3-2016

Learning to find topic experts in Twitter via different relations

Wei WEI
Huazhong University of Science and Technology

Gao CONG
Nanyang Technological University

Chunyan MIAO
Nanyang Technological University

Feida ZHU
Singapore Management University, fdzhu@smu.edu.sg

Guohui LI
Huazhong University of Science and Technology

DOI: https://doi.org/10.1109/TKDE.2016.2539166

Published in IEEE Transactions on Knowledge and Data Engineering, 2016 March, Volume 28, Issue 7, Pages 1764-1778
https://doi.org/10.1109/TKDE.2016.2539166

Learning to Find Topic Experts in Twitter via Different Relations

Wei Wei, Gao Cong, Chunyan Miao, Feida Zhu, and Guohui Li

Abstract—Expert finding has become a hot topic along with the flourishing of social networks, such as micro-blogging services like Twitter. Finding experts in *Twitter* is an important problem because tweets from experts are valuable sources that carry rich information (e.g., trends) in various domains. However, previous methods cannot be directly applied to *Twitter* expert finding problem. Recently, several attempts use the relations among users and *Twitter List*s for expert finding. Nevertheless, these approaches only partially utilize such relations. To this end, we develop a probabilistic method to jointly exploit three types of relations (i.e., *follower* relation, *user-list* relation, and *list-list* relation) for finding experts. Specifically, we propose a *S*emi-*S*upervised *G*raph-based *R*anking approach (SSGR) to offline calculate the *global authority* of users. In SSGR, we employ a normalized Laplacian regularization term to jointly explore the three relations, which is subject to the supervised information derived from Twitter crowds. We then online compute the *local relevance* between users and the given query. By leveraging the *global authority* and *local relevance* of users, we rank all of users and find top-N users with highest ranking scores. Experiments on real-world data demonstrate the effectiveness of our proposed approach for *topic-specific* expert finding in *Twitter*.

Index Terms—Expert search, micro-blogging, Twitter, list, graph-based ranking

1 Introduction

Expert finding (a.k.a., *expert search* [9]), which aims at identifying people with the relevant expertise or experiences on a given topic query, has been studied broadly in domains such as enterprise [9], [10], question answering [15], [43], Web [13] and academic society [14], [16].

Recently, *expert finding* problem has gained increasing attention in social media [17], [18], such as micro-blogging services like *Twitter*, a new type of social media in providing a publicly available channel for users to publish 140-character short messages (i.e., *tweets*). *Twitter* has gained huge popularity and gathered a tremendous amount of *tweets* in recent years[1]. These *tweets* cover extremely wide and diverse topics, such as routine activities or experiences, top news, technology, and myriad of other highly specialized areas, etc. Correspondingly, users in *Twitter* have rich expertise on various topics and finding these *topic-specific* experts paves a way to enable others to retrieve or follow the *relevant* and *trustworthy* information on a specific topic in micro-blogging services [3], [4], [5]. For example, if a *Twitter* user wants to follow expert users for receiving *tweets* that are highly relevant for an *event* topic like "Boston Marathon bombings", or follow users whose *tweets* are worthy of reading for a *domain-specific* topic like "machine learning". In addition, identifying such users is also a *preprocessing* step towards many applications like opinion mining [6] and name entity recognition (NER) [3], [7], [8]. For instance, opinions mined from beauticians' tweets are more likely to favor a cosmetic manufacturer (e.g., *Dior*) than those from common users.

Nevertheless, the problem of *Twitter* expert search differs from the conventional expert search problem [9], [10], [13], [14], [16], which generally relies on the assumption that all the documents associated with the candidate experts contain tacit knowledge related to the expertise of individuals [9], [10]. However, this might not be true in *Twitter*, as users' published tweets might not be directly related to their expertise, such as a rumormonger [1], [2], who is not an expert, but may publish/retweet a substantial amount of tweets containing the topic words. Therefore, the problem of expert finding in Twitter is more challenging.

There exist several attempts for the *Twitter* expert finding problem. For example: (i) Several traditional methods like *PageRank*-based method [3] and *clustering*-based method [4], make use of the follower relations as well as users' bios and tweets to infer the *general* influence of users on different topics; and (ii) a recent study (*Cognos* [5]) proposes to identify *topic-specific* experts by mining the meta-data of *Twitter List*s. A *Twitter* list is usually created by a user to group her followings according to a criterion, e.g., having expertise on "data mining". Intuitively, the meta-data (e.g., title) of a list can be viewed as the *crowdsourced* topical annotations of users in that list [5]. For instance, a *user* involved in a list named "machine learning" is likely to have expertise on machine learning. Hence, a user contained in many lists on a theme is very likely

1. As of July 2013, on average 58 million *tweets* were posted daily by more than 550 million active *Twitter* users, http://www.statisticbrain.com/twitter-statistics/, accessed on 01/10/2013.

- *W. Wei and G. Li are with the School of Compute Science and Technology, Huazhong University of Science and Technology, Luoyu Road, Wuhan 430074, P.R. China. E-mail: weiwei8329@gmail.com, guohuili@hust.edu.cn.*
- *G. Cong and C. Miao are with the School of Computer and Engineering, Nanyang Technological University, Singapore 639798. E-mail: {gaocong, ascymiao}@ntu.edu.sg.*
- *F. Zhu is with the School of Information Systems, Singapore Management University, Singapore 178901. E-mail: fdzhu@smu.edu.sg.*

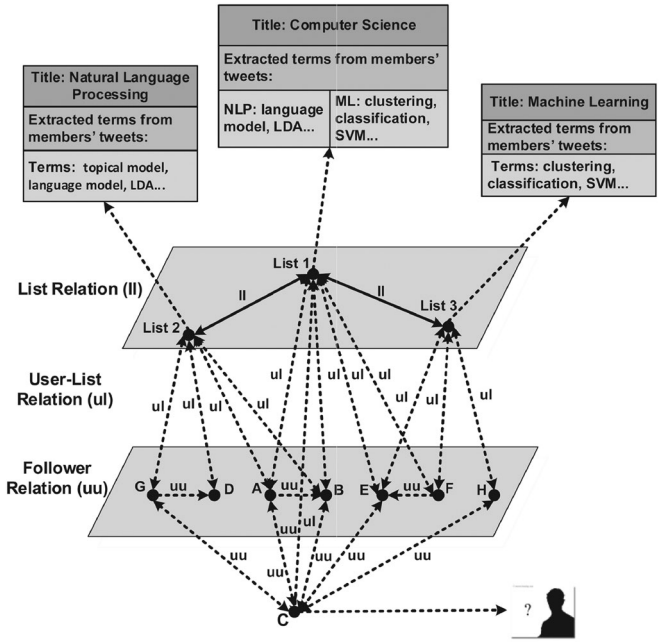


Fig. 1. *Example*. An illustration of different types of relations among users and lists.

to be an expert on that topic [5]. It has been found that *Cognos* [5] utilizing *user-list* relation is more effective for inferring the expertise of users than previous methods based on *follower* relations as well as users' bios and tweets.

However, existing approaches [3], [4], [5] only partially utilize either *follower* relation or *user-list* relation *alone*, and they are thus insufficient for *Twitter expert finding* problem. To this end, we propose to jointly exploit such relations for accurately inferring the users' domain of expertise. We illustrate this with an example in Fig. 1. Without loss of generality, suppose we infer the expertise of user C only depending on the *user-list* relation between user C and List1 ("Computer Science, CS"), and thus we only deduce that user C has expertise on *CS*. However, if we exploit more available relations (e.g., *follower relation* or *list-list* relation), we can more accurately identify the expertise of user C, such as: (1) The majority of users (e.g., user A, B and D) in List2 ("Natural Language Processing, NLP") are the followers of user C. Hence, user C is likely to be an expert on *NLP*. Furthermore, the semantic similarity between List1 and List2 can be used to strengthen this possibility. (2) Similarly, we can infer that user C may have expertise on *ML* ("Machine Learning"). Consequently, by means of more relations, we not only certify that user C is more likely an expert on *CS*, but also refine her possible specific expertise on *CS*, i.e., *NLP* and *ML*.

Therefore, we propose an approach to jointly exploit the different types of relations among users and lists for improving the accuracy of finding experts on a given topic in *Twitter*. Specifically, we take into account two types of information to target *Twitter* expert finding problem, namely: (i) *Local Relevance*, the similarity between users' published *tweets* and the given query; and (ii) *Global Authority*, the *global* expertise scores of users on a given topic in *Twitter*.

Consequently, we estimate the probability of each user being an expert on a given topic from two aspects. First, we propose an innovative *S*emi-*S*upervised *G*raph-based *R*anking approach, called SSGR, to compute the *global authority* of users on a given topic, by jointly exploiting different types of relations in Twitter Lists and follower graphs. To the best of our knowledge, this is the first attempt that targets *expert finding* problem in *Twitter* by utilizing all of these relations. In particular, SSGR employs: (a) a normalized *Laplacian* regularization term to smooth the ranking of users and lists on three different *topic-specific* graphs; and (b) a loss term to ensure the *global authority* of users is in accordance with the wisdom of *Twitter crowds*. Second, we also propose a Gaussian-based method to estimate the *local relevance* of candidates for arbitrary topical queries. The extensive experiments conducted on a real-world data set demonstrate the effectiveness of our proposed method over the state-of-the-art baselines.

Roadmap. The remaining of the paper is organized as follows. In Section 2, we review the related work. The proposed method is presented in Sections 3, 4, and 5, followed by the experimental results in Section 6. Finally, Section 7 concludes this paper.

2 Related Work

Early Work. The task of *expert finding* has attracted extensive attention of information retrieval community since it was included in TREC enterprise track [21]. Most of work [9], [10], [19], [23] on *organization expert search* problem generally falls into two categories, namely *profile*-centric and *document*-centric methods. These methods assume that individuals' published documents are relevant to their expertise with different degrees of match, and they focus on modeling the associations between *documents* and *candidate experts*. In contrast, *tweets* are clearly attached to their publishers. Additionally, there exist many methods of *expert finding* for other domains, such as academic search [14], [16], Web [13] and Question Answering [15].

Influential user identification. Most of existing work on expert identification in social networks focuses on finding influential users from different types of social networks, such as Community Question Answering [11], [12], [43], Blog [28], Academic Social Network [24], Twitter [34], [44], and other social networks [39], [42]. However, these methods do not consider topical dimension, which thus cannot identify *topic-specific* experts as we do in this work.

Topic-specific Expert Finding. Several proposals [5], [3], [4] study the problem of identifying *topic-specific* experts in Twitter. Weng et al. [3] propose an approach called *TwitterRank*, which works in two steps. First, it employs Latent Dirichlet Allocation (LDA) model [31] to detect the topics of individuals based on their tweets. Second, for each topic, it builds a weighted graph by considering both the topical similarity between two users and follower graph, and then employ *PageRank* algorithm [35] to find topic-specific influential users. *TwitterRank* tends to select the well-known *user*s that are highly visible over the follower graph. Another approach proposed by Pal et al. [4], extracts users' features from the follower graph and users' posted tweets, and then employs a *Gaussian*-based mixture model to cluster users for ranking. One fundamental difference of our approach from *TwitterRank* and Pal's work is that we utilize the wisdom of *Twitter* crowds, enclosed in Twitter Lists, as

the supervised information to infer the topical expertise of *users*. Correspondingly, our approach handles the *user-user* relation, *user-list* relation and *list-list* relation while *TwitterRank* and Pal's work only consider *user-user* relation.

Ghosh et al. [38] propose to utilize *Twitter List* to analyze the attributes of Twitter users. In their subsequent work, they develop a system named *Cognos* [5] to infer the topical expertise of *users* by utilizing only user-list relation in Twitter *Lists*, which captures the wisdom from *Twitter* crowds. Cognos represents each user by the meta-data of Twitter lists that contain the user, and then employs a similarity measure [32] to compute the similarity score between each user and a topical query, which is used to rank users for search. Intuitively, Cognos tends to choose users that are contained in many lists whose meta-data contain the query. The experimental results show that *Cognos* outperforms the approach using social relations [4]. In contrast, our method is able to make use of three types of relations for identifying experts.

