviewer selection and find the better solutions. Jian Ma et al. proposed a hybrid knowledge and model approach for reviewer assignment (Sun et al. 2008). And then a hybrid approach using knowledge rule and genetic algorithm to group the proposals (Fan et al. 2009). Knowledge rules are designed to deal with proposal identification and proposal classification, and the genetic algorithm is developed to search for the expected groupings. Decision support approaches (Sun et al. 2008; Fan et al. 2009; Xu et al. 2010) were often proposed to identify valid proposals and reviewers, and it increased overall grouping quality. Protasiewicz et al. (2016) proposed the architecture of the content-based recommender system for selection of reviewers, which was supported by various techniques of information retrieval. The recommendation were based on the combination of cosine similarity between keywords and full-text index. Tayal et al. proposed a new method of solving reviewer assignment problem using type-2 fuzzy sets to represent the expertise levels of the various reviewers in the different domains. This method considered the four important aspects: workload balancing of reviewers, avoiding conflicts of interest, considering individual preferences by incorporating bidding and mapping multiple keywords of a proposal (Tayal et al. 2014). And it can calculate the expertise level of each reviewer in different domains.

In the past, many methods of paper-reviewer recommendation considered the keywords of papers or the profiles of reviewers. Word is the most basic unit of carrying semantics in a sentence or a document. Properly transforming words into a continuous vector space is a commonly used method in many natural language processing (NLP) tasks or other text-related works. Word embedding may be the most popular technique for word representation learning over billions of words and skip-gram model (Mikolov et al. 2013a) is a widely employed method to compute such an embedding. In embedding spaces, the neighbors of every word are generally semantically related (Mikolov et al. 2013b), especially we can also use a similarity function to predict how likely a word is given its context. Many researchers used this property to solve challenging problems, such as Q/A matching (Shen et al. 2017), sentiment classification (Ren et al. 2016), query expansion (Diaz et al. 2016).

For many approaches, few of them can exploit the semantic information of reviewers and papers, and capture the semantic similarity between reviewers and papers. Besides, making full use of the research field label information of reviewers and papers can improve the quality of the reviewer recommendation. Hence, in this study, a method is proposed to measure the deeper research relationship between the reviewers and papers, and find the better solutions with the help of label information.

Problem formulation

In this section, we first define fundamental concepts, including papers (i.e. submissions), reviewers, research field labels and field relationship. Subsequently, we give the formulations of problems and the solutions to the problems we face.

We assume that the paper and the reviewer are described as different numbers of tags, denoted as $p_i = \{k_{i1}, k_{i2}, \ldots, k_{in_i}$ (for paper p_i) and $r_j = \{k_{j1}, k_{j2}, \ldots, k_{jn_j}\}$ (for reviewer r_j), respectively. For instance, different tags k_{in_i} are the different keywords of the submission p_i, and tags k_{jn_j} are research interests of the reviewer r_j. The paper submitted has the research field label l_i and then the paper set can be represented as $P = \{(p_1, l_1), (p_2, l_2), \ldots, (p_n, l_n)\}$. Let $L = \{L_1, L_2, \ldots, L_t\}$ be the set of $|L|$ = t (t < n) different field labels and all $l_i \in L$. It's notable that research field relationship distance between a paper p_i and a reviewer r_j is important in paper-reviewer recommendation task. To make reviewer recommendation more effective and accurate, we transform this task into a classification issue.

According to the above notations, the problem is formulated as follows.

Problem

Given a paper set $P = \{(p_1, l_1), (p_2, l_2), \ldots, (p_n, l_n)\}$, where $p_i = \{k_{i1}, k_{i2}, \ldots, k_{in_i}\}$ is a paper described by n_i tags, and l_i refers to the label of this paper. And given a reviewer candidate set $R = \{r_1, r_2, \ldots, r_m\}$, where $r_j = \{k_{j1}, k_{j2}, \ldots, k_{jn_j}\}$ is a reviewer candidate described by n_j research interests. The goal of the problem is to predict the $y_j \in L$ for each r_j.To achieve the goal mentioned before, there are two main sub-problems we should solve first. We define these sub-problems addressed in this paper:

Sub-problem 1

Tags representation The complex lexical gap between k_{in_i} and k_{jn_j} is the obstacle of mining the deeper relationship of p_i and r_j. To overcome this problem, we represent k_{in_i} and k_{jn_j} as vectors v_i and v_j which can capture distributional syntactic and semantic information in a corpus.

Sub-problem 2

The deeper research field relationship The semantic relationship of k_{in_i} and k_{jn_j} can be attained from a distance function and the tag-level relationship contributes to the research field relationship between p_i and r_j. It's not accurate to learn the deeper relationship by easily considering the semantic distance $d(v_i, v_j)$ of two keyword (tag) vectors. To tackle this limitation, the keyword-level correlation and the weight of keyword are used to optimize the deeper relationship between p_i and r_j. A minimum cumulative distance $\text{distance}(p_i, r_j)$ is computed to represent the deeper relationship. The $d(\cdot,\cdot)$ and $\text{distance}(\cdot,\cdot)$ are formulated by two distance methods, which would be introduced in next section.

With above two sub-problems solved, it's highly desirable but challenging to transform this recommendation task into a classification problem. The research field labels l_i of each p_i are seldom utilized to guide the reviewer recommendation as supervised information. In this paper, we use a constructive classification methods to achieve our goal. The field labels of P are considered to predict $y_j \in L$ for a given r_j. After field label prediction, r_j can be assigned to review p_i owing to $l_i = y_j$.

This paper-reviewer recommendation task is transformed into a classification issue, and we proposed a method WMD–CCA to solve it. We first represent the tags of p_i and r_j as distributional vectors. Then the deeper research field relationship between p_i and r_j is calculated, and finally, we predict the field label for the reviewer candidate according to the field relationship. The details of the proposed method are introduced in "WMD–CCA method for reviewer recommendation" section.

WMD–CCA method for reviewer recommendation

Transforming the reviewer recommendation problem into a classification issue is a challenging task. In the context of paper-reviewer recommendation, it is crucial to accurately mine the relationship between the paper and the reviewer candidate. Actually, a reviewer can review a paper when they belong to the same research field. In another word, they are in the same category. In this paper, we propose novel classification method WMD–CCA, which takes the research field relationship and the research field information into consideration. In Algorithm 1, we introduce WMD–CCA, and we denote the "paper or proposals" as "paper" in this section. The details of the proposed method are described as followed and an overview is presented in Fig. 1.

