Identifying Knowledge Brokers and Their Role in Enterprise Research through Social Media

Zhe Xu, Jay Ramanathan, and Rajiv Ramnath,
The Ohio State University

Information searches based on expert-seeking technology can prove time-consuming or unsuccessful if search terms do not turn up extrinsic identifiers in profiles and saved documents. In many such cases, knowledge brokers function as "humans in the loop," providing intrinsic enterprise knowledge to mediate between information seekers and expert sources—a fact that future collaborative information-seeking system designs should take into account.

Those seeking information within an organization to meet a task goal must first learn the appropriate repositories of expertise for that information. This "meta-knowledge" includes experts within the organization they might ask, blogs or other electronic resources they might consult, prior projects that might provide models, and so on. These intermediate searches often involve a time-consuming series of enquiries that meet a dead end, with information seekers no closer than they were when they started to having their questions answered. Early stages of information seeking like these have been termed "wandering," particularly when specific information queries have yet to be precisely formulated.[1]

To provide some insight into how such searches might better succeed, we examined a large insurance company's in-house social media message thread postings initiated by employees seeking information to solve a problem. Our analysis confirmed that individuals often failed to find the sources of information they needed. But we also discovered that in a significant fraction of successful searches, early-stage information seeking was assisted by human *knowledge brokers*, or k-brokers, who mediated access to knowledge sources. This finding suggests that collaborative information seeking (CIS) systems could be greatly improved by provisioning for k-broker expert recommendations.

THE ROLE OF KNOWLEDGE BROKERS

Consider the following general scenario. An individual, whom we term a requestor, has a goal to meet but lacks the particular knowledge required to achieve that goal; moreover, the requestor lacks the necessary "social capital"[2]—that is, the requisite network of social relationships—to find appropriate sources for that information. In such cases, a knowledge broker is often instrumental in connecting the requestor to an expert (or expert group) or another source for the needed information. The k-broker mediates by bridging the "structural holes"[3] in the enterprise's social network.

For example, a developer new to an enterprise seeking the best practice for server capacity planning, posts a query to that effect on an in-house social media site. Someone else accessing the site, mediates by responding,

Published by the IEEE Computer Society

"C might be the person to start with. She just completed an assessment of capacity planning within the testing group." Thus, B brokers a collaboration based on intrinsic knowledge both of A's goal and of the social network; B, as k-broker, acts as the "human in the loop"[3] connecting sub-networks.

Previous research has focused primarily on the advantages of this social capital to individual brokers themselves. Our aim here is to characterize the k-broker's value in terms of his or her social capital to others within an enterprise's collaborative infrastructure—and, by extension, to the enterprise's goals. Our experience shows that facilitating the identification and availability of k-brokers can be greatly beneficial. Considerable tacit knowledge of projects, for example, builds up within any enterprise over time and is often distributed across numerous experts. To find a k-broker who can help them bridge the structural holes they face, requestors might have to cast a wide net that extends across several organizations. Helping "wandering" requestors connect with k-brokers can make their information searches more productive and less time-consuming.

MEDIATION IN CIS

CIS studies have encompassed various search environments—for example, libraries, digital archives, and Web browsers and search engines—and proposed various different cognitive and social analytic frameworks to examine how mediation occurs. We use the following funnel framework[1] in order to focus on the early stages of information searching when a k-broker can provide the most help, guiding the requestor's transition to later search stages.

1. *Wandering with general goals.* In this stage, the requestor needs information (for example, about capacity planning within the enterprise) but does not know where to find it. Here, a k-broker can suggest where to access information to help narrow the search (find someone with assessment documents of capacity planning within the testing group).
2. *Exploring with specific goals.* The requestor has a more specific goal (find assessment documents to guide capacity planning), and again the k-broker can lead the requestor to a clearer focus (find out about allocation of virtual machines) and sources for answers.
3. *Seeking with information goals.* The requestor has started to map out information queries (How are virtual machines used?), but these are still open-ended. At this point, k-broker mediation has guided the requestor to appropriate experts or specific search technologies, so that he or she can move on to the next stage.
4. *Asking specific queries.* The requestor has been funneled from a general goal to a specific query (What is the accepted algorithm for allocating applications to virtual resources and manage capacity?)—a query that can solve the initial need for information.

