



Finding experts in online forums for enhancing knowledge sharing and accessibility



Chih-Ping Wei^{a,*}, Wen-Ben Lin^b, Hung-Chen Chen^d, Wen-Yu An^c, Wei-Chang Yeh^b

^a Department of Information Management, College of Management, National Taiwan University, Taipei, Taiwan, ROC

^b Department of Industrial Engineering and Engineering Management, College of Engineering, National Tsing Hua University, Hsinchu, Taiwan, ROC

^c Systems Engineering Department, VMware, Inc., Taipei, Taiwan, ROC

^d Emerging Wireless Technology Research Department, Information and Communications Research Laboratories, Industrial Technology Research Institute (ITRI), Hsinchu, Taiwan, ROC

ARTICLE INFO

Article history:

Available online 27 May 2015

Keywords:

Expert identification
Data mining
Graph mining
PageRank
Online forum
Knowledge sharing

ABSTRACT

Online forums have been extensively used in many organizational knowledge management practices as well as virtual communities for sharing knowledge and opinions. Identifying experts in certain domains is essential for improving knowledge sharing and accessibility through online forums. Existing expert identification techniques can broadly be classified into two major approaches: *content-based* and *link-based*. Although the link-based approach has shown its superiority over the content-based approach, it incurs some limitations when applying to the task of identifying experts in online forums. In this study, we propose an expert identification technique that relies on the opinion ratings from the members in an online forum. Specifically, we extend PageRank and propose the ExpRank algorithm, which considers both positive and negative agreement relations among the members of the online forum. Using two datasets (pertaining to different product categories, books and music) collected from a well-known product-review website (i.e., Epinions.com), our empirical evaluation results show that our proposed ExpRank algorithm outperforms its benchmark technique (i.e., PageRank). Our evaluation results also highlight that the incorporation of negative agreement relations can improve the effectiveness of expert identification.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

To compete in emerging knowledge-centric economy, organizations around the globe have undertaken various initiatives that aim to manage their most valuable yet volatile asset: knowledge. Knowledge management (KM) denotes the systematic process for creating, retaining, organizing, sharing, reusing, and assimilating tacit and/or explicit knowledge in order to support the learning processes within organizations, thereby leading to improved organizational performance and adaptability (Curado & Bontis, 2011; Davenport, DeLong, & Beers, 1998; Davenport & Prusak, 1998; Wei, Cheng, & Pai, 2006; Wei, Hu, & Chen, 2002). Knowledge management initiatives fall into four broad types: creating knowledge repositories, improving knowledge access, enhancing knowledge environment, and managing knowledge as an asset (Davenport & Prusak, 1998). Among them, knowledge access improvement focuses on facilitating knowledge sharing among individuals, specifically from knowledgeable individuals to others, because knowledgeable individuals often can answer questions, point to

definitive sources or specialists, as well as perform needed functions requiring special knowledge, skills, and experiences (Maybury, 2006; Wang, Jiao, Abrahams, Fan, & Zhang, 2013; Zhang & Ackerman, 2005). However, finding individuals with the knowledge or expertise (i.e., experts) for a specific need is often a difficult task (Davenport et al., 1998). To enhance the accessibility of knowledge, expert identification systems (also called expert finding or expert recommender systems) that can automatically identify experts for a given domain are essential for knowledge access improvement initiatives (Balog, Azzopardi, & de Rijke, 2009; Jackson & Tedmori, 2004; Wang et al., 2013).

Online forums have been extensively used in many organizational KM practices as well as virtual communities for sharing knowledge and opinions. A member of an online forum can share his or her opinions as posts in the forum. In some online forums, members can reply to or comment on the opinions (i.e., posts) contributed by other members or even can give positive or negative ratings (e.g., helpful or not helpful) to other members' posts. Thus, in addition to the posts contributed by members, the interactions among members (i.e., replies or opinion ratings) also provide valuable information for expert identification tasks (i.e., identifying experts in certain domains) in online forums. Fig. 1 shows an

* Corresponding author. Tel.: +886 2 3366 1181; fax: +886 2 3366 9455.

E-mail address: cpwei@ntu.edu.tw (C.-P. Wei).

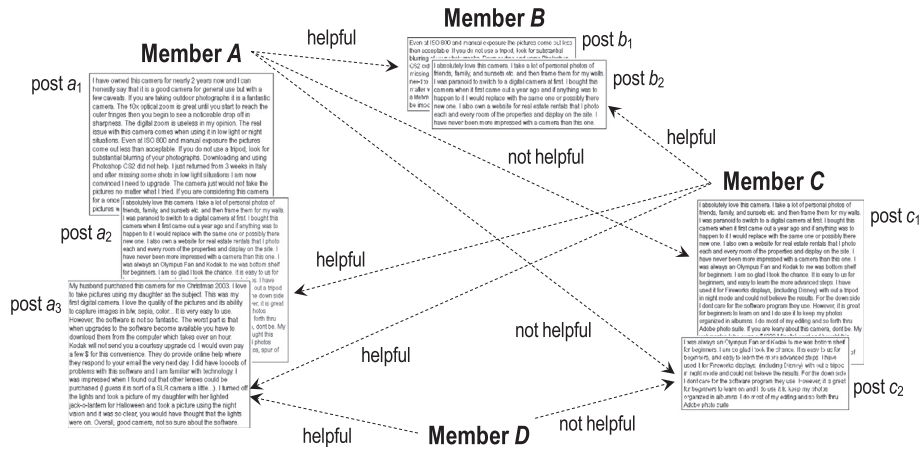


Fig. 1. An illustrative example of members' behavior in an online forum.

