Connecting People with Questions to People with Answers

Mehmet H. Göker, Cynthia Thompson, Simo Arajärvi, Kevin Hua

The Connection Machine helps PricewaterhouseCoopers LLP (PwC) partners and staff to solve problems by connecting people to people. It allows information seekers to enter their question in free text, finds knowledgeable colleagues, forwards the question to them, obtains the answer and sends it back to the seeker. In the course of this interaction, the application unobtrusively learns and updates user profiles and thereby increases its routing accuracy. The Connection Machine combines aspects of expertise locators, adaptive case-based recommender systems and question answering applications.

1 The Power of Connected People

Information, knowledge and experience are key success factors and the most important competitive advantage for any business. However, most of this core corporate asset is in the heads of the employees and cannot be easily accessed, shared or distributed. Capturing and protecting it in documents (electronic or otherwise) is not only cumbersome, but the documents become rapidly outdated and the maintenance effort required to keep document collections up-to-date is formidable.

Furthermore, in the complex business scenarios of today's world, problem solving requires an increasingly large amount of specialized knowledge. It is nearly impossible for one individual to be an expert in every aspect of a company's business and deliver comprehensive solutions. Problem solving requires co-operation and the sharing of ideas and information. The size of a corporation and the collective knowledge of its employees are only valuable if these employees can share their information and cooperate. We believe that the best way to provide the most up-to-date and accurate information to those who seek it is by putting them directly in touch with the experts and implemented this idea in the PwC Connection Machine.

2 Directory Systems, Expertise Locators and the Connection Machine

Most firms allow their employees to search for other colleagues by means of directories or expertise locator systems. Directory systems are essentially computerized versions of phone-books and help users to find the contact information of the person they are looking for using the name or departmental information. Expertise locators extend the data contained in the employee directories with work experience and areas of specialization. They sometimes provide the ability to search for experts using free text and typically permit users to select and contact the suggested experts from within the application [1, 2]. Both directory systems and expertise locators return *contact information* of potentially knowledgeable people as a result of a query. It is left to the seeker to contact the experts and obtain the answer.

The Connection Machine extends directory systems and expertise locators beyond the pure search for contact information and helps PwC partners and staff to get *answers* to their questions and to solve problems together. The Connection Machine matches incoming questions to the expertise profiles of users, routes questions to the experts whose profiles are of the highest similarity until one of them answers, collects the answer and relays it back to the seeker. Experts can also refer the question to a more knowledgeable colleague from their personal network. Users can interact with the Connection Machine by email or through its web interface.

By providing answers to questions rather than just locating people, the Connection Machine acts as a virtual, adaptive expertise provider. The application merges technologies from expertise locators (e.g. expertise profiles, routing of questions) with adaptive case-based recommender systems (e.g. user modeling, selection of experts to route the question to) and question answering applications. From the seeker's perspective, the application encapsulates all experts in the firm and provides real answers to real questions.

3 Overview of the Connection Machine

Figure 1 provides a general overview of the interaction between the information Seeker, potential Providers and the

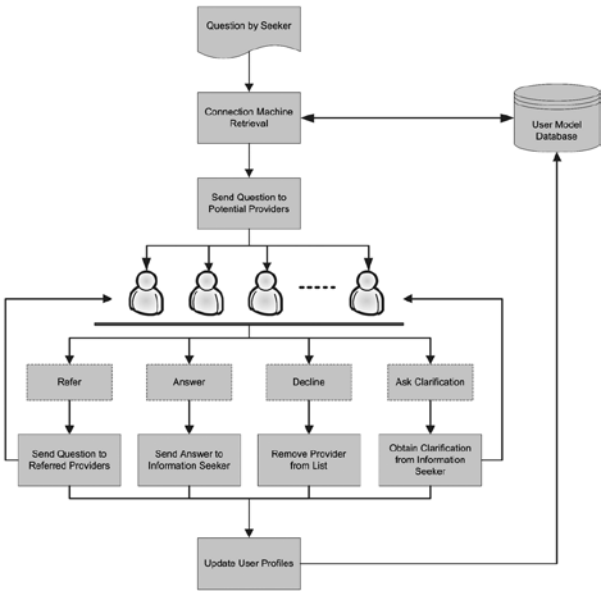

Figure 1: Overview of the application workflow in the PwC Connection Machine

Connection Machine. An interaction with the Connection Machine starts with an Information Seeker entering a question in free text format, as if he/she were asking a colleague a question via email. The Seeker is also able to specify the urgency of the question, the name of a client the question relates to as well as additional, optional, structured information (e.g. knowledge domain, line of service, industry) to be used to locate appropriate potential Providers (Figure 2).