Note that the k-broker intervenes in the early stages (1 and 2) of information seeking. Without the k-broker's help, the requestor has access only to *extrinsic* knowledge—knowledge represented as network nodes in profiles, such as those on LinkedIn, accessible to expert-finding systems (which can answer limited questions like "Who is an expert in what?") and knowledge contained in emails and other historically preserved conversations. But in most cases, a requestor has no way to filter or narrow extrinsic knowledge sources; available search technologies are limited in this regard.

Machines and algorithms do not operate with the kind of intrinsic knowledge often required for successful knowledge mediation.

K-brokers, on the other hand, have *intrinsic* knowledge about sources that cannot be captured in networks or in profiles and preserved documents. They can mediate between requestors and these knowledge sources, guiding requestors to those that are most useful. No technology can achieve this as effectively as "human-in-the-loop" mediation, a form of "human computation" that "solves problems … the computer cannot yet solve."[4] A k-broker does not have to actually solve the problem; rather, he or she helps the requestor find the appropriate avenues for doing so.

In thinking of k-brokers as "humans in the loop," we also include electronic communities and platforms that provide requestors with access to multiple individual brokers—for example, a social media forum on capacity planning. Members of such communities who respond in subsequent conversation threads to problems and questions posted by requestors often provide connections to appropriate experts that requestors could not easily access through current expert-finding technologies.

Machines and algorithms do not operate with the kind of intrinsic knowledge often required for successful knowledge mediation. Moreover, if some particular expertise is not recorded or documented mechanically, expert-finding systems will necessarily fail to discover it.[5]

RELATED WORK

Previous related research has focused on two main areas: social matching and expert-finding technology, and search collaboration.

At a high level, social matching and expert-finding technology research emphasizes optimizing algorithms to meet certain performance criteria, such as communication cost, shortest communication distance, and overall coverage.[6] In addition, social matching researchers are interested in the requestor's question "whom can I connect with?" Methods such as link prediction[6] and expert searching[5] answer this question based on extrinsic knowledge, represented as network edges or in profile-content mining; the result set is consequently too extensive or contains multiple irrelevancies. So, from a broader perspective, research has looked at ontologies and Web-mining techniques for building expert profiles.[5] Our k-brokering approach, however, is concerned less with the "what" of finding appropriate experts than the "how"—that is, gaining an understanding of the microprocesses by which brokers play a connection-making role.

Our k-brokering approach is concerned less with the "what" of finding appropriate experts than the "how"—that is, gaining an understanding of the microprocesses by which brokers play a connection-making role.

Research on collaborative searching is relatively broad-based, beginning with studies of how introducing collaboration to library science and information retrieval makes more effective use of searching systems.[7] Also of interest are a survey of common collaborative Web search activities showing the lack of necessary tool support in current systems[8] and a presentation of algorithmic mediation as a search system feature to let multiple users conduct collaborative searches simultaneously.[9] Specifically related to our concept of k-brokers is activity theory, which uses human action as the unit for analyzing human–computer interaction.[10]

K-BROKERS WITHIN AN ENTERPRISE: A DATA-BASED ANALYSIS

To explore the role of k-brokers within a real-world enterprise, we examined social media microblog postings from a Fortune 500 insurance company from September 2012 to February 2013. Our dataset covers public message threads within the company, with each thread consisting of a conversation initiated when someone posted a message and others subsequently replied.

We used these threads to establish evidence of k-brokers' existence and also to identify the specific type of help they provided to requestors. We defined a thread to be successful if, during the course of the conversation, the requestor posted that he or she located the knowledge needed. Based on this definition, we developed the following new-thread measure:

success rate of a set of threads = number of successful threads/ total number of threads

A goal can be achieved with or without k-brokers, as we shall see. Note, however, that since a k-broker's intrinsic knowledge is not captured in documents, traditional information retrieval measures such as recall and precision do not apply.

Figure 1 provides an overview of the process we used to identify k-brokers and the help they provided to requestors in the dataset.

Narrowing the dataset

The total dataset consisted of 45,133 threads. Of these, we were interested only in threads related to problem solving, so we filtered the threads by looking for those in which the first message contained words such as "problem," "help," "need," or "issue"; this yielded 4,391 threads. To identify threads that were, in fact, related to problem solving, we randomly selected 2,000 of 4,391 to read manually for validation. We then removed threads revealing false positives—phrases like "No problem," for example—yielding 152 threads that could be positively identified as problem-oriented. All our analysis focused on these 152 threads.

