WILEY | Hindawi

Research Article

A Novel Expert Finding System for Community Question Answering

Nan Zhao ⓘ,[1] **Jia Cheng,**[1] **Nan Chen,**[1] **Fei Xiong,**[2] **and Peng Cheng**[3]

[1]*State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China*
[2]*Key Laboratory of Communication & Information Systems, Beijing Jiaotong University, Beijing 100044, China*
[3]*Space Star Aerospace Technology Applications Limited Corporation, Xi'an 710077, China*

Correspondence should be addressed to Nan Zhao; zhaonan@xidian.edu.cn

With the popularity of community question answering (CQA) sites, the research on identifying the expert users in online communities attracted increasing attention. We present a novel expert ranking algorithm based on the quality of user posts and the authority of user in community, and the similarity between the knowledge tags of users and questions in CQA sites is adopted in our scheme. Experimental results show that our scheme has better performance and accuracy under the same background with an amount of data samples.

1. Introduction

The number of Internet users is growing rapidly, along with the fast development of the network applications and infrastructure. In the enjoyment of convenience the network brings, it becomes difficult for the users to obtain the effective acquisition and screening of information [1, 2]. It follows then that community question answering (CQA) sites spring up [3, 4]. CQA sites are online knowledge communities, specializing in knowledge sharing and seeking, such as Stack Overflow [5, 6] and Yahoo Answers [7]. CQA sites provide a network platform for users to ask and answer questions and achieve information transfer and knowledge sharing among Internet users. Due to various topics and abundance content in CQA sites, network users prefer CQA sites to conventional web pages when seeking topic-specific information or solving problems [8]. The quality of information provided by CQA sites has been greatly improved in recent years. However, with the increasing number of community users, online communities amass an enormous amount of knowledge, which contains many useless answers inevitably in the community. Therefore, it is crucial to identify and recommend the experts in different fields of the CQA sites for the community

operation and extension services [9–11]. Meanwhile, network users can gain the accurate and high quality experience. Therefore, expert finding technique is of significance to improve the accuracy and efficiency of information acquisition in the CQA sites [12–14].

The existing expert finding techniques [15, 16] are divided into three major categories in generally. A directed graph is built based on the interaction between network users in the community, and the users are ranked by adopting the link analysis algorithm [17, 18] in the first category. In the second category, the text data in the community is analyzed based on the topic models [19, 20], and the results are applied to expert recommendation [21, 22]. In the last category, the hybrid models are built for expert finding with the methods mentioned above. A number of strategies are proposed unceasingly, but there are still some imperfections. Most of the traditional expert finding techniques ignore the user's activeness in the community. It may lead to the expert users not providing timely response. In addition, the comprehensive factors are not considered completely in some methods when users' expertise is evaluated. Finally, it may lead to the limited authority of recommended experts.

In our study, a more complete expert finding system that includes the expert ranking and the expert recommendation

is proposed. A new expert ranking algorithm is presented in this paper, named as Exp-rank. Exp-rank considers not only the authority of users in the community but also the quality of content published by users. On the basis of expert ranking, we calculate the similarity between the new questions and the knowledge tags of the users. According to the calculated results, we recommend experts more accurately to the new question. The rest of the paper is organized as follows. Section 2 briefly introduces the related work. Section 3 presents the proposed Exp-rank algorithm and the expert recommendation. Section 4 provides the dataset, performance evaluation, results, and discussion. Finally, section 5 summarizes the full task.

2. Related Work

The expert finding in online community is a widely investigated problem [23, 24]. In the development of CQA sites, a large number of users register for the community and participate in topic discussions. Meanwhile, the community has accumulated a lot of content, including a lot of useless information. Expert finding techniques can help identify and recommend the expert users in the CQA sites and avoid the adverse effects caused by spam and useless information [25–27]. The results of expert finding can be applied to the information management of the CQA sites and are helpful to provide users with more efficient and accurate question and answer service.

Link analysis algorithm is significantly adopted in the research of the expert ranking [28]. Zhang et al. [29] propose and evaluate several link-based expert ranking algorithms. They reveal that PageRank-based expert ranking algorithms outperform other algorithms in the online community. Yang and Wu [30] adopt weighted HITS algorithm to find experts in CQA sites. Link analysis algorithm can reflect the authority of users in the community. However, the link-based expert finding techniques focus merely on the link structure among individuals. They ignore the impact of useless replies and advertising accounts.