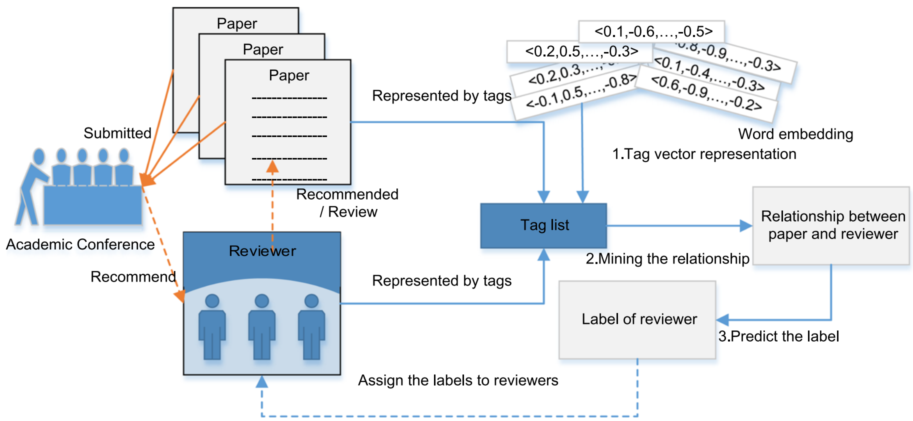


Fig. 1 Overview of WMD–CCA method for reviewer recommendation

---

Algorithm 1: WMD-CCA

Input: A paper set $P = \{(p_1, l_1), (p_2, l_2), \dots, (p_n, l_n)\}$; a reviewer set $R = \{r_1, r_2, \dots, r_m\}$.

Output: Research field label $y_j \in L$ for each $r_j \in R$.

Step 1. **Tags Representation**. Represent the tags $\{k_{i1}, k_{i2}, \dots, k_{in_i}\}$ of p_i, $\{k_{j1}, k_{j2}, \dots, k_{jn_j}\}$ of r_j as vectors $\{v_{i1}, v_{i2}, \dots, v_{in_i}\}$ and $\{v_{j1}, v_{j2}, \dots, v_{jn_j}\}$.

Step 2. **Mining Relationship**. For each $p_i \in P$, $r_j \in R$, obtain their deeper relationship.

Step 2.1. Compute keyword level similarity $d(k_{in_i}, k_{jn_j}) \leftarrow Euclidean(v_i, v_j)$ by equation (**1**) in Section 4.1.

Step 2.2. Optimize minimum $distance_{wmd}(p_i, r_j) \leftarrow \mathrm{WMD}(p_i, r_j)$ by equation (**2**) in Section 4.2.

Step 3. **Label Prediction**. Predict y_j for $r_j \in R$.

Step 3.1. Compute relationship between every two papers p_i and $p_{i'}$ in P.

Step 3.2. Form a relationship matrix where $Q \in \mathbb{R}^{n \times n}$ and $Q_{i,i'} = \mathrm{WMD}(p_i, p_{i'})$.

Step 3.3. Construct cover set $C \leftarrow \mathrm{CCA}(P, Q)$.

Step 3.4. Predict $y_j \leftarrow Predict(C, distance_{wmd}(p_i, r_j))$.

In Algorithm 1 above, $\mathrm{Euclidean}(v_i, v_j)$ is Euclidean Norm to measure the similarity between two tag vectors. It will be introduced in "Tags representation of submission and reviewer" section. The $\mathrm{distance}_{wmd}(p_i, r_j)$ is the deeper relationship between a paper p_i and a reviewer r_j calculated by the operation function $\mathrm{WMD}(p_i, r_j)$. It will be introduced in "Mining relationship between paper and reviewer" section. In Step 3, a relationship matrix Q is computed, and the element is the relationship $\mathrm{WMD}(p_i, p_{i'})$. According to the matrix Q and the paper set P, the classification algorithm CCA is used to form a cover set C. Finally, the research field label is predicted by the function $Predict(C, distance_{wmd}(p_i, r_j))$ based on the cover set C and $\mathrm{distance}_{wmd}(p_i, r_j)$. The details of Step 3 will be introduced in "Predict the label of reviewer" section.

Tags representation of submission and reviewer

Representing the tags of submission and reviewer is a fundamental task. The keywords of a paper are important tags in an academic repository, generally summarizing the core information of paper. The aim of our method is to transform keywords into a continuous space and carry the syntactic and semantic information. For this purpose, techniques from Word2vec package[1] are used, which contains two main models, namely skip-gram and CBOW. They could produce word embedding based a training corpus. For comparison experiments, we use the Google News (Mikolov et al. 2013b) as public corpus and choose the skip-gram model to obtain the distribution representation of tags (keywords). For example, a brief description of tags representation for a paper (or a reviewer) is presented in Fig. 2. From the trained neural network based language model, the tags $\{k_{i1}, k_{i2}, \ldots, k_{in_i}\}$ are represented as vectors $\{v_{i1}, v_{i2}, \ldots, v_{in_i}\}$. All of tag vectors have the dimensionality of 300.

Mining relationship between paper and reviewer

Mining the field relationship between paper and reviewer is crucial to the performance of reviewer recommendation. The easily-measured similarity between the word embeddings cannot describe the deeper field relationship. An optimization problem on the relationship between paper and reviewer is considered in our proposed method. We utilize a distance metric-Word Mover's Distance to measure the minimum distance as the deeper relationship between them.

Word Mover's Distance (WMD) (Kusner et al. 2015) is enlightened by a metric Earth Mover's Distance (EMD) (Rubner et al. 1998) in Computer Vision and it's a novel effective metric to measure the similarity or dissimilarity for a pair documents. An assumption that the distance (or similarity) between two words is a natural building block to measure the distance between two documents. In this paper, we regard the paper or the reviewer as documents, which consist of some tags. Further, the distance calculated from WMD can be the deeper relationship between a paper and a reviewer.