Graph-based Ranking. Graph-based ranking methods [25], [27] have been used for expert finding, such as Hyperlink-Induced Topic Search (HITS) based expert authority [26], PageRank-based user influence [3], and probabilistic random walk on expertise graphs [19]. However, these methods heavily rely on a single type of relations and are usually topic-irrelevant. In contrary, we take into account three different types of relations to identify the topic-specific experts in *Twitter*.

Other work on Twitter List. In addition, Twitter lists have also been used for other purposes such as entity link (ZenCrowd) [33] and news curators finding [34]. Welch et al. [41] utilize the *Lists* features as a context to find the source of topic information. However, the purposes of these studies are different from the current work and thus will not be discussed in detail.

3 Overview of Proposed Approach

We first give the statement of our *expert search* problem, in Section 3.1, and then present an overview of our proposed approach.

3.1 Problem Statement

Let the set of *Twitter* users be $\mathcal{U} = \{u_i\}_n$ (n is the number of users), which are candidate experts. User u_i's posted tweets, bio and the meta-data of lists containing u_i are concatenated and form a pseudo-document, which is referred to as the context of u_i, denoted by d^{u_i}.

A topic query is defined as Q comprising several terms, namely $Q = \{t_1, \ldots, t_{|Q|}\}$, where $|Q|$ is the number of terms. Then, the *topic-specific expert finding* problem is to rank a set of candidate experts $\mathcal{U}$ based on the relevance of their expertise to the topic query Q. The probability of user u_i in $\mathcal{U}$ being an expert on Q can be estimated via Bayes theorem by following previous work on expert search:

$$\Pr(u_i|Q) = \frac{\Pr(Q|u_i)\Pr(u_i)}{\Pr(Q)} \propto \Pr(Q|u_i)\Pr(u_i), \tag{1}$$

where $\Pr(Q)$ is the prior probability of Q; $\Pr(u_i)$ is the prior probability of candidate u_i. As $\Pr(Q)$ is the same for all candidate experts and $\Pr(u_i)$ is generally assumed uniform over $\mathcal{U}$ [2], they do not affect the rankings of candidate experts, and thus are ignored.

Therefore, the problem is transformed to estimate the probability of a query Q given candidate u_i, i.e., $\Pr(Q|u_i)$. Many language models are proposed for this task [9], [10], [19]. By following the work [3], we treat each term t in Q as a potential topic, and adopt the query likelihood model to approximately estimate the probability $\Pr(Q|u_i)$,

$$\begin{aligned} &\Pr(Q|u_i) = \prod_{\substack{j=1 \\ t_j \in Q}}^{|Q|} \Pr(t_j|u_i), \\ &\qquad \Rightarrow \\ &L(Q,u_i) = \sum_{\substack{j=1 \\ t_j \in Q}}^{|Q|} \log\big(\Pr(t_j|u_i)\big) \propto \sum_{\substack{j=1 \\ t_j \in Q}}^{|Q|} \log\big(\Pr(u_i|t_j)\Pr(t_j)\big), \end{aligned} \tag{2}$$

where $L(Q,u_i) \equiv \log\big(\Pr(Q|u_i)\big)$; $\Pr(u_i|t_j)$ indicates the probability of candidate u_i being an expert on t_j over $\mathcal{U}$; $\Pr(t_j)$ is the prior probability of t_j. Similarly, it is uniform for all candidates and thus is ignored.

An expert to query Q should not only be the authority on Q, but also publish many relevant *tweets* containing the terms of Q. To characterize the two aspects, we incorporate two factors in estimating $L(Q,u_i)$: (i) *global authority*; and (ii) *local relevance*.

Global Authority. It indicates the *global* expertise score of a user u_i on a potential topic t_j in *Twitter*, i.e., $\Pr(u_i|t_j)$. By following the work [19], it can be calculated as follows,

$$\Pr(u_i|t_j) = \frac{\mathcal{H}(u_i,t_j)}{\sum_{u' \in \mathcal{U}} \mathcal{H}(u',t_j)}, \tag{3}$$

where $\mathcal{H}(u_i,t_j)$ is the scoring function that assigns a score to user $u_i \in \mathcal{U}$ proportional to u_i's *global authority* on t_j. We will present our proposed method of computing the global authority in Section 4.

Local Relevance. It denotes the *local* similarity between user u_i and the given query Q over the context d^{u_i} of u_i. To consider the sequence of terms in Q, we take each two adjacent terms in Q for computing the local relevance, $\mathcal{K}(t_j, t_{j+1}; d^{u_i})$. For example, given query $Q = \{t_1 t_2 t_3\}$, we compute $\mathcal{K}(t_1, t_2; d^{u_i})$ and $\mathcal{K}(t_2, t_3; d^{u_i})$.

Consequently, Eq. (2) is converted as follows:

$$\begin{aligned} L(Q,u_i) &\propto \sum_{\substack{j=1 \\ t_j \in Q}}^{|Q|} \log\big(\Pr(u_i|t_j)\big) \\ &\propto \frac{1}{2} \sum_{\substack{j=1 \\ t_j \in Q}}^{|Q|-1} \big(\log(\Pr(u_i|t_j)) + \log(\Pr(u_i|t_{j+1}))\big)\mathcal{K}(t_j, t_{j+1}; d^{u_i}) \\ &\propto \frac{1}{2} \sum_{\substack{j=1 \\ t_j \in Q}}^{|Q|-1} \big(\Pr(u_i|t_j)\Pr(u_i|t_{j+1})\big)\mathcal{K}(t_j, t_{j+1}; d^{u_i}). \end{aligned} \tag{4}$$

In particular, when $|Q| = 1$, $L(Q,u_i) \propto \log\big(\Pr(u_i|t_1)\big)$. The remaining problem is how to compute the *global*

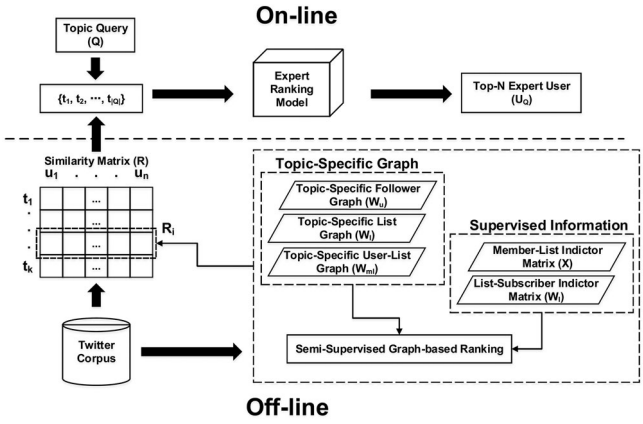


Fig. 2. Overview of proposed approach.

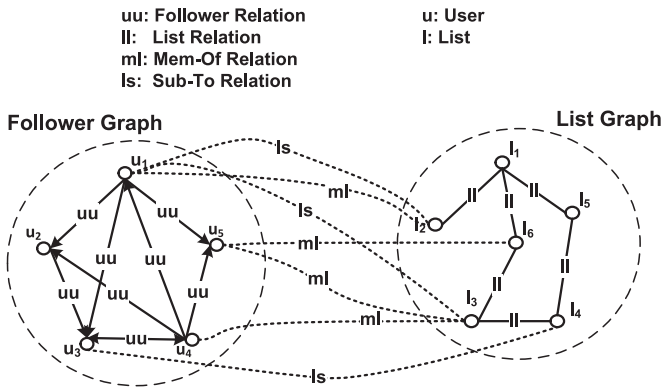


Fig. 3. *Example*. An illustration of different types of relations between users and lists.

authority score $\mathcal{H}(u_i, t_j)$ and *local relevance* $\mathcal{K}(t_j, t_{j+1}; d^{u_i})$. We will detail them in the following sections, respectively.

3.2 Overview

Next, we present an overview (as shown in Fig. 2) of our approach to addressing the *topic-specific expert finding* problem. Specifically, it consists of two components, namely, an *offline* graph-based ranking algorithm (called SSGR, detailed in Section 4) to learn the *global authority* of each candidate and an *online* ranking model (named RM, detailed in Section 5.2) to select top-N relevant experts on the given query. In particular, each term[2] t in *Twitter* is treated as a potential topic by following the work [3].

- We first construct an authority matrix (similar to the inverted index) $\mathbf{R}$ over the *Twitter* corpus. Specifically, each row $\mathbf{R}_i \in \mathbf{R}$ is *offline* computed by SSGR for each term t in Twitter, in which we jointly exploit the three different relations of users and *Twitter* lists for inferring the *global authority* of each candidate on t in *Twitter*.
- For a given topic query $Q = \{t_1, \ldots, t_{|Q|}\}$, we use an *online* ranking model (i.e., RM), based on the corresponding rows in $\mathbf{R}$ for terms contained in Q, to select top-N users as experts on Q, by taking into account the *global authority* and *local relevance* of candidates (*rf*. Eq. (4)).

Remark. As the learning of the *global authority* of candidates is computed *offline*. Hence, the correlation between terms is considered in the *online* ranking model for *multiple term query* (*rf*. Eq. (4)).

4 Learning the *Global* Authority

To learn the *global authority* of candidate users on a *single term* topic query[3] (denoted by Q_t), we present a novel semi-supervised graph-based ranking method, called SSGR. It is capable of exploiting the different relations (i.e., *follower relation*, *user-list relation* and *list-list relation*) among users and lists to mutually reinforce the ranking of users and lists for inferring the *global expertise scores* of users on Q_t. We present the User-List Interaction (ULI) graph to model the different relations in Section 4.1, followed by the intuitive benefits of jointly exploiting the three relations for calculating the *global authority* of candidates on a given topic in Section 4.2. Then we present the proposed method SSGR in Section 4.3.

4.1 User-List Interaction Graph (ULI)

In this section, we present the definition of ULI graph (as shown in Fig. 3). For clarity, some notations and their definitions are listed in Table 1.

Let $G = \{\mathcal{U} \cup \mathcal{L}, E\}$ be the ULI graph, where $\mathcal{U} = \{u_i\}_n$ and $\mathcal{L} = \{L_i\}_m$ denote a set of n users and a set of m lists, respectively, and E denotes the edge set which comprises three different relations, namely (i) *follower* relation, an edge between a user and her follower (denoted by e_{uu}); (ii) *user-list* relation, an edge between a user and a list, which consists of two types of edge: a) MEM-OF relation: an edge between a member user and her included list (denoted by e_{ml}); and b) SUB-TO relation: an edge between a subscribe user and her subscribing list (denoted by e_{ls}); (iii) *list-list* relation, an edge between two lists (denoted by e_{ll}). Correspondingly, there are three types of *topic-specific* graphs related to a given topic (Q_t), which are:

- $\mathbf{W_u}$: a n-by-n symmetric topic-specific follower graph, in which w_{ij}^u denotes the similarity between user u_i and her follower u_j for a given topic. Note each entry in $\mathbf{W_u}$ only considers the symmetric relation between two users, i.e., u_i and u_j follow each other;
- $\mathbf{W_l}$: a m-by-m symmetric topic-specific list graph, which is generated based on the *mutual k-nearest neighbor graph* [40], in which w_{ij}^l denotes the similarity between two different lists, i.e., L_i and L_j, for a given topic;
- $\mathbf{W_{ml}}$: a n-by-m topic specific user-list graph, in which w_{ij}^{ml} denotes the similarity between user u_i and list L_j containing u_i for a given topic. Each entry in $\mathbf{W_{ml}}$ refers to MEM-OF relation, i.e., user u_i is included in list L_j.

Similarity Measure. Given a ULI graph, one way to compute the similarity w_{ij} (e.g., w_{ij}^u, w_{ij}^l or w_{ij}^{ml}) between two objects, denoted by d_i and d_j (e.g., user u or list L) under Q_t is given in Eq. (5)

2. We removed non-English characters, stopwords, punctuation as well as the high-frequency words in Twitter (e.g., RT), and keep the left.
3. The topic query corresponds to a single term in *Twitter*.

