Received 21 March 2019
Revised 12 October 2019
Accepted 16 October 2019

Expert recommendation in community question answering: a review and future direction

Zhengfa Yang
Central University of Finance and Economics, Beijing, China
Qian Liu
China Center for Internet Economy Research,
Central University of Finance and Economics, Beijing, China
Baowen Sun
Central University of Finance and Economics, Beijing, China, and
Xin Zhao
School of Economics and Management, Xi'an University of Technology, Xi'an, China

Abstract

Purpose – This paper aims to make it convenient for those who have only just begun their research into Community Question Answering (CQA) expert recommendation, and for those who are already concerned with this issue, to ease the extension of our understanding with future research.

Design/methodology/approach – In this paper, keywords such as "CQA", "Social Question Answering", "expert recommendation", "question routing" and "expert finding" are used to search major digital libraries. The final sample includes a list of 83 relevant articles authored in academia as well as industry that have been published from January 1, 2008 to March 1, 2019.

Findings – This study proposes a comprehensive framework to categorize extant studies into three broad areas of CQA expert recommendation research: understanding profile modeling, recommendation approaches and recommendation system impacts.

Originality/value – This paper focuses on discussing and sorting out the key research issues from these three research genres. Finally, it was found that conflicting and contradictory research results and research gaps in the existing research, and then put forward the urgent research topics.

Keywords Knowledge sharing, Community question answering, Expert finding, expert recommendation

Paper type General review

1. Introduction

Development of Web 2.0 has led to the increasing popularity of systems based on user-generated content. Community Question Answering (CQA) websites such as Yahoo!

International Journal of Crowd Science
Vol. 3 No. 3, 2019
pp. 348-372
Emerald Publishing Limited
2398-7294
DOI 10.1108/IJCS-03-2019-0011

Answers (answers.yahoo.com), Quora (www.quora.com) and Stack Overflow (stackoverflow.com) have become quite prominent in the past few years. CQA is a Web-based service that leverages the "wisdom of crowds" in which people can seek information by asking a question and simultaneously share knowledge by providing answers on questions asked by the rest of the community (Jie *et al.*, 2015). CQA systems are able to harness tacit knowledge (embedded in their diverse communities) or explicit knowledge (embedded in all of the resolved questions) in answering an enormous number of new questions posted each day (Xiang *et al.*, 2017). The question-answering process causes a flow of knowledge from more experienced users to less experienced users who can gain new knowledge by reading, asking and answering questions (Rostami and Neshati, 2019). In addition, users are able to elaborate solutions to solved problems through discussions attached to questions or answers (Aritajati and Narayanan, 2013).

A CQA website may have tens of thousands of questions posed every day. The growing number of new questions could induce two problems for CQA systems without appropriate collaboration support. First, it becomes more difficult for a general answerer to find the appropriate question to answer (Chang and Pal, 2013). In addition, the quality of answers to is uncontrollable because of the uncertain question-answering process as expertise and education levels vary a lot among answerers. An increasing failure rate (i.e. the proportion of questions that remain unanswered) and growing amount of low-quality content would cause a high churn rate (i.e. the proportion of users that leave the community), which significantly hamper the long-term sustainability of CQA systems (Srba and Bielikova, 2016).

To resolve the problems proposed above, expert recommendation recommend questions to potential answerers who are the most likely to provide satisfying answers. However, the focus of existing expert recommendation in literature varies a lot among understanding profile modeling, recommendation approaches and recommendation system impacts. To make it convenient for those who have only just begun their research into CQA expert recommendation, and to depict the current trends and highlight the areas that require further attention from the research community, in this paper, we perform an extensive survey on expert recommendation in CQA.

To maximize our survey coverage, we have paid significant attention to the collection of research articles. During the initial search phase, we used search tools provided by major digital libraries that contain computer science articles (i.e. ACM DL, IEEE Xplore, Springer Link and ScienceDirect). More specifically, the following search queries were used: "Community Question Answering", "CQA", Social Question Answering", "expert recommendation", "question routing" and "expert finding". The articles we obtained have given us an interesting overview of conferences and journals where CQA approaches are most often published. A significant number of articles were published at major international conferences such as the ACM International World Wide Web Conference (WWW), the ACM Conference on Research and Development in Information Retrieval (SIGR), the ACM Conference on Computer Supported Cooperative Work (CSCW), the ACM Conference on Human Factors in Computing Systems (CH), the Hawaii International Conference on System Sciences HICSS and the IEEE/ACM Conference on Advances in Social Networks Analysis and Mining (ASONAM). We also enriched the list of relevant articles with additional publications identified from their related works. Our final sample includes a list of 83 relevant articles authored in academia as well as industry that have been published from January 1, 2008 to March 1, 2019.

Based on a comprehensive review and a classification of approaches employed in CQA systems, our main contributions are as follows:

- A proposal of a general descriptive framework. We propose a general descriptive framework, which categorizes extant studies into three broad areas of CQA expert recommendation research: understanding profile modeling, recommendation approaches and recommendation system impacts.
- A comprehensive understanding of characteristics of entities (both users and content) of modeling expert recommendations. More specifically, we categorize the profile of content into four types: textual features, non-textual features, thread features and topic statistics, while the profile of users is categorized as question answering (QA) features and non-QA features. Based on an elaborate classification of the various possible types, future researchers can get a better understanding of state-of-the-art expert recommendation approaches.
- A review of representative approaches. We summarize and compare the advantages and shortcomings of state-of-the-art techniques based on different characteristics of both the entities of users and content for expert recommendation in CQA.
- Summary of impacts of expert recommendation applications on content, users and the community, to help researchers better understand the mechanisms necessary to best promote the management and development of CQA.

Through a comprehensive review of existing literature on expert recommendation, this paper propose four challenges, which outlook the promising directions for future research. First, the existing recommendation methods ignore users' willingness to continue to contribute within an online knowledge community, such as the motivational affordances of Q&A (Chen *et al.* , 2019), especially with the characters of pay-per-question (Jan *et al.*, 2018). Second, similar to recommendations in other fields, lack of sufficient information in users' profiles is the primary obstacle toward identifying potential experts. Third, it is promising to recommend experts as a collaborative group instead of finding knowledgeable individuals, which can largely improve the recommended answer rate. Last, the dynamicity-related research for CQA is still at a preliminary stage and it is important to consider self-evolution of existing recommendation methods, with new users joining and leaving, users' interests changing, users' roles transforming, users' mutual interactions evolving.

The remainder of this paper is organized as follows. Section 2 gives an overview of the expert recommendation problem. Section 3 categorizes the profiles of both users and content to construct an expert recommendation model. Section 4 presents the classification and introduction of some state-of-the-art expert recommendation approaches. In Section 5, we discuss the impacts of expert recommendation applications in CQA that is followed by a highlight of several promising research directions in Section 6.

2. Expert recommendation system

The expert recommendation issue is defined as the question routing, expert finding, answerer recommendation problem or expert recommendation. According to the research of (Srba *et al.*, 2015), among 400 K users of Stack Overflow, more than 169 K did not answer any question and more than 270 K did not answer more than four questions. This suggests that low-level CQA activity can be improved if users receive better recommendations that efficiently match their interests, motivations and expertise. A complete question-answering process is characterized by the presence of three crucial domain entities: a question, an answer and a user, who can play two main roles: asker and answerer.

The expert recommendation problem can be formalized as follows: given a newly posted question q we need to create an ordered list of top k users $u_1, u_2, \ldots, u_k$ who are the most suitable to answer question q. This list is usually ordered by the probability that user u would answer a given question q. To compressively understand how and why an expert recommendation is employed, it is necessary to solve four subproblems: construction of a question profile, which represents the questions' topics; construction of a user profile, which represents a user's expertise and/or interests, and optionally also additional characteristics (e.g. motivations); matching between the profile of a new question and all relevant users' profiles; and the impact of expert recommendation employment on users, content and community (Figure 1).

The basic inputs of an expert recommendation problem include the characteristic of the domain entities of both users (i.e. requesters and answerers) and user-generated content (i.e. the questions raised by requesters and the answers provided by answerers).

The profile of questions and answers can be categorized into four groups: textual features, which relate to the textual body of an analyzed question or answer such as length, structure, complexity, quality level and style/readability; non-textual features, which capture important metadata about questions and answers such as community feedback (i.e. votes, best answer selection and rating) and temporal (the time when the question was posted and the time it takes for someone to the first answer); thread features, which describe the context of questions and answers such as the relevance/similarity and thread statistics (e.g. the number and position of answers); and 4) topic features, which capture the meaning of the content such as a user assigned topic, the language/topic model and topic statistics (e.g. the number of questions assigned to the same topic and an average score of questions with the same topic).