We begin with the word-level distance between two keywords k_{in_i} and k_{jn_j} in p_i and r_j respectively according to their learned word embedding v_i and v_j. Here we use Euclidean Norm to compute the relationship between two distributional vectors. The relationship is defined as:

$$\mathrm{d}(k_{in_i}, k_{jn_j}) = \left|v_i - v_j\right| = \mathrm{Euclidean}(v_i, v_j) \tag{1}$$

We convert a transportation problem to an optimization problem as Eq. (**2**). The research field correlation between p_i and r_j can be obtained owing to the keyword-level distance $\mathrm{d}(k_{in_i}, k_{jn_j})$ is a distance metric.

[1] https://code.google.com/p/word2vec/.

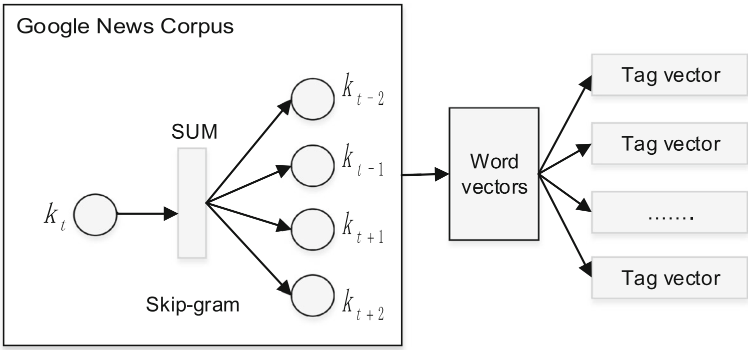


Fig. 2 Tags representation

$$
\begin{aligned}
distance_{wmd}\left(p_i, r_j\right) &= \min_{W \geq 0} \sum_{i,j=1}^{\boldsymbol{n}} W_{ij} \cdot d(k_{in_i}, k_{jn_j}) \\
&\text{Subject to:} \\
&\sum_{j=1}^{\boldsymbol{n}} W_{ij} = weight_{k_i} \forall i \in \{1, \ldots, \boldsymbol{n}\} \\
&\sum_{i=1}^{\boldsymbol{n}} W_{ij} = weight'_{k_j} \forall j \in \{1, \ldots, \boldsymbol{n}\}
\end{aligned} \tag{2}
$$

The core of our work as Eq. (**2**) showed is to calculate the deeper research field relationship between p_i and r_j. We first regard p_i or r_j as a keyword list. Then we denote $\text{weight}_{k_i} = tf_i / \sum_{i'}^{n} tf_{i'}$ as the weight of keyword k_i if k_i appears tf_i times in p_i, where $\boldsymbol{n}$ is the number of words in the list. And $\text{weight}'_{k_j} = tf_j / \sum_{j'}^{n} tf_{j'}$ is the weight of k_j in r_j. However, word embedding and Euclidean Norm can incorporate the semantic information between two individual words into the paper-reviewer level distance. Then the relationship between p_i and r_j is mined out by the minimum amount of summing up keyword-level distances, which are the similarities between the keywords in p_i and the keywords in r_j. Let $\mathrm{W} \in \mathrm{R}^{\boldsymbol{n} \times \boldsymbol{n}}$ be a flow matrix, where $W_{ij} \geq 0$ represents how much weight of word k_i in p_i moves to k_j in r_j, and $\boldsymbol{n}$ is the number of unique words in p_i and r_j. Theoretically, WMD allows every word travel into any word in total or in parts. When entirely moving p_i (all keywords in p_i) to r_j (all keywords in r_j), we ensure that the entire outgoing flow from word k_i should equal to weight_{k_i} and the entire incoming flow from word k_j should equal to weight'_{k_j}.

The $\text{distance}_{wmd}(p_i, r_j)$ represents the deeper research field relationship between p_i and r_j after an optimization. It's a good solution to mining the correlation by the similarities among their tag vectors. Therefore, we incorporate the deeper field relationship further in our method to predict the research field label for an unlabelled reviewer.

Predict the label of reviewer

According to the deeper research field relationship above, we transform the reviewer recommendation task into a classification issue. In another word, the solution is to predict the research field labels for reviewer candidates using a classification method. The fast and constructive learning method- Constructive Covering Algorithm (CCA) is applied to our recommendation task. CCA can construct some covers adaptively by computing the cover radius based on the relationship between given paper samples and the location of sample distribution space. During the constructive learning process of CCA (Wang 2008), the cover radius is related to the distances among the papers. We regard the relationship between papers as this distance, so the relationship mined from WMD are used rightly here. In our method, we only utilize CCA to predict the label of the reviewer candidate due to the relationship obtained from the WMD.

Before the CCA process, we first compute the relationship $\mathrm{WMD}(p_i, p_{i'})$ between every two papers p_i and $p_{i'}$. Then a relationship matrix Q is formed and the element $Q_{i,i'}$ is $\mathrm{WMD}(p_i, p_{i'})$. Given a training paper set $P = \{(p_1, l_1), (p_2, l_2), \ldots, (p_n, l_n)\}$ and the relationship matrix Q. CCA first trains the covers according to Q and then predicts the labels in the testing process for the reviewer candidates. The main steps of CCA is presented in Algorithm 2 below.

Algorithm 2: Constructive Covering Algorithm

Input: training set $P = \{(p_1, l_1), (p_2, l_2), \ldots, (p_n, l_n)\}$, relationship matrix Q, reviewer candidate set $R = \{r_1, r_2, \ldots, r_m\}$.

Output: Research field labels $\{y_1, y_2, \ldots, y_m\}$ for R.

Step 1. Select a p_i as the center of a cover randomly.

Step 2. Compute the cover radius **R** by equation (**3**).

Step 2.1. Compute $\mathbf{R}_{max} \leftarrow \mathbf{min}_{l_i \neq l_{i'}} Q_{i,i'}$.

Step 2.2. Compute $\mathbf{R}_{min} \leftarrow \mathbf{max}_{l_i = l_{i'}} \{Q_{i,i'} \mid Q_{i,i'} < \mathbf{min}_{l_i \neq l_{i'}} Q_{i,i'}\}$.

Step 2.3. Cover radius $\mathbf{R} \leftarrow (1-\mu) \cdot \mathbf{R}_{\mathbf{min}} + \mu \cdot \mathbf{R}_{\mathbf{max}}$.

Step 3. Construct a cover according to p_i and **R**.

Step 4. Go back to Step 1 till all samples are covered.

Step 5. Cover set $C = \{C_1^1, C_1^2, \ldots, C_1^{n_1}, C_2^1, C_2^2, \ldots, C_2^{n_2}, \ldots, C_k^1, \ldots, C_k^{n_k}\}$ (where C_i^j is one of all covers of i_{th} label) are formed.