Graph-based algorithms are also applied in the research of expert finding. Zhao et al. [31] consider the problem of expert finding from the viewpoint of missing value estimation. The performance of the expert finding in CQA systems is improved because they employ users' social networks for inferring user model. Aslay et al. [32] propose Competition-Based Expertise Networks (CBEN), a novel community expertise network structure based on the principle of competition among the answerers of a question.

On the other hand, some researches reveal experts by analyzing online community content and user profile. Because of the complexity and diversity of information, the related strategies are varied. Shao and Yan [33] propose a model with two prediction methods that include the traditional feature-based method and LDA method. Specifically, when a new question arises, the model adopts LDA method to label and classify the question according to the latent semantic and content features. Then, with the traditional features of the question and the asker information, the model can recommend the appropriate expert users to answer this new question. Lu et al. [34] use semantic information extracted from user interaction to identify expert users. They construct the user question-answer interaction graph through direct semantic links and potential links extracted from the records of question session records and user profiles. After that, they employ the semantic information in the propagation link analysis method and in the language model. Faisal et al. mainly adopt the reputation of users in the community and the quality of users' answers as the main experts' evaluation indexes. On this basis, they combine voter reputation, voting rate, and other characteristics to measure the user expertise [35–37]. However, not every CQA site provides users with services like reputation system. Therefore, it is not conducive to extend these researches to other network communities.

Expert finding techniques based on social network features are rare. Most studies take the features of social network as one of the indicators to evaluate experts and propose a hybrid model for expert finding [38]. Wang et al. [39] consider both the relevance of documents and the authority of users in the community to assess the level of experts. Rafiei and Kardan [40] propose a hybrid method for expert finding in online communities, which is about the content analysis and the social network analysis. The content analysis is based on the concept map and the social network analysis is based on PageRank algorithm. Zhou et al. [41] present a topic-sensitive probabilistic model, which is an extension of PageRank algorithm to find experts in CQA. Compared with the conventional link analysis technology, their method considers not only the link structure, but also the topic similarity between different users. In fact, most of the previous works focus only on the static ranking or matching of domain experts without considering comprehensive factors that influence the user's expertise. In particular, our work serves as a method of dynamic expert finding system that combines expert ranking and expert recommendation.

3. Materials and Methods

We propose a new expert finding system containing expert ranking and expert recommendation for CQA sites. We adopt the cumulative the quality factor and the authority of users as the expert evaluation indicators and recommend experts to the new questions. A generic overview of the proposed scheme is given in Figure 1. In CQA site, a group of questions are $Q = \{q_1, q_2, \ldots, q_n\}$, where the question q_i owns answers $A = \{a_1, a_2, \ldots, a_m\}$ from users $U = \{u_1, u_2, \ldots, u_i\}$. In particular, we identify the expert users based on the expertise and the authority from U. In more detail, we evaluate the expertise of the users by analyzing their past performance. Meanwhile, the link-based ranking algorithm is employed to calculate the authority of the users. Then, we combine the expertise and the authority of the users to identify expert users. For recommending experts more accurately for the new questions, we extract the knowledge tags from top-ranking experts, and obtain appropriate recommended expert users by calculating the
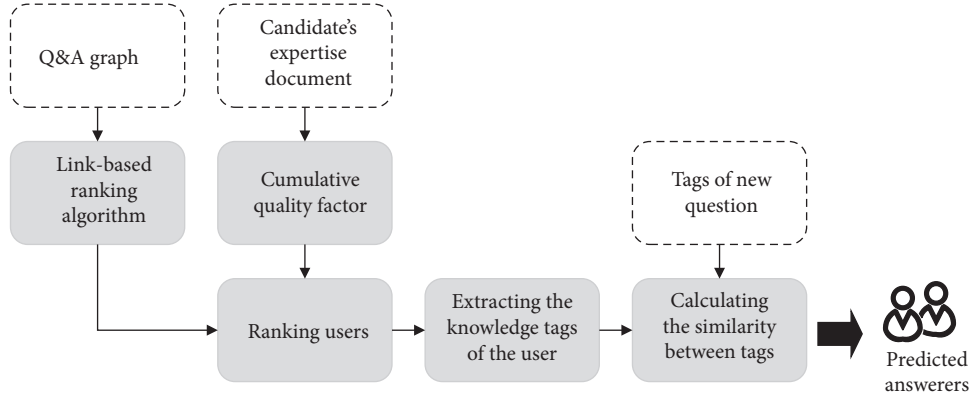
Figure 1: Architecture of proposed expert finding system.

similarity between the user knowledge tags and the new questions in the community.