illustrative example in which members can give opinion ratings to other members' posts. Member A contributes three posts, a_1 , a_2 , and a_3 , and member B shares her opinions as posts b_1 and b_2 . Furthermore, post b_1 receives a positive opinion rating (i.e., 'helpful' vote) from member A and post b_2 also receives a positive opinion rating ('helpful' vote) from member C. Although member D does not share any opinion in this forum, he gives two opinion ratings to the posts written by other members, i.e., a positive rating ('helpful' vote) to post a_3 contributed by member A and a negative rating ('not helpful' vote) to post c_2 by member C.

Existing expert identification techniques can be classified into two major categories: *content-based* and *link-based*. Some recent studies have also investigated hybrid expert identification techniques, which combine the content-based and link-based approaches (Lin, Xu, Ding, & Liu, 2013; Liu, Chen, Kao, & Wang, 2013; Wang et al., 2013). The content-based approach considers expert identification as an information retrieval (IR) or text mining task (Wang et al., 2013). It typically extracts and builds for each individual an expertise profile automatically from the set of documents related to (e.g., authored by) him or her and then finds experts through a series of matching processes between an expertise query and the previously extracted expertise profiles (Balog et al., 2009; Fu, Xiang, Liu, Zhang, & Ma, 2007; Yang, Chen, Lee, & Ho, 2008). The content-based approach assumes the set of documents related to an individual to be a reliable indication of his or her knowledge and expertise areas. For example, a researcher's academic publications provide strong evidences for his or her expertise in certain areas and thus can be employed to extract his or her expertise profile (Yang et al., 2008). However, this assumption may not be reasonable in some contexts, e.g., online forums. Due to the openness of online forums, the opinions posted by individuals vary in quality. If we put together all posts contributed by an individual and perform the content-based analysis on these posts, the profile extracted for a specific individual may not truly reflect his or her knowledge and expertise areas.

In contrast, the link-based approach focuses on analyzing the linkages among members instead of dissecting the content of documents. For example, Campbell, Maglio, Cozzi, and Dom (2003) used email communications within an organization for finding experts, because they considered people receiving frequent communication requests to be knowledgeable as those individuals are common destinations for questions. Existing link-based expert identification techniques generally use graph-based ranking mechanisms, which often are based on how frequently a person is linked in a network (Campbell et al., 2003; Dom, Eiron, Cozzi, & Zhang, 2003; Jurczyk & Agichtein, 2007a,b; Lu, Quan, Ni, Liu, & Xu,

2009). Although link-based expert identification techniques have advantages over content-based ones (Campbell et al., 2003; Dom et al., 2003), existing link-based expert identification techniques still suffer from some limitations when applied to an online forum setting. For example, links between individuals (or between individuals and posts) in an online forum may not always be positive indications of expertise scores. As Fig. 1 illustrates, member A considers that the two posts contributed by member C (i.e., post c_1 and c_2) are not helpful and thus gives negative opinion ratings to these posts. Likewise, member D also gives a 'not helpful' vote to post c_2 written by member C. Evidently, these negative opinion ratings to the posts contributed by member C indicate that member C may not be an expert in the domain under discussion. Existing link-based expert identification techniques do not model such negative edges in their ranking mechanisms, which will undermine their effectiveness of identifying experts in online forums.

In response to the limitations of existing link-based expert identification techniques, we attempt to propose an expert identification technique based on opinion ratings given by members in online forums. Specifically, we extend the PageRank algorithm (Page, Brin, Motwani, & Winograd, 1998), a graph-based ranking mechanism commonly employed in existing link-based expert identification techniques, to develop an ExpRank algorithm, which considers both the positive and negative opinion ratings commonly observed in online forums. The remainder of this article is organized as follows. In Section 2, we review the existing link-based expert identification techniques and discuss their limitations to justify our research motivation. We then depict the design of our proposed ExpRank algorithm in Section 3. In Section 4, we describe our empirical evaluations and discuss some important empirical evaluation results. Finally, we conclude our study with a summary and some further research directions in Section 5.

2. Literature review

To identify experts, the link-based expert identification approach focuses on analyzing the links among individuals instead of the content of the documents related to (e.g., authored by) each individual. Prior studies adopting or involving this approach utilize one of the following connections or links for expert identification: email communications (Campbell et al., 2003; Dom et al., 2003), question–answering relationships (Jurczyk & Agichtein, 2007a,b; Liu et al., 2013; Lu et al., 2009; Wang et al., 2013), or citation (or co-citation) networks of academic articles (Ding, Yan, Frazho, &

Download English Version:

<https://daneshyari.com/en/article/10312568>

Download Persian Version:

<https://daneshyari.com/article/10312568>

[Daneshyari.com](https://daneshyari.com)