Step 6. Predict $\{y_1, y_2, \ldots, y_m\}$ for R.

Radius computing is very important in covers training process, and we can clearly understand from Fig. 3. For current learning paper $p_{current}$ which is regarded as a cover center in a cover construction. Cover radius **R** is the compromise radius $\mathbf{R}_{\mathbf{com}}$ (Zhang et al. 2013),which is computed from Eq. (3) with a balanced parameter μ. $\mathbf{R}_{\mathbf{max}}$ is the minimum distance (i.e. relationship) between the cover center and the dissimilar (i.e. different category) papers, while $\mathbf{R}_{\mathbf{min}}$ is the maximum distance between the cover center and the similar (i.e. same category) papers. The $\mathbf{R}_{\mathbf{max}}$ and the $\mathbf{R}_{\mathbf{min}}$ are described as Eqs. (4) and (5) respectively. Notably, $\mathbf{R}_{\mathbf{min}}$ cannot beyond $\mathbf{R}_{\mathbf{max}}$, and three kinds of radius (i.e. $\mathbf{R}_{\mathbf{max}}$,$\mathbf{R}_{\mathbf{min}}$ and $\mathbf{R}_{\mathbf{com}}$) are three cover radius computing methods. We choose $\mathbf{R}_{\mathbf{com}}$ in our proposed method. Based on the relationship and distribution of the papers, compromise radius $\mathbf{R}_{\mathbf{com}}$ is computed and the covers of the different category are constructed adaptively.

$$\mathbf{R}_{\mathbf{com}} = (1-\mu) \cdot \mathbf{R}_{\mathbf{min}} + \mu \cdot \mathbf{R}_{\mathbf{max}} \tag{3}$$

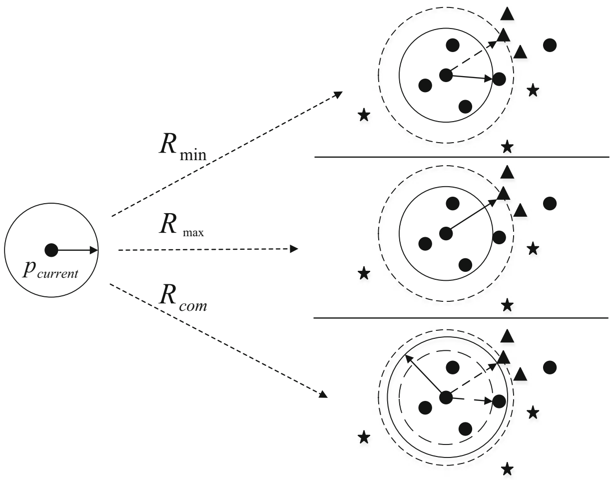


Fig. 3 Computing the new radius. There are three kinds of cover radius computing methods and the $\mathbf{R}_{com}$ we choose in this paper

$$\mathbf{R}_{\mathbf{max}} = \min_{l_i \neq l_{i'}} \boldsymbol{Q}_{i,i'} \quad i = 1, 2, \ldots, n \tag{4}$$

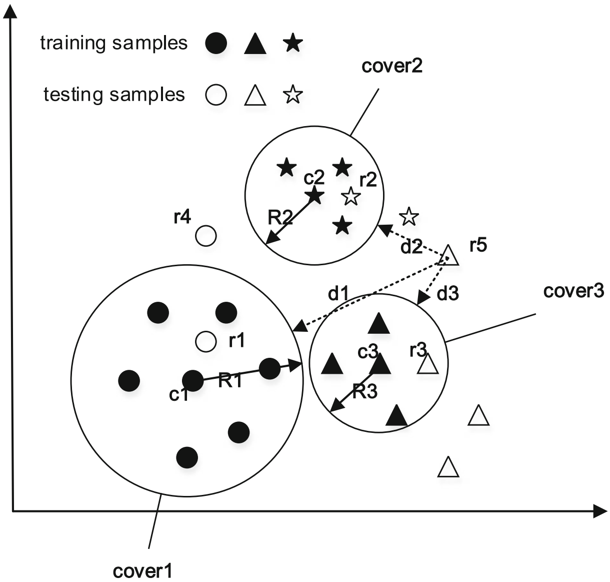
$$\mathbf{R}_{min} = \max_{l_i = l_{i'}} \{ \boldsymbol{Q}_{i,i'} | \boldsymbol{Q}_{i,i'} < \min_{l_i \neq l_{i'}} \boldsymbol{Q}_{i,i'} \} \quad i = 1, 2, \ldots, n \tag{5}$$

After the training process, a cover set $C = \{C_1^1, C_1^2, \ldots, C_1^{n_1}, C_2^1, C_2^2, \ldots, C_2^{n_2}, \ldots, C_k^1, \ldots, C_k^{n_k}\}$ are formed. Based on the cover set, reviewer candidates are treated as testing samples to predict the field labels.

Furthermore, we let the reviewer candidate set $R = \{r_1, r_2, \ldots, r_m\}$ be the testing set. The samples in P and R are located in the same distribution space. To be simple, we take the two-dimensional dataset as an example to describe the testing process. From the description in Fig. 4, there are three covers $\mathrm{cover}_1, \mathrm{cover}_2, \mathrm{cover}_3$ constructed which consist of three different categories and c_1, c_2, c_3 are the cover centers, R_1, R_2, R_3 are the cover radiuses respectively. The samples in black are training samples, and the samples in white are testing samples. We have two strategies for prediction process according to the location of the testing sample. First, for the reviewer r_1, r_2, r_3 which are in a cover, their labels are similar to the cover center of these covers. Second, for some samples which are not in the covers, such as r_4, r_5, etc. A simple method, which is according to the shortest distance from the sample to the different cover boundary, is applied to the prediction process. For example, d_1, d_2, d_3 are the distances from r_5 to the three cover boundaries. The label of r_5 is similar to c_3 because of the inequality $d_3 < d_2 < d_1$.

