Chapter 14

NETEXPERT: AGENT-BASED EXPERTISE LOCATION BY MEANS OF SOCIAL AND KNOWLEDGE NETWORKS.

Ramon Sangüesa, Josep M. Pujol
Software Department
Universitat Politècnica de Catalunya (UPC)
Barcelona, SPAIN
sanguesa@lsi.upc.es, jmpujol@lsi.upc.es

Abstract Locating expertise sources in a community of interest or practice is a critical need for distributed organizations. Usually expertise location is done through the use of personal social or knowledge networks and involves aspects such as trust and reputation. However, and especially in distributed organizations relying on communication technologies for cooperation, each member of a community is just aware of its own personal social or knowledge network. This makes difficult to get to know other potential experts in the community which may pertain to other members' networks. *NetExpert* is an agent-based expertise location system that replicates the process of social and knowledge network building at community or organization level. In so doing it is able to connect several networks and put into contact expertise that otherwise would remain hidden.

Keywords: Software agents, multiagent systems, assistant agents, expertise location, social networks, knowledge networks, collaborative systems, knowledge mapping, knowledge management.

1. INTRODUCTION

Knowledge Management recognizes the importance of locating knowledge as it originates and evolves within a community of people inside an organization with a common set of goals. Any KM process has to incorporate some aspects of leveraging existing knowledge, sharing it and distributing it to the relevant people within the organization. This latter aspect also makes clear

that there is a special type of knowledge, i.e., the knowledge about *who knows what* and *who knows who knows what* that is important for at least two reasons. Firstly, because it helps deciding who may be interested in new knowledge generated in the community. Secondly, and even more important, because it is necessary to decide which people may cooperate in a given area. Current IT systems for KM support (see (Borghoff and Pareschi, 1998) for an extensive survey and assessment) often are more related to other aspects of KM such as finding good representations of knowledge or devising searching mechanisms for locating it. We address the problem of finding experts as well as the paths to reach them. We devised an agent-based system that works on behalf of each expert in the community and builds and maintains his network of close experts and recover the whole network of a community. It provides an effective means for mapping and locating experts within a community.

Experts are recognized as such thanks to their activity within a community that, in response to their performance, grant them the "expert" level. In order to locate an expert we need to combine two operations: finding him, and knowing how to reach him. The problem can be broken down, then, into two parts:

- The problem of finding who has the necessary knowledge about the central topic of the problem (Contractor et al., 1998) amounts to answering the question *Who knows what?*
- The problem of knowing how we can succeed in reaching the expert. Knowing i.e. *Who knows who knows what?*

Both problems have to do with the issue of trusting the result of a *finding* operation. Experts appearing in the results of a *finding* operation should be recognized as such by the community. Trust and reputation are established through different and subtle mechanisms (Wenger, 1999). However, one simple way to establish trust about a previously unknown expert is to rely on his or her relationship with experts that are already known to us and whose expertise has been tested in previous situations. The problem of trust can be reduced to the problem of finding relationships with experts known to the person who is in need of an expert. This relationship can be induced by peer-to-peer assessment or through an analysis of the different personal networks of each member of the community. Two types of networks in a community are relevant for this:

- **The Knowledge Network:** represents people who share knowledge. If two people that are connected it means that they have similar knowledge about a given set of topics. These networks have been used, for example, in *IKNOW* (Contractor et al., 1998).
- **The Social Network:** relates people who know each other. Two people are connected if they are acquainted. These networks have been used, for example, in *ReferralWeb* (Kautz et al., 1997).

Any person has only a partial and biased awareness of the social structures around him or her. He or she only knows his or her own Knowledge and Social Networks. People relating to other members of a community by means of computer systems have a reduced knowledge of other people's networks. Conversely, the number of people that relate through new communication technologies is even greater than in previous cooperative situations. Both aspects diminish the perception of other social and cognitive networks. In the following sections we describe *NetExpert*, an agent-based system for locating expertise within a community which relies on the dynamic construction and management of such networks.

