INTEGRATED INFORMATION SEARCH IN THE WWW AND A HUMAN GROUP

Seiji Yamada Motohiro Mase

CISS, IGSSE, Tokyo Institute of Technology
4259 Nagatsuta, Midori, Yokohama 226-8502, JAPAN
E-mail: yamada@ymd.dis.titech.ac.jp

ABSTRACT

We propose a framework for searching information through both the WWW and a human group. We can currently obtain much information by searching the WWW with a search engine as well as by asking persons in our group. Thus we integrate the information search in the WWW with that in a human group, and develop the information search system as multi-agent system. Knowledge profiles for users and search engines are build, and the system selectively searches information resources consisting the WWW and personal knowledge when a user gives a query. As a result, a user acquires integrated results of a hit list from the WWW and a audio/video connection to a person who can answer the query.

1 Introduction

The accessible information through the WWW(World Wide Web)is increasing explosively. In such a situation, the WWW has become a useful information resource because it potentially includes huge knowledge and is updating worldwide. Usually we utilize a search engine for information retrieval in the WWW. We give a search engine a query which is concerned with our target information, and it returns a hit list indicating the relevant Web pages retrieved from its database.

Though the information retrieval using a search engine in the WWW is very useful, we can not acquire local, implicit and specific information like knowledge on configure files of LAN in our laboratory, knowledge in personal memorandum on application install and so on. We call this information *closed information* contrasted with *open information* which is opened for public in the WWW. Most of the closed information is owned by a person and is not explicitly described in text. Thus the best way to access the closed information is to directly ask a person who knows the target information. Furthermore a hit list of a search engine often includes many irrelevant Web pages because adequate filtering to user's intention is very hard. Hence searching closed information through a human group is still significant as well as the information search in the WWW. Since a user knows neither where target information is in the WWW nor who knows the target information in a human group, we need an integrated information system which can search the target information through both the WWW and a human group.

Thus, in this paper, we propose an integrated information search system which can search the target information through both the WWW and a human group. After a user gives a query to a system, it broadcasts the query to information resources, gathers the results and indicates the integrated results to a user. We develop this framework as a multi-agent system consisting of user agents and search engine agents, and utilizes contract net protocol to communication among the agent. Each of the agents maintains a profile describing the properties on the information resource: a person and a search engine. An agent which received a query becomes a manager and other agent become contractors. The manager broadcasts the query as task announcement to contractors, and they reply bid messages including their knowledge profiles. After investigating the bid messages, the manager selects the contractors winning a bid. As a result, the manager actually gives the query to the selected contractors as a task and they return retrieved information from their resources. Finally a system integrates the returned information and indicates it to a user using a hit list and audio/video

applications connected to a person who can answer the query.

Collaborative filtering has been applied to sharing personal knowledge on Web pages, stored in a bookmark file[5][3]. Mori and Yamada proposed Bookmark Agent: a multi-agent system in which an agent constructs a profile from the user's bookmark file by analyzing the html files fetched using the links[3]. An agent having a query from a user broadcasts it to other agents, and indicates the results to a user. Unfortunately their system can deal with only the restricted closed information as bookmark files, and does not provide the connection among persons.

Connection among persons in a virtual space has been studied. For example, FreeWalk[4] provides the opportunity for a casual meeting in a computer network. However such researches do not support the integrated framework for searching information both of the WWW and human community.

2 Multi-agent system for integrated information search in the WWW and a human group

System overview

Fig.1 shows the system overview. Information search framework in which information is acquired through both the WWW and a human group. In the framework, after a user gives a query to a system, it broadcasts the query to information resources, gathers the results and indicates the integrated results to a user. We apply a multi-agent system to this framework.

The multi-agent system consists of *user agent*s and *SE(search engine) agent*s, and utilizes contract net protocol to communication among the agent. An user agent and a SE agent are assigned to a user and a SE agent, and each of the agents maintains a *user profile* and a *SE profile* respectively. A profile consists of two types of sub-profiles: a *knowledge profile* and a *resource profile*. The knowledge profile contains stemmed terms standing for concepts which a user or a search engine acquired. The resource profile contains other properties on the information resource.