The field labels of all reviewer candidates are predicted based on two strategies here. The prediction is related to the reviewer candidate location, and it's an adaptive process. From the descriptions of Constructive Covering Algorithm in our method, the advantages are fast and constructive. First, every sample in the training set is learned only once. Based on the distribution of the current learned sample, it can either be learned immediately or memorized (the information of the cover construction can be memorized). If a sample is learned, it is removed from the training set immediately after learning. Not only that, the samples which are covered during training process are also removed and they cannot be

Fig. 4 Label prediction for reviewer



selected as a cover center anymore. Second, CCA begins without covers. It computes the radius and adds new covers according to the location of the sequentially coming data. This learning process can be applied to low or high dimensional datasets. Besides, it has 100% recognition rate for the test sample which is located in a cover. That's to say, it can increase the classification accuracy to some extent.

Experiments

In this section, we evaluate the efficiency of our method proposed in "WMD–CCA method for reviewer recommendation" section. We conducted experiments compared with nine methods, and we used four public datasets and one Synthetic dataset.

Experimental datasets and comparison methods

In this paper, the reviewer recommendation problem is transformed into a classification issue. We compare against nine methods and apply them to the text classification problem on four public datasets. Furthermore, to show the efficiency and the potential on reviewer recommendation for submissions, we evaluate our method on a synthetic paper-reviewer dataset.

Public dataset

TWITTER (Sanders 2011): a set of tweets labelled with sentiments 'positive', 'negative', or 'neutral'. *OHSUMED*: a collection of medical abstracts categorized by different cardiovascular disease groups (we use the 3rd class and the 7th class). **3-class** *MOV*[2] (Bo and Lee

[2] http://www.cs.cornell.edu/people/pabo/movie-review-data/.

2005): a collection of movie-review documents labelled subjective rating. **4-class** *MOV*[2] (Bo and Lee 2005): a collection of movie-review documents labelled subjective rating. The statistical description of public datasets is given in Table 1.

Synthetic dataset

The experimental data used in academic expert recommendation or co-author prediction is often from the scientific paper repository. The papers in this repository are often matched to the wrong authors due to duplication of full names and abbreviations of the researchers (i.e. reviewers) are very common. As a result, the recommended reviewers are not suitable in fact, and this problem has a great effect on the evaluation results (Tang et al. 2012). To solve it, we construct an experimental dataset from the Baidu Scholar[3] according to the Member List of the *Program Committee in NCIIP2017*[4] (The 6th National Conference on Intelligent Information Processing in 2017). The reviewers (experts) on this list are authentic and distinct due to the names and organizations.

Referring to the fifth edition of Chinese Library Classification (CLC), we randomly select 152 papers, which are published in latest years and belongs to three subareas in TP (*the field automation technology & computer technology*) with respective CLC code TP181 (*the subfield of Automatic reasoning & machine learning*), TP301 (*the subfield of theory & method*) and TP391 (*the subfield of information processing*). Then we randomly choose 38 reviewers from **NCIIP2017** and crawl their 5 papers published in latest years. For each of all 342 papers, we collect the keywords the author written and the CLC Codes. So we regard the 152 publications as submissions, and every reviewer in our experiment has five CLC Codes $r_{\emptyset} = \{y_1, y_2, y_3, y_4, y_5\}$. The description of the synthetic dataset is given in Table 2. Finally, we regard the 152 submissions and 38 reviewers as the training and the testing in our method, respectively.

We preprocess all datasets by removing all words in the SMART stop word list (Salton 1971). For all submissions and reviewers above, we regard the CLC Code as their research field label. The word embedding utilized in our experiments is the open-available word2vec word embedding, which obtained from the trained corpus (Google News) in the approach in Mikolov et al. (2013b).

Comparison methods

We compare the following nine methods with our proposed method. These methods are combinations of a text representation method and a classifier. However, LDA (Blei et al. 2003) (latent dirichlet allocation), LSI (Deerwester et al. 1990) (Latent Semantic Indexing), KNN (K-nearest neighbor), SVM, GaussianNB we used in experiments.

LDA–KNN It calculates the text topic distributions by LDA and KNN are used as classifiers based on *Euclidean Norm*

LDA–CCA It calculates the text topic distributions by LDA and CCA are used as classifiers based on *Euclidean Norm*

LSI–KNN It calculates the text topic distributions by LSI and KNN are used as classifiers based on *Euclidean Norm.*

[3] http://xueshu.baidu.com/.

[4] http://www.htu.edu.cn/nciip2017.

Table 1 Statistical description of public dataset

Name	Number	Doc length (AVG)	|Class|
TWITTER	3108	9.9	3
OHSUMED	953	88.5	2
3-class MOV	1027	103.1	3
4-class MOV	1027	103.1	4

Table 2 Statistical description of the synthetic dataset

Description	Value
Number of reviewers	38
Number of reviewer papers	190
Number of submitted papers	152
CLC Code number of a reviewer	5
CLC Code kinds of submitted papers	3
CLC Code kinds of reviewer papers	27

LSI–CCA It calculates the text topic distributions by LSI and CCA are used as classifiers based on *Euclidean Norm.*

WMD–KNN It calculates the text relationship by WMD and KNN is used as a classifier based on *Euclidean Norm.*

LDA–GaussianNB It calculates the text topic distributions by LDA and then feeds them to GaussianNB as features.

LDA–SVM It calculates the text topic distributions by LDA and then feed them to SVM as features.

LSI–GaussianNB It calculates the text topic distributions by LSI and then feed them to GaussianNB as features.

LSI–SVM It calculates the text topic distributions by LSI and then feed them to SVM as features.

For all methods, we split the datasets into 80/20 train/test. It's worth emphasizing that we set the neighborhood size (K $\in \{1,\ldots,5\}$) of KNN and the balanced parameter ($\mu \in \{0.70, 0.71, \ldots, 0.89\}$) in CCA as a result of empirical experience. The first five compared methods need a similarity function after text representation, and we all use *Euclidean Norm* here. For the last four compared methods, we first calculate the text representations and feed them to the classifier as features directly.

Evaluation metrics

We transform the recommendation problem as a classification issue, and here we evaluate the performance with *accuracy*, *precision* and *recall.*

$$\text{Accuracy} = \frac{1}{|R|}\sum_{j=1}^{|R|} 1^{\{y_{pre}=y_i\}} \tag{6}$$

Generally, the fact that a reviewer can review a paper as long as his research field is similar to this paper should be addressed. Given a reviewer r who has 5 publication label $r_\emptyset = \{y_1, y_2, y_3, y_4, y_5\}$ (5 published papers have 5 CLC codes, i.e. 5 publication labels), we predict the field label y_{pre} for reviewers and recommend them to the papers as their field label is similar.