The profile of requesters and answerers can be categorized into two groups: QA features, which come from the question-answering process itself such as activity level (e.g. the number of questions asked or the number of answers/best answers posted), expertise level (e.g. a ratio of those answers selected as the best) and temporal (e.g. the time from registration); and non-QA features, which describe a user on the basis of information that does not emerge directly from the question-answering process such as internal (e.g. an "about me" description and the number of followers) and external (e.g. the connected accounts at social networking sites).

Despite various possible types of inputs, which include characteristics of both content and users, only a subset of them might be available in a specific application scenario. Therefore, researchers may define the expert recommendation problem differently according to the inputs. By taking into account the different types of inputs, outputs and

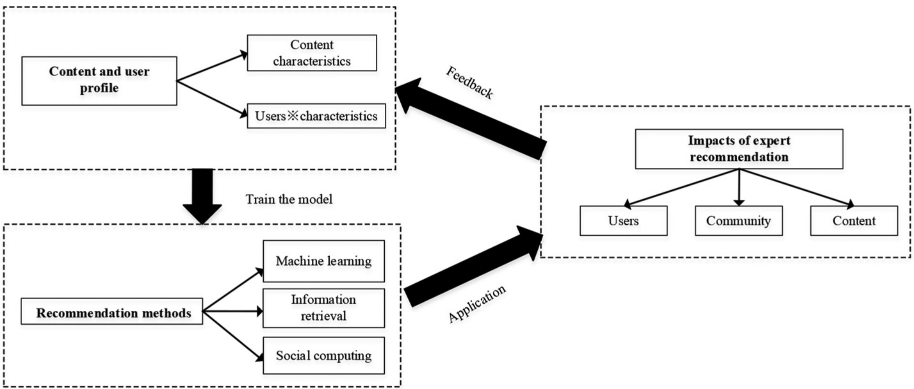

Figure 1.
A general descriptive framework of expert recommendation in CQA

recommendation method principle, expert recommendation research can be categorized into three groups: information retrieval (IR), machine learning and the social computing perspective.

3. Content and user profile

In the case of adaptive support methods, high-level profiles are employed to describe domain entities (both the content and users). Profile of enterprise and online community, is a general classification of expert input finding proposed by Zuhair *et al.* (2018) in which enterprise information can be classified as: self-disclosed information (Li *et al.*, 2016; Hu *et al.*, 2013), documents (Hu *et al.*, 2013; Petkova and Croft, 2006) and social networks (Zhang *et al.*, 2007). Expert finding in a social network suggests that online communities can be extracted from two sources: social networks (Aslay *et al.*, 2013; Anderson *et al.*, 2012; Hu *et al.*, 2013)) and documents (Jiang *et al.*, 2008; Souza *et al.*, 2013). However, the existing profiles of different expert recommendation methods has not been classified accurately nor deeply. In this paper, we classify high-level features of the content profile and user profile. These profiles are commonly filled with characteristics calculated by content/user modeling approaches (e.g. user topical expertise). To make the orientation in the large set of low-level features easier, we created their simple categorization (Table I).

Content profile modeling refers to the question of how to identify the "information needed" from a question? Questions and answers have two groups of features: textual features and non-textural features.

Textual features relate to the textual body of an analyzed question or answer such as its length, structure, complexity and style/readability. The profile of question textual features includes traditional question answering features such as the words and two-word phrases in a question, the "wh"-type (e.g. what or where), and the length of the subject (title) and detail (description) of the question. For example, (Jiang *et al.*, 2009) leverage characteristics of questions (e.g. subject length and detail length), answers (e.g. posting time, question stars and number of answers) and users, and their connection from a CQA-network to develop a semi-supervised coupled mutual reinforcement framework for simultaneously calculating content quality and user reputation. Our evaluation demonstrates that their methods are more effective than previous approaches for finding high-quality answers, questions and

Type	Content	Features	Articles
Content profile modeling	Textual features	Length Structure Complexity Style/readability	(Jiang *et al.*, 2009); (Pal *et al.*, 2012b); (Dijk *et al.*, 2015)
	Non-textual features	Community Feedback Temporal Thread features Topic features	(Jeon *et al.*, 2006); (Tomasoni and Huang, 2010); (Wu *et al.*, 2008); (Pedro and Karatzoglou, 2014); (Liu and Agichtein, 2011)
User profile modeling	QA features	Activity level Expertise level Temporal	(Dom and Paranjpe, 2008); (Xiong *et al.*, 2018); (Bouguessa *et al.*, 2008); (Qu *et al.*, 2009); (Wu *et al.*, 2008); (Lei *et al.*, 2009); (Liu and Agichtein, 2011); (Rybak *et al.*, 2014); (Xiong *et al.*, 2018); (Mukherjee *et al.*, 2016)
	Non-QA features	Internal External	(He *et al.*, 2014); (Pal *et al.*, 2012b); (Pal *et al.*, 2012a); (Srba *et al.*, 2015)

Table I.
Categorization of low-level features describing main domain entities

users. (Li and King, 2010) combined expertise-aware QLL with the Jelinek-Mercer smoothing model that leveraged multiple metadata features such as answer length, question-answer length, the number of answers for each question, the answerer's total points and the answerer's best answer ratio to measure expertise estimation with answer quality. (Pal *et al.*, 2012b) measured the existing value of prior answers by considering the number of answers, votes received, and answer status (i.e. unsolved, wrongly solved, partially solved and solved) to explore the question selection preferences among community experts and potential experts. (Dijk *et al.*, 2015) proposed a semi-supervised machine learning approach that uses textual, behavioral and time-aware features to measure whether a user shows signs of early expertise for a given topic.

Non-textual features capture important metadata about questions and answers including community feedback, temporal information, thread features and topic features. The profile of question non-textual features includes the specific features of communities such as temporal (i.e. the time when the question was posted and the time up until the first answer), community feedback (i.e. the votes, best answer selection and rating), thread features (the relevance and similarity), thread statistics (e.g. the number and position of an answer) and topic features (e.g. the number of questions assigned to the same topic, an average score of the questions with the same topic and question type). For example, one common feature of most CQA systems is the presence of community feedback tools, which serve as a crowd-sourced and distributed curation mechanism. Users can easily vote, positively or negatively, for questions (if they consider them interesting for future reference) or answers (if they correctly solve the problem stated in the associated question). (Wu *et al.*, 2008) proposed an algorithm that can take into account both users' positive and negative feedback. If a user gives a high rating to a recommended item, they regard it as positive feedback, otherwise it is negative feedback to make a real time expert recommendation. (Pedro and Karatzoglou, 2014) proposed RankSLDA, where both community feedback and text content topics are jointly modeled for ranking users according to their relevance for new questions. Hai *et al.* (2015) proposed a tag-LDA model to determine the user topic distribution and predict the topic distribution of new questions. They considered user post contents, answer votes, the ratio of best answers and user relations to find an appropriate user to answer a new question. Another important example of non-textual features refers to topic features. Liu and Agichtein (2011) analyzed the factors that may affect users' decisions of which questions to answer, which included the question category (topic), question position in the list shown to users and the surface patterns in the question text.

User profile modeling refers to the question of how to accurately profile a user's knowledge from various information sources for expert recommendation. Users have two groups of features: QA features and non-QA features.

QA features are derived from the question-answering process itself (i.e. the activity level, expertise level and temporal situation). More specifically, the profile of user QA features includes the activity level (e.g. the number of asked questions or posted answers/best answers), expertise level (e.g. a ratio of the answers selected as the best) and temporal (e.g. the time from registration). For example, (Xiong *et al.*, 2018) captured users' features of user type by leveraging a user activity model that predicted an influential long-term contributor by analyzing the honor system of Stack Overflow to train models. (Bouguessa *et al.*, 2008) used the number of best answers to model the expertise of users in Yahoo Answers. They proposed models that could automatically find the number of users that should be chosen as experts in the community.