TABLE 1
Notations and Definitions

Notation	Definition
Query	
Q	Given topic query, $Q = \{w_1, \ldots, w_{\|Q\|}\}$ and $\|Q\| \geq 1$
Q_t	Single term topic query, $Q_t = \{t\}$ and $\|Q_t\| = 1$
Node Set	
$\mathcal{U}$	User set, $\mathcal{U} = \{u_i\}_{i=1}^{n}$, n is the number of users
$\mathcal{L}$	List set, $\mathcal{L} = \{L_i\}_{i=1}^{m}$, m is the number of lists
Context	
d^{u_i}	Context of u_i, including u_i's bio, posted tweets and the meta-data of lists containing u_i
d^{L_i}	Meta-data of L_i
Graph	
$\mathbf{W}_u$	$\mathbf{W}_u = \{w_{ij}^u\}_{n\times n}$, w_{ij}^u denotes the similarity between user u_i and her follower u_j
$\mathbf{W}_l$	$\mathbf{W}_l = \{w_{ij}^l\}_{m\times m}$, w_{ij}^l denotes the similarity between list L_i and list L_j
$\mathbf{W}_{ml}$	$\mathbf{W}_{ml} = \{w_{ij}^{ml}\}_{n\times m}$, w_{ij}^{ml} denotes the similarity between user u_i and the list containing that user

$$w_{ij} = \frac{N(Q_t, d_i, d_j)}{N(d_i, Q_t) + N(d_j, Q_t)}, \tag{5}$$

where $N(Q_t, d_i, d_j)$ is the co-occurrences of Q_t in d_i and d_j, and $N(Q_t, d_i)$ is the number of occurrences of Q_t in d_i. However, Eq. (5) might be problematic in some cases, for example, for a topic query like "travel", assume that user u_i and user u_j have a high overlap between their published *tweets* on "travel", but if u_i never mentions the term "travel" in her *tweets*, $N(Q_t, d_i, d_j)$ will be 0. Moreover, Eq. (5) also ignores the similarity between u_i and u_j. To address the problem, we compute the similarity by

$$w_{ij} = \frac{\Pr(Q_t|d_i) + \Pr(Q_t|d_j)}{2} Cosine(d_i, d_j), \tag{6}$$

where $Cosine(d_i, d_j)$ is the cosine similarity of two objects d_i and d_j, each word probability vector (i.e., unigram) of document d_i (or d_j) is based on the TF-IDF [37] method, and $\Pr(Q_t|d_i) = \frac{N(Q_t,d_i)}{\sum_{w\in d_i} N(w,d_i)}$.

4.2 Intuitions

Recall that the information on users alone might be insufficient for measuring the global authority of candidates on the given topic. We propose to jointly exploit three different types of relations (i.e., follower relation, list-list relation and user-list relation) for inferring the *global authority* of candidates on Q_t. The motivation is based on the intuitions as follows.

- *Intuition 1 (Follower Relation)*. Users that are socially connected are more likely to share similar interests (Homophily [3], [29]). Hence (a) if a user is followed by another user with high *global authority* on Q_t, this user is more likely an expert on Q_t; and (b) the more followers of a user are experts on a topic, the more likely that the user is an expert on that topic;
- *Intuition 2 (User-list Relation)*. In-depth analysis of *user-list* relation is helpful to infer the expertise of users [5]. We explore two types of *user-list* relations: (a) MEM-OF *relation*, i.e., a set of users are included in a list. A list is built by a user to group her followings sharing a common characteristic. Hence, intuitively, i) if a user is relevant to the lists containing her, the user is likely to an expert on topic Q_t that is relevant to the lists; ii) if a user is contained in many lists relevant to Q_t, the user is likely to be an expert on Q_t. (b) SUB-TO *relation*, i.e., a set of users subscribe to a list. It is analogous to *follower relation*, i.e., this relation is a strong indicator that the users subscribing to a list are interested in the topic of that list. Intuitively, the more subscribers are experts on a topic, the more likely the subscribed list is relevant to that topic. In particular, we use (a)-ii) and (b) of *user-list* relation as the supervised information in our proposed model.
- *Intuition 3 (List-List Relation)*. If a list L_i is highly similar (i.e., w_{ij}^{ll}) to another list L_j that is relevant to Q_t, list L_i is also likely to be relevant to Q_t. Note that exploring the similarity between lists aims to find relevant lists for Q_t, which can be used to enhance the relevance of users in such lists to topic Q_t.

4.3 Semi-Supervised Graph-Based Ranking

Based on these intuitions, we propose a *semi-supervised* graph-based ranking method, named SSGR, for computing the *global authority* of a user on the given topic Q_t.

4.3.1 Graph-Based Regularization Framework

In this section, we present the proposed graph-based regularization framework, which comprises two terms: (i) *regularization term*, which is used to smooth the ranking scores on the graph; (ii) *loss term*, which aims to ensure the ranking scores are consistent with the supervision information.

Let n-dimensional vector $\hat{\mathbf{f}} = [f_1, \ldots, f_n]^\top$ be the ranking scores of users and m-dimensional vector $\hat{\mathbf{g}} = [g_1, \ldots, g_m]^\top$ be the ranking scores of lists. In particular, the ith entry of $\hat{\mathbf{f}}$ (i.e., f_i) denotes the *global authority* of user u_i on the given topic, and the ith score in $\hat{\mathbf{g}}$ (i.e., g_i) denotes the relevance between list L_i and the given topic. In fact, we are only interested in the ranking scores of users to identify topic-specific experts. We also consider the ranking scores for lists in our framework because the ranking scores of users and lists will reinforce each other mutually as explained in the intuitions in Section 4.2. Formally, the ranking framework is formulated as the following optimization problem

$$\langle \hat{\mathbf{f}}, \hat{\mathbf{g}} \rangle = \underset{\hat{\mathbf{f}}\geq 0, \hat{\mathbf{g}}\geq 0}{\arg\min} \Big(\lambda \mathcal{F}(\mathbf{W_u}, \mathbf{W_l}, \mathbf{W_{ml}}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) + (1-\lambda)\ell(\hat{\mathbf{f}}, \hat{\mathbf{g}}) \Big), \tag{7}$$

where $\mathcal{F}$ is a regularization term to smooth the expertise scores (i.e., *global authority*) of users; the affinity matrices (i.e., $\mathbf{W_u}, \mathbf{W_l}, \mathbf{W_{ml}}$) are computed in a *topic-specific* manner; ℓ is a loss term that aims to ensure the expertise scores of users are consistent with the wisdom of *Twitter* crowds; and λ is a parameter to trade-off the contributions of regularization term $\mathcal{F}$ and loss term ℓ.

4.3.2 Regularization Term

Within the framework, the regularization term aims to give similar ranking scores for similar users (and similar lists) by considering three different types of similarities of users and lists, namely, the similarity between a user and her followers, the similarity between a user and the lists containing her, and the similarity between two lists. The regularization term $\mathcal{F}$ is defined as follows, based on the principle of *normalized Laplacian regularization*,

$$\begin{aligned}\mathcal{F}(\mathbf{W_u}, \mathbf{W_l}, \mathbf{W_{ml}}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \alpha_1 \sum_{i,j=1}^{n} w_{ij}^{u} \Big(\frac{1}{\sqrt{[\mathbf{D}_u]_{ii}}} f_i - \frac{1}{\sqrt{[\mathbf{D}_u]_{jj}}} f_j \Big)^2 \\ + \alpha_2 \sum_{i,j=1}^{m} w_{ij}^{l} \Big(\frac{1}{\sqrt{[\mathbf{D}_l]_{ii}}} g_i - \frac{1}{\sqrt{[\mathbf{D}_l]_{jj}}} g_j \Big)^2 \\ + \alpha_3 \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}^{ml} \Big(\frac{1}{\sqrt{[\mathbf{D}_{ml}^{u}]_{ii}}} f_i - \frac{1}{\sqrt{[\mathbf{D}_{ml}^{l}]_{jj}}} g_j \Big)^2,\end{aligned} \tag{8}$$

where $\mathbf{D}_u$ and $\mathbf{D}_{ml}^{u}$ are n-by-n diagonal matrix, $\mathbf{D}_l$ and $\mathbf{D}_{ml}^{l}$ are m-by-m diagonal matrix. The (i,i)-element of $\mathbf{D}_u$, $\mathbf{D}_l$, $\mathbf{D}_{ml}^{u}$ and $\mathbf{D}_{ml}^{l}$ equals to the sum of ith row of $\mathbf{W}_u$, the sum of ith row of $\mathbf{W}_l$, the sum of ith row of $\mathbf{W}_{ml}$, and the sum of ith column of $\mathbf{W}_{ml}$, respectively. In addition, α_i ($\alpha_i \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$) is the fusing weight.

Next, we illustrate the first term in the right-hand side of Eq. (8). Minimizing the first term aims to ensure that a user u_i and her follower u_j should be assigned similar normalized scores (e.g., $\frac{f_i}{\sqrt{[D_u]_{ii}}} = \frac{f_j}{\sqrt{[D_u]_{jj}}}$) while they are similar to each other, i.e., their similarity w_{ij}^{u} is high. Hence, u_i's normalized score will be high if follower u_j is an expert on Q_t (*Intuition 1* (a)); meanwhile if u_i is similar to many followers specialized on Q_t, she will be assigned a high ranking score (i.e., f_i), as f_i is proportional to D_{ii}, which is the sum of similarities between u_i and her followers over the given topic Q_t (*Intuition 1* (b)). Similar to the first term, the second term has the same purpose for lists (*Intuition 3*).

The third term (*rf.* Eq. (8)) is for the mutual ranking of users and lists. From the *user* perspective, if a user u_i is similar to her associated list L_j, then user u_i and list L_j should be assigned similar normalized ranking scores, and the normalized score of u_i should be increased if L_j is relevant to Q. If many lists containing u_i are relevant to Q and similar to u_i (i.e., D_{ml}^{u} is large), u_i's score (i.e., f_i) should be high (*Intuition 2* (a)-i)). From the *list* perspective, we can give similar analysis for lists.

By introducing the matrics $\mathbf{S}_u = (\mathbf{D}_u)^{-\frac{1}{2}} \mathbf{W}_u (\mathbf{D}_u)^{-\frac{1}{2}}$, $\mathbf{S}_l = (\mathbf{D}_l)^{-\frac{1}{2}} \mathbf{W}_l (\mathbf{D}_l)^{-\frac{1}{2}}$, and $\mathbf{S}_{ml} = (\mathbf{D}_{ml}^{u})^{-\frac{1}{2}} \mathbf{W}_{ml} (\mathbf{D}_{ml}^{l})^{-\frac{1}{2}}$, we can convert Eq. (8) into *matrix-vector* form as follows:

$$\begin{aligned}\mathcal{F}(\mathbf{W_u}, \mathbf{W_l}, \mathbf{W_{ml}}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \alpha_1 \hat{\mathbf{f}}^{\top} (\mathbf{I}_n - \mathbf{S}_u) \hat{\mathbf{f}} + \alpha_2 \hat{\mathbf{g}}^{\top} (\mathbf{I}_m - \mathbf{S}_l) \hat{\mathbf{g}} \\ + \alpha_3 (\hat{\mathbf{f}}^{\top} \hat{\mathbf{f}} + \hat{\mathbf{g}}^{\top} \hat{\mathbf{g}} - 2 \hat{\mathbf{f}}^{\top} (\mathbf{S}_{ml}) \hat{\mathbf{g}}),\end{aligned} \tag{9}$$

where $\mathbf{I}_n$ is a n-by-n identity matrix.

4.3.3 Loss Term

The *regularization* term does not incorporate the supervision information derived from the wisdom of *Twitter* crowds in the ranking process. Here, we proposed to use two types of relations as the supervised information for our problem, namely, MEM-OF *relation* and SUB-TO *relation*. The former is viewed as the supervision from the *creators* of lists. Intuitively a user listed in many relevant lists under a given topic is very likely to be an expert on that topic (*Intuition 2* (a)-ii)). Similarly, a list that has many subscribers who are experts on a given topic is likely related to that topic (*Intuition 2* (b)). Hence, we introduce two different indicator matrices to encode these two relations respectively.