Due to the practicality and particularity of our study, the evaluation metric *accuracy* is calculated according to that fact mentioned before. When y_{pre} is included in the $r_{\emptyset}$, the y_{pre} is equal to y_i. The *accuracy* is a real number in [0, 1] and $1^{\{\cdot\}}$ in this function is the indicator function. When y_{pre} is included in the $r_{\emptyset}$, it suggests that this reviewer has studied the research about the field y_{pre} and it's reasonable to assign this reviewer to review the papers which are labelled with y_{pre}.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

The metrics *precision* and *recall* are described as Eqs. (7) and (8), where TP denotes the number of true positives, FP denotes the number of false positives and FN denotes the number of false negatives.

Experiment results and discussion

The test accuracies of first five methods and WMD–CCA on 4 text classification datasets is presented in Fig. 5. Violin plot shows the result distribution of 6 methods with different K values and μ values. The average accuracy and the best accuracy of each method are presented clearly. On all datasets except OHSUMED, our method WMD–CCA outperformed other 5 methods. Although WMD–CCA was not better than WMD–KNN on OHSUMED, it was still better than other 4 methods. It's worth noting that WMD–CCA, LSI–CCA, LDA–CCA can almost obtain a better performance than WMD–KNN, LSI–KNN, LDA–KNN respectively. The performance of KNN is limited to the parameter K, but CCA is a constructive learning algorithm which can adaptively construct the covers based on different datasets. Besides, CCA has the 100% recognition rate for the learned samples which are in some covers. So the WMD–CCA can perform better than these five methods.

Combined with KNN or CCA, WMD–CCA attained the higher average accuracy and best accuracy than LDA–CCA, LDA–KNN, LSI–CCA and LSI–KNN on all 4 datasets. It can be deduced that WMD can deeply measure the relationship between each pair texts. WMD captures distributional syntactic and semantic information. Thus it can easily describe the real relationship between two text samples.

Additionally, Fig. 6 shows the performance of WMD–CCA on the synthetic paper-reviewer dataset. Based the fact mentioned, WMD–CCA performed much better than other methods on the real paper dataset. The average accuracy can approximately reach 0.73 and the best accuracy reached 0.84. Notably, the average accuracy increased by 4% than WMD–KNN and 8% promotion for the best accuracy. In Fig. 7, the accuracy all surpass 0.5 with different balance parameter μ values in WMD–CCA. Especially, for more than half of μ values, the accuracy exceeds the average accuracy 0.73. The higher accuracy value of WMD–CCA method indicates that the reviewer recommendation problem is more desirable to transform into the classification issue. However, our method has the promising efficiency and potential on the reviewer recommendation task.

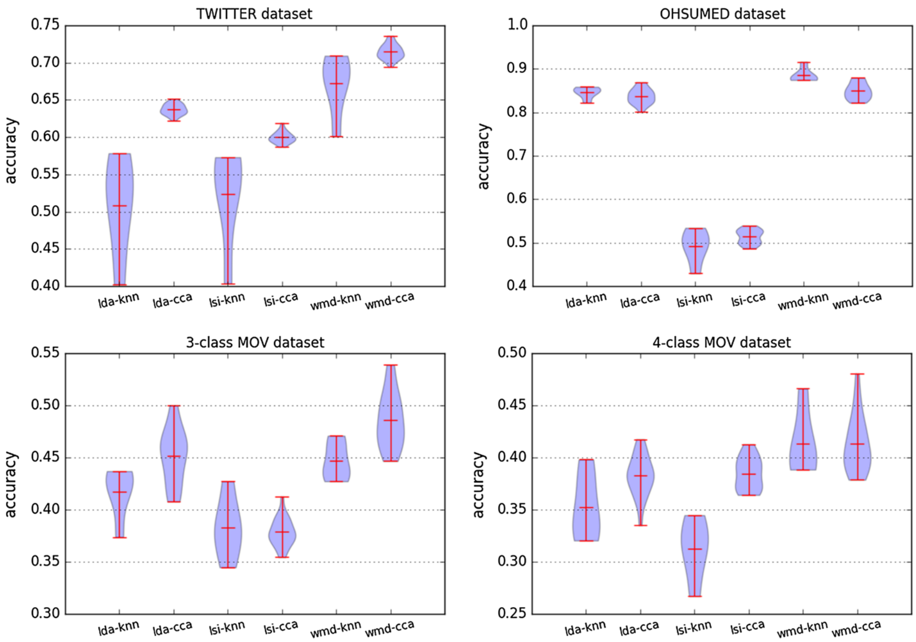


Fig. 5 Test accuracy results on 4 public datasets, compared to other 5 methods

Fig. 6 Accuracy of WMD–CCA on SYNTHETIC dataset

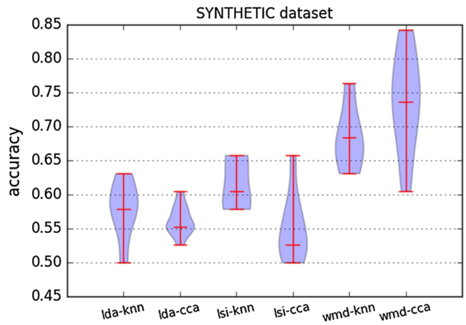


Fig. 7 Accuracy of different μ values on SYNTHETIC dataset

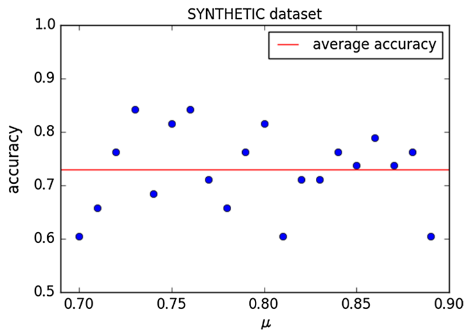


Table 3 Test accuracy results compared to 4 no-similarity function based methods

Dataset / Method	TWITTER	OHSUMED	3-class MOV	4-class MOV	SYNTHETIC
LSI-GuassianNB	0.162	0.571	0.364	0.442	0.605
LSI-SVM	0.685	0.555	0.417	0.432	0.605
LDA-GuassianNB	0.563	0.838	0.437	0.360	0.632
LDA-SVM	0.687	**0.880**	0.476	0.445	0.684
WMD-CCA	**0.736**	**0.880**	**0.539**	**0.481**	**0.842**