Of the 152 problem-oriented threads, we identified 114 as having been successfully solved—that is, the original requestor reported locating the needed information. For the remaining 38 threads, the requestor provided no evidence that a solution was found. (This lack of evidence, however, does not necessarily mean that any of these 38 problems was solved without a k-broker's help. The requestor could have talked directly with a k-broker, thus leaving no traceable messages in the thread.)

We classified the 114 successful threads as "Expert knowledge found" and the 38 other threads as "Expert knowledge missed." We also distinguished three categories of problems with different subtopics:

- Business-related—health policy, bank, claim, agent, identity theft, HR, rapid alignment, sales, retirement;
- IT-related—collaboration tool, android, VPN, Java, Lotus, website, Linux, Excel; and
- Other—photography, event, pets, miscellaneous.

To determine the existence of k-brokers, we first looked for threads that met three criteria:

- an individual responded to the original request for information or help solving a problem;
- the requestor received the information or help requested; and
- this outcome was stated explicitly in the thread.

Next, we ascertained that the individual who helped the requestor was indeed a k-broker as follows. If there was no documentary evidence in social media or enterprise websites that the individual who provided help was

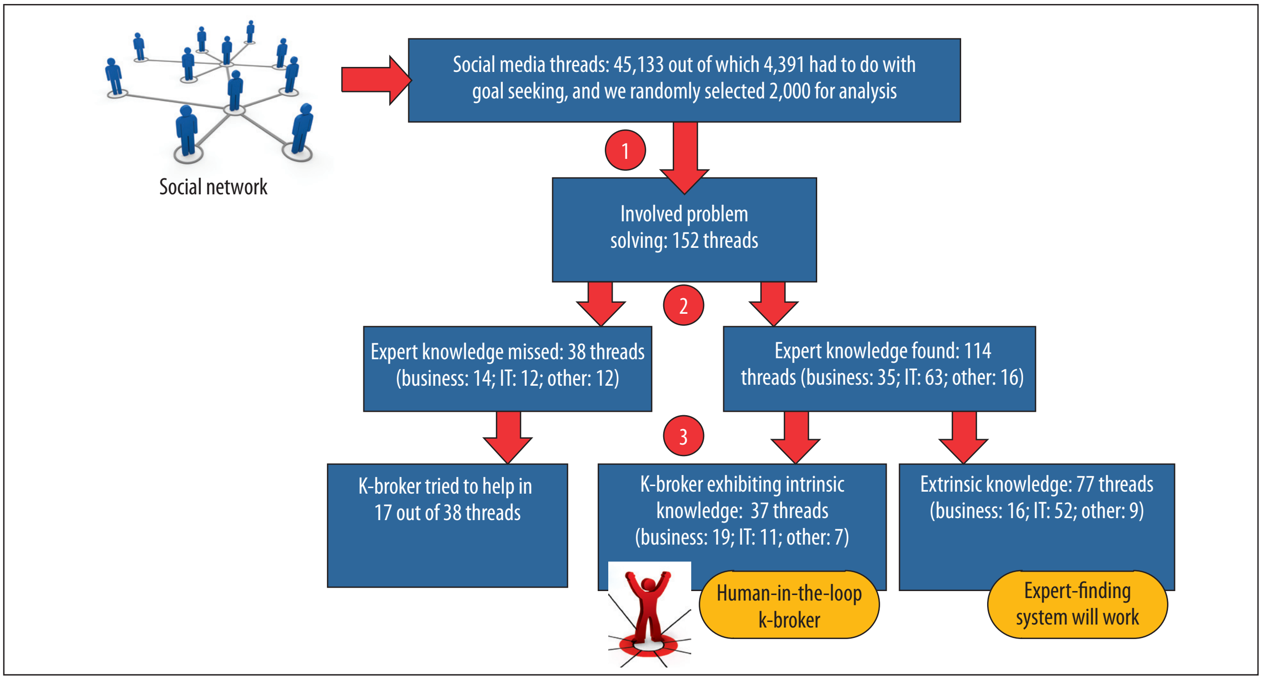

Figure 1. Breakdown of total social media threads from an insurance enterprise between September 2012 and February 2013 for purposes of distinguishing problem-solving threads and analyzing knowledge broker (k-broker) influence on successful information seeking.

associated with the expert knowledge needed to solve the requestor's problem, then we identified that individual as a k-broker—in other words, a "human in the loop" who provided intrinsic knowledge that led the requestor to the information needed. Otherwise, we assumed that that the requestor could have identified or located the information via an expert-finding system—that is, extrinsically.