3.1. Expert Authority Ranking Algorithm.

We build a particular network that indicates the interaction of community members to determine the social influence of the users in the online community. We adopt Q&A graph to represent a social network based on the interaction of the users. In Q&A graph, the nodes represent different users in the community, and a directed edge is built between two users when they are the inquirer and the responder about the same question, respectively, as shown in Figure 2. Q&A graph describes the interactions of user in the online community. Link analysis algorithms include PageRank and HITS and the authority of nodes can be measured by link analysis algorithms based on Q&A graph. Lü et al. [42] proposed LeaderRank algorithm based on PageRank. As shown in Figure 3, they consider a network of N nodes and M directed links. Nodes correspond to users and links are established according to the relations among leaders and fans. The idea of LeaderRank algorithm is to add a ground node which connects to every user through bidirectional links (see Figure 3 for an illustration). The network thus becomes strongly connected and consists of $N + 1$ nodes and $M + 2N$ links. The out-degree or in-degree of all nodes is greater than zero, which avoids isolated nodes in complex networks and ensures the convergence of the algorithm. Moreover, LeaderRank algorithm is an adaptive parameter free algorithm. Comparing with PageRank, LeaderRank has higher accuracy and robustness in mining important network nodes.

Figure 2 illustrates the user relationship in the CQA sites. Nodes represent users and links are established according to the relations among inquirer and the responder. In the discussions of the online community, the question from u_1 is answered by u_2, can gain, u_2, a vote of support from u_1. If u_3 and u_4 answer the question from u_2, the vote of support from u_2 is evenly distributed to them. The expert authority ranking algorithm is based on the fact that a user owns more authority than the user whose question is answered by him.

In particular, the community has some user groups whose members rarely communicate with users outside the group, and these user groups usually contribute less to the mainstream topics in the community. However, more internal links may exist in these user groups. These internal links are worthful to improve the quality of ranking, but they are usually ignored in link analysis algorithms. Thus, we propose an expert authority ranking algorithm to measure the authority of the users based on the above. E consists of directed links formed by the relationship between question and answer from users in the community. We present u_{ji} to indicate the contribution of user i to user j:

$$u_{ji} = \begin{cases} 0, & \text{if } (j, i) \notin E, \\ n(j, i), & \text{if } (j, i) \in E, \end{cases} \quad (1)$$

$$u'_{ji} = u_{ji} - \beta u_{ij}, \quad (2)$$

$$a_{ji} = \frac{u'_{ji}}{\sum_{k=1}^{N} u'_{jk}}, \quad (3)$$

where $n(j, i)$ is the number of times that user i has answered j. β denotes a damping factor, and the range of β is from 0 to 1. The value of u'_{ji} equals the value that the backlinks between users are subtracted from u_{ji}. This helps to eliminate the effect of internal links in user groups on ranking. When $u'_{ji} < 0$, we set u'_{ji} equal to zero. N is the total number of users that have answered user j, and a_{ji} represents the vote of support from user j to user i. The authority value of user i at time t is $AU_i(t)$, and we have

$$AU_i(t + 1) = \sum_{j=1}^{N+1} a_{ji} AU_j(t). \quad (4)$$

$AU_i(0) = 1$ represents the initial score of all user nodes except the ground node, and $AU_g(0) = 0$ represents the initial score of the ground node. AU_i tends to be stable at t_c, and the final authority score of user i is

$$AU_i = AU_i(t_c) + \frac{\theta_i AU_g(t_c)}{N}, \quad (5)$$

$$\theta_i = \theta * e^{-(d - d_0)}. \quad (6)$$

$AU_g(t_c)$ is the score of the ground node when it reaches the steady state. θ_i is the time attenuation of user i, θ is the attenuation coefficient, and d_0 represents the user's last post
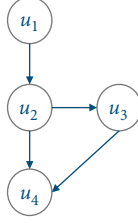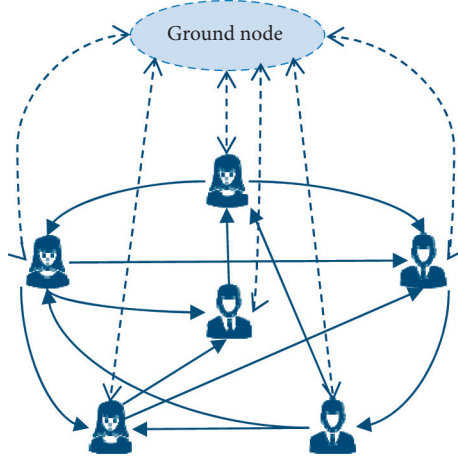
FIGURE 2: User-user relationship graph.