Let n-by-m indicator matrix (i.e., $\mathbf{X}$) encode the MEM-OF relations for supervising the ranking of users, and m-by-n indicator matrix (i.e., $\mathbf{Y}$) encode the SUB-TO relations for supervising the ranking of lists. In particular, each element (x_{ij}) of $\mathbf{X}$ is set by $x_{ij} = \frac{1}{|u_i|}$ if u_i is the member of L_j ($|u_i|$ is the number of lists containing u_i), and each element (y_{ij}) of $\mathbf{Y}$ is set by $y_{ij} = \frac{1}{|L_i|}$ if u_j is a subscriber of L_i ($|L_i|$ is the number of users who subscribe to L_i). Then, the *loss term* is defined as,

$$\begin{cases} (a)\ (f_i - \sum_j x_{ij} g_j)^2,\ f_i \in \hat{\mathbf{f}},\ x_{ij} \in \mathbf{X},\ g_j \in \hat{\mathbf{g}}, \\ (b)\ (g_i - \sum_j y_{ij} f_j)^2,\ g_i \in \hat{\mathbf{g}},\ y_{ij} \in \mathbf{Y},\ f_j \in \hat{\mathbf{f}}, \end{cases} \tag{10}$$

Here *loss term* (a) aims to ensure a user should be ranked higher if most of lists containing that user are related to the given topic. Similarly, *loss term* (b) aims to ensure a list should be assigned a higher ranking score if most of subscribers of that list are relevant to the given topic, which in return enhances the ranking scores of the associated members. Therefore, the loss term can be formulated as follows:

$$\ell(\mathbf{X}, \mathbf{Y}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \gamma \sum_{i=1}^{n} (f_i - \sum_{j=1}^{m} x_{ij} g_j)^2 + (1-\gamma) \sum_{i=1}^{m} (g_i - \sum_{j=1}^{n} y_{ij} f_j)^2, \tag{11}$$

where γ ($0 \leq \gamma \leq 1$) is a non-negative coefficient to trade-off the two different loss terms. Correspondingly, Eq. (11) can also be transformed into a matrix-vector form as follows:

$$\ell(\mathbf{X}, \mathbf{Y}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \gamma \|\hat{\mathbf{f}} - \mathbf{X}\hat{\mathbf{g}}\|_2^2 + (1-\gamma) \|\hat{\mathbf{g}} - \mathbf{Y}\hat{\mathbf{f}}\|_2^2, \tag{12}$$

where $\|\hat{\mathbf{v}}\|_2$ denotes ℓ_2-norm of vector $\hat{\mathbf{v}}$.

For ease of explanation, we use $J(\hat{\mathbf{f}}, \hat{\mathbf{g}})$ to denote the objective function in Eq. (7). By substituting Eqs. (9) and (12), the optimization problem of this paper is formulated as follows:

$$\begin{aligned}J(\hat{\mathbf{f}}, \hat{\mathbf{g}}) &= \lambda \Big(\alpha_1 \hat{\mathbf{f}}^{\top} (\mathbf{I}_n - \mathbf{S}_u) \hat{\mathbf{f}} + \alpha_2 \hat{\mathbf{g}}^{\top} (\mathbf{I}_m - \mathbf{S}_l) \hat{\mathbf{g}} + \alpha_3 (\hat{\mathbf{f}}^{\top} \hat{\mathbf{f}} + \hat{\mathbf{g}}^{\top} \hat{\mathbf{g}} \\ &\quad - 2 \hat{\mathbf{f}}^{\top} (\mathbf{S}_{ml}) \hat{\mathbf{g}}) \Big) + (1-\lambda) \Big(\gamma \|\hat{\mathbf{f}} - \mathbf{X}\hat{\mathbf{g}}\|_2^2 + (1-\gamma) \|\hat{\mathbf{g}} - \mathbf{Y}\hat{\mathbf{f}}\|_2^2 \Big) \\ &= \alpha_1 \hat{\mathbf{f}}^{\top} (\mathbf{I}_n - \mathbf{S}_u) \hat{\mathbf{f}} + \alpha_2 \hat{\mathbf{g}}^{\top} (\mathbf{I}_m - \mathbf{S}_l) \hat{\mathbf{g}} + \alpha_3 \Big(\hat{\mathbf{f}}^{\top} \hat{\mathbf{f}} + \hat{\mathbf{g}}^{\top} \hat{\mathbf{g}} \\ &\quad - 2 \hat{\mathbf{f}}^{\top} (\mathbf{S}_{ml}) \hat{\mathbf{g}} \Big) + C_1 \|\hat{\mathbf{f}} - \mathbf{X}\hat{\mathbf{g}}\|_2^2 + C_2 \|\hat{\mathbf{g}} - \mathbf{Y}\hat{\mathbf{f}}\|_2^2,\end{aligned} \tag{13}$$

where $C_1 = \frac{(1-\lambda)\gamma}{\lambda}$ and $C_2 = \frac{(1-\lambda)(1-\gamma)}{\lambda}$. When $\lambda = 0$, $C_1 = \gamma$ and $C_2 = 1 - \gamma$.

Remark. Learning with regard to the loss term is in a supervised manner, and in contrast, learning with regard to the regularization term is in an unsupervised manner. We therefore call our proposed method SSGR a *semi-supervised* graph-based ranking method. Next, we detail how to solve the optimization problem stated in Eq. (13).

4.3.4 Solving the Optimization Problem

We proceed to present a solution to solving the optimization problem stated in Eq. (13). We first introduce the following theorem.

Theorem 1. *The objective function* $J(\hat{\mathbf{f}}, \hat{\mathbf{g}})$ *in Eq. (13) is a convex function* w.r.t. *any* $\hat{\mathbf{f}}$ *and* $\hat{\mathbf{g}}$.

Proof. By introducing *normalized Laplacian* $\mathbf{H}_u = \mathbf{I}_n - \mathbf{S}_u = \mathbf{D}_u^{-\frac{1}{2}}(\mathbf{D}_u - \mathbf{W}_u)\mathbf{D}_u^{-\frac{1}{2}}$, and $\mathbf{H}_l = \mathbf{I}_m - \mathbf{S}_l = \mathbf{D}_l^{-\frac{1}{2}}(\mathbf{D}_l - \mathbf{W}_l)\mathbf{D}_l^{-\frac{1}{2}}$, the regularization term in Eq. (9) can be converted as follows:

$$\mathcal{F}(\mathbf{W}_u, \mathbf{W}_l, \mathbf{W}_{ml}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} \alpha_1 \mathbf{H}_u & 0 \\ 0 & \alpha_2 \mathbf{H}_l \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + \alpha_3 \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} \mathbf{I}_n & -\mathbf{S}_{ml} \\ -\mathbf{S}_{ml}^\top & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \mathbf{H} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix},$$

where $\mathbf{H} = \begin{pmatrix} (\alpha_1 \mathbf{H}_u + \alpha_3 \mathbf{I}_n) & -\alpha_3 \mathbf{S}_{ml} \\ -\alpha_3 \mathbf{S}_{ml}^\top & (\alpha_2 \mathbf{H}_l + \alpha_3 \mathbf{I}_m) \end{pmatrix}$.

As $\mathbf{H}_u$ and $\mathbf{H}_l$ are both symmetric matrices (follow the symmetry of $\mathbf{W}_u$, $\mathbf{D}_u$, $\mathbf{W}_l$, $\mathbf{D}_l$), the matrix $\mathbf{H}$ is clearly a symmetric matrix, namely $\mathbf{H} = \mathbf{H}^\top$. In addition, since $w_{ij}^u \geq 0$, $w_{ij}^l \geq 0$, and $w_{ij}^{ml} \geq 0$, the regularization term $\mathcal{F}(\cdot)$ should be non-negative (Eq. (8)), i.e., $\mathcal{F}(\cdot) \geq 0$. Hence, $\mathbf{H}$ is a symmetric and positive semi-definite matrix, which demonstrates the regularization term $\mathcal{F}$ is a convex function.

Similarly, the loss term ℓ can also be transformed as the following block structure:

$$\begin{aligned} \ell(\mathbf{X}, \mathbf{Y}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \gamma \Bigg(& \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} \mathbf{I}_n & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \\ & \begin{pmatrix} 0 & -\mathbf{X} \\ -\mathbf{X}^\top & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} 0 & 0 \\ 0 & (\mathbf{X}^\top \mathbf{X}) \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} \Bigg) \\ + (1-\gamma) \Bigg(& \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \\ & \begin{pmatrix} 0 & -\mathbf{Y}^\top \\ -\mathbf{Y} & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \begin{pmatrix} (\mathbf{Y}^\top \mathbf{Y}) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} \Bigg). \end{aligned}$$

Subsequently, the above equation is transformed into

$$\ell(\mathbf{X}, \mathbf{Y}; \hat{\mathbf{f}}, \hat{\mathbf{g}}) = \gamma \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \mathbf{H}_1 \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix} + (1-\gamma) \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix}^\top \mathbf{H}_2 \begin{pmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{pmatrix},$$

where and $\mathbf{H}_1 = \begin{pmatrix} \mathbf{I}_n & -\mathbf{X} \\ -\mathbf{X}^\top & (\mathbf{X}^\top \mathbf{X}) \end{pmatrix}$ and $\mathbf{H}_2 = \begin{pmatrix} (\mathbf{Y}^\top \mathbf{Y}) & -\mathbf{Y}^T \\ -\mathbf{Y} & \mathbf{I}_m \end{pmatrix}$.

We can easily prove that the matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are both symmetric and positive semi-definite matrices, which demonstrate that the loss term ℓ is also a convex function. Accordingly, the object function $J(\hat{\mathbf{f}}, \hat{\mathbf{g}})$ is a convex function, which completes the proof. □

Theorem 1 guarantees that objective function J has *first-order* partial derivatives, which means the objective function at least has an optimal solution. Thus we introduce the *gradient descent* method to minimize $J(\hat{\mathbf{f}}, \hat{\mathbf{g}})$. The partial derivatives of $J(\hat{\mathbf{f}}, \hat{\mathbf{g}})$ with respect to $\hat{\mathbf{f}}$ and $\hat{\mathbf{g}}$ can be calculated as follows:

$$\begin{aligned} \nabla \hat{\mathbf{f}} = \frac{\partial J}{\partial \hat{\mathbf{f}}} = & \big((1 - \alpha_2 + C_1)\mathbf{I}_n - \alpha_1 \mathbf{S}_u + C_2 \mathbf{Y}^\top \mathbf{Y}\big)\hat{\mathbf{f}} \\ & - (\alpha_3 \mathbf{S}_{ml} + C_1 \mathbf{X} + C_2 \mathbf{Y}^\top)\hat{\mathbf{g}}, \end{aligned} \tag{14}$$

$$\begin{aligned} \nabla \hat{\mathbf{g}} = \frac{\partial J}{\partial \hat{\mathbf{g}}} = & \big((1 - \alpha_1 + C_2)\mathbf{I}_m - \alpha_2 \mathbf{S}_l + C_1 \mathbf{X}^\top \mathbf{X}\big)\hat{\mathbf{g}} \\ & - (\alpha_3 \mathbf{S}_{ml}^\top + C_1 \mathbf{X}^\top + C_2 \mathbf{Y})\hat{\mathbf{f}}. \end{aligned} \tag{15}$$

Accordingly, we can iteratively compute $\hat{\mathbf{f}}$ and $\hat{\mathbf{g}}$ by using *gradient descent* method, namely,

$$\hat{\mathbf{f}}^{(s+1)} = \hat{\mathbf{f}}^{(s)} - \rho \nabla \hat{\mathbf{f}}^{(s)}, \tag{16}$$

$$\hat{\mathbf{g}}^{(s+1)} = \hat{\mathbf{g}}^{(s)} - \rho \nabla \hat{\mathbf{g}}^{(s)}, \tag{17}$$

where ρ is the step size that is allowed to change at each iteration, and s is the iteration number. When J is a convex function, all local minima can also be treated as the global minima. Consequently, our method can converge to a *global* solution. The learning process of *global authority* of users on a given topic is summarized in Algorithm 1.