The best performance is denoted by bold font on different datasets

Table 4 Description of reviewer 8, reviewer 16, reviewer 37

	Publication labels	Predicted label	Field description
Reviewer 8	TP391, TP391, TP751, TP391, TP309	TP391	Information processing
Reviewer 16	TP301, TP301, TN911, TP18, TP391	TP301	Theory & method
Reviewer 37	TP181, TP301, TP181, TP181, TP391	TP181	Automatic reasoning & machine learning

(a) **(b)** **(c)**

Fig. 8 Tags (keywords) of reviewer 8, reviewer 16, reviewer 37, which are showed in sub-figures (**a**), (**b**), (**c**), respectively

Besides, Table 3 shows the best accuracies of last four compared methods and WMD–CCA. From Table 3, our method still has a better performance on all five datasets. Especially for the SYNTHETIC dataset, the accuracy can reach 0.842. From the comparisons with the first five methods, WMD can calculate the deeper relationship between two texts and CCA has high recognition rate. From the combination view, our proposed method WMD–CCA can perform better than LSI–GuassianNB, LSI–SVM, LDA–GuassianNB and LDA–SVM. It also shows the quite optimistic performance when compared with SVM.

To be more convincing, Table 4 shows the part of experimental results when $\mu = 0.76$ and the accuracy reach 0.84. We choose three reviewers (Reviewer 8, Reviewer 9 and Reviewer 37) from Member List in *NCIIP2017* and their tags (keywords extracted from publications) presented in Fig. 8. The research field information of three reviewers is easy to know from their word cloud. For example, we can easily judge that Reviewer 37 is related to the field of machine learning from the keyword tags such as “learning”, “supervised”, “multi”, “machine” and “label”. While the prediction of research field using WMD–CCA is TP181, which is related to the field of automatic reasoning & machine learning. The Reviewer 8 and Reviewer 16 are easily judged from Table 4 and Fig. 8.

To measure the quality of proposed method further, Tables 5 and 6 show the experimental results on precision and recall respectively. We report the best value of all methods on four public datasets. From the table of precision, WMD–CCA can rank first on three datasets and rank second on two datasets. WMD–CCA also rank first on four datasets and rank second on one dataset in the table of recall value. Generally, WMD–CCA doesn’t

Table 5 Experimental results on precision

Dataset / Method	TWITTER	OHSUMED	3-class MOV	4-class MOV	SYNTHETIC
LDA-KNN	0.574	0.862	0.446	0.403	0.713
LDA-CCA	0.595	0.859	0.491	0.396	0.715
LSI-KNN	0.540	0.534	0.434	0.314	0.671
LSI-CCA	0.548	0.549	0.421	0.361	0.697
LDA-GaussianNB	0.575	0.838	0.433	0.358	0.717
LDA-SVM	0.532	0.880	0.486	0.414	0.751
LSI-GaussianNB	0.672	0.563	0.426	0.380	0.535
LSI-SVM	0.506	0.544	0.292	0.475	0.599
WMD-KNN	**0.734**	**0.916**	0.491	0.469	0.745
WMD-CCA	0.729	0.881	**0.544**	**0.491**	**0.828**

The best performance is denoted by bold font on different datasets

Table 6 Experimental results on recall

Dataset / Method	TWITTER	OHSUMED	3-class MOV	4-class MOV	SYNTHETIC
LDA-KNN	0.578	0.859	0.437	0.398	0.405
LDA-CCA	0.650	0.859	0.481	0.408	0.408
LSI-KNN	0.572	0.534	0.427	0.345	0.406
LSI-CCA	0.606	0.545	0.422	0.393	0.385
LDA-GaussianNB	0.563	0.838	0.437	0.359	0.400
LDA-SVM	0.687	0.880	0.476	0.447	0.430
LSI-GaussianNB	0.188	0.571	0.364	0.442	0.339
LSI-SVM	0.682	0.555	0.417	0.432	0.337
WMD-KNN	0.727	**0.916**	0.471	0.466	0.435
WMD-CCA	**0.748**	0.880	**0.549**	**0.495**	**0.461**

The best performance is denoted by bold font on different datasets only measure the deeper relationship between two texts, but also predict the label accurately.

Conclusions

In this paper, we transform the reviewer recommendation problem into a classification issue by proposing a novel classification method named Word Mover' s Distance–Constructive Covering Algorithm (WMD–CCA). Firstly, the keywords of submissions and reviewers are represented as word embeddings. These word embeddings incorporate the complex semantic relationship, which is the basis of the deeper relationship further. Secondly, the deeper field relationship between submission and reviewer is computed by an optimization WMD from the keyword-level relationship. Thirdly, we learn from submissions with the field label information by a constructed learning algorithm CCA. It makes full use of the field information of papers, and it transforms the recommendation task into the classification issue ingeniously. Furthermore, we can assign the same field research reviewer to the submissions. Comparing to nine methods presents that WMD can deeply compute the relationship and CCA can conduct an accurate prediction process. Our experiment results on five public datasets present that our method has the potentiality in reviewer recommendation.

Regarding the paper-reviewer recommendation problem as a classification issue is an interesting and novel idea. But this classification issue has two important features which are multi-label and multi-granular. In the future, a multi-labels problem will be studied, which would improve the applicability of the real reviewer recommendation. From the coarse granular to the thin granular, we aim to recommend the more suitable reviewer to review the submission and improve the accuracy. Also, our proposed method is planned to be evaluated on different academic datasets from other research fields, including Digital Bibliography & Library Project (DBLP) in computer and information science. In addition, under the training of big data, reviewer or expert recommendation for scientific and technology projects is challenging to study.

Acknowledgements This work was partially supported by National Natural Science Foundation of China (Grants #61402006, #61602003 and #61673020), National High Technology Research and Development Program (863 Plan)(Grant #2015AA124102), Innovation Zone Project Program for Science and Technology of China's National Defense (Grant No. 2017-0001-863015-0009), the National Key Research and Development Program of China (2017YFB1401903), the Provincial Natural Science Foundation of Anhui Province (Grants #1508085MF113 and #1708085QF156), Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry (Forty-ninth batch) and the Recruitment Project of Anhui University for Academic and Technology Leader.

References

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). *Latent dirichlet allocation*. JMLR.org.