The Connection Machine processes the query, finds a set of matching potential Providers, contacts them via email (Figure 3) and places a visual indicator in the Summary page of the web interface. In addition to the question, the potential Providers are informed of the Seeker's contact information (e.g. name, line of service) and of the timeframe in which the question needs to be answered.

After receiving a question, the potential Providers may choose to respond either via web interface or via email. Potential Providers may

- offer an answer to the question;
- request additional information from the Seeker;
- refer the question to other potential Providers; or
- decline to answer.

Once one of the *potential* Providers offers an answer or requests additional information, he/she becomes the *Provider* for the interaction. From this point on, the Connection Machine facilitates communicates between the information Seeker and the Provider. It notifies and removes other potential Providers from the problem solving conversation.

If a Provider chooses to answer the question, the Seeker is notified via email and a visual indicator in the web application. Upon receiving an answer, a Seeker can choose to accept it and close the request, ask a clarification question about the answer, or reject the answer and request a second opinion.

If a (potential) Provider decides that someone else from his/her personal network is better suited to answer the question, he/she may choose to refer the question. In this way the (virtual) personal network of the seeker is expanded and the Connection Machine can learn about users who may have been missing from its initial set of profiles. If the question is declined by all contacted potential Providers in the first set, it is sent to a second set. If these do not respond either, the question is sent to the Domain Manager of the domain for further processing.

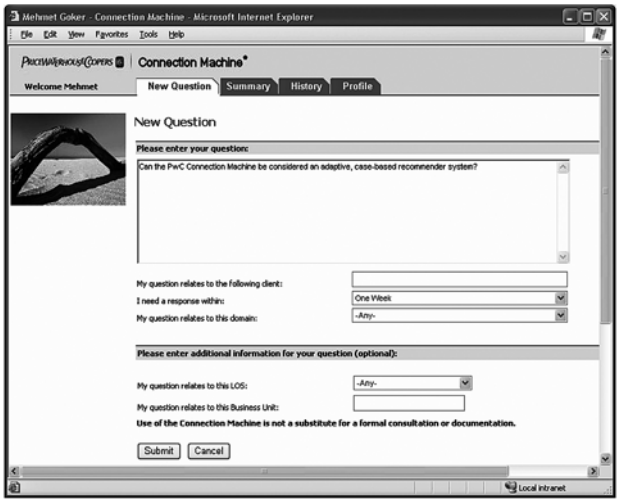

Figure 2: Web interface of the PwC Connection Machine

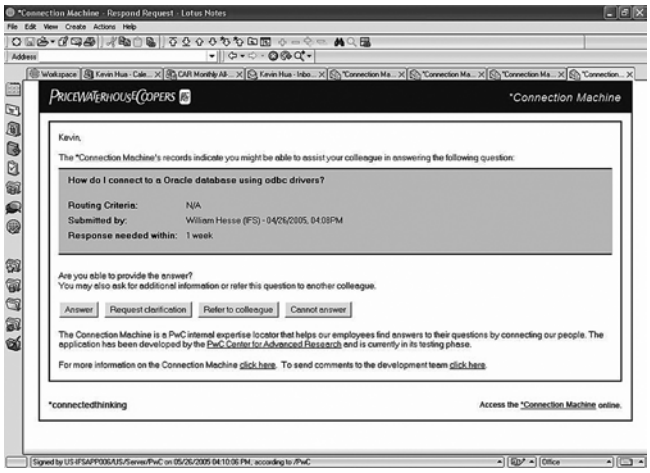

Figure 3: Sample email from the PwC Connection Machine

4 User Modeling and Retrieval of Potential Providers

To execute the workflow described above, the Connection Machine needs to be able to determine who is potentially able to answer the question by matching a seeker's query against information it stores about other users. These user profiles can be initialized from documents the user authored, a resume, prior engagement histories or similar. To achieve consistently high accuracy over a long period of time, the user profiles are continuously updated with appropriate sections of the interaction with the Connection Machine (Figure 4).

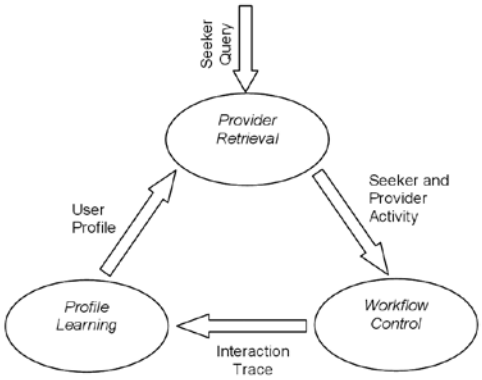

Figure 4: Profile maintenance and usage in the Connection Machine

A user model in the Connection Machine contains three types of user profiles, each of which captures one aspect of a user's preferences or capabilities:

- The *interest profile* denotes the topics a user is interested in. It is updated from the questions a user asks and the associated clarifications.
- The *expertise profile* represents the topics that the user is knowledgeable in. It is updated with the questions a user could answer, the answers he/she provided, and all associated clarification conversations.
- The *referral profiles* represent the topics in which the user is able to refer questions. It is updated with referred questions, any clarification conversation associated with them, and any referral comments.