Qu *et al.* (2009), Qu *et al.* (2009) and Qu *et al.* (2009) applied probabilistic latent semantic analysis (PLSA) to capture users' interests in terms of topics based on their answering history. Wu *et al.* (2008) proposed an algorithm that can take into account both the users' long-term and short-term interests. The long-term interests are reflected from all of the questions that users have already asked, while short-term interests are reflected from the new questions that users most recently asked. (Lei *et al.*, 2009) analyzed the daily activity patterns of users' contributions in knowledge-sharing online social networks to assess temporal features. Their work revealed that users' activity patterns follow a stretched exponential distribution. Similarly, (Hong and Shen, 2009) showed that users' temporal activities can be used to model changes in the network structure associated with the users. Compared to static graph analysis, their temporal model was able to better recognize users' common interests and make predictions about users' future activities. (Liu and Agichtein, 2011) analyzed the time of the day during which users preferred answering questions and proposed question routing schemes that would take the users' timing preferences into account to ensure that a question gets answered in a timely manner. (Rybak *et al.*, 2014) introduced the concept of a hierarchical expertise profile as a weighted tree. They defined a temporal expertise profile as a series of time-stamped hierarchical profiles and compared them to characterize important changes. (Xiong *et al.*, 2018) leveraged a temporal behavior model that used the answering time interval, the answering time rank and the wall clock time of the answer to extract users' timing features. In addition, (Mukherjee *et al.*, 2016) used the coupling between users' experiences, the interest in specific item facets, writing style and rating behavior to capture users temporal evolution and proposed an individual recommendation approach that takes into account users' maturity levels.

Non-QA features describe a user on the basis of information that does not emerge directly from the question-answering process. The profile of users' QA features includes internal (e.g. an "about me" description, the number of followers, user type and motivations) and external description (e.g. the connected accounts at social networking sites). According to (Pal *et al.*, 2012b), users with different roles in the Q&A community differ in their question selection heuristics. A disproportionately higher number of experts tend to answer questions with lower existing value than ordinary users. Similarly, (Liu and Jansen, 2014) built a classifier leveraging the non-QA characteristics of an answerer's profile and style of posting coming from Wenwo, which shows that characteristics of the questioner such as gender, popularity and activeness on social networking sites (SNS) are not that important to potential responders in social Q&A. The motivations and interests vary among potential answers. Since a Q&A site's popularity is based on the breadth of the questions posed and answers provided, it is important for the site to make sure that users who post questions look at the site as reliable and can reach the high-quality content that they are looking for as efficiently as possible while motivating those who provide high-quality answers to continue doing so. Pal and Konstan (2010) used the relative temporal series of the number of answers, best answers, and the number of answers and best answers to train support vector machines for the task of expert identification. They found that experts had more accurate identifications compared to a model that completely ignores the temporal aspect.

With the prevalence of online social networks today, it is easy to find CQA users (e.g. Facebook, Twitter, etc.). For example, more than one-third of the users in Quora have a twitter account (Zhou *et al.*, 2015). A social relationship between two users provides strong evidence for them to have common backgrounds. To recommend new questions to a wider part of a community such as newcomers or lurkers, (Srba *et al.*, 2015) proposed a question routing method that analyzes users non-QA data from CQA and external services (e.g. blogs, microblogs or social networking sites) as a supplement to QA activities in their estimation of users expertise.

4. Expert recommendation approaches

Considerable efforts have contributed to the expert recommendation research, which is the major technique to facilitate effective CQA. Information retrieval (IR), machine learning and social computing perspectives have each delivered fruitful results. In this paper, we classify state-of-the-art expert recommendation methods into five categories and review these methods by category in the following subsections.

4.1 Classification methods

The problem of identifying experts, as a particular class of users among all users, can be easily transformed into a classification problem that aims to distinguish such a particular class of expert users from other users. Classification methods can easily apply multiple aspects of features from the perspective of a user, question, answer or user-user interaction, to the expert recommendation problem. For example, (Pal *et al.*, 2012b) present a measure to capture users' question selection preferences in CQA that has high effectiveness in the identification of community experts and potential experts. Their result shows that experts have a tendency to answer low existing value (EV) questions. As a supplement, (Pal *et al.*, 2012a) use the relative temporal series of the number of answers and best answers, as well as the number of answers and best answers, to train Support Vector Machines for the task of expert identification. They find that experts were more accurate compared to the models that completely ignore the temporal aspect. Similarly, to predict potential responders and non-responders, (Liu and Jansen, 2014) built a binary classifier by combining profile-based features (e.g. the total number of days on Weibo and the number of followers) and posting style-based features (e.g. the total number and percentage of posts sharing) from both the questioner's and answerer's perspectives extracted from Wenwo. They found that newer and more active users were more willing to respond to routed questions. (Patil and Lee, 2016) analyzed the behavior of expert and non-expert people according to their level of activities and linguistic characteristics, and found that there was a considerable difference between them. (Dijk *et al.*, 2015), in their analysis of the effectiveness classification algorithm, showed that Random Forest outperformed Gaussian Naive Bayes and the Linear Support Vector Classification on the indices of the F1 score by considering three types of features: textual, behavioral and time-aware.

4.2 Language models

In the Q&A community, the interests of answer providers reflect stability over long periods of time, and an answer provider is usually interested only in questions that fall within specific topics. The query likelihood language model is implemented to measure the degree of interest in the question (Zheng *et al.*, 2012). Language models represent both question and user profiles as a bag of words, use a generative approach to compute the word-based relevance of a user's previous activities to the given question, and in turn, to predict the possibility of a user answering the question. In these traditional language models, data sparseness can lead to word mismatch between the routed question and user profiles, which can be caused by the co-occurrence of random words in user profiles or questions (Zhou *et al.*, 2012).

4.2.1 Expertise-aware QLL. Assuming that the responses of an answer provider for a given topic exhibit a greater probability of being high-quality answers (e.g. the Baidu Knows community answers that exhibit high-adoption rates), it can be inferred that the answer provider possesses sufficient expertise in a particular field (Zheng *et al.*, 2012).(Li and King, 2010) proposed a question routing framework that considers user expertise, answer quality and user availability to provide answers in a range of time. Query likelihood language (QLL) was leveraged to expertise estimation without answer quality, which assumed that a user has high expertise on a new question if they had previously answered many similar questions. By combining the Jelinek-Mercer smoothing model, which estimated answer quality as the weighted average of previous answer quality and the autoregressive model, which measured users' availability to answer a given question during a given period, the mean reciprocal rank (MRR) of expertise-aware QLL has been greatly improved.

4.3 Topic models

Since language models are based on exact word matching, they are most effective when they are used within the same topic. However, traditional language models are based on exact word matching, and thus they are not able to capture more advanced semantics and solve the problem of the lexical gap between the posted questions and users' profiles (Liu *et al.*, 2010). As a result of this limitation, topic models measure their relationship in the topic space rather than in the word space, which significantly outperforms language models.

The Latent Dirichlet Allocation (LDA) model (Blei *et al.*, 2003) is probably the most widely used topic model among all of the existing topic models developed to date. Topical expertise and authority ranking approaches rank users according to their expertise on particular topics (rather than their overall expertise). Moreover, some of these approaches also consider community feedback, which was neglected in previous approaches. (Zhu *et al.*, 2011) exploited information not only from the target category but also from other relevant categories, which are identified by a similarity measure based on an LDA topic model (Zhu *et al.*, 2011). The LDA model cannot take advantage of the internal structure of users' profiles, as each answered question can relate to a different topic. (Riahi *et al.*, 2012) proposed a segmented topic model (STM) that can discover the hierarchical structure of topics, and thus, instead of grouping all users' questions under one topic, allows each question to have a different topical distribution. (Momtazi and Naumann, 2013) used LDA to induce probabilistic topics from documents. In the first step, the LDA method has been used to extract topics from each document. The extracted topics show the connection between expert candidates and user queries. In the second step, the topics are used as a bridge to find the probability of selecting each candidate for a given query. The candidates are then ranked based on these probabilities. (Yi and Godavarthy, 2014) proposed a predictive language model to solve the future expert finding problem. Their method probabilistically estimates the association between a candidate e and a topic m in a future time t_2; while the method described in (Momtazi and Naumann, 2013) estimates the mentioned probability according to their association in current time t_1. (Srba *et al.*, 2015) employ a probability model based on latent topics identified by LDA for the expertise estimation employed in QA-based approaches with non-QA sources of data to estimate users' knowledge early and more accurately for users with low levels of QA activity.