FIGURE 3: An illustration of the LeaderRank algorithm.

before the deadline d. If the last post from the user emerged a year ago, the attenuation coefficient of the user is 0. We measure the activity of the user according to the time attenuation and assign the score of the ground node through the activity of the users. Through the formula above, we can get the authority score of the users.

3.2. Cumulative Quality Factor. Most of the online communities possess malicious registered accounts, which are generally active in the communities, disseminating advertising or spam information. These accounts cannot be screened out when the authorities of users are calculated. Although some users actively participate in discussions to improve their authority value, the posts of these users are not professional and have little reference value for other users.

To solve these problems, we propose the concept of cumulative quality factor. In the process of data acquisition, we remove the users that rarely speak. Then, we summarize the scores or likes (the positive comments) of all the posts produced by users in the community and calculate the cumulative quality factor AS of the users:

$$AS_i = \frac{1}{N} \sum_{j=1}^{N} \ln(\delta_j + 1). \qquad (7)$$

The total number of answers posted by user i is N. δ_j represents the score for the answer from user i, and $\delta_j + 1$ is to avoid the situation that the score is zero.

We can estimate the past performance of the users by calculating the cumulative quality factor. Moreover, it helps

us to remove useless accounts and identify expert users with the expertise and the continuous excellent performance.

3.3. Exp-Rank. In order to evaluate expert users comprehensively, we combine the cumulative quality factor and the authorities of the users into an expert ranking standard. It is expressed as follows:

$$\text{Exp}_i = AU_i \cdot AS_i^{\lambda}. \qquad (8)$$

The expert score of user i is calculated with the cumulative quality factor AS_i and the authority value AU_i of the user. λ denotes a weighting factor, and the range of λ is from 0 to 1. In the expert finding of knowledge community, we think that the weight of AS_i should be slightly less than AU_i, and therefore the value of λ is set to 0.9. At last, we rank the candidates according to their expert scores Exp and obtain the results of expert ranking.

3.4. Expert Recommendation. We establish user files for the top ranking candidates. The user files are composed of questions and answers posted by users. The low-score answers posted by users are not adopted in the user files. Moreover, we extract keywords from user files by applying RAKE [43] algorithm. RAKE algorithm adopts punctuations to divide a file into several clauses, and the stop words are as delimiters to divide the clauses into several phrases, which are the candidates for the final extracted keywords. Each phrase can be split into several words by spaces, and every word can be given a score expressed as follows:

$$\text{word Score}\,(w) = \frac{\text{word Degree}\,(w)}{\text{word Frequency}\,(w)}. \qquad (9)$$

In the formula, the value of word Degree (w) consists of two portions, which are the number of times the word w forms a phrase with other words and the total number of times the word w occurs in the file, and word Frequency (w) represents the total number of times the word w occurs in the file. The word Score (w) of each word is calculated by formula. The score of every phrase can be obtained by accumulating the scores of words. RAKE algorithm extracts phrases that score in the top third as keywords.

As shown in Table 1, we take the keywords of user files as the knowledge tags of the user. We employ cosine similarity to calculate the similarity between the knowledge tags of user and the tags of the new question because all the questions in Stack Exchange website possess their own tags [44]. The expert users are recommended to the new question according to the similarity scores.

4. Experiments

We adopt the dataset of Stack Exchange website to simulate and compare the results with other algorithms for verifying the effectiveness of our proposed method.

4.1. Dataset. Stack Overflow is an online knowledge community, originally designed for programmers and computer engineers. It was founded in 2008 by two programmers, Joel Spolsky and Jeff Atwood. Users can post and answer questions, discuss with each other, and retrieve information from previous questions in the website. With the popularity of Stack Overflow, the founders of the website apply the same pattern to other fields, such as cooking and photography. Each CQA site is called Stack Exchange. Stack Exchange covers a wide range of topics.

Stack Exchange owns a large amount of Q&A data, and website operators regularly expose their data for the purpose of research. Based on the Q&A datasets, Correa and Sureka's [45] study deleted questions in the website to remind community members not to ask low-quality questions. Beyer and Pinzger [46] studied Stack Overflow tags and looked for the similar tags and merging them to avoid tag overflow. Meanwhile, the datasets from Stack Overflow have been used in expert finding and expert recommendation research. Faisal et al. [35] applied the g-index to expert ranking. Yang et al. [47] propose Topic Expertise Model (TEM) for expert finding. In order to verify our method, we adopt the dataset under the coffee topic of Stack Exchange website to simulate. The statistics of the dataset are shown in Table 2.