2. A MULTIAGENT SYSTEM FOR EXPERTISE LOCATION

NetExpert is a multiagent system that locates experts in a virtual community. It was originally devised for a community formed by researchers. Its use is now being extended to other communities. *NetExpert* discovers which members in a community are experts and how they can be reached through chains of intermediate experts. It builds and exploits social and knowledge networks. We describe *NetExpert* from the outside in, starting from the view a typical user has of it.

2.1 Functionalities

NetExpert offers the following functionalities.

1 **Searching:** *NetExpert* searches an expert using keywords. There are four possible types of search by: name, knowledge profile, web, resources.

2 **Communication:** Once an expert is located it can be contacted by: videoconference, mail, chat or a WAP phone call.

3 **Exploring Knowledge and Social Networks:** *NetExpert* allows exploring the Knowledge and Social Networks existing around any expert by showing which people share knowledge or are acquainted with the selected expert.

4 **Finding the optimal path to an expert:** The optimal path to an expert is the best *referral chain* (Kautz et al., 1997) that goes from one member of the community to another through their shared Social Network. An expert *A*, would like to meet *C* but *A* does not know who *C* is. If *A* uses the Social Network, he can discover the acquaintances of *C*. Let us suppose that one of these acquaintances of *C* is *B* who happens to know *A* and *C*. *A* could use his acquaintance to meet *C*. The "shortest"

path is calculated by using a heuristic that takes into account the number of edges in the graph between two people and the strength of their relationship.

3. INFORMATION SOURCES FOR EXPERTISE LOCATION

NetExpert uses a lot of information to generate its knowledge including knowledge generated by the rest of a larger system, the *Collaboratory* (Vázquez et al., 2001; Sangüesa et al., 2000). We describe here this possible interaction, although *NetExpert* can also work in independent fashion. Users are expected to contribute and consult documents, URLs, and messages. What the *Collaboratory* offers them is a recommender system that issues recommendations on new contributions to the common document base. Users can rate these recommendations following a *collaborative filtering* schema (Resnick and Varian, 1997; Shardanand and Maes, 1995; Terveen et al., 1997). Information that has some relationship to a user's competence and interests can be extracted from the contents of these documents and from document ratings as well as other sources.

3.1 Complementary sources of information

Another way for the system to represent the knowledge and the interests of a member is through the analysis of his or her personal webpage. In personal pages, information that is not representative of the real knowledge of the author could exist. This could introduce noise in the process of building the corresponding knowledge model. However, the knowledge model extracted from a personal webpage is crude but also wide. This is an excellent quality in order to search for an expert afterwards. The actual *Knowledge profile* built by *NetExpert* contains a description of his expertise areas described in terms of the most important terms appearing in the documents contributed by the user. These documents are treated as if they had received a positive vote from the user. A contribution is an implicit positive vote. The creation and maintenance of this profile is done automatically. A *knowledge model* is built for each page or resource, using an elaboration of TFIDF (Salton and McGill, 1983). Personal webpages usually reflect the knowledge and interests of its authors. Experts tend to include in their personal pages publications and other sources directly related to the knowledge they generate. To analyze webpages *NetExpert* uses *WebMining* (Pujol, 2000), an independent system that analyzes the web and creates generic content models.

3.2 Generating knowledge for expertise location

Social and knowledge networks can be built from the profiles described previously. In order to build this network we have devised some heuristics that assess the degree of acquaintance between two members of a community. These heuristics use the following data:

- **Documents published within the community:** All authors of a published document are assumed to know each other. The more quantity of published documents in common they have, the stronger the relationship between their authors is.
- **Relations among personal webpages:** Personal webpages are a source of useful information in order to know more about their author and his relation with other members of the community. In a personal webpage we can find references such as email addresses of other members of the community, links towards the main page of personal webs, links towards resources, that are included in other personal webpages. In this last case, the depth of the resource referenced by both personal homepages is taken into account to fine tune the degree of mutual relationship.

4. INTERNAL ARCHITECTURE

NetExpert is built as a distributed multiagent architecture. Each agent and each non-agent software component could be located in heterogeneous platforms. Figure 14.1 shows the internal architecture of *NetExpert*. Two kinds of users can exist in *NetExpert*.