We also make a distinction between *positive* and *negative* profiles. Negative profiles contain information that the user does not want to be associated with.

Auszug aus: Künstliche Intelligenz, Heft 4/2007, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**

The current version of the Connection Machine does not utilize the interest and referral profiles for routing purposes. These profiles are collected for future applications such as intelligent document routing and the analysis of social networks.

The representation that is used to store the content of the user models has a significant impact on the capabilities, flexibility, maintainability and learning abilities of the system. For details on our approach to user modeling please refer to [6].

Once the system starts being used, the interest, expertise and referral profiles are updated directly from the interactions. The user is also able to manually update and manage his/her profiles by adding relevant documents or keywords. The profile changes caused by a user's interactions with the Connection Machine are visible in the profile section of the application as well and can be removed by users if they should choose the do so.

The technology we used to implement these functions in the Connection Machine is similar to User Adaptive, Case-Based Recommender Systems [3, 4, 5]. It is worthwhile to note that the items in the case-base of the Connection Machine are continuously evolving user models where each model contains multiple profiles. Additionally, rather than being the final goal, the retrieval process is an intermediate step and users, whose expertise profile matched the query, are utilized in the workflow to route questions to. The resulting interaction between the Seeker and Provider is the desired outcome for the application. This is obviously not the case for standard recommender systems where the retrieved items are suggested to the user as potential solutions.

5 Current Status

The Connection Machine has been in pilot with approximately 5000 users from five user groups (domains) within PricewaterhouseCoopers (U.S. firm only) for one year (July 2006 to July 2007). After the successful completion of the pilot phase, the application has been deployed on PwC's knowledge gateway and is currently accessible to all 30,000 US employees of the firm.

To initiate a pilot with a user group, we collected all available and relevant directory information, resumes, and documents that represent user's experience. Once the profiles had been populated with this initial data, we sent an announcement to the members of each group and give them access to the application. Each domain is also assigned a Domain Manager who ensures that the questions flow smoothly.

Between July 2006 and July 2007, the Connection Machine received 838 questions. 244 of these were subsequently withdrawn due to various reasons (e.g. test questions, queries intended for search engines, or questions which were answered through alternative channels). Of the remaining 593 questions, 437 were answered (74%), and 105 were waiting for a response (18%). For 36 questions (6%), the potential provider requested clarification of the question and 15 seekers (2%) had provided such clarification.

At first look, the number of questions and answers in the Connection Machine may seem to be low compared to standard web applications or search engines. However, the interactions in the Connection Machine are of an entirely different quality: they contain concise questions related to a business problem and highly specific answers to these. Considering the drastically different nature of the domain and the business value of answers of this nature, the numbers cannot be compared to utilization numbers of search engines or similar application.

We also noted that, based on their experience with standard search engines, users are accustomed to express their problem in keywords and feel awkward to ask their real question. For our pilot, we had to remind users that they can phrase their question as if they were asking a colleague and to provide all information a colleague might need to know to provide a good answer. With the Connection Machine, users can enter what they actually want to know rather than abstracting their question artificially. We consider support for this ability to be a very important and real gap in knowledge management technology.

The application was able to suggest a potential provider to send the question to 99% of the time. This high number is due to the fact that we always have domain managers in the loop to help in situations where the Connection Machine was unable to locate a provider. 91.3% of the time one of the contacted potential providers started to interact with the seeker to provide an answer and 88% of the answers provided were deemed sufficient and no second opinion was requested.

The seekers have the option to specify the timeframe in which they need to have their questions answered. To allow time for a potential second set of providers and for the domain manager to act, the actual time each provider has to answer a question is one third of the time the seeker states. Table 1 shows the percentage of questions for which the interaction was started within the expected timeframe as well as the percentage of questions answered on time. As one would expect, the more time the experts have, the more likely it is that they will answer.