Identifying k-brokers

More specifically, we tried to answer the following questions: Is there evidence for the existence of k-brokers? Can we identify cases in which an individual helped solve the requestor's problem and no expert-to-knowledge association existed explicitly?

To do so, as already noted, we first looked at threads classified as "Expert knowledge found" due to individuals who exhibited intrinsic knowledge in the thread. We then examined these individuals' enterprise profiles and historical messages accessible within the enterprise. If no explicit keyword for the information requested was associated with them, we posited that there was little likelihood of any expert-finding system locating this particular expert for the requestor. However, if any relevant keyword was associated with the expert, we identified this knowledge as extrinsic and therefore identifiable by an expert-finding system. Only in the former cases did we identify the expert as a k-broker. We made our keyword matching as permissive as possible to obtain the performance upper bound of expert-finding systems.

Among the 114 "Expert knowledge found" threads, we discovered that a k-broker appeared in 37 threads, or 32.5 percent—in other words, for a third of successful threads, success in achieving the requestor's goal appeared to result from k-broker mediation. And even when we looked at the "Expert knowledge missed" threads, in 17 of the 38 total, or 44.7 percent, we found evidence of k-brokers who attempted to help the requestor, although the requestor did not report finding the information sought.

Of the threads in which a k-broker played a role in solving the requestor's problem, it seems clear that human computation provided help which expert-finding technology could not have discovered. In fact, our analysis indicates that an ideal expert-finding system would achieve a success rate of only 67.5 percent with this dataset. This suggests a place for a "beneficial recommender" feature within any larger CIS infrastructure to introduce k-brokers to information seekers.

K-broker benefits

In addition to determining the existence and role of k-brokers, we analyzed the data to assess their usefulness in the search process: Did requestors find knowledge sources more effectively depending on the appropriateness of the k-broker group from which they sought help?

Requestors in the threads we analyzed had a choice of specific discussion groups within the enterprise social media for posting their initial information request. In addition to a general group called "All Company," where each

Table 1. Success rates for information requests posted to correct and incorrect k-broker groups.

		Business	IT	Other	Total
Expert knowledge found	Correct group	26	52	14	92
	Incorrect group	9	11	2	22
Expert knowledge missed	Correct group	6	4	8	18
	Incorrect group	8	8	4	20
	Total	49	75	28	152

employee can both post and view messages, the in-house social media site also includes groups that target a subset of the company—for example, the "Insurance Policy" subgroup for those interested in policy discussions. Individuals do not have to join a particular group to view its threads, but without joining will not get threads from that group automatically sent to their messages boxes.

We found that posting a question to an appropriate k-broker group significantly increased the requestor's likelihood of finding the knowledge being sought. We examined all 152 problem-solving threads to determine those that were and those that were not originally posted to a subgroup clearly related to the requested information. Our criterion was whether the initial request's content matched the description and historical threads of the group to which it was posted. For example, we considered a request for information about Java programming posted to the group titled "Java programming" as belonging in that k-broker group. Table 1 shows the results of our analysis.

We found that requests posted to the correct k-broker group had an 83.6 percent (92/(92+18)) probability of finding expert knowledge—much higher than the 52.4 percent (22/(20+22)) success rate of threads posted to the "All Company" group or to an incorrect group, where apparently no brokering occurred. From this we conclude that access to k-brokers improves outcomes for seeking expert knowledge.

For successful information requests, the difference in the number posted to correct groups (92) and incorrect groups (22) was significant; for the 38 unsuccessful requests, the difference between postings to correct (18) and incorrect (20) groups was much smaller.

Finally, it is worth noting that brokers occasionally served an additional function: identifying messages posted to an incorrect group and directing the requestor to a more appropriate group.

Clearly, enterprises can benefit from systems that connect information seekers with k-brokers who have the intrinsic knowledge to provide guidance in the quest for information. Of the nearly 300 problem-oriented threads we analyzed, mediation by k-brokers occurred in 32.5 percent of requests that resulted in a problem being solved. Understanding that k-brokers exist and the function they play in CIS offers a step toward designing search schemas that identify and involve a "human in the loop."