5. Results and Discussion

Stack Exchange possesses a reputation system, and each user owns a reputation score. Actually, the reputation system of Stack Exchange community has strict evaluation standards. Generally, if users want to improve their reputation score, they need to post valuable questions in their professional fields for a long time and provide high-quality answers or comments for other questions. The reputation system of Stack Exchange community helps the community to stimulate the potential of users and form a virtuous circle. Therefore, the reputation score of Stack Exchange community users has great reference value for evaluating expert users. Reputation score comes from the comprehensive performances of the user. We compare and analyze the results of the experts ranking with reputation scores of the users.

In the experiment, we choose two expert ranking techniques for comparison. The first one is Expertise Rank [29], which is an expert ranking method based on link analysis. The users are ranked in Expertise Rank according to the Q&A relationship among the users. The second one is LeaderRank [42], which is an improvement of Page Rank algorithm. In addition, LeaderRank is an adaptive parameter-free algorithm. We adopt the accuracy P to measure the difference of expert ranking algorithms. P is calculated as

$$P = \frac{r_{\text{list}} \cap \ e_{\text{list}}}{\text{num}_{\text{list}}}, \qquad (10)$$

where r_{list} denotes the user reputation ranking list, e_{list} denotes the experts ranking list, and num_{list} presents the number of experts in the list. Adopting the reputation ranking list as a benchmark, we select the top 30, 50, and 100 users respectively from the experts ranking lists obtained by different methods to calculate the accuracies, the results are shown in Figure 4. Exp-Rank model combines the persistent performance and authority of users. The result shows that the expert list ranked by Exp-Rank has a high correlation with the user reputation ranking list. It indicates that expert users selected by Exp-Rank are generally recognized. In addition, the performances of both Expertise Rank and LeaderRank algorithms are not satisfied, principally because the evaluation indexes of these techniques are not comprehensive.

In the section of the expert recommendation, we obtain the top 100 expert users from the expert ranking list of Exp-rank and extract their knowledge tags. In the Stack Exchange dataset, more than 90% of the questions possess five answers at most, and the high-score answers are rare. Therefore, we select five questions with the most answers as the new questions, and these questions have their own tags. Though comparing the similarity between the new question tags and the knowledge tags from the top 100 expert users, the expert users can be further screened according to the similarity score. We compare the recommended expert users with the users who actually answered these questions and calculate the accuracy of the expert recommendation. The accuracy is calculated as

$$P = \frac{R_{\text{list}} \cap \ A_{\text{list}}}{R_{\text{list}}}, \qquad (11)$$

$$\overline{P} = \frac{\sum_{i=1}^{N} P_i}{N}, \qquad (12)$$

where R_{list} indicates recommended list of experts and A_{list} indicates the list of users who actually answered the

TABLE 1: User knowledge Tags descriptions.

User ID	User knowledge tags
1	Moka-pot aeropress latte-art bean-varieties egg-coffee
2	Diner-coffee caffeine cholesterol percolator soya-milk
3	French-press latte espresso-machine space kahlúa
4	Cafetiere Indian coffee-maker French-press acidic

TABLE 2: Stack Exchange dataset statistics.

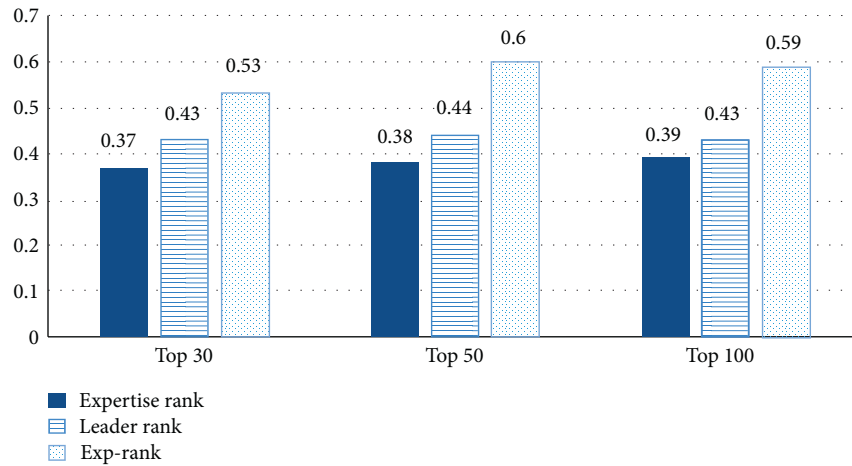| Users	7510
Questions	1059
Answers	2107
Comments	4602