- **Human users:** they acces *NetExpert*, either via WWW or via WAP pages. Both use lightweight clients, a web browser or a mobile phone.
- **Agent users:** artificial agents roaming the web that can obtain knowledge and information from *NetExpert* by interacting with the Graph Agent. This is the basis for expanding the expertise location utilities across several communities.

Grey components in the diagram architecture are agents. Here is the description of each one of them.

- **Interface Agent:** it is an adaptor between UI and Graph Agent. Each user within the system uses a UI, and each UI needs an Interface Agent to communicate with the Graph Agent. Each simultaneous user of *NetExpert* has a UI and Interface Agent that are in the client side.
- **Graph Agent:** it is the one that encapsulates all *NetExpert*'s knowledge and information. To obtain information from the system a request must be addressed to the Graph Agent.

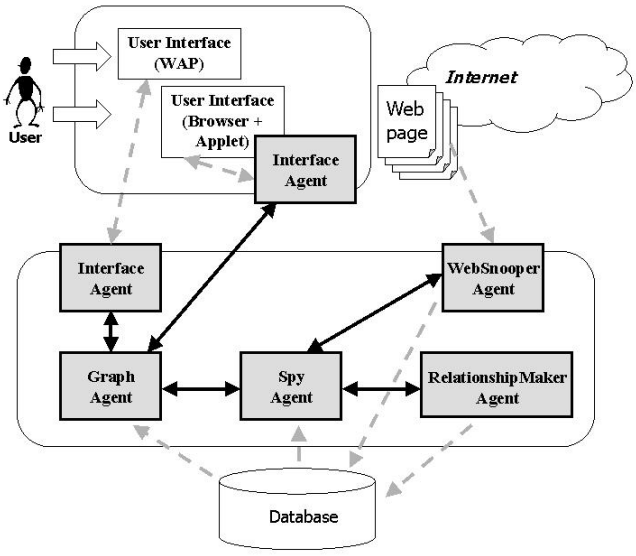

Figure 14.1. Internal architecture of NetExpert

- **Spy Agent:** it is in charge of monitoring any change in the Community repository.
- **RelationshipMaker Agent:** it builds and keeps the relationships among community members.
- **WebSnooper Agent:** This agent analyzes the web pages of community members. It uses the *WebMining* system (Pujol, 2000).

Agents have been developed with JATLite (JATLite, 2001) with FIPA ACL (FIPA, 2001) as a communication language.

5. EXPERIMENTS

NetExpert has been tested with a user group formed by the faculty of the Software Department at *Universitat Politècnica de Catalunya* (UPC). This test only intended to assess the construction of Social Networks as Knowledge networks needed information created in the *Collaboratory* and at the moment of testing not enough people had contributed documents to ensure a significant sample. The UPC Software Department has 170 members, 104 of whom have a personal webpage. The sample was then 104 individuals.

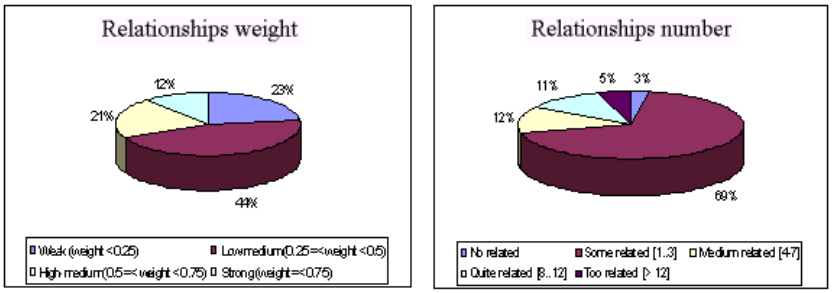

Figure 14.2. Relationships weight and per individual

NetExpert found 190 relationships among the 104 members of the Software Department. Relationships between people have a weight between 0 and 1, [0..1]. On the left side of figure 14.2 the weight of the relationships can be seen. Most of them are low-medium ones, between 0.25 and 0.50. The percentages of weak relationships and high-medium relationships are almost the same. The right side of Figure 14.2 shows the number of relationships per individual. Most people, 69 percent, are classified as *somewhat related*. They have between 1 and 3 relationships with other members of the Department. Only 3 percent of the individuals have no relation at all, and 5 percent of individuals can be considered to have many relationships. *NetExpert* found relationships that correspond to real relationships among 92 percent of members of Software Department. Information about published documents in the community was not used in this test.