An agent which received a *query* becomes a *manager* and other agent become *contract*s. The manager agent broadcasts a *task announcement* including the query to the contractors, and they reply *bid message*s made from their profiles to the manager. After investigating the bid messages, the manager determines the contractors which win a bid. As a result, the manager agent actually gives the query as *client message*s to the selected contract agents and they return *result messages* including the retrieval information. Finally the manager integrates them and indicates it to the a user with a hit list and an audio/video application connected to a person who can answer the query.

Task announcement and bid message

In this subsection, we define a task announcement, a bid message, a client message and a result message. The following lists stand for the attributes and the values for a task announcement and a bid message.

Task announcement

- *Type* : task_announcement
- *From* : user_agent_1
- *Task name* : information_search_1
- *Query* : [cscw, www, information_ retrieval]
- *Deadline* : Tue Mar 28 07:44 2000

Bid message

- *Type* : bid_message
- *From* : SE_agent_2
- *To* : user_agent_1
- *Relevance* : [0, 1] The relevance between a query and a knowledge profile is computed with the term frequency of the query in a knowledge profile.

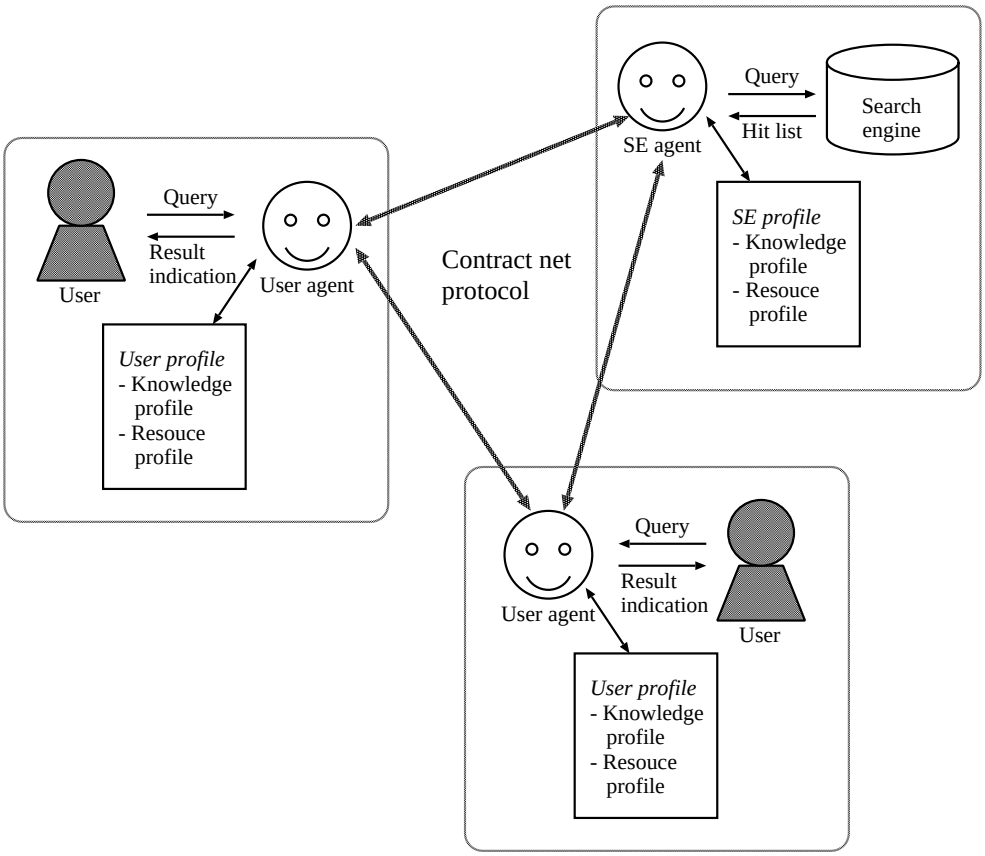

Figure 1 System overview

- *Profile* : Resource profile of a user profile or a SE profile.