We note, however, that of the 2,000 threads we examined closely, only 152, or 7.6 percent, related to problem solving. This small percentage could be due in part to the fact that social media as a venue for problem solving is still in its infancy. Thus, increasing the beneficial impact of k-brokers requires more widespread use of social media for this purpose—for example, by expanding the role of social media in task-oriented information seeking and integrating k-brokers with social media through game-type interfaces that provide situational access to knowledge sources.[11]

Other potential interdisciplinary research goals include obtaining a better understanding of k-brokers and their role to incorporate social networks and related methods into a CIS infrastructure; conducting organizational research to find ways to engage more experts to become active k-brokers on social media; and designing search features to direct requestors efficiently to the right k-broker groups.

Ultimately, what is required is an integration of social networking technology and CIS infrastructures that facilitates connecting k-brokers with requestors in their search for information.

Acknowledgments

This research is supported by the National Science Foundation under grant No. 0753710 and the CERCS IUCRC Center for Enterprise Transformation and Innovation (CETI). We also acknowledge sponsors and members of the Center.

References

1. D. Rose, "The Information Seeking Tunnel," *Nat'l Science Foundation Workshop Report on Information Seeking Support Systems*, 2009, pp. 32–35; http://ils.unc.edu/ISSS/ISSS_final_report.pdf.
2. R.D. Putnam, *Democracies in Flux: The Evolution of Social Capital in Contemporary Society*, Oxford Univ. Press, 2002.

3. R.S. Burt, *Structural Holes: The Social Structure of Competition*, Harvard Univ. Press, 1992.
4. A.J. Quinn and B.B. Bederson, “Human Computation: A Survey and Taxonomy of a Growing Field,” *Proc. 2011 SIGCHI Conf. Human Factors in Computing Systems* (CHI 11), 2011, pp. 1403–1412.
5. I. Becerra-Fernandez, “Searching for Experts on the Web: A Review of Contemporary Expertise Locator Systems,” *ACM Trans. Internet Technology*, vol. 6, no. 4, 2006, pp. 333–355.
6. J. Zhang and M.S. Ackerman, “Searching for Expertise in Social Networks: A Simulation of Potential Strategies,” *Proc. 2005 Int’l ACM SIGROUP Conf. Supporting Group Work* (GROUP 05), 2005, pp. 71–80.
7. M.B. Twidale and D.M. Nichols, “Collaborative Browsing and Visualisation of the Search Process,” *Proc. Assoc. Information Management*, vol. 48, nos. 7–8, 1996, pp. 177-182; http://eprints.lancs.ac.uk/53455/1/twidale_aslib_96.pdf.
8. M.R. Morris, “A Survey of Collaborative Web Search Practices,” *Proc. 2008 SIGCHI Conf. Human Factors in Computing Systems* (CHI 08), 2008, pp. 1657–1660.
9. J. Pickens et al., “Algorithmic Mediation for Collaborative Exploratory Information Retrieval,” *Proc. 31st Int’l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR 08), 2008, pp. 315–322.
10. B.A. Nardi, ed., *Context and Consciousness: Activity Theory and Human-Computer Interaction*, MIT Press, 1995.
11. J. Ramanathan et al., “Sense-Respond Cloud Mediator Architecture for Services Evolution,” *Proc. 2011 ACM Symp. Applied Computing* (SAC 11), 2011, pp. 162–169.

Zhe Xu *is a PhD candidate in the Department of Computer Science and Engineering at the Ohio State University. His research interests include text mining of social media and machine learning algorithms in fields such as financial risk prediction, sentiment analysis, and expert search. Contact him at xu.246@buckeyemail.osu.edu.*

Jay Ramanathan *is an associate research professor in the Department of Computer Science and Engineering at the Ohio State University. She is also director of the CERCS for Enterprise Transformation and Innovation (CETI), a consortium engaged in industry-focused research, practice, and education. Her research interests include enterprise solutions using workflow and middleware technology. Ramanathan received a PhD in computer science from Rice University. She is a member of IEEE and ACM. Contact her at ramanathan.2@osu.edu.*

Rajiv Ramnath *teaches software development at the Ohio State University and is director of practice with the CERCS for Enterprise Transformation and Innovation (CETI). His research interests include developing industry-facing programs for applied R&D, classroom and professional education, and technology transfer; wireless sensor networking; and pervasive computing for business-IT alignment, enterprise architecture, e-government, and work-management systems. Ramnath received a PhD in computer science from the Ohio State University. Contact him at ramnath@cse.osu-state.edu.*