Several personal interviews were carried out with users in this community and in the wider community using the *Collaboratory* in order to assess the validity of the recovered networks. In general, people agreed on the structure of their personal networks and their relationship to other close members' networks. Further work is on the way to quantify the degree of correctness of the recovered networks and the ranking of experts. Initial results seem to point that the information used by *NetExpert* is a good starting basis to map a community knowledge as well as stablishing its users level of competence (Pujol and Sangüesa, 2001).

The Software Department network is a small world (Watts and Strogatz, 1998) as it has been found when analysing other virtual communities (Adamic, 1999). We are exploring the implications of this for the improvement of Social and Knowledge Network building and searching.

6. DISCUSSION AND FURTHER RESEARCH

NetExpert is an agent-based expertise location system. It exploits social and knowledge networks to facilitate finding experts within a community. The difference between *NetExpert* and systems like *IKNOW* (Contractor et al., 1998) or *ReferralWeb* (Kautz et al., 1997) is that it merges and uses each type of network that these systems use (in the first case *only* social networks and in the second one *only* knowledge networks). *IKNOW* and *ReferralWeb* resort to webpages to create profiles but *NetExpert* uses also documents that each expert publishes and consults within the community. Moreover, in analyzing webpages *NetExpert* not only uses the actual contents of pages but also information about links, which represents an additional enhancement with respect to other systems.

Other systems have been devised for expertise location. *ExpertFinder* (Vivacqua, 1999) locates experts in Java Programming) from the code written by each programmer and his contributions to discussion lists. Although the system seems to give good results it is argueable that it could be extended to other domains. Ackerman *Answer Garden* (Ackerman and McDonald, 1996; McDonald and Ackerman, 1998) uses FAQ in to answer user questions. An expertise finding functionality (*Expert Recommender* (McDonald and Ackerman, 2000)) has been built upon this system. In contrast to *NetExpert*, *Expert Recommender* uses a social network that has been built manually after personal interviews. Some authors (Liao et al., 1999) claim that profiles built only with the aid of significant words may be too shallow for describing adequately the expertise of a member of the community. One alternative solution is to have a fully developed ontology describing competences in a community and use it to build and analyze user models. Up to now, these ontologies have to be built manually through a tiresome process. Although going through such a process may result in a more precise search results, the present performance of *NetExpert* seems to indicate that our approach has sufficient retrieval quality. Nevertheless, we are presently investigating ways to create ontologies semi-automatically from the analysis of webpages and documents (Fava, 2001).

Another difficulty of our present approach lies in assessing the level of expertise of each member of the community solely from the current information sources. This involves trust, confidence and reputation issues. In order to assess these aspects some feedback from the people that have used an expert's information may be needed. Ranking experts on the basis of the quality of their responses lies in the basis of some systems, even commercial ones such as *http://www.ask.me*. We are currently trying to connect an expert ranking system based on user feedback (Garcia, 2001) with *NetExpert* in order to complement the current avaluation of the expertise members in the community. With no feedback, an alternative way of measuring quality or reputation is to

use metrics about his position in the network and of the network's topology. We have performed some experiments using as a test group a research community and compared the obtained rankings against a publication impact index, which is a recognized way of establishing reputation and expertise within the scientific community and requieres no feedback. The results show a good correlation between *NodeRanking*, which is the ranking method used by *NetExpert*'s, and the established scientific publication ranking. For details about techniques for reputation extraction using social network topology see (Pujol and Sangüesa, 2001).

The fact that *NetExpert* has been implemented as a multiagent system implies that reconfigurations and improvements as the ones that have been mentioned in the last paragraph can be done in a stepwise fashion with ease. The current implementation is FIPA-compliant (FIPA, 2001) which means that it can cooperate with other agents developed under this framework. This could allow the federation of several expert location systems serving similar communities.