Algorithm 1. Learning *Global* Authority Algorithm

OFF-LINE CONSTRUCTION OF MATRIX **R**

Input: $\mathcal{T} = \{t_i\}$: *Twitter* corpus, t_i denotes a appeared term t_i in *Twitter*; $\mathcal{U}$: Candidate user set; $\mathcal{L}$: List set; $\mathbf{E}_u$: Follower graph, each entry in E_u denotes two users that follow each other; $\mathbf{E}_{ml}$: Member graph, each entry in E_{ml} denotes a member and her included list; $\mathbf{E}_{ls}$: Subscriber graph, each entry inE_{ls} denotes a subscriber and her subscribing list; $\mathbf{E}_{ll}$: List Graph, each entry in E_{ll} denotes two lists that are similar to each other; α_i, λ, γ, ρ, ε

Result: Authority matrix **R**

foreach $t_i \in \mathcal{T}$ **do**
  Initialize: Compute $f_k^{(0)} \leftarrow \frac{N(t_i, d_k^u)}{\sum_{u_j \in \mathcal{U}} N(t_i, d_j^u)}$ and $g_k^{(0)} \leftarrow \frac{N(t_i, d_k^l)}{\sum_{L_j \in \mathcal{L}} N(t_i, d_j^l)}$;
  Construct *topic-specific* matrices $\mathbf{W}_u$, $\mathbf{W}_l$ and $\mathbf{W}_{ml}$ by Eq. (6) and indicator matrices **X** and **Y**;
  Set $s = 0$;
  while *true* **do**
    Compute $\nabla \hat{\mathbf{f}}^{(s)} \leftarrow \frac{\partial J}{\partial \hat{\mathbf{f}}}\big|_{\hat{\mathbf{f}} = \hat{\mathbf{f}}^{(s)}}$ and $\nabla \hat{\mathbf{g}}^{(s)} \leftarrow \frac{\partial J}{\partial \hat{\mathbf{g}}}\big|_{\hat{\mathbf{g}} = \hat{\mathbf{g}}^{(s)}}$ by Eqs. (14), (15);
    Update: $\hat{\mathbf{f}}^{(s+1)}$ and $\hat{\mathbf{g}}^{(s+1)}$ by Eq. (16), (17);
    Normalize: $f_i^{(s+1)} \leftarrow \frac{f_i^{(s+1)}}{\sum_{j=1}^n f_j^{(s+1)}}$ and $g_i^{(s+1)} \leftarrow \frac{g_i^{(s+1)}}{\sum_{j=1}^m g_j^{(s+1)}}$;
    Calculate: $J(\hat{\mathbf{f}}^{(s+1)}, \hat{\mathbf{g}}^{(s+1)})$ by Eq. (13);
    if $|J(\hat{\mathbf{f}}^{(s+1)}, \hat{\mathbf{g}}^{(s+1)}) - J(\hat{\mathbf{f}}^{(s)}, \hat{\mathbf{g}}^{(s)})| \leq \varepsilon$ **then**
      $\hat{\mathbf{f}}^* \leftarrow \hat{\mathbf{f}}^{(s+1)}$, $\hat{\mathbf{g}}^* \leftarrow \hat{\mathbf{g}}^{(s+1)}$ and Stop
  $\mathbf{R} \leftarrow < t_i, \hat{\mathbf{f}}^* >$;
Return **R**;

Time Complexity. The learning of *global authority* of users on a given topic is calculated *offline*. In fact, the computation overhead of SSGR is not high due to the sparsity of the matrices $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{S}_u$, $\mathbf{S}_l$, and $\mathbf{S}_{ml}$. It can be solved with the iterative update rules (i.e., Eqs. (14) and (15)) until convergence. We note that the entries of $\mathbf{X}^\top\mathbf{X}$ and $\mathbf{Y}^\top\mathbf{Y}$ remain unchanged for any topic, which is solely related to the graph structure, and thus they can be *pre-computed*. Therefore, the time complexity is only $O(Kmn^2)$, where n is the number of users, m is the number of lists, K is the number of iterations (which usually converges in fewer than 30 on our experimental datasets).

5 Online Ranking for Expert Finding

We propose a Gaussian-based method to estimate the *local relevance* of candidates on a given topic in Section 5.1. Then, in Section 5.2 we give an *online* ranking model, called RM, to address the *expert finding* problem by leveraging the *global authority* and *local relevance* of candidates on any topic.

5.1 *Local* Relevance Estimation

According to Eq. (4), for each pair of adjacent terms, t_j and t_{j+1}, in query Q, we estimate the local relevance of a user u_i to them as follows:

$$\mathcal{K}(t_j, t_{j+1}; d^{u_i}) = \Pr(t_j|d^{u_i}) \Pr(t_{j+1}|d^{u_i}) \cdot e^{-\frac{\|\Pr(t_j|d^{u_i}) - \Pr(t_{j+1}|d^{u_i})\|^2}{2}}, \tag{18}$$

where $\Pr(t|d) = \frac{N(t,d)}{\sum_{t' \in d} N(t',d)}$, and $N(t,d)$ denotes the number of occurrences of term t in document d.

The first two terms in the right hand of Eq. (18) are used to favor users who frequently use the terms of query Q. However, they cannot model the co-occurrences of t_i and t_{j+1}. Hence, a *Gaussian*-based function (*rf.* the last term of Eq. (18)) is used to favor users who might frequently use two consecutive terms of query Q.

5.2 Online Ranking Model

Next, we will present an *online* ranking model to address the Twitter *expert finding* problem for the *arbitrary* topic query (i.e., $Q = \{t_1, \ldots, t_{|Q|}\}$), which is defined based on Eq. (4) as follows,

$$U_Q \leftarrow \arg\max_{u_k \in \mathcal{U}}^{N} \begin{cases} \frac{\mathcal{H}(u_k,t)}{\sum_{u_j \in \mathcal{U}} \mathcal{H}(u_j,t)}, & \text{if } |Q| = 1; \\ \sum_{\substack{i=1 \\ t_i, t_{i+1} \in Q}}^{|Q|-1} \Pr(u_k|t_i) \Pr(u_k|t_{i+1}) \mathcal{K}(t_i, t_{i+1}; d^{u_k}), & \text{otherwise}, \end{cases} \tag{19}$$

where $\mathbf{U}_Q$ denotes the retrieved top-N experts that are most relevant to query Q; $\mathcal{H}(u_k, t_i) = R_{i,k}$, which is an entry of the *authority* matrix $\mathbf{R}$. It indicates the *global authority* of user u_k on t_i, computed by SSGR; function $\mathcal{K}$ is computed by Eq. (18). The *online expert finding* algorithm is given in Algorithm 2.

Time Complexity. Note that each row $\hat{\mathbf{R}}_i$ in $\mathbf{R}$ is computed *offline* and the probability of term t in the context d^u of each user (i.e., $\Pr(t|d^u)$) can also be pre-computed *offline*. The time complexity of RM is $O(n|Q|)$, where n is the number of users and $|Q|$ is the length of the given topic query. In our experiments conducted on a real-world *Twitter* user set (about 0.5M users), the average running time of finding experts on a given topic is less than 0.01 seconds. The experiments are completed on a modest commodity desktop that is equipped with a Intel-i5 Dual-core 2.8 GHz CPU and 8 GB RAM. It shows that our proposed *online expert finding* model is computationally feasible for the real-time online *Twitter expert finding* applications.

TABLE 2
Statistics of the User-List Relations in the Crawled Data

	# Lists	# Lists (No Description)	Ratio
Mem-of (A1)	5,410,831	3,096,655	67.26%
Sub-to (A2)	635,852	340,616	58.27%
A1 ∪ A2	5,503,155	3,437,271	73.44%

Algorithm 2. *Online* **Expert Finding Algorithm**

On-Line Expert Finding
Input: $\mathbf{R}$: Authority matrix;
$\mathbf{Q}$: Given topic query;
$\mathcal{U}$: User set
Result: Topic-specific expert set U_Q
Set $\hat{\mathbf{f}} \leftarrow 0$;
foreach $u_k \in \mathcal{U}$ **do**
  if $|Q| = 1$ **do**
    $f_k \leftarrow \frac{\mathcal{H}(u_k,Q)}{\sum_j \mathcal{H}(u_k,Q)}$; break;
  for $i \leftarrow 1$ **to** $|Q| - 1$ **do**
    $f_k \leftarrow f_k + \Pr(u_k|t_i) \Pr(u_k|t_{i+1}) \mathcal{K}(t_i, t_{i+1}; d^{u_k})$;
$\mathbf{U}_Q \leftarrow \arg\max_{u_i \in \mathcal{U}}^{N} f_i \ \ (f_i \in \hat{\mathbf{f}})$;
Return $\mathbf{U}_Q$;

6 Experiments

6.1 Data Set

Users. The data set used in this paper was crawled via Twitter API[4] from April 4, 2013 to June 10, 2013. For each user in *Twitter*, we crawled five types of data, i.e., user profiles, followers, tweets,[5] user-list membership information, and user-list subscribe information. In particular, we used a user-centric strategy to collect data as a brute-force crawling of all users for all lists would be prohibitively expensive and would not scale[6] [5]. More specifically, to be unbiased to the users, we randomly crawled the information of users by utilizing a publicly available user collection[7] as the seed set. Consequently, we obtained about 5.5M lists and 770,235 users who had subscribed to (or been members in) at least one list. In the 5.5M lists, 73.44 percent lists only had a List name while the others had a description (detailed statistics is in Table 2). In addition, the dataset contained a mixture of different languages, e.g., Chinese, English, German,

4. https://dev.twitter.com/
5. For each user, we collect a set of the most recent ($\leq$1,000) posted tweets.
6. Twitter normally rate-limited the number of API requests from a single machine (IP Address) to 150 per hour, i.e., 3,600 user profile crawls per day.
7. https://wiki.engr.illinois.edu/display/forward/Dataset-UDI-TwitterCrawl-Aug2012

TABLE 3
Statistics of TwL: Each User in TwL has Subscribed to (or Been Members in) at Least One List

User			
Total #	**# $\mathcal{M}^1$**	**# $\mathcal{S}^2$**	**# $\mathcal{M} \cap \mathcal{S}$**
491,622	452,119	129,449	89,946
List			
Total #	**# Lists ($\mathcal{M}$ are in)**	**# Lists ($\mathcal{S}$ sub. to)**	**# Lists ($\mathcal{M} \cap \mathcal{S}$)**
4,486,954	4,412,514	461,780	387,340
Relation			
	Avg. Degree		**Avg. Degree**
Mem-Of ($\mathcal{M} - \mathcal{L}$)	33.41	**Sub-To ($\mathcal{S} - \mathcal{L}$)**	6.36
Mem-Of ($\mathcal{L} - \mathcal{M}$)	3.32	**Sub-To ($\mathcal{L} - \mathcal{S}$)**	1.73
Follower Relations	135.07	**Mutual Follower Relations**	79.39

Note: [1]$\mathcal{M}$ *denotes the members of lists;* [2]$\mathcal{S}$ *indicates the subscribers of lists.*

Italian and etc. We filtered non-English characters, stopwords, punctuation as well as the high-frequency words in *Twitter* (e.g., "RT"), and employed Porter's stemmer [36] for remaining words. After the processing, the users without any context information are removed; finally we obtained 491,622 *users* (with 61.6M *tweets* and 4.4M *lists*) out of 770,235 *users* as the experimental dataset, named TwL (i.e., Twitter-List). The details about TwL are shown in Table 3.

Queries. We use 28 sample queries for evaluation, whose topics are from general to specific, e.g., a general personal hobby like "traveling" or a specific TopNews like "Boston Marathon bombings", which can be used to comprehensively evaluate the effectiveness of our proposed approach. In the work on Cognos [5], only general queries are used for evaluation.

6.2 Experimental Setting

6.2.1 Ground Truth & Evaluation Metrics

Ground Truth.[8] To evaluate the quality of the expert search results of different methods, we follow the evaluation strategy in [4], [5]. That is, we aggregate the top-10 users returned by each evaluated method, and then nine graduate students (whose research areas are not in text processing area) are invited for labeling. The annotators are divided into three groups (three annotators in each group) to label each suggested user (as shown in Fig. 4). Each user is labeled to be relevant (score 1) or irrelevant (score 0) with respect to the given query by evaluators. In each group if conflicts happen, the third annotator determines the final result of each group, and the majority vote of groups is used as the label of the user. Each evaluator is required to label the relevance of users based on the contents of their posted tweets (e.g., whether including the URLs related to the given query), users' bios and the meta-data of lists containing that user.

Evaluation Metrics. To evaluate the expert finding performance of different approaches, we adopt the following evaluation metrics: (i) *Precision* [30] (P@N). It measures the percentage of relevant user in the top-N returned users. (ii) *Normalized Discounted Cumulative Gain* [30] (NDCG@N).