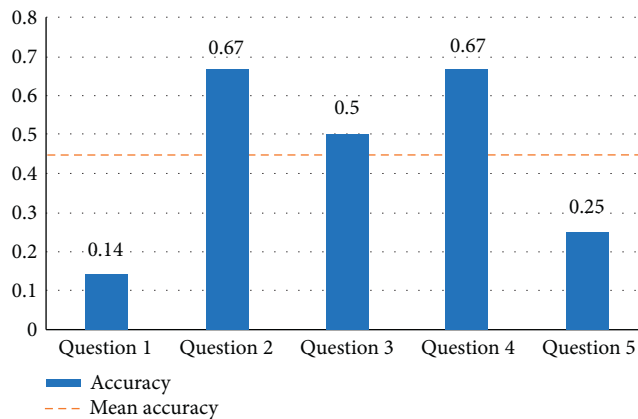FIGURE 4: Accuracy of the top users.
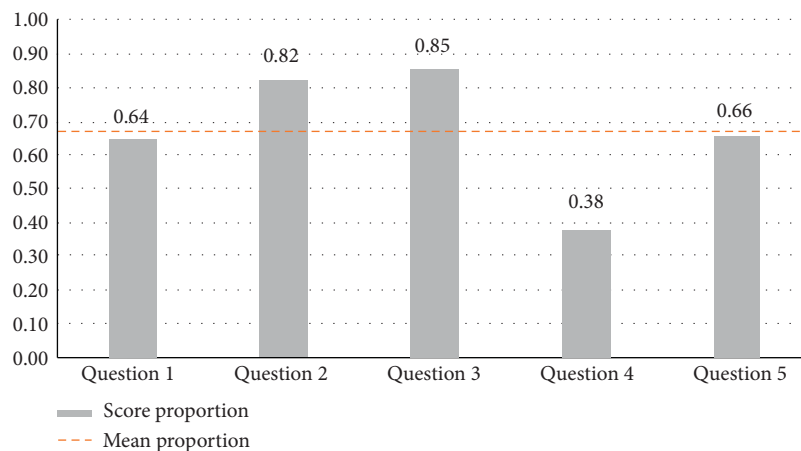

FIGURE 5: Accuracy score of recommended experts.

FIGURE 6: Percentage of the score for the expert's answers.

question. $\overline{P}$ is the mean accuracy. As shown in Figure 5, we recommend experts for the top five questions. The average accuracy is 0.44. That means when we recommend 10 experts, at least 4 of them will answer questions with high probability. As shown in Figure 6, for the five questions, the average proportion of expert users' answers scores is 0.67, which indicates that the answers of expert users are recognized by other users and have a higher professional level.

In conclusion, the experimental results prove that a better expert ranking result will be obtained by combining the authority of users with the continuous performance of users. In addition, recommending experts according to the similarity of the new questions can improve the accuracy of expert finding system.

6. Conclusion and Future Work

In order to identify the expert users in the complex online community, we propose a novel expert finding system based on the characteristics of the CQA sites. In our scheme, we propose an expert ranking algorithm named Exp-Rank, which considers the continuous performance and the authority of users and gives a more objective and comprehensive ranking of experts. Furthermore, we recommend experts according to the similarity between the new question and the knowledge tags of expert users. In particular, we can obtain some better results when we recommend 10–20 users. It should be noted that the evaluation indexes adopted in our method are common in CQA sites, so it can be widely applied in different types of online communities, such as Yahoo Answers and Zhihu. Consequently, we will try to enhance the performance of schemes with some more complex factors, including the user activity and the cold start problems of new users.

Data Availability

Previously reported Stack Exchange dataset was used to support this study and is available at https://archive.org/download/stackexchange. The dataset is freely available for research and has been used for finding experts and quality answers.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] J. Bian, Y. D. Liu, D. Zhou et al., "Learning to recognize reliable users and content in social media with coupled mutual reinforcement," in *Proceedings of the Web Conference*, pp. 51–60, Madrid, Spain, April 2009.