Bo, P., & Lee, L. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In *Meeting on association for computational linguistics, University of Michigan, USA, 2005* (pp. 115–124).

Charlin, L., Zemel, R. S., & Boutilier, C. (2012). A framework for optimizing paper matching. *Paper presented at the Proceedings of the twenty-seventh conference on uncertainty in artificial intelligence,* Barcelona, Spain.

Conry, D., Koren, Y., & Ramakrishnan, N. Recommender systems for the conference paper assignment problem. In *ACM conference on recommender systems, New York, 2009* (pp. 357–360).

Deep, K., & Das, K. N. (2008). Quadratic approximation based hybrid genetic algorithm for function optimization. *Applied Mathematics and Computation, 203*(1), 86–98.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391.

Diaz, F., Mitra, B., & Craswell, N. (2016). Query expansion with locally-trained word embeddings. *Paper presented at the proceedings of the 54th annual meeting of the association for computational linguistics*, ACL 2016, Berlin.

Fan, Z. P., Chen, Y., Ma, J., & Zhu, Y. (2009). Decision support for proposal grouping: A hybrid approach using knowledge rule and genetic algorithm. *Expert Systems with Applications, 36*(2), 1004–1013.

Fang, H., & Zhai, C. X. Probabilistic models for expert finding. In *European conference on Ir research, Rome, 2007* (pp. 418–430).

Jin, J., Geng, Q., Zhao, Q., & Zhang, L. Integrating the trend of research interest for reviewer assignment. In *International conference on world wide web companion, Perth, 2017* (pp. 1233–1241).

Karimzadehgan, M., & Zhai, C. X. Constrained multi-aspect expertise matching for committee review assignment. In *ACM conference on information and knowledge management, Hong Kong, 2009* (pp. 1697–1700).

Karimzadehgan, M., Zhai, C. X., & Belford, G. Multi-aspect expertise matching for review assignment. In *ACM conference on information and knowledge management, Napa Valley, 2008* (pp. 1113–1122).

Kolasa, T., & Krol, D. (2011). A survey of algorithms for paper-reviewer assignment problem. *IETE Technical Review, 28*(2), 123–134.

Kolasa, T., & Król, D. (2010). *ACO-GA approach to paper-reviewer assignment problem in CMS*. Berlin: Springer.

Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. From word embeddings to document distances. In *International conference on machine learning, Lille, 2015* (pp. 957–966).

Li, K., Cao, Z., & Qu, D. Fair reviewer assignment considering academic social network. In *Asia-Pacific Web, 2017* (pp. 362–376).

Liu, O., Wang, J., Ma, J., & Sun, Y. (2016). *An intelligent decision support approach for reviewer assignment in R&D project selection*. Amsterdam: Elsevier.

Liu, X., Suel, T., & Memon, N. (2014) A robust model for paper reviewer assignment. In *Eighth ACM conference on recommender systems, RecSys '14, Foster City, Silicon Valley, CA, USA* (pp. 25–32).

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013a). Efficient estimation of word representations in vector space. *Computer Science*. arXiv:1301.3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems, 26,* 3111–3119.

Mimno, D., & Mccallum, A. Expertise modeling for matching papers with reviewers. In *ACM SIGKDD International conference on knowledge discovery and data mining, San Jose, California, USA, August, 2007* (pp. 500–509).

Peng, H., Hu, H., Wang, K., & Wang, X. (2017). Time-aware and topic-based reviewer assignment. *Paper presented at the database systems for advanced applications—22nd international conference*, Suzhou, China.

Protasiewicz, J., Pedrycz, W., Kozłowski, M., Dadas, S., Stanisławek, T., Kopacz, A., et al. (2016). A recommender system of reviewers and experts in reviewing problems. *Knowledge-Based Systems, 106*(C), 164–178.

Ren, Y., Zhang, Y., Zhang, M., & Ji, D. Improving Twitter sentiment classification using topic-enriched multi-prototype word embeddings. In *Thirtieth AAAI conference on artificial intelligence, 2016* (pp. 3038–3044).

Rubner, Y., Tomasi, C., & Guibas, L. J. A metric for distributions with applications to image databases. In *International conference on computer vision, Bombay, 1998* (pp. 59–66).

Salton. (1971). *The SMART retrieval system—experiments in automatic document processing*. Upper Saddle River: Prentice-hall Inc.

Sanders, N. J. (2011). Sanders-twitter sentiment corpus.

Shen, Y., Rong, W., Jiang, N., Peng, B., Tang, J., & Xiong, Z. (2017). Word embedding based correlation model for question/answer matching. *Paper presented at the proceedings of the thirty-first AAAI conference on artificial intelligence*, San Francisco, California.

Shon, H. S., Han, S. H., Kim, K. A., Cha, E. J., & Ryu, K. H. (2017). Proposal reviewer recommendation system based on big data for a national research management institute. *Journal of Information Science, 43*(2), 147–158.

Sun, Y. H., Ma, J., Fan, Z. P., & Wang, J. (2008). A hybrid knowledge and model approach for reviewer assignment. *Expert Systems with Applications, 34*(2), 817–824.

Tang, J., Fong, A. C. M., Wang, B., & Zhang, J. (2012). A unified probabilistic framework for name disambiguation in digital library. *IEEE Transactions on Knowledge and Data Engineering, 24*(6), 975–987.

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. ArnetMiner:extraction and mining of academic social networks. In *ACM SIGKDD international conference on knowledge discovery and data mining, Las Vegas, Nevada, USA, 2008* (pp. 990–998).

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, D. (2008). Fast constructive-covering algorithm for neural networks and its implement in classification. *Applied Soft Computing Journal, 8*(1), 166–173.

Xu, Y., Ma, J., Sun, Y., Hao, G., Xu, W., & Zhao, D. (2010). A decision support approach for assigning reviewers to proposals. *Expert Systems with Applications, 37*(10), 6948–6956.

Yang, K. H., Kuo, T. L., Lee, H. M., & Ho, J. M. A Reviewer recommendation system based on collaborative intelligence. In *Ieee/wic/acm international joint conference on web intelligence and intelligent agent technology, Milan, 2009* (pp. 564–567).

Zhang, Y., Xing, H., Zou, H., Zhao, S., & Wang, X. A three-way decisions model based on constructive covering algorithm. In *International conference on rough sets and knowledge technology, Halifax, 2013* (pp. 346–353).