4.3.1 Probabilistic latent semantic analysis. PLSA, proposed by (Hofmann, 1999) was developed based on latent semantic indexing (LSI) (Deerwester *et al.*, 1990), which uses singular value decomposition (SVD) to map high-dimensional count vectors to a lower dimensional representation in a so-called latent semantic space. PLSA captures underlying topics to represent documents and model the data generation process as a Bayesian network to leverage the semantics between words in documents, reduce the document representation space dimension and make up for the lack of semantic analysis of LSA.

PSLA is based on the observation that users' preferences and item characteristics are often governed by a few latent semantics. More specifically, PLSA introduces a latent variable, and decouples the probabilistic dependency between users and items into the dependency between users and latent semantics, and the dependency between the latent semantics and items, both in a probabilistic way (Wu *et al.*, 2008). Qu *et al.* (2009) treated all of the questions that a user accesses as one document and leveraged the distribution of users and their answered questions to contract a user-word aspect model to overcome the problem of data sparsity. The expectation maximization (EM) algorithm is generally used to find a local maximum of the log-likelihood of the question collection and to learn the model parameters. The incremental automatic expert recommendation framework based on PLSA proposed by Wu *et al.* (2008) overcame the problem that PLSA lacks incremental ability (i.e. it cannot handle new data arriving in a stream) and made sure that the system could reach a real-time online update. They considered not only users' long-term and short-term interests but also users' negative and positive feedback and questions as documents.

4.3.2 Latent Dirichlet allocation. LDA was proposed by Blei *et al.* (2003) to deal with the weakness of PLSA that the PLSA model is estimated for only those documents appearing in the training set, where LDA parameters were estimated by the approximate inference algorithms, such as variational EM and Gibbs Sampling. (Pedro and Karatzoglou, 2014) proposed a novel learning-to-rank extension to supervised LDA, and provided the derivation of a Gibbs sampler to perform inference. Their method is based on a Bayesian inference framework that extends the LDA model to account for the authorship of questions and answers as well as for community feedback. The proposed model combines the semantic content modeling benefits of LDA with supervised ranking learning, to model the observed community scores based on the latent topics assigned to each question.

4.4 Network-based methods

Network-based methods evaluate users' authoritativeness in a user-user network formed by their asking-answering relationships and recommend the most authoritative users as experts for a new question. The network-based approaches model is the underlying domain in the form of an expertise graph where the nodes represent the domain entities (i.e. experts and nonexperts) and the edges between the nodes represent some notion of expertise (e.g. influence, prominence, authoritativeness, etc.). This method supposes that if two users have a strong connection in the social network, they may qualify to answer similar questions. The simplest technique that can be used to measure the authority of CQA is the InDegree (Zhang *et al.*, 2007). Degree distribution is a function describing the number of users in the network with a given degree (i.e. the number of neighbors). An interesting common feature of many known complex networks is their scale-free nature. In a scale-free network, the majority of nodes are each connected to just a handful of neighbors, but there are a few hub nodes that have a disproportionately large number of neighbors. The InDegree technique measures the authority of a node (user) by the number of nodes that link to this node. A node with a high InDegree is likely to be a good authority (Aslay *et al.*, 2013). There are two basic types of community expertise networks: the asker-replier network and the asker-best answerer network. While the first one contains the edges weighted by a number of all of the provided answers and ignores the best answers, the second one considers only the best answers and ignores other non-best answers.

4.4.1 Expertise rank. Zhang *et al.* (2007) proposed ExpertiseRank, a slight variant of PageRank, which not only considers how many people one helped but also whom they helped. It supposes that If B is able to answer A's questions, and C is able to answer B's questions, then C should receive a high authority score, since C is able to answer the questions of someone who has some expertise. Their research also proposed a feature-based measure called the Z score, which is measured based on the number of answers (a) and the number of questions (g) as Zscore $= \frac{a-q}{\sqrt{a+q}}$. A user with a higher Zscore is more likely to be an expert than a user with a lower Zscore. This implies that experts answer a lot of questions and ask very few questions (often zero).

4.4.2 hits. Different from the PageRank method, the HITS algorithm distinguishes nodes into two types: hubs, which link to authoritative nodes; and authorities, which provide useful information on the given topics (Jurczyk and Agichtein, 2007). HITS assigns each node two scores: a hub score and an authority score. A hub score represents the quality of the outgoing links from the nodes while an authority score represents the quality of the information located on these nodes.

4.4.3 Other network-based methods. Zhu *et al.* (2011) take into consideration both content and user interaction-based category similarities to measure the relevancies between categories in an extended category link graph for ranking user authority. Liu *et al.* (2017) propose a question routing method from the viewpoint of knowledge graph embedding, which integrates topic representations with the network structure into a unified Knowledge Graph Question Routing framework to overcome the sparsity of CQA data.

Summary The output of network-based ranking algorithms is a ranked list of users based on their degree of authority on the subjects of interest. Based on this list, the top K users are considered to be the most authoritative. The weakness of such an approach resides in the unprincipled selection of the value of K. In general, the value of K is chosen solely on the basis of the specific knowledge of an application, while an inappropriate choice of the value of K can have a very negative impact on the quality of the service (Zhang *et al.*, 2007).

4.5 Deep learning

Deep learning has proven successful in many applications, and various deep learning models such as those based on autoencoders and neural auto regressive models have been applied to recommender systems (Lin *et al.*, 2017). Deep learning models have the advantage of using multimodal heterogeneous features and thus have the potential to solve complex problems such as the expert recommendation problem on a large scale. Until now, a convolutional neural network (CNN) has been the only deep learning model applied to recommending experts for a given question in CQA (Chen *et al.*, 2017). (Zhou *et al.*, 2016) consider the problem of expert finding from the viewpoint of ranking metric network embedding to overcome the sparsity of CQA data. They integrate both the semantic representation of the questions and heterogeneous CQA network structure learning into a unified Ranking Metric Network Learning framework, and develop a random-walk based learning method with deep recurrent neural networks to learn ranking metric embedding for questions and users in the proposed heterogeneous CQA network. Dargahi Nobari *et al.* (2017) have proposed two translation models to solve the vocabulary gap problem in expert finding to resolve the challenge of term mismatch between query words and candidates' documents Their first model is based on mutual information and their second model is a learning method based on a neural network. Dehghan *et al.* (2019) propose a new method long short-term memory (LSTM) deep neural network for T-shaped expert finding that is based on temporal expert profiling to resolve the challenge of the dynamicity and variability of CQA networks over time.

4.6 Hybrid methods

4.6.1 Network-based method+ language model-based question routing. Zhou *et al.* (2015) integrate both the social relation of users and their past question-answering activities into one common framework for the problem of expert finding and propose the graph regularized matrix completion method for estimating the missing values in the rating matrix with the social relationships of users. Zhou *et al.* (2014) combine a graph-based PageRank with an LDA semantic model to take into account not only link structure but also the topical similarity between askers and answerers.

4.6.2 Language model + topic model. Liu and Agichtein (2011) employed an integration of the language model and the LDA model to measure the relationship between an answerer and a question, which also considered user activity and authority information.

4.6.3 Network-based method + clustering. Following an idea similar to the geo-social community discovery in the point of interest (POI) recommendation, which incorporated clustering methods with network-based measures for the expert recommendation, Bouguessa *et al.* (2008) considered the number of best answers as an indicator of the authoritativeness of a user in a user-user network, where there is an edge from every requester to each of the corresponding best answers. Each edge was weighted by the number of best answers in-between. In particular, they modeled the authority scores of users as a mixture of their gamma distribution and used the Fuzzy C-Means algorithm to partition users into different numbers of clusters. They further used Bayesian information criteria (BIC) to estimate the appropriate number of mixtures. Finally, users were classified into two classes, one representing authoritative users with high indegrees and the other representing non-authoritative users with low indegrees. In this way, the method can automatically surface the number of experts in a community rather than produce a ranked list of users.

To ease illustration, we first summarize the typical inputs and outputs of existing expert recommendation methods in Table II. Then, we summarize the advantages and disadvantages of expert recommendation methods in Table III. Then, based on the input/output list, we further present a comparison of the representative methods with respect to their inputs and outputs in Table IV. Some methods may use derived features from the original inputs as additional inputs. For example, classification methods may use the length of questions (implied by question content), the total question number of users (implied by users' question histories) and the total answer number of users (implied by users' answer histories) as additional features to train their models.