[2] M. Bouguessa, B. Dumoulin, S. R. Wang et al., "Identifying authoritative actors in question-answering forums: The case of Yahoo! answers," in *Proceedings of the Knowledge Discovery and Data Mining*, pp. 866–874, Las Vegas, NV, USA, August 2008.

[3] A. Kardan, M. Garakani, and B. Bahrani, "A method to automatically construct a user knowledge model in a forum environment," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 717-718, Geneva, Switzerland, July 2010.

[4] F. Liu, S. Xue, J. Wu et al., "Deep learning for community detection: progress, challenges and opportunities," in *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, pp. 4981–4987, Yokohama, Japan, July 2020.

[5] A. Barua, S. W. Thomas, and A. E. Hassan, "What are developers talking about? An analysis of topics and trends in stack overflow," *Empirical Software Engineering*, vol. 19, no. 3, pp. 619–654, 2014.

[6] F. Calefato, F. Lanubile, and N. Novielli, "How to ask for technical help? Evidence-based guidelines for writing questions on stack overflow," *Information and Software Technology*, vol. 94, pp. 186–207, 2018.

[7] L. Adamic, J. Zhang, E. Bakshy et al., "Knowledge sharing and Yahoo answers: Everyone knows something," in *Proceedings of the Web Conference*, pp. 665–674, Beijing, China, April 2008.

[8] H. Lin, W. Fan, L. Wallace et al., "An empirical study of web-based knowledge community success," in *Proceedings of the Hawaii International Conference on System Sciences*, p. 178, Big Island, HI, USA, January 2007.

[9] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4186–4199, 2020.

[10] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, pp. 1–12, Article ID 6325654, 2019.

[11] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, pp. 1–13, 2018.

[12] J. Wu, X. Zhu, C. Zhang, and P. S. Yu, "Bag constrained structure pattern mining for multi-graph classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 10, pp. 2382–2396, 2014.

[13] J. Wu, S. Pan, X. Zhu et al., "Boosting for multi-graph classification," *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 416–429, 2015.

[14] B. Yang and S. Manandhar, "Exploring user expertise and descriptive ability in community question answering," in *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, pp. 320–327, Beijing, China, August 2014.

[15] S. Lin, W. Hong, D. Wang, and T. Li, "A survey on expert finding techniques," *Journal of Intelligent Information Systems*, vol. 49, no. 2, pp. 255–279, 2017.

[16] A. M. Zuhair, K. Seifedine, and O. A. Isiaka, "Understanding expert finding systems: Domains and techniques," *Social Network Analysis & Mining*, vol. 8, no. 1, pp. 1–9, 2018.

[17] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1–7, pp. 107–117, 1998.

[18] R. Divya, C. Aswani Kumar, S. Saijanani et al., "Deceiving communication links on an organization email corpus,"

Malaysian Journal of Computer Science, vol. 24, no. 1, pp. 17–33, 2011.

[19] D. M. Blei, A. Y. Ng, M. I. Jordan et al., "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[20] A. Tehmina, D. Ying, D. Ali et al., "Topic-based heterogeneous rank," *Scientometrics*, vol. 104, no. 1, pp. 313–334, 2015.

[21] A. El-korany, "Integrated expert recommendation model for online communities," *International Journal of Web & Semantic Technology*, vol. 4, no. 4, pp. 19–29, 2013.

[22] J. Yang, A. Bozzon, G. J. Houben et al., "E-WISE: An expertise-driven recommendation platform for web question answering systems," in *Proceedings of the International Conference on Web Engineering*, pp. 691–694, Rotterdam, The Netherlands, June 2015.

[23] A. Bozzon, M. Brambilla, S. Ceri et al., "Choosing the right crowd: expert finding in social networks categories and subject descriptors," in *Proceedings of the 16th International Conference on Extending Database Technology*, pp. 637–348, Genoa, Italy, March 2013.

[24] Z. Jing, T. Jie, and J. Li, "Expert finding in a social network," in *Proceedings of the Database Systems for Advanced Applications*, pp. 1066–1069, Bangkok, Thailand, April 2007.

[25] F. Yi, S. Luo, and A. P. Mathur, "Discriminative probabilistic models for expert search in heterogeneous information sources," *Information Retrieval*, vol. 14, no. 2, pp. 158–177, 2011.

[26] Y. P. Fu, R. J. Xiang, Y. Q. Liu et al., "Finding experts using social network analysis," in *Proceedings of the Web Intelligence*, pp. 77–80, Silicon Valley, CA, USA, November 2007.