8. The annotated results can be accessed from the following link: https://www.dropbox.com/s/47up1xqjcrbr7zz/annotated_new.txt?dl=0

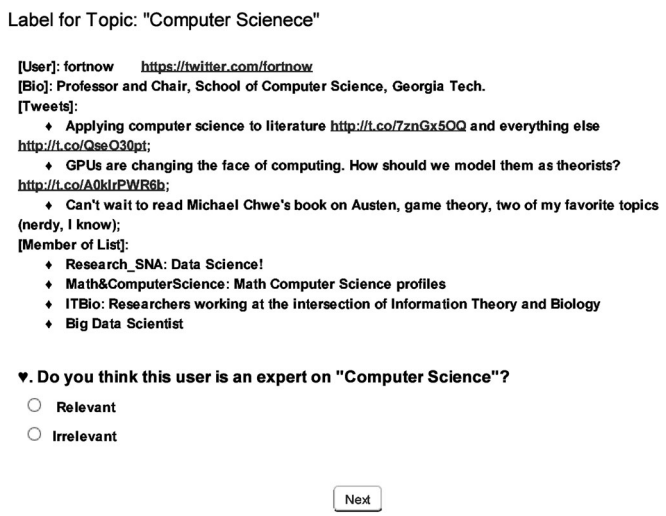


Fig. 4. Non-anonymous label screen shows bio, tweets (≤ 3) and lists (≤ 4) of a user and asks evaluators to label for relevant or irrelevant to a given topic query.

It measures the performance of expert finding system based on the relevance (i.e., relevant (1)/irrelevant (0)) of the selected experts, which is the normalization of Discounted Cumulative Gain (DCG) at each position for a chosen value of k. In our experiments, we use P@5, P@10, NDCG@5 and NDCG@10.

6.2.2 Baseline Methods & Parameter Setting

We compare our approach with *TwitterRank* [3] and *Cognos* [5]. In [5], *Cognos* is demonstrated to outperform the other previous state-of-the-art methods, such as [4] that relies on the user's bio or tweets, and *WTF*[9] (Twitter Who To Follow) that is the official *Twitter* expert search service. We evaluated 7 expert finding methods listed in Table 5.

TwitterRank [3]. This method first employs Latent Dirichlet Allocation (LDA [31]) model to identify users' interested topics from their tweets, then builds a topic-specific graph for each detected topic to compute a PageRank [35] vector of users, and finally linearly combines the PageRank vectors of different topics of a given query for finding topic-specific influential users. The damping factor of PageRank algorithm is set at 0.85 by following [3].

Cognos [5]. This method employs a topic vector to represent each *Twitter* user by the *Twitter List* information of users (Mem-of relation) and uses *cover density ranking* (CDR [32]) method to identify topic-specific experts. The length of covers ($\mathscr{K}$) is selected from $(4, 6, 8, 10, 12, 14, 16)$ and $\mathscr{K}$ with the best performance is used to report the final comparison results.

Our Method. Our proposed approach chooses the topic-specific authorities by jointly exploiting users' profiles and the meta-data of lists containing users, as well as the three different types of relations, i.e., *follower relation*, *list-list relation* and *user-list relation*. Our approach contains a graph-based ranking method (i.e., SSGR) and a ranking model (i.e., RM). Correspondingly, our method is denoted by SSGR + RM.

To analyze the different impacts of the two parts of our approach, the online ranking model (RM) is formulated as a

9. https://twitter.com/who_to_follow/suggestions.

TABLE 4
Sample Queries Used for Evaluation

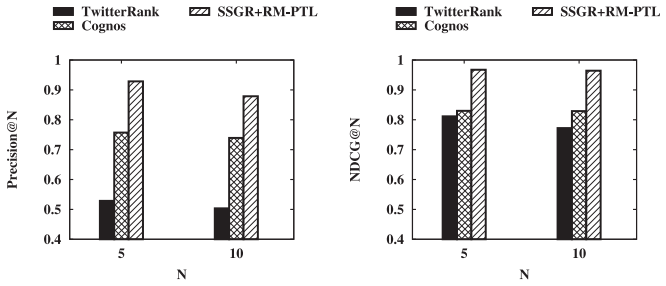
Catetory	Sample Queries
News	Egypt Balloon Explosion, Iran Nuclear Program, Curiosity on Mars, Boston Marathon bombings, Fukushima nuclear leak
Sports	football, soccer
Hobbies	traveling, photography, cooking, classical music
Science	biology, computer science
Entertainment	classical music
Lifestyle	dining, wine, health, fashion
Technology	smartphone, data mining, apple app, linux, cloud computing, iphone
Business	stock, finance, markets, energy

baseline method by replacing the *global authority* score ($R_{i,k}$) computed by SSGR with $R'_{i,k}$, which is calculated by a straightforward method. Here, $R'_{i,k}$ is the similarity between the ith term t_i and the kth user in the corpus, and is computed by $R'_{i,k} = \frac{N(t_i,d_k)}{\sum_{u_j \in \mathcal{U}} N(t_i,d_j)}$, where $N(t,d)$ denotes the number of occurrences of term t in document d. In addition, we have three different methods of constructing the information of users: (i) **PT**, user's bios and tweets; (ii) **PL**, user's bios and the meta-data of lists containing those users; and (iii) **PTL**, user's bios and tweets, as well as the meta-data of lists containing users. Correspondingly, different variants of SSGR + RM and RM are listed in Table 5.

In addition, there are some other works on the detection of influential users in social networks, such as Community Question Answering [11], [12], [43], Blog [28], Academic Social Network [24], Twitter [34], [44], and other social networks [39], [42]. However, these methods utilize different domain features (e.g., the category information in CQA domain) or do not consider topical dimension, which are thus not appropriate to make a comparison of them with the current work.

TABLE 5
Different Comparison Methods

Notations	Description
Different strategies for constructing the information of users	
PT	User's profiles (bios and tweets).
PL	User's bios and the meta-data of lists containing users.
PTL	User's profiles (bios and tweets) and the meta-data of lists containing users.
Methods For Comparison	
Cognos	*Cover density ranking* [5] (PL)
TwitterRank	*PageRank-based method* [3] (PT)
RM − PL	*Online Ranking model (RM) based on* $R'_{i,k}$ (PL)
RM − PTL	*Online Ranking model (RM) based on* $R'_{i,k}$ (PTL)
SSGR + RM − PT	*Online Ranking model (RM) based on SSGR* (PT)
SSGR + RM − PL	*Online Ranking model (RM) based on SSGR* (PL)
SSGR + RM − PTL	*Online Ranking model (RM) based on SSGR* (PTL)



(a) Precision@N (b) NDCG@N

Fig. 5. Comparison of overall expert finding performance.

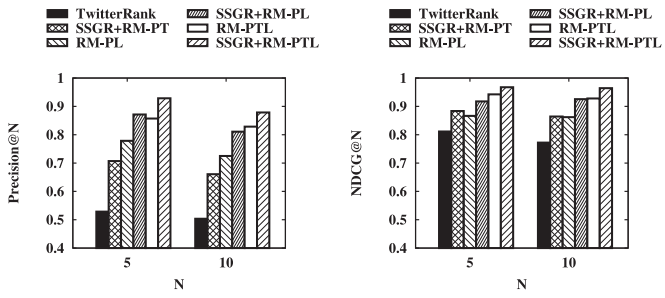
The parameters for our proposed method are empirically set as follows: $\alpha_1 = 0.33,\ \alpha_2 = 0.33,\ \alpha_3 = 1 - \alpha_1 - \alpha_2$, $\theta = 0.1$, $\delta = 10^{-8}$, $\rho = 0.1$, $\varepsilon = 10^{-12}$, and $\lambda = \gamma = 0.5$. As the construction of *list* graph is based on *mutual k-nearest neighbor graph* [40], we empirically set $k = 50$ in this paper.

6.3 Evaluation Results and Analysis

Comparison of the Expert Finding Performance. This experiment is to evaluate the effectiveness of finding the topic-specific experts by our approach, i.e., SSGR + RM − PTL. In this work, we compare SSGR + RM − PTL with the baseline methods *TwitterRank* and *Cognos*. Figs. 5a and 5b shows P@5, P@10, NDCG@5 and NDCG@10 of each method.

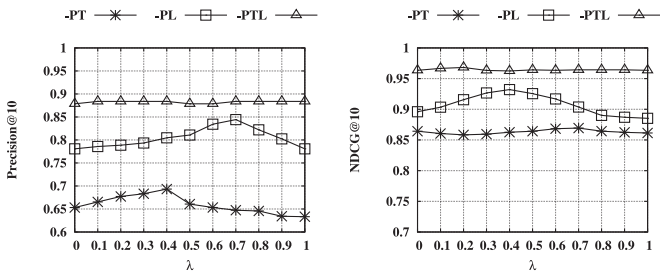
From Figs. 5a and 5b, we observe that: *First*, *Cognos* performs better than *TwitterRank* on all metrics. For example, *Cognos* outperforms *TwitterRank* by 30.19 percent (p-value ≤ 0.005) in terms of Precision@5. This is because *Cognos* utilizes *Twitter List* relation to find the topic-specific authorities while *TwitterRank* is based on the propagation (reciprocity in follower relations [3]) of the topical importance of users in follower graph. The results demonstrate that methods using Twitter Lists (*Cognos*) is more effective than approaches to utilizing follower graph and user's tweets (*TwitterRank*). *Second*, our proposed method SSGR + RM − PTL consistently outperforms the two baseline methods. The improvements are statistically significant on all metrics (p-value ≤ 0.005). For example, SSGR+ RM − PTL outperforms *TwitterRank* by 75.68 percent (p-value ≤ 0.000001) and 19.23 percent (p-value ≤ 0.005), *Cognos* by 22.64 percent (p-value ≤ 0.01) and 16.58 percent (p-value ≤ 0.005), in terms of Precision@5 and NDCG@5, respectively. The reason might be due to two facts: (i) Unlike *TwitterRank* that employs PageRank algorithm [35] on *follower* relation graph or *Cognos* that employs a similarity measure [32] to rank users based on user-list relation, SSGR + RM − PTL effectively exploits three different types of relations among users and lists. Specifically, it employs a normalized Laplacian regularization to take into account different relations (i.e., *follower* relation, *user-list* relation and *list-list* relation) for ranking, and utilizes *user-list* relations reflecting the wisdom of Twitter crowds to supervise ranking users. (ii) SSGR + RM − PTL makes use of two types of user-related information to model user's domain of expertise, i.e., user's profiles and List information. On one hand, some queries almost do not appear in Lists, such as the top news "Boston Marathon bombings",[10] which is more likely

10. http://en.wikipedia.org/wiki/Boston_Marathon_bombings, occurred on 15/04/2013



(a) Precision@N (b) NDCG@N

Fig. 6. The impact of graph-based ranking method.



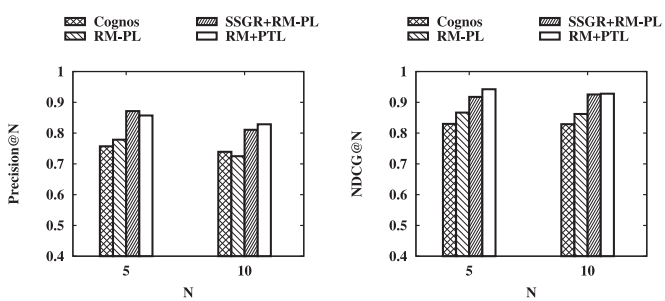
(a) Precision@10 (b) NDCG@10

Fig. 8. Illustration of the effect of the parameter λ in different variants of SSGR + RM (i.e., -PT, -PL, -PTL).

contained in the user's tweets. On the other hand, Lists are usually carefully built according to the wisdom of *Twitter* crowds, which are trustworthy for identifying the topical expertise of users contained in the lists. Our method SSGR + RM − PTL is able to make use of both types of information. The results demonstrate the effectiveness and superiority of our proposed method as compared to the state-of-the art method *Cognos* and *TwitterRank*.