Type	Category	ID	Input/output name
Input	Question profile	I0	Content and category of the given question
	User profile	I1	Users' question histories
		I2	Users' answer histories
		I3	Users' historical viewing and answering activities
		I4	Timestamps of users' answering activities
	Historical questions and answers	I5	Question content
		I6	Question category information
		I7	Question tags
		I8	Answer content
		I9	Best answer information
	Social profile	IA	Community feedback
		IB	User reputation
	Network profile	IC	Question-answer relationships among users
		IE	External links (website) to CQA
Output	Recommended experts	O1	Unranked group of experts
		O2	Ranked list of experts

Table II. Typical inputs and outputs of expert recommendation methods

Table III
Summary of advantages and disadvantages of expert recommendation methods

Category	Representative method	Advantages	Disadvantages	Articles
Language models	QLL Category-sensitive QLL Expertise-aware QLL	Generative approach Based on the same topic Word-based relevance of a user's previous activities to the given question	Data sparseness Co-occurrence of Random words in user profiles or questions	(Zheng *et al.*, 2012)) (Zhou *et al.*, 2012)) (Li and King, 2010)) (Mandal *et al.*, 2015))
Topic models	PLSA LDA STM DRM TagLDA	Captures more advanced semantics and measures their relationship in the topic space	Cannot leverage the internal structure of users' profiles and cannot explain users' relationships and their interactions	Liu *et al.* (2010) (Zhu *et al.*, 2011)) (Riahi *et al.*, 2012)) (Momtazi and Naumann, 2013)) (Yi and Godavarthy, 2014)) (Srba *et al.*, 2015)) (Qu *et al.*, 2009)) (Pedro and Karatzoglou, 2014))
Network-based methods	Indegree, PageRank, HITS, Expertise-aware Methods, Reputation-aware methods, Category-sensitive methods and Graph-embedding method	Leverage the internal structure of users' profiles, and explain users' relationships and their interactions	Cannot leverage specific syntax and semantics of textual profile, and ignores the topics and categories of questions	(Zhang *et al.*, 2007)) (Aslay *et al.*, 2013)) (Zhang *et al.*, 2007)) (Jurczyk and Agichtein, 2007)) (Zhu *et al.*, 2011)) (Liu *et al.*, 2017))
Classification methods	SVM RF GBDT LTR	Easily apply multiple aspects of features from the perspective of the user, question, answer or user–user interaction	Lack of personalization in the recommendation results that ignores the chance of a user noticing the question and ignores a user's willingness to answer the question	(Pal *et al.*, 2012b)) (Pal *et al.*, 2012a)) (Dijk *et al.*, 2015))

(continued)

Table III

Category	Representative method	Advantages	Disadvantages	Articles
Hybrid methods	QLL+LDA, Topical PageRank TEL, LDA+TF Topical PageRank+Expertise QLL+LDA+userActivity+Indegree, Indegree+Clustering	Comprehensively take into account multiple aspects of clues Combine different aspects of concern techniques and different techniques, and	Output depends on the objective function and may deal with data sparseness The recommendation algorithm is complex	(Zhou *et al.*, 2015)) (Zhou *et al.*, 2014)) (Bouguessa *et al.*, 2008))
Deep learning	Factorization machines (FM) Artificial neural network Ensemble learning	promising research direction utilizing multimodal heterogeneous features;have the potential of solving complex problems such as the expert recommendation problem on a large scale	Data points are required to be abundant	(Lin *et al.*, 2017)) (Chen *et al.*, 2017)) (Zhou *et al.*, 2016))

Table IV.
Comparison of inputs and outputs of representative expert recommendation methods

Category	Representative method	I0	I1	I2	I3	I4	I5	I6	I7	I8	I9	IA	IB	IC	IE	O1	O2
Language models	QLL	✓		✓			✓	✓			✓	✓					✓
	Category-sensitive QLL	✓		✓			✓										✓
	Expertise-aware QLL	✓		✓			✓										✓
Topic models	PLSA, LDA and STM	✓	✓	✓			✓		✓								✓
	DRM	✓		✓			✓										✓
	TagLDA	✓		✓													✓
Network-based methods	Indegree, PageRank and HITS	✓	✓	✓			✓	✓		✓	✓		✓	✓	✓		✓
	Expertise-aware Methods	✓		✓							✓		✓	✓	✓		✓
	Reputation-aware methods													✓	✓		✓
	Category-sensitive methods													✓	✓		✓
	Graph-embedding method														✓		✓
Classification methods	SVM, RF and GBDT	✓		✓	✓		✓		✓		✓	✓	✓	✓		✓	✓
	LTR	✓		✓	✓		✓		✓		✓	✓	✓	✓			
Hybrid methods	QLL + LDA, Topical PageRank,	✓		✓		✓	✓				✓			✓		✓	✓
	TEL and LDA + TF	✓		✓			✓				✓			✓			✓
	Topical PageRank + Expertise	✓		✓			✓				✓						✓
	QLL + LDA + userActivity + Indegree Indegree+Clustering	✓		✓			✓										
Deep learning	Factorization machines (FM)	✓	✓	✓	✓		✓		✓		✓	✓	✓	✓			✓
	Artificial neural network	✓		✓			✓		✓		✓			✓			✓
	Ensemble learning	✓		✓			✓				✓						✓

5. Impacts of expert recommendation

5.1 Impacts on content

The quality of user-generated content in CQAs is not uniform for all users (Yang and Manandhar, 2014). Answerers usually have varying amounts of interest and expertise in different topics and knowledge domains, which means the quality of the answers given by different background answerers on the same question may vary a lot. An extreme case is that answerers may give irrelevant answers that distract other users without thinking seriously (Agichtein *et al.*, 2009). In addition, the time required for preparing answers and the intention of answering also affect the quality of users' responses. Instead of receiving an answer instantly, users in CQA may need to wait a long time until a satisfactory answer appears (Procaci *et al.*, 2016)). According to the study of (Li and King, 2010), many questions on real-world CQA websites cannot be resolved adequately, meaning the requesters do not recognize the best answers to their questions within 24 hours. Thus, the quality of expert finding algorithms that depend on the quality of documents (i.e. questions and answers) may be indirectly affected.

It is worth noting that different scholars evaluate the quality of content problems in different ways. For example, (Jiang *et al.*, 2009) and (Li *et al.*, 2012) measure question quality as a question's effectiveness at attracting high-quality answers, while (Ponzanelli *et al.*, 2014) measure question quality from community feedback (i.e. deleted or closed questions with negative vote counts are defined as low quality). (Ravi *et al.*, 2014) measure question quality from the popularity of a question, and the ratio of the upvotes of a question and its number of visits is used to evaluate the quality of the question. Suzuk and Joho (2011) investigated how various contextual information included in a question can lead to better answers. They found that including a contextual factor to the question can improve a questioner's assessm-ent of the quality of the answer. Meanwhile, (Blooma *et al.*, 2012) measure answer quality from the social features of the person, which means that expert recommendation is based on the network method that could, to some extent, enhance answer quality. (Toba *et al.*, 2014) point out that the hybrid hierarchy-of-classifiers method can help detect high-quality answers. In addition, by combing non-textual features and unlabeled data (Liu *et al.*, 2015) and the summary-style (i.e. novelty and redundancy) (Wei *et al.*, 2016), it is easy to detect potential high-level quality answers.

Another important impact refers to content that enhances question answerability. Unanswered questions, which are not a rare phenomenon, reached a total of 13 per cent of the questions in the Yahoo! Answers dataset in our study, and users whose questions remained unanswered were considerably more prone to churn from the CQA service (Dror *et al.*, 2012). Since an asker has little or no influence on the answerer's behavior, (Dror *et al.*, 2012) introduced a novel task of predicting the number of expected answers for a question before it is posted, extracting various attributes of the question metadata, and questioning the content and user data to train a classification model. Their results showed that questions are more often answered toward the end of the week, and that the fraction of unanswered questions is negatively correlated with the average number of answers per question.

5.2 Impacts on users

Asker satisfaction plays a crucial role in the growth of the decay of a question answering community. If many of the askers in CQA are not satisfied with their experiences, they will not post new questions and they will rely on other means of finding information. Expert recommendation improves the experience of other community members who have less knowledge, reduces user waiting time and conserves system resources (Zheng *et al.*, 2012). Agichtein *et al.* (2009) proposed an Asker Satisfaction Prediction system (ASP) to predict information seeker satisfaction in collaborative question answering communities by considering both the offline and online setting, and leveraging the features of questions, answers, question-answer pairs, users and categories. Their supervised classification methods show that the question-answer relationship, answerer reputation and answerer user history have less effect on asker satisfaction, while asker satisfaction varies with the past experience of the asker and textual features.