References

Ackerman, M. S. and McDonald, D. W. (1996). Answer Garden 2: Merging Organizational Memory with collaborative help. In *Computer Supported Cooperative Work*, pages 97–105.

Adamic, L. (1999). The Small World Web. In Abiteboul, S. and Vercoustre, A.-M., editors, *Proc. 3rd European Conf. Research and Advanced Technology for Digital Libraries, ECDL*, number 1696. Springer-Verlag.

Borghoff, U. M. and Pareschi, R. (1998). *Information Technology Support for Knowledge Management*. Springer-Verlag.

Contractor, N., Zink, D., and Chan, M. (1998). IKNOW: A tool to assist and study the creation, maintenance, and dissolution of knowledge networks. In Ishida, T., editor, *Community Computing and Support Systems*, number 1519 in Lecture Notes in Computer Science, pages 201–217, Berlin. Springer-Verlag.

Fava, M. (2001). Construcció automàtica de thesaurus mitjantcant xarxes bayesianes. Master's thesis, Technical University of Catalonia.

FIPA (2001). FIPA. Foundation for Intelligent Physical Agents. Available at http://www.fipa.org/.

Garcia, R. (2001). Extensió col.laborativa del servei de localització d'expertesa. Master's thesis, Technical University of Catalonia.

JATLite (2001). Java Agent Template, Lite. Stanford University. Available at http://java.stanford.edu.

Kautz, H., Selman, B., and Shah, M. (1997). The hidden web. *AI Magazine*, (18).

Liao, M., Hinkelmann, K., Abecker, A., and Sintek, M. (1999). A Competence Knowledge Based System as Part of the Organizational Memory. In on Artificial Intelligence, L. N., editor, *Proceedings of the 5th Biannual German Conference on Knowledge-Based Systems*, number 1570, pages 125–137.

McDonald, D. W. and Ackerman, M. S. (1998). Just Talk To Me: A Field Study of Expertise Location. In *Proc. Conference on Computer-Supported Cooperative Work (CSCW'98), Seattle*, pages 315–324. ACM Press.

McDonald, D. W. and Ackerman, M. S. (2000). Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the ACM 2000 Conference on Computer Supported Cooperative Work, CSCW*, pages 231–240, Philadelphia.

Pujol, J. M. (2000). NetExpert: Localitzador d'expertesa. Master's thesis, Technical University of Catalonia.

Pujol, J. M. and Sangüesa, R. (2001). Reputation Measures for Expertise Location Agents by means of Social Networks. In *Proceedings of the 4th Congrés Català de Intel.ligència Artificial*, pages 205–213.

Resnick, P. and Varian, H. (1997). Recommender systems. *Communications of the ACM*, 40(3):56–58.

Salton, G. and McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. McGraw-Hill.

Sangüesa, R., Vázquez, A., and Vázquez, J. (2000). ACE: a multiagent recommender system using mixed collaborative and cognitive filtering. In *Proceedings of the WARS 2000, Workshop on Agent-based Recommender Systems*.

Shardanand, U. and Maes, P. (1995). Social Information Filtering: Algorithms for automating "word of mouth". In *Proceedings of the Conference in Human Factors in Computer Systems (CHI'95)*, pages 210–217.

Terveen, L., Hill, W., Amento, B., McDonald, D., and Creeter, J. (1997). Phoaks: a System for Sharing Recommendations. *Communications of the ACM*, 40(3):59–62.

Vázquez, A., Vázquez, J., Barrio, I., Pujol, J. M., and Sangüesa, R. (2001). The Collaboratory. Technical Report LSI-01-22-R., Technical University of Catalonia.

Vivacqua, A. S. (1999). Agents for expertise location. In *Proceedings of the 1999 AAAI Spring Symposium on Intelligent Agents in Cyberspace*.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, (393):440–442.

Wenger, E. (1999). *Communities of Practice: Learning, Meaning, Identity*. Cambridge University Press.