Our proposed approach consists of two parts, i.e., a graph-based ranking method (SSGR) and a ranking model (RM). Next we use two groups of experiments to further evaluate the impact of each part in our approach. The first group is to evaluate the impact of the graph-based ranking method SSGR; and the second group is to evaluate the impact of the ranking model RM. *The Impact of Graph-based Ranking Method*. This is to evaluate the effectiveness of the graph-based ranking method in our approach. Since the graph-based ranking method cannot work alone for expert finding, we compare the methods utilizing the ranking scores computed by SSGR (i.e., SSGR + RM − PT, SSGR + RM − PL and SSGR + RM − PTL) and methods without using such ranking scores (i.e., RM − PL and RM − PTL), as well as *TwitterRank*. The Precision@N and NDCG@N of each method are plotted in Figs. 6a and 6b.

As can be observed from Figs. 6a and 6b: (i) Methods utilizing the ranking scores of SSGR outperform methods without using such ranking scores. In terms of Precision@5, SSGR + RM − PL and SSGR + RM − PTL outperform RM − PL and RM − PTL by 11.93 percent (p-value ≤ 0.05) and 8.33 percent (p-value ≤ 0.05), respectively. (ii) Even without using the meta-data of *Twitter Lists* to model user's information, our proposed method (i.e., SSGR + RM − PT) still outperforms *TwitterRank*. The improvements are statistically significant on all metrics (p-value ≤ 0.05). For example, SSGR + RM − PT outperforms *TwitterRank* by 31.21 percent (p-value ≤ 0.005) and 11.93 percent (p-value ≤ 0.05) in terms of Precision@10 and NDCG@10, respectively.



(a) Precision@N (b) NDCG@N

Fig. 7. The impact of online ranking model.

The observation (i) and (ii) demonstrate the effectiveness of our proposed graph-based ranking method SSGR, in exploiting different relations among users and lists for identifying the users' domain of expertise.

The Impact of Online Ranking Model. This experiment is to evaluate the effectiveness of the ranking model (RM) in our approach. Figs. 7a and 7b present the Precision@N and NDCG@N of each method.

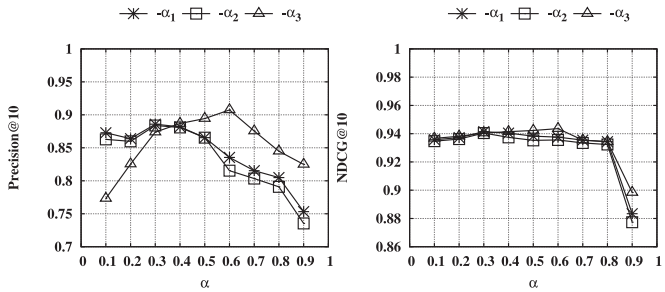
As shown in Figs. 7a and 7b, even utilizing the same strategy for constructing the information of users (i.e., PL), RM − PL performs better than *Cognos* in terms of Precision@5, NDCG@5 and NDCG@10 (except Precision@10). The reason might be RM considers not only the *local relevance* between a user and the given query, but also the *global authority* score (i.e., $R'_{i,k}$) of that user on the given query. Additionally, RM is also capable of effectively finding experts by: (i) combining with other *global authority* scores. For example, by replacing $R_{i,k}$ with the ranking score of SSGR, RM (i.e., SSGR + RM − PL) consistently outperforms *Cognos* on all metrics; and (ii) utilizing other information. For instance, with PTL, RM (i.e., RM − PTL) outperforms *Cognos* by 13.56 percent (statistically significant, p-value ≤ 0.01) in terms of NDCG@5.

6.4 On the Sensitivity of Parameter

In this section, we study the impact of parameters in our method.

First, we study the impact of parameter λ, which is used in Eq. (13) to trade-off the regularization term and the loss term. We compare 3 different variants of our proposed methods when varying λ from 0 to 1, i.e., SSGR +RM− PT, SSGR + RM − PL and SSGR + RM − PTL. As shown in Figs. 8a and 8b, the performance of our methods do not significantly change with varying λ. We observe that our proposed methods achieve the best result when λ is within the range of $[0.3 - 0.5]$, and simultaneously making use of both regularization term and loss term outperforms the extreme cases when only the regularization term ($\lambda = 1$) or the loss term ($\lambda = 0$) is used.

Second, we study the importance of different types of relations (i.e., *user-user* relation, *user-list* relation and *list-list* relation) in our approach. In particular, as our method is to mutually reinforce the ranking of users and lists by means of the three relations, to learn the *global authority* of users on a given topic. Hence, we fix the parameter λ at 1, and vary each of the three parameters α_1, α_2 and α_3 to evaluate the



(a) Precision@10 (b) NDCG@10

Fig. 9. Illustration of the effect of different parameter α (i.e., α_1, α_2, and α_3) in our method SSGR + RM − PTL.

impact of each type of relations in our method (i.e., SSGR + RM − PTL). In each time, the other two parameters are set at $\frac{1-\alpha}{2}$. For example, the value of parameter α_1 and α_3 are set at $\frac{1-\alpha_2}{2}$ when varying α_2. Figs. 9a and 9b shows Precision@10 and NDCG@10 of our method (i.e., SSGR + RM − PTL) with respect to different α (i.e., α_1, α_2 and α_3) ranging from 0.1 to 0.9 with an increment of 0.1. As shown in Fig. 9, when varying one of the three parameters (i.e., α_1-α_3) respectively, the performance first increases and then decreases. The average improvements of varying α_3 over varying α_1 and varying α_2 are (27.89, 2.98 percent) and (38.63, 5.28 percent), in terms of Precision@10 and NDCG@10. This demonstrates the importance of *user-list* relation, which contributes more to improve the performance of our method, as the correlation between *user-user* relation and *list-list* relation are bridged by *user-list* relation. From Figs. 9a and 9b, we can observe that our proposed method achieves the best performance with setting α_3 at 0.6, and the best parameter setting for our method is: $\alpha_1 = 0.3$, $\alpha_2 = 0.1$ and $\alpha_3 = 0.6$.

6.5 In-Depth Analysis of Expert Search Results

In this section, we present an in-depth analysis of expert search results of our proposed approach.

Single Term Query. We observe that both SSGR + RM − PTL and *Cognos* achieve comparable performance over most of single term queries while TwitterRank performs slightly worse. For example, for query "astronomy", both SSGR + RM − PTL and *Cognos* find the same user "NASA" and *TwitterRank* selects "hubblescience" on the top of their ranking lists. However, for seven queries on hot topics (e.g., "stock", "finance", etc.), *TwitterRank* and SSGR + RM − PTL perform worse than *Cognos*. This is because single term queries on hot topics are frequently contained in many users' tweets, and our method utilizing user's tweets and user's follower relations, i.e., SSGR + RM − PTL, would choose these users if they post many relevant tweets and have many relevant followers; however these users may not be experts on such topics. For example, for query "stock", SSGR + RM − PTL selects user 'Bertieis" as a top-10 result, who publishes many tweets containing "stock", and has 229 followers (most of them publish tweets containing "stock") in our experimental data set TwL. Interestingly, she is also included in a list named "Stocks". However, she is not an expert on stock. In contrast, *Cognos* is less affected as it chooses users contained by many relevant lists. *Multiple Term Query*. We observe that SSGR + RM − PTL consistently outperforms *Cognos* and *TwitterRank* for all multiple term queries,[11] which are usually more specific than single term queries. For example, for query "classical music", *Cognos* selects many pop singers (e.g., "AvrilLavigne") and *TwitterRank* chooses users with many music-related followers (e.g., "theglowradio") while SSGR + RM − PTL selects users like "nyphil", which is the official Twitter account of New York Philharmonic.

To better illustrate the effectiveness of our proposed method on the multiple term queries, the top-6 (due to the space limitation) selected users returned by SSGR + RM − PTL, *Cognos* and *TwitterRank* for two representative queries are shown in Tables 6 and 7, respectively. From the results, we observe that: (i) SSGR + RM − PTL is able to effectively choose the topic-specific experts for queries with specific topics. For example, in addition to "*mwahby*", all other users found by our method in Table 6 have posted the tweets regarding the topic "Egypt Balloon Explosion[12]". In contrast, none of users found by *Cognos* or *TwitterRank* has reported that topic. (ii) Unlike *Cognos*, which tends to select users listed by many other users, our approach fairly treats all *Twitter* users for search. For example, user "JoAnnaScience" in Table 7, who has only four *List*s, is also found by SSGR + RM − PTL, and it is however not selected by *Cognos*. and (iii) Our proposed method is more robust than *Cognos* and *TwitterRank*. For example, in Table 7, *Cognos* chooses some irrelevant users (e.g., *SmlTwnEchelon*) for the topic "Curiosity on Mars". This is because there are many *Twitter* lists are built by the fans of an American rock band named "Thirty Seconds to Mars[13]". However, all of users chosen by SSGR + RM − PTL are relevant to the given query. Similarly, in Table 6, *TwitterRank* selects several users relevant to "Balloon" for query "Egypt Balloon Explosion", as *TwitterRank* employs a linear function to combine the PageRank scores of users on different topics in the given query. In contrast, most of users (exclude "mwahby") found by SSGR + RM − PTL are relevant to the query.

7 Conclusion

In this paper, we address the problem of topic-specific expert finding in *Twitter*. We successfully integrate different types of user-related information (i.e., the crowdsourced *Lists* information, follower graph and users' profiles) into a unified ranking framework for accurately inferring the topical expertise of users. To the best of our knowledge, this is the first attempt that targets expert finding problem in Twitter by utilizing all of such information. Specifically, within the framework, we develop a semi-supervised graph-based ranking method, comprising a regularization term and a loss term. Our method aims to assign similar ranking scores to the similar users and lists, and meanwhile the ranking scores are subjected to the supervised information from the wisdom of *Twitter* crowds. Based on the computed ranking scores, we select the top-N relevant users for any given topic. The experiments conducted on real-world *Twitter*

11. Thirty-nine percent queries are multiple term query in Table 4.
12. http://edition.cnn.com/2013/02/27/world/meast/egypt-balloon-deaths/, occurred on 26/02/2013
13. http://en.wikipedia.org/wiki/Thirty_Seconds_to_Mars

TABLE 6
Top-6 Experts Selected for Query "*Egypt Balloon Explosion*", along with Users' Extracted Bios and Tweets in TwL

SSGR + RM − PTL Results		
User	**Extracted Bio (B) and related tweets (T)**	**# List**
Bennu	**[B]: Delivering daily breaking news about...Egypt... [T]: Hot air balloon flights resume in Luxor Egypt...: http://t.co/ybzXgXuW9y.**	**371**
scorpion–kiss	**[B]: Documentary photographer... [T]: Hot air balloon crash in Egypt kills 19 ... tourists...http://t.co/9lWSKIueYY**	**23**
Alternativ–Egypt	**[B]: Egypt travel guide ... [T]: ... balloon disaster ... http://t.co/xhxeZXxrsb**	**34**
Breaking–NZ	**[B]: Biggest alerts ... Updates delayed by 10 minutes... [T]: ... 19 foreign nationals are dead following a hot air balloon crash in Egypt ...**	**57**
Egyptian–Texts	**[B]: The latest Egyptian news and discoveries... [T]: ... Balloon flights resume Luxor Forum TripAdvisor: ... http://t.co/hlmrmoMbxD**	**113**
mwahby	[B]: ... Microsoft Egypt ... Programs Manager....	4
***Cognos* Results**		
shadihamid	[B]: Director of Research ... for Middle East Policy...	1,838
ahramonline	[B]: ... Egypts largest news organization....	1,231
Linaattalah	[B]: Journalist	498
NevineZaki	[B]: ... Writer & host of 'Motion Pictures' ...	376
Elazul	[B]: I do not claim to be objective, subjective...	244
stevenacook	[B]: ... senior fellow for Middle Eastern studies	464
***TwitterRank* Results**		
CarbinCopy	[B]: The Balloon Bandit of Amusement ...	49
BalloonFNDN	[B]: ...build membership for the ABQ Balloon Museum.	49
fustat	[B]:	155
ncnearspace	[B]: NC Near Space Research.	10
VOAArrott	[B]: Cairo Bureau chief and regional correspondent, Voice of America. News of the Arab Spring	27
musicnever–heard	[B]: Norwegian music artist Helge Krabye (Homeless Balloon)...	6

Relevant experts are highlighted in bold font.