5.3 Impacts on community

It is important that researchers of CQA understand why some experts leave the system, and what measures can be used to retain them in the community and explore the effectiveness of question routing schemes. In CQA services, it is hard to convince users to start answering questions, or even visit the Web-service for the first time, and the cost of acquiring new customers is higher than the cost of keeping existing customers (Wei and Chiu, 2002). Thus, it is important to decrease the churn rate of newbies and enhance the lifespan of prolific users that already actively answer questions in the system. (Dror *et al.*, 2012) researched the task of churn prediction in new users. Their results validate the effectivity of potential expert recommendation for new users as users over the first week of their "life" as answerers who post more answers, are much less likely to churn. It is essential to recommend relevant questions to both newbies and veterans with more activity in a specific frequency since the time gap between subsequent posts is the most significant indicator of diminishing user interest (Pudipeddi *et al.*, 2014). Karumur *et al.* (2016) overcame the inherent problems of the lack of available information on new users with little previous activity history and often incomplete demographic information, by collecting data from MovieLens to assess newcomer retention. They found activity diversity and that user participation is positively related to newcomer retention. (Pal *et al.*, 2012a), to find users for a community task, question-routing and providing stimulus to improve users' participation, classified the evolution of experts into three distinctive patterns: experts who are consistently active in the community, experts who are initially very active but become passive over time and experts who are initially passive but become very active over time. Core answerers are the primary drivers of answer production in CQA. According to the study by Anderson *et al.* (2012), there are many highly dedicated domain experts who aim to satisfy requesters' queries but, more importantly, provide answers with high-lasting value to a broader audience in Stack Overflow and Quora. Expert recommendation inherently encourages the fast acquisition of higher-quality answers, it potentially increases the participation rates of users, improves the visibility of experts and fosters stronger communities in CQA.

6. Future directions

6.1 Realistic user modeling

Expert recommendation relies on effective user modeling. Intuitively, two aspects of concerns exist that affect whether a user gives a high-quality answer to a question in a real Q&A scenario: the chance of a user noticing the question and a user's willingness to answer the question.

Chance of a user noticing the question. As a user may not have an opportunity to see a question, the user may not be an answerer to this question even though the user is an expert. However, the expert recommendation problem in CQA is based on a different assumption from real-world scenarios (i.e. what is the possibility that a user would answer a question and meanwhile provide a high-quality answer to that question if the user is invited to answer the question?). Due to the above difference, when using real-world labeled data to train recommendation models, the recommendation methods should take into account the possibility that a user may not have answered a question simply because the user did not have the chance to notice the question. The likelihood that a user would see a question in real-world scenarios depends on various factors such as user availability (e.g. how often a user is online and available to answer questions), users' behaviors (e.g. whether users look for new questions to answer actively) and other users' activities related to the question (e.g. how widespread the question is among users).

User's willingness to answer the question. Even if a user has noticed a question, the user may choose not to answer it. A user's willingness to answer a question also depends on various factors such as how well the question fits the user's interest, user's self-confidence in the quality of their answers and a user's expected gains from answering the question. In addition, the existing recommendation methods ignore users' willingness to continue to contribute within an online knowledge community, such as the motivational affordances of Q&A (Chen *et al.*, 2019). In addition, the characters of pay-per-question (Jan *et al.*, 2018), and the historical records of payed questions and answers have not yet been taken into existing recommendation arithmetic, in which aspects of state-of-the-art research in the expert recommendation in CQA can be improved and inspire additional future research in this area. (Pal *et al.*, 2012a) found that some users are consistently active, some start active but end passive over time, and that some start passive but become active over time. According to (Khansa *et al.*, 2015) active participation can be understood as the setting, pursuit and automatic activation of goals – what an effective design mechanism CQA can adopt to promote active online participation and how to verify the effect is unknown.

6.2 Early expert detection

Detecting topical expertise is a well-studied problem, which relates to expertise finding and retrieval under the premise that the historical profile of users and content is available and affluent (Wang, 2013; Hu *et al.*, 2013). The main challenge of early expert detection lies in inherent data sparsity issues and the cold start problem (i.e. how to profile an expert given only a handful of data points such as questions, answers and comments). The lack of sufficient information in users' profiles is the primary obstacle toward identifying early-career experts since the more complicated methods usually require a larger train set. For example, classification methods generally perform better under high-dimensional features given sufficient training data. When the training data is limited, these methods need to restrain the dimensionality to avoid over-fitting.

Few researchers have studied the discovery of potential experts at an early stage of CQA. However, there are great differences in the definition and operationalized measurement of early expert recognition in existing studies (Pal *et al.*, 2011). Pal *et al.* (2012a) define early expert identification in terms of the consideration of the temporal dynamics and interactions among experts. More specifically, they look at the first few weeks of their activity in the community. Pal *et al.* (2012b) address this identification problem between both current experts and potential experts based on the question selection preferences (QSP) of CQA users. Dijk *et al.* (2015) define the early expertise shown by a user as between the moment of joining and becoming an expert, based on the best answers provided. Among the existing methods, in this paper, we identify two promising categories of methods that can potentially better detect experts early. One is the semi-supervised machine learning approach proposed by (Dijk *et al.*, 2015), which leverages the characteristics of textual, behavioral and time-aware information to predict whether a user will become an expert in the long term.

The existing literature has researched early detection of high quality on CQA (Yuan *et al.*, 2015; Ponzanelli *et al.*, 2014; Arora *et al.*, 2015; Neshati, 2017), which can improve the process of question routing, reduce the number of questions with no answers, improve users' experiences and promote the content quality of a CQA by rejecting low-quality content. Despite the significance of the early question quality detection issue, the effect of high-quality content prediction on the performance of expert finding and expert profiling tasks is unknown. This is an emerging problem calling for researchers to endeavor to solve.

6.3 Recommending experts as a collaborative group

Rather than finding knowledgeable individuals, sometimes locating a group in an organization with appropriate skills and knowledge is of great importance to the success of a project being undertaken. An ideal expert group should satisfy the following conditions. First, the question must appeal to all of the group members so that they are likely to answer the question. Second, the group members should be compatible with one another so that the existence of one user in the group will not discourage another user to answer a question. Third, it is desirable for the group members to complement one another in the knowledge domains required to address a question, given that users may have different sets of skills and different levels of expertise on different skill aspects. For example, based on probabilistic language modeling techniques, (Liang and Rijke) present five general strategies for this group finding task, given a heterogenous document repository. Three types of variables, which include groups (G), queries (Q) and documents (D) are leveraged so that first, evidence of whether a group is knowledgeable about the topic via the experts in the group (G) is collected, then whether each expert in the group has expertise on the topic via documents (D) is determined, and finally whether a document addresses the given query (Q) topic is understood.

Intuitively, those users who have frequently answered similar questions are likely to be compatible with one another (Feng *et al.*, 2018). A possible solution following this insight is to propose an expert recommendation scheme that aims at selecting the best subset (e.g. of a size of k) of collaborative users by simultaneously learning their co-occurrence in the same thread and topical expertise. The selection of a group of collaborative users could also borrow ideas from two closely related topics, namely optimal task decomposition (Tong *et al.*, 2018) and user group evaluation. Task decomposition is the opposite approach of group formation, which aims to break the knowledge requirement into sub-requirements and find a user for every sub-requirement to comprise the final group. User group evaluation aims to set better heuristics to promote improved recommendation results and answers to any given question.

6.4 Dynamicity of expert finding

Finding potential experts in CQA is beneficial to several problems such as the low participation rate of users, long waiting times to receive answers and the low quality of answers. CQAs are dynamic environments because of the massive daily posts, joining or the addition of new users, users who change their activities and interests, emerging new topics and the uptrend or downtrend of topics (Neshati *et al.*, 2017). However, most of existing literature take expertise into consideration in a single snapshot of the environment (i.e. at the query time) and ignore the evolution of personal expertise over the time of the expert finding problem. Importantly, the real-world question of answering websites is dynamic, with new users joining and leaving, users' interests changing, users' roles transforming, users' mutual interactions evolving, and the content on the website continuously being updated. (Patil and Lee, 2016) studied users on Quora and identified three types of expert users based on the weekly changes in the number of answers they provided. They also used temporal features including daily changes in the number of followees, followers, edits, questions and answers to improve the precision of expert detection. introduced the new problem of Future Expert Finding to predict the best ranking of experts in the future given the expertise evidence at the current time. Four factors: topic similarity, emerging topics, users' behaviors and topic transition were used in modeling expertise in a dynamic environment. Their temporal profile based model (TPBM) improved the mean average precision (MAP) measure up to 39.7 per cent in comparison with our best baseline method. Yeniterzi and Callan (2015) proposed adapting temporal discounting models to expertise estimation methods for question routing. Two widely used expert finding approaches, Answer Count and Zscore, were modified to use the available temporal information. They used available temporal information in CQA sites to make these existing approaches more effective for the task of question routing. Yi and Godavarthy (2014) propose a new probabilistic model to characterize how people change or stick with their expertise and a predictive language model is derived to estimate the distribution of the expert's words in their future publications.