Response Needed in	4 Hrs	1 day	3 days	1 week
Interaction started by Provider	50.43%	63.04%	76.42%	91.71%
Answered on time	35.04%	40.58%	54.72%	71.50%

Table 1: Response Times in the Connection Machine

6 Future Work and Summary

As next steps, we are planning to utilize the user models of the Connection Machine for tasks such as targeted content distribution to interested parties, routing content to experts for verification, personalization of portals, as well as the creation of communities of interest and expertise. By analyzing interest, expertise and referral profiles for the entire organization, analyses of gaps in knowledge could be performed and areas of concentrated expertise or interest highlighted.

In summary, the PricewaterhouseCoopers Connection Machine allows information seekers to enter their question in free text, finds knowledgeable colleagues, forwards the question to them, obtains the answer and sends it back to the seeker. In the course of this interaction, the Connection Machine unobtrusively updates and refines the interest, expertise and referral profiles of each user. Rather than just locating *people*, it extends the concepts of directory systems and expertise locators and acts as a virtual (adaptive) expertise provider and *answers questions*.

References

1. Ackerman, M., Pipek V., Wulf V. (Eds), "Sharing Expertise: Beyond Knowledge Management", MIT Press, Cambridge, Massachussetts and London, England, 2003
2. American Productivity & Quality Center (APQC), "Expertise Locator Systems: Finding the Answers", Benchmarking Report, APQC Publications, Houston, Texas, 2003 (www.apqc.org)
3. Thompson C., Göker M., Langley P., „A Personalized System for Conversational Recommendations", Journal of Artificial Intelligence Research (JAIR), Volume: 21 (2004), pp 393-428
4. Bridge, D., Göker, M., McGinthy, L., Smyth, B., "Case-based Recommender Systems", The Knowledge Engineering Review, 2006, Cambridge University Press, to appear.
5. Burke, R., "Hybrid Recommender Systems: Survey and Experiments". User Modeling and User-Adapted Interaction. 12(4), pages 331-370.
6. Göker, M., Thompson, C., Hua, K. , Arajärvi, S, „The PwC Connection Machine, An Adaptive Expertise Provider", in T.R. Roth-Berghofer et al. (Eds.): Advances in Case-Based Reasoning, Proceedings of the 8th European Conference on Case-Based Reasoning, ECCBR 2006, Fethiye, Turkey, September 4-7, 2006. LNAI 4106, pp. 549 – 563, Springer-Verlag Berlin Heidelberg 2006

Contact
Mehmet H. Göker, Cynthia Thompson,
Simo Arajärvi, Kevin Hua
PricewaterhouseCoopers LLP
Center for Advanced Research
Ten Almaden Blvd, Suite 1600
San Jose, CA 95113
Email: {mehmet.goker, cynthia.thompson, simo.arajarvi, kevin.k.hua}@us.pwc.com

Simo Arajärvi is a senior Java/JEE developer with 12 years of experience and various certificates from Sun Microsystems and Oracle Corporation. Originally from Helsinki, Finland, Simo studied economics prior to pursuing software. He graduated from Florida Southern College with a BS in CIS, and holds an MBA from the University of Florida. Recently, Simo has concentrated on Service Oriented Architecture and other strategic initiatives within PricewaterhouseCoopers.

Mehmet H. Göker leads the Connection Machine Project at the PricewaterhouseCoopers Center for Advanced Research in San Jose, California. His main research focus is on using Case-Based Reasoning and User-Adaptive technologies to develop intelligent, human centric knowledge management systems that further the state of the art in applied information technology. Mehmet received his Ph.D. in Mechanical Engineering from the Department of Machine Elements and Engineering Design at the Darmstadt Institute of Technology in Darmstadt, Germany. His M.Sc. degree in Aerospace Engineering is from the University of Michigan, and his M.Sc. degree in Computer Engineering from the Bogaziçi University in Istanbul, Turkey.

Kefeng Hua joined the Center for Advanced Research of PricewaterhouseCoopers as Knowledge Mgmt. Systems Developer/Researcher to work on the Connection Machine in 2004. Kefeng received his Ph.D. in Computer Science (Artificial Intelligence) in 1994 from Swiss Federal Institute of Technology at Lausanne, a MS in Computer Science (Artificial Intelligence) from Beijing Institute of Technology and a BS in Computer Science from Hohai University in China.

Cynthia Thompson joined PricewaterhouseCoopers Center for Advanced Research in 2004. She is the project leader for Magic Lens, which expedites the review and comparison of even paper-based documents and previously worked on the Connection Machine Project as a Knowledge Representation researcher and developer. Cindi received her Ph.D. and M.A. in Computer Sciences from the University of Texas at Austin, and her B.S. in Computer Science from North Carolina State University.

Auszug aus: Künstliche Intelligenz, Heft 4/2007, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**