[27] K. Balog and R. M. De, "Determining expert profiles (with an application to expert finding)," in *Proceedings of the International Joint Conference on Artifical Intelligence*, pp. 2657–2662, Hyderabad, India, January 2007.

[28] P. Jurczyk and E. Agichtein, "HITS on question answer portals: An exploration of link analysis for author ranking," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 845-846, Amsterdam The Netherlands, July 2007.

[29] J. Zhang, M. Ackerman, and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in *Proceedings of the International Conference on World Wide Web*, pp. 221–230, Alberta, Canada, May 2007.

[30] R. R. Yang and J. H. Wu, "Study on finding experts in community question-answering system," *Applied Mechanics and Materials*, vol. 513–517, pp. 1760–1764, 2014.

[31] Z. Zhao, L. Zhang, X. He, and W. Ng, "Expert finding for question answering via graph regularized matrix completion," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 4, pp. 993–1004, 2015.

[32] Ç. Aslay, N. O'Hare, L. Aiello et al., "Competition-based networks for expert finding," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1033–1036, Dublin Ireland, August 2013.

[33] B. Shao and J. Yan, "Recommending answerers for stack overflow with LDA model," in *Proceedings of the Conference on Computer Supported Cooperative Work*, pp. 80–86, Portland, OR, USA, March 2017.

[34] Y. Lu, X. Quan, J. Lei et al., "Semantic link analysis for finding answer experts," *Journal of Information Science and Engineering*, vol. 28, no. 1, pp. 51–65, 2012.

[35] M. S. Faisal, A. Daud, and A. Akram, "Expert ranking using reputation and answer quality of co-existing users," *International Arab Journal of Information Technology*, vol. 14, no. 2, pp. 118–126, 2017.

[36] A. U. Akram, K. Iqbal, C. M. S. Faisal et al., "An effective experts mining technique in online discussion forums," in *Proceedings of the International Conference Computing Electronic and Electrical Engineering*, pp. 95–99, Piura, Peru, 2016.

[37] F. M. Shahzad, D. Ali, A. A. Usman et al., "Expert ranking techniques for online rated forums," *Computers in Human Behavior*, vol. 100, pp. 168–176, 2018.

[38] W. C. Kao, D. R. Liu, and S. W. Wang, "Expert finding in question-answering websites: a novel hybrid approach," in *Proceedings of the ACM Symposium on Applied Computing*, pp. 867–871, Sierre, Switzerland, March 2010.

[39] G. A. Wang, J. Jiao, A. S. Abrahams, W. Fan, and Z. Zhang, "ExpertRank: A topic-aware expert finding algorithm for online knowledge communities," *Decision Support Systems*, vol. 54, no. 3, pp. 1442–1451, 2013.

[40] M. Rafiei and A. A. Kardan, "A novel method for expert finding in online communities based on concept map and PageRank," *Human-centric Computing and Information Sciences*, vol. 5, no. 1, pp. 1–18, 2015.

[41] G. Zhou, S. Lai, K. Liu et al., "Topic-sensitive probabilistic model for expert finding in question answer communities," in *Proceedings of the Conference on Information and Knowledge Management*, pp. 1662–1666, Maui, HI USA, October 2012.

[42] L. Lü, Y. Zhang, C. H. Yeung, and T. Zhou, "Leaders in social networks, the delicious case," *PLoS One*, vol. 6, no. 6, Article ID e21202, 2011.

[43] S. J. Rose, D. Engel, N. Cramer et al., "Automatic keyword extraction from individual documents," *Text Mining: Applications and Theory*, pp. 1–20, 2010.

[44] A. Furtado, N. Oliveira, and N. Andrade, "A case study of contributor behavior in Q&A site and tags: The importance of prominent profiles in community productivity," *Journal of the Brazilian Computer Society*, vol. 20, no. 16, p. 5, 2014.

[45] D. Correa and A. Sureka, "Chaff from the wheat: characterization and modeling of deleted questions on stack overflow," in *Proceedings of the Web Conference*, pp. 631–642, Toulouse, France, July 2014.

[46] S. Beyer and M. Pinzger, "Grouping android tag synonyms on stack overflow," *Mining Software Repositories*, vol. 11, pp. 430–440, 2016.

[47] L. Yang, M. Qiu, S. Gottipati et al., "CQARank: Jointly model topics and expertise in community question answering," in *Proceedings of the Conference on Information and Knowledge Management*, pp. 99–108, San Francisco, CA. USA, October 2013.