In summary, all of the above dynamic aspects suggest an expert recommendation method suitable to self-evolve in an online fashion. However, none of the above methods is designed to be online-friendly, and it could take a tremendous amount of time to retrain the new model when new information becomes available, which is unacceptable in practice as most Q&A systems in the real-world involve massive amounts of data. Therefore, a promising research direction is to introduce novel methods that are capable of incrementally learning about users and continuously adapting their recommendation behaviors efficiently and effectively over time. There is also a paucity of related work because we believe that the dynamicity-related research for CQA is still at a preliminary stage, as most of the methods used are relatively simple and predict different aspects of consideration such as user availability, user interest, and user expertise, separately. Moreover, they have not considered the possible correlations among these aspects. Therefore, another potential point for future research is to predict different aspects of features simultaneously using a single, comprehensive model for better results.

7. Conclusion

Based on relevant articles authored in academia as well as in industry published from Jan 1st, 2008 to March 1st, 2019, we find that the literature on expert recommendation in CQA is fragmented and lacks an overarching framework to systematically guide research and integrate findings. Therefore, we propose a general descriptive framework in this survey, which categorizes extant studies into three broad areas of CQA expert recommendation research: understanding profile modeling, recommendation approaches and recommendation system impacts, followed by the identification of the open issues and the promise of future research directions. Our aim in this survey research paper has been to make it convenient for those who have only just begun their research in this area by providing a summary of state-of-the-art expert recommendation approaches, and for those who are already concerned with CQA expert recommendation to ease the extension of our understanding of this issue with future research.

References

Agichtein, E., Liu, Y. and Bian, J. (2009), "Modeling information-seeker satisfaction in community question answering", *Acm Transactions on Knowledge Discovery from Data*, Vol. 3 No. 2, pp. 1-27.

Anderson, A., Huttenlocher, D., Kleinberg, J., *et al.* (2012), "Discovering value from community activity on focused question answering sites: a case study of stack overflow", *Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*.

Aritajati, C. and Narayanan, N.H. (2013), "Facilitating students' collaboration and learning in a question and answer system", *Conference on Computer Supported Cooperative Work Companion*.

Arora, P. Ganguly, D. and Jones, G.J.F. (2015), "The good, the bad and their kins: identifying questions with negative scores in stackoverflow",

Aslay, Ç., O'hare, N., Aiello, L.M., *et al.* (2013), "Competition-based networks for expert finding", *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, pp.1033-1036*.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3 No. Jan, pp. 993-1022.

Blooma, M.J., Goh, D.H.L. and Chua, A.Y.K. (2012), "Predictors of high-quality answers", *Online Information Review*, Vol. 36 No. 3, pp. 383-400. (318).

Bouguessa, M., Dumoulin, B. and Wang, S. (2008), "Identifying authoritative actors in question-answering forums: the case of Yahoo! answers", *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 866-874.

Chang, S. and Pal, A. (2013), "Routing questions for collaborative answering in community question answering", *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.

Chen, L., Baird, A. and Straub, D. (2019), "Why do participants continue to contribute? Evaluation of usefulness voting and commenting motivational affordances within an online knowledge community", *Decision Support Systems*, Vol. 118, pp. 21-32.

Chen, Z., Zhai, S. and Zhang, Z. (2017), "A deep learning approach for expert identification in question answering communities".

Dargahi Nobari, A., Sotudeh Gharebagh, S. and Neshati, M. (2017), "Skill translation models in expert finding", *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1057-1060.

Deerwester, S., Dumais, S.T., Furnas, G.W., *et al.* (1990), "Indexing by latent semantic analysis", *Journal of the American Society for Information Science*, Vol. 41 No. 6, pp. 391-407.

Dehghan, M., Biabani, M. and Abin, A.A. (2019), "Temporal expert profiling: with an application to t-shaped expert finding", *Information Processing and Management*, Vol. 56 No. 3, pp. 1067-1079.

Dijk, D.V., Tsagkias, M. and Rijke, M.D. (2015), "Early detection of topical expertise in community question answering", *International Acm Sigir Conference on Research and Development in Information Retrieval*.

Dom, B. and Paranjpe, D. (2008), "A Bayesian technique for estimating the credibility of question answerers", *Proceedings of the 2008 SIAM International Conference on Data Mining*, pp. 399-409.

Dror, G. Dan, P. Rokhlenko, O. *et al.* (2012), "Churn prediction in new users of Yahoo! answers".

Feng, W., Zhu, Q., Zhuang, J., *et al.* (2018), "An expert recommendation algorithm based on pearson correlation coefficient and fp-growth", *Cluster Computing No*, Vol. 3, pp. 1-12.

He, X., Gao, M., Kan, M.Y., *et al.* (2014), "Predicting the popularity of web 2.0 items based on user comments", *International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 233-242.

Hofmann, T. (1999), "Probabilistic latent semantic analysis", *Fifteenth Conference on Uncertainty in Artificial Intelligence*.

Hong, D. and Shen, V.Y. (2009), "Online user activities discovery based on time dependent data", *International Conference on Computational Science and Engineering*.

Hu, X., Gurnani, H. and Wang, L. (2013), "Managing risk of supply disruptions: incentives for capacity restoration", *Production and Operations Management*, Vol. 22 No. 1, pp. 137-150.

Jan, S.T.K., Wang, C., Zhang, Q., *et al.* (2018), "Pay-per-question: towards targeted q&a with payments", *Acm Conference on Supporting Groupwork*.

Jeon, C., Bruce, W., *et al.* (2006), "A framework to predict the quality of answers with non-textual features", *International Acm Sigir Conference on Research and Development in Information Retrieval*.

Jiang, B., Agichtein, E., Liu, Y., *et al.* (2009),"Learning to recognize reliable users and content in social media with coupled mutual reinforcement", *International Conference on World Wide Web*.

Jiang, B., Liu, Y., Agichtein, E., *et al.* (2008), "Finding the right facts in the crowd: factoid question answering over social media", *International Conference on World Wide Web*.

Jie, Y. Bozzon, A. and Houben, G.J. (2015), "E-wise: an expertise-driven recommendation platform for web question answering systems",

Jurczyk, P. and Agichtein, E. (2007), "Discovering authorities in question answer communities by using link analysis", *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pp. 919-922.

Karumur, R.P., Nguyen, T.T. and Konstan, J.A. (2016), "Early activity diversity: assessing newcomer retention from first-session activity", *Acm Conference on Computer-supported Cooperative Work and Social Computing*.

Khansa, L., Xiao, M., Liginlal, D., *et al.* (2015), "Understanding members' active participation in online question-and-answer communities: a theory and empirical analysis", *Journal of Management Information Systems*, Vol. 32 No. 2, pp. 162-203.

Lei, G., Tan, E., Chen, S., *et al.* (2009), "Analyzing patterns of user content generation in online social networks", *Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*.

Li, B. and King, I. (2010), "Routing questions to appropriate answerers in community question answering services".

Li, B., Tan, J., Lyu, M.R., *et al.* (2012), "Analyzing and predicting question quality in community question answering services", *International Conference on World Wide Web*.

Li, W. Eickhoff, C. and Vries, A.P.D. (2016), "Probabilistic local expert retrieval",

Liang, S. and Rijke, MD. (2016), "Formal language models for finding groups of experts", *Information Processing and Management*, Vol. 52 No. 4, pp. 529-549.

Lin, S., Hong, W., Wang, D., *et al.* (2017), "A survey on expert finding techniques", *Journal of Intelligent Information Systems*, Vol. 49 No. 2, pp. 1-25.

Liu, B., Jian, F., Ming, L., *et al.* (2015), "Predicting the quality of user-generated answers using co-training in community-based question answering portals", *Pattern Recognition Letters*, Vol. 58, pp. 29-34.