TABLE 7
Top-6 Experts Selected for Query "*Curiosity on Mars*", along with Users' Extracted Bios and Tweets in TwL

SSGR + RM − PTL Results		
User	**Extracted Bio (B) and related tweets (T)**	**# List**
martian–soil	**[B]: All the fresh dirt on the planet Mars... [T]: Curiosity: ... on Mars http://t.co/JZLO4TKkOe...**	**199**
NASA	**[B]: Explore the universe and discover ... planet ... [T]: ... landing on Mars ... http://t.co/lNkrfGRaTu...**	**57,001**
JoAnna–Science	**[B]: Science nerd and future science writer. [T]: Wendel: #Curiosity rover brings hope of future journeys to #Mars: http://t.co/MemA6rYq...**	**4**
Mars–SanDiego	**[B]: ... Mars Society-San Diego ... Space Society [T]: ... Curiosity's arm... http://t.co/tlXtpiGJPW**	**34**
mars–today	**[B]: News about Mars and its moons... [T]: NASA Curiosity Rover Wins Prestigious Awards http://t.co/7n69f8ugnB#MSL.**	**303**
TheSpace–Trap	**[B]: ... Collecting .. amazing photos ... on our universe. [T]: .. Curiosity ... poised ... http://t.co/dObTOu2Mvi...**	**187**
***Cognos* Results**		
ShannonLeto	[B]:	5,275
vckbee	[B]:	905
NASA	**[B]: Explore the universe and discover ... planet ... [T]: ... landing on Mars ... http://t.co/lNkrfGRaTu...**	**57,001**
jaredletosays	[B]: ...inspire you with all the BEST...	385
Veroni–caMcG	**[B]: NASA-JPL news... Also tweeting... @MarsCuriosity [T]: ...NASA's Curiosity Rover Soars at This Year's 'Shorty Awards'... http://t.co/PGm0ZRbi3p...**	**695**
SmlTwn–Echelon	[B]: ... a small town girl .. Listening to way too much 30 Seconds to Mars ...	167
***TwitterRank* Results**		
curiositychat	[B]: Founder+managing partner of Curiosity Inc...	28
ShopCurious	[B]: ShopCurious = style with brains - for lovers...	45
JoAnna–Science	**[B]: NASA-JPL news... Also tweeting... @MarsCuriosity [T]: Wendel: #Curiosity rover brings hope of future journeys to #Mars: http://t.co/MemA6rYq...**	**695**
martian–soil	**[B]: All the fresh dirt on the planet Mars... [T]: Curiosity: ... on Mars http://t.co/JZLO4TKkOe...**	**199**
spareair	[B]: Follow SpareAir to get AirAlerts via Twitter...	75
foundon–mars	**[B]: ... editor of http://FoundonMars.com... [T]: ... #Curiosity launch ... http://t.co/UDFLOnm1**	**228**

Relevant experts are highlighted in bold font.

data set demonstrate that our method significantly outperforms the state-of-the art methods.

The following potential directions: First, we would like to improve the efficiency of the learning of the *global authority* of users for expert search; Second, we also interest in studying the diversity issue in the expert finding problem.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61332001, Grant 61173049, Grant 61300045 and Grant 61572215, and in part by the National Research Foundation hosted at Media Development Authority of Singapore under Grant MDA/

IDM/2012/8/8-2 VOL 01. The authors would also like to thank the anonymous reviewers for their comments on improving the quality of this paper. G.-H. Li is the corresponding author.

REFERENCES

[1] V. Qazvinian, E. Rosengren, D.-R. Radev, and Q.-Z. Mei, "Rumor has it: Identifying misinformation in microblogs," in *Proc. Conf. Empirical Methods Natural Language Process.*, 2011, pp. 1589–1599.

[2] L. Chen, Z.-Y. Liu, and M.-S. Sun, "Expert finding for microblog misinformation identification," in *Proc. Int. Conf. Comput. Linguistics*, 2012, pp. 703-712.

[3] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "Twitterrank: Finding topic-sensitive influential Twitterers," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2010, pp. 261–270.

[4] A. Pal and S. Counts, "Identifying topical authorities in microblogs," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2011, pp. 45–54.

[5] S. Ghosh, N. Sharma, F. Benevenuto, N. Ganguly, and K. Gummadi, "Cognos: Crowdsourcing search for topic experts in microblogs," in *Proc. 35th Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2012, pp. 575–590.

[6] X. Meng, F. Wei, X. Liu, M. Zhou, S. Li, and H. C. Wang, "Entity-centric topic-oriented opinion summarization in twitter," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 379-387.

[7] X. Liu, S. Zhang, F. Wei, and M. Zhou, "Recognizing named entities in tweets," in *Proc. 49th Annu. Meet. Assoc. Comput. Linguistics: Human Language Technol.*, 2011, pp. 359-367.

[8] P. Desislava and W.-B. Croft, "Proximity-based document representation for named entity retrieval," in *Proc. 16th ACM Conf. Inf. knowl. Manag.*, 2007, pp. 731-740.

[9] Y. Fang, S. Luo, and O. Etzioni, "Discriminative models of integrating document evidence and document-candidate associations for expert search," in *Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2010, pp. 683-690.

[10] K. Balog, L. Azzopardi, and M. De Rijke, "Formal models for expert finding in enterprise corpora," in *Proc. 29th Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2006, pp. 43–50.

[11] X. Liu, W. B. Croft, and M. Koll, "Finding experts in community-based question-answering services," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2005, pp. 315–316.

[12] A. Pal and J. A. Konstan, "Expert identification in community question answering: Exploring question selection bias," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2010, pp. 1505–1508.

[13] A. Pal and J. A. Konstan, "Co-occurrence-based diffusion for expert search on the web," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 5, pp. 1001–1014, May 2013.

[14] M. David and A. Andrew, "Expertise modeling for matching papers with reviewers," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2007, pp. 500–509.

[15] Y.-H. Zhou, G. Cong, B. Cui, C.-S. Jensen, and J.-J. Yao. "Routing questions to the right users in online communities," in *Proc. IEEE 25th Int. Conf. Data Eng.*, 2009, pp. 700–711.

[16] H. Deng, I. King, and M.-R. Lyu, "Formal models for expert finding on DBLP bibliography data," in *Proc. Int. Conf. Data Mining*, 2008, pp. 163–172.

[17] J. Zhang, J. Tang, and J. Li, "Expert finding in a social network. Advances in databases: Concepts, systems and applications," in *Proc. 12th Int. Conf. Database Syst. Adv. Appl.*, 2007, pp. 1066–1069.

[18] J. Tang, J. Sun, C. Wang, and Z. Yang, "Social influence analysis in large-scale networks," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2009, pp. 807–816.

[19] P. Serdyukov, H. Rode, and D. Hiemstra, "Modeling multi-step relevance propagation for expert finding," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2008, pp. 1133–1142.

[20] T.-H. Davenport and L. Prusak, *Working Knowledge: How Organizations Manage What They Know*. Boston, MA, USA: Harvard Business Press, 1998.

[21] N. Craswell, A. P. de Vries, and I. Soborof, "Overview of the TREC 2005 enterprise track," in *Proc. Text Retrieval Conf.*, 2005, pp. 199–205.

[22] D. Yimam-Seid and A. Kobsa, "Expert-finding systems for organizations: Problem and domain analysis and the DEMOIR approach," *J. Org. Comput. Electron. Commerce*, vol. 13, no. 1 pp. 1–24, 2003.

[23] K. Balog and M. de Rijke, "Non-local evidence for expert finding," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2008, pp. 489–498.

[24] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Arnetminer: Extraction and mining of academic social networks," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2008, pp. 489–498.

[25] W. Wei, B. Gao, T.-Y. Liu, T.-F. Wang, H.-G. Li and H. Li. "A ranking approach on large-scale graph with multidimensional heterogeneous information," *IEEE Trans. Cybern.*, vol. pp, no. 99, pp. 1–15, Apr. 2015.

[26] C. S. Campbell, P.-P. Maglio, A. Cozzi, and B. Dom, "Expertise identification using email communications," in *Proc. ACM Conf. Inf. Knowl. Manag.*, 2003, pp. 528–531.

[27] B. Gao, T.-Y. Liu, W. Wei, T.-F. Wang, and H. Li, "Semi-supervised ranking on very large graphs with rich metadata," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 96–104.

[28] N. Agarwal, H. Liu, L. Tang, and P.-S. Yu, "Identifying the influential bloggers in a community," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2008, pp. 207–218.

[29] M. McPherson, L. Smith-Lovin, and J.-M. Cook, "Birds of a feather: Homophily in social networks," *Annu. rev. Sociology*, vol. 27, pp. 415–444, 2001.

[30] R. A. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Boston, MA, USA: Addison-Wesley, 1999.

[31] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learning Res.*, vol. 3, pp. 993–1022, 2003.

[32] C. L. Clarke, G. V. Cormack, and E. A. Tudhope, "Relevance ranking for one to three term queries," *Inform. Process. Manage.*, vol. 36, no. 2, pp. 291–311, 2000.

[33] G. Demartini, D. E. Difallah, and P. Cudré-Mauroux, "Zencrowd: Leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 469–478.

[34] J. Lehmann, C. Castillo, M. Lalmas, and E. Zuckerman, "Finding news curators in Twitter," in *Proc. Int. Conf. World Wide Web*, 2013, pp. 469–478.

[35] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," *Stanford Digit. Libr. Technol. Project*, Stanford, CA, USA, Tech. Rep. 1999-66, Nov. 1999.

[36] M. F. Porter, "An algorithm for suffix stripping," *Program: Electron. Library Inf. Syst.*, vol. 14, no. 3, pp. 130–137, 1980.

[37] G. Salton and C. Buckley, "Term weighting approaches in automatic text retrieval," *Inform. Process. Manage.*, vol. 24, no. 5, pp. 513–523, 1987.

[38] N. K. Sharma, S. Ghosh, F. Benevenuto, N. Ganguly, and K. Gummadi, "Inferring who-is-who in the Twitter social network," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 533–538, 2012.

[39] E. Smirnova, "A model for expert finding in social networks," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2011, pp. 1191–1192.

[40] U. Von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.

[41] M. J. Welch, U. Schonfeld, D. He, and J. Cho, "Topical semantics of twitter links," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2011, pp. 327–336.

[42] J. Zhang, M. S. Ackerman, and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in *Proc. 16th Int. Conf. World Wide Web*, 2007, pp. 221–230.

[43] Z. Zhao, L.-J. Zhang, X.-F. He, and W. Ng, "Expert finding for question answering via graph regularized matrix completion," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 4, pp. 993–1004, Apr. 2015.

[44] Z.-Y. Cheng, J. Caverlee, H.-Barthwal, and V. Bachani, "Who is the barbecue king of Texas?: A geo-spatial approach to finding local experts on Twitter," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2014, pp. 335–344.

[45] R. Yeniterzi and J. Callan, "Analyzing bias in CQA-based expert finding test sets," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2014, pp. 967–970.

Wei Wei received the PhD degree from the Huazhong University of Science and Technology, Wuhan, China, in 2012. He is currently a lecturer with the School of Computer of Science and Technology, Huazhong University of Science and Technology. He was a research fellow with Nanyang Technological University, Singapore, and Singapore Management University, Singapore. His current research interests include information retrieval, data mining, social computing, and natural language processing.

Gao Cong received the PhD degree from the National University of Singapore. He is an associate professor at Nanyang Technological University, Singapore. He previously worked at Aalborg University, Microsoft Research Asia, and the University of Edinburgh. His current research interests include geo-textual and mobility data management, data mining, social media mining, and POI recommendation.

Chunyan Miao received the PhD degree from Nanyang Tchnological University, in 2003. She is an associate professor at Nanyang Technological University, Singapore. Her research interests include agent, multiagent systems, agent-oriented software engineering, and AgentWeb/Grid/Cloud.

Feida Zhu received the PhD degree from the University of Illinois at Urbana-Champaign, in 2009. He is an assistant professor at Singapore Management University, Singapore. His research interests include large-scale data mining, text mining, graph/network mining, and social network analysis.

Guohui Li received the PhD degree from Huazhong University of Science and Technology, in 1999. He is a professor at Huazhong University of Science and Technology, China. His research interests include location-based data management, social computing, big data processing, and real-time computing.