Liu, Q. and Agichtein, E. (2011), "Modeling answerer behavior in collaborative question answering systems", *European Conference on Information Retrieval*.

Liu, Z. and Jansen, B.J. (2014), "Predicting potential responders in social q&a based on non-qa features", *CHI' 14 Extended Abstracts on Human Factors in Computing Systems*, pp. 2131-2136.

Liu, Z., Li, K. and Qu, D. (2017), "Knowledge graph based question routing for community question answering", *International Conference on Neural Information Processing*, pp. 721-730.

Mandal, D.P., Kundu, D. and Maiti, S. (2015), "Finding experts in community question answering services: a theme based query likelihood language approach", *Computer Engineering and Applications*,

Momtazi, S. and Naumann, F. (2013), "Topic modeling for expert finding using latent dirichl*et al.* location", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 3 No. 5, pp. 346-353.

Mukherjee, S., Lamba, H. and Weikum, G. (2016), "Experience-aware item recommendation in evolving review communities", *IEEE International Conference on Data Mining*.

Neshati, M. (2017), "On early detection of high voted q&a on stack overflow", *Information Processing and Management*, Vol. 53 No. 4, pp. 780-798.

Neshati, M., Fallahnejad, Z. and Beigy, H. (2017), "On dynamicity of expert finding in community question answering", *Information Processing and Management*, Vol. 53 No. 5, pp. 1026-1042.

Pal, A., Chang, S. and Konstan, J.A. (2012a), "Evolution of experts in question answering communities", *Sixth International AAAI Conference on Weblogs and Social Media*.

Pal, A., Farzan, R., Konstan, J.A., *et al.* (2011), "Early detection of potential experts in question answering communities", *International Conference on User Modeling, Adaptation, and Personalization*, pp. 231-242.

Pal, A., Harper, F.M. and Konstan, J.A. (2012b), "Exploring question selection bias to identify experts and potential experts in community question answering", *Acm Transactions on Information Systems*, Vol. 30 No. 2, pp. 1-28.

Pal, A. and Konstan, J.A. (2010), "Expert identification in community question answering: exploring question selection bias", *Acm International Conference on Information and Knowledge Management*.

Patil, S. and Lee, K. (2016), "Detecting experts on quora: by their activity, quality of answers, linguistic characteristics and temporal behaviors", *Social Network Analysis and Mining*, Vol. 6 No. 1, pp. 1-11.

Pedro, J.S. and Karatzoglou, A. (2014), "Question recommendation for collaborative question answering systems with Rankslda",

Petkova, D. and Croft, W.B. (2006), "Hierarchical language models for expert finding in enterprise corpora", *IEEE International Conference on Tools with Artificial Intelligence*.

Ponzanelli, L., Mocci, A., Bacchelli, A., *et al.* (2014),."Improving low quality stack overflow post detection", *IEEE International Conference on Software Maintenance and Evolution*.

Procaci, T.B., Nunes, B.P., Nurmikko-Fuller, T., *et al.* (2016), "Finding topical experts in question and answer communities", *IEEE International Conference on Advanced Learning Technologies*.

Pudipeddi, J.S., Akoglu, L. and Tong, H. (2014), "User churn in focused question answering sites: characterizations and prediction", *Sheridanprinting Com*, pp. 469-474.

Qu, M., Qiu, G., He, X., *et al.* (2009), "Probabilistic question recommendation for question answering communities", *International Conference on World Wide Web*.

Ravi, S., Pang, B., Rastogi, V., *et al.* (2014), "Great question! Question quality in community q&a", *Eighth International AAAI Conference on Weblogs and Social Media*.

Riahi, F. Zolaktaf, Z. Shafiei, M. *et al.* (2012), "Finding expert users in community question answering".

Rostami, P. and Neshati, M. (2019), "T-shaped grouping: expert finding models to agile software teams retrieval", *Expert Systems with Applications*, Vol. 118, pp. 231-245.

Rybak, J., Balog, K. and Nørvåg, K. (2014), "Temporal expertise profiling", *European Conference on Information Retrieval*.

Souza, C., Magalhães, J., Costa, E., *et al.* (2013), "Social query: a query routing system for twitter", *Proc. 8th International Conference on Internet and Web Applications and Services (ICIW)*, pp. 147-153.

Srba, I. and Bielikova, M. (2016), "Why is stack overflow failing? Preserving sustainability in community question answering", *IEEE Software*, Vol. 33 No. 4, pp. 80-89.

Srba, I., Grznar, M. and Bielikova, M. (2015), "Utilizing non-qa data to improve questions routing for users with low qa activity in cqa", *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.

Toba, H., Ming, Z.Y., Adriani, M., *et al.* (2014), "Discovering high quality answers in community question answering archives using a hierarchy of classifiers", *Information Sciences*, Vol. 261 No. 5, pp. 101-115.

Tomasoni, M. and Huang, M. (2010), "Metadata-aware measures for answer summarization in community question answering", *Meeting of the Association for Computational Linguistics*.

Tong, Y., Lei, C., Zhou, Z., *et al.* (2018), "Slade: a smart large-scale task decomposer in crowdsourcing", *IEEE Transactions on Knowledge and Data Engineering*, Vol. PP No. 99, pp. 1-1.

Wei, C.P. and Chiu, I.T. (2002), "Turning telecommunications call details to churn prediction: a data mining approach", *Expert Systems with Applications*, Vol. 23 No. 2, pp. 103-112.

Wei, W., Ming, Z.Y., Nie, L., *et al.* (2016), "Exploring heterogeneous features for query-focused summarization of categorized community answers", *Information Sciences*, Vol. 330 No. C, pp. 403-423.

Wu, H. Wang, Y. and Cheng, X. (2008), "Incremental probabilistic latent semantic analysis for automatic question recommendation",

Xiang, C., Zhu, S., Su, S., *et al.* (2017), "A multi-objective optimization approach for question routing in community question answering sevices", *IEEE Transactions on Knowledge and Data Engineering*, Vol. PP No. 99, pp. 1-1.

Xiong, X., Min, F., Min, Z., *et al.* (2018), "Visual potential expert prediction in question and answering communities ☆", *Journal of Visual Languages and Computing*,

Yang, B. and Manandhar, S. (2014), "Exploring user expertise and descriptive ability in community question answering", *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.

Yeniterzi, R. and Callan, J. (2015), "Moving from static to dynamic modeling of expertise for question routing in cqa sites".

Yi, F. and Godavarthy, A. (2014), "Modeling the dynamics of personal expertise", *International Acm Sigir Conference on Research and Development in Information Retrieval*.

Yuan, Y., Tong, H., Tao, X., *et al.* (2015), "Detecting high-quality posts in community question answering sites", *Information Sciences*, Vol. 302 No. C, pp. 70-82.

Zhang, J., Ackerman, M.S. and Adamic, L. (2007), "Expertise networks in online communities: Structure and algorithms", *International Conference on World Wide Web*.

Zheng, X., Hu, Z., Xu, A., *et al.* (2012), "Algorithm for recommending answer providers in community-based question answering", *Journal of Information Science*, Vol. 38 No. 1, pp. 3-14.

Zhou, G., Lai, S., Kang, L., *et al.* (2012), "Topic-sensitive probabilistic model for expert finding in question answer communities", *Acm International Conference on Information and Knowledge Management*.

Zhou, G., Zhao, J., He, T., *et al.* (2014), "An empirical study of topic-sensitive probabilistic model for expert finding in question answer communities", *Knowledge-Based Systems*, Vol. 66 No. 9, pp. 136-145.

Zhou, Z., Yang, Q., Deng, C., *et al.* (2016), "Expert finding for community-based question answering via ranking metric network learning", *International Joint Conference on Artificial Intelligence*.

Zhou, Z., Zhang, L., He, X., *et al.* (2015), "Expert finding for question answering via graph regularized matrix completion", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 27 No. 4, pp. 993-1004.

Zhu, H., Cao, H., Hui, X., *et al.* (2011), "Towards expert finding by leveraging relevant categories in authority ranking", *Acm International Conference on Information and Knowledge Management*.

Zuhair, A.T.M., Seifedine, K. and Isiaka, O.A. (2018), "Understanding expert finding systems: Domains and techniques", *Social Network Analysis and Mining*, Vol. 8 No. 1, p. 57.

Corresponding author
Zhengfa Yang can be contacted at: yzf_cufe@163.com