Client message

- *Type* : client_message
- *Contract label* : contract_11
- *To* : search_engine_agent_2
- *From* : user_agent_1
- *Task name* : information_search_1

Result message

- *Type* : result_message
- *Contract label* : contract_11
- *From* : SE_agent_2
- *To* : user_agent_1
- *Current precision* : [0, 1]
- *Task results* :

  A Search engine
  - *Hit list* : a list of relevant Web pages.

  A person
  - *Person name* : Seiji_Yamada
  - *Port for vic*:
  - *Port for vat*:

Procedure of a agent

The procedure of a user agent and a SE agent is described in the followings. As mentioned earlier, it has two modes, a manager mode and a contractor mode. The procedure of a user agent and that of a SE agent are almost the same except the ways to construct a profile and a bid message. The procedure of a agent is described in the followings.

- *Manager mode*: When a query is given to a user agent, it starts to work as a manager.
  1. Broadcast a task announcement to all the other agents as contractors.
  2. Obtain the bid messages from the contractors.
  3. Evaluate the bid messages, select adequate contractors, and make contracts with the selected contractors by sending client messages to them.
  4. Receive the result messages from contracted agents.
  5. Integrate the results and indicate them to a user.
- *Contractor mode*: When an agent receives a task announcement, it starts to work as a contractor.
  1. Compute the similarity between the *Query* in the task announcement and its user profile or SE profile.
  2. If the similarity is over the threshold τ, introduce an user profile or a SE profile to a bid message and reply the bid message to the manager. Otherwise this procedure finishes.
  3. If it receives a client message from the manager, investigate the *Current precision*, *Task results* in a result message and replay the result message to the manager agent.

3 Constructing a user profile and a SE profile

We describe a user profile and a SE profile, and develop methods to compute the properties.

User profile

A user profile consists of the following sub-profiles.

- *Knowledge profile*
- *Resource profile*
  - *Presence* : [0, 1]
  - *Cognitive_load* : [0, 1]
  - *Social_relation* : $[-1, 1]$

Knowledge profile Basically a user constructs his/her knowledge profile by his/herself. A Web browser with CGI is indicated to a user when he/she spends much time in editing a particular file, seeing a particular Web page using a browser and so on. The indicated Web browser has input forms by which a user can input terms into his/her knowledge profile. A knowledge profile is described as a table including stemmed terms and its occurrence frequency. Note that a system allows a user to input the same terms into a knowledge profile several times. This makes the frequency of a term correspond to a amount of knowledge on the term.

Presence A user agent checks the possibility of user's presence using a vision algorithm like face recognition with a CCD camera used to video communication.

Cognitive load A user agent can estimate the cognitive load by investigating user's usage frequency of an text editor, Web browser and other application which he/she needs cognitive load to use.

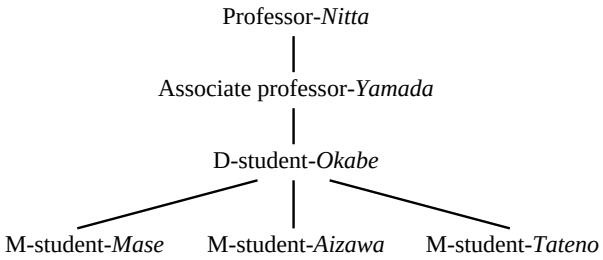

Figure 2 Ontology in my laboratory.

Social relation A user agent can compute the social relation using ontology among users in their organization. Fig.2 shows an example of ontology. If a contractor's user has a higher position than a manager's user in the organization, the social relation takes a negative value. In the opposite case, the social relation takes a positive value. This makes the direct audio/video communication between a manager assigned to a high-position person and a contractor assigned to a low-position person easier.

SE profile

- *Knowledge profile*
- *Resource profile*
  - *Network_load* : [0, 1]

Knowledge profile We apply the statistic method developed in Savvy Search[1] to construct a SE profile.

Network load This is estimated by sending *ping* command to the search engine and measuring the response time.

4 Selecting information resources and indicating results

At step-3 in a procedure for the manager mode, a manager determines the contractors with which it make contract as followings.

1. Using the following equations, a manager computes evaluations E_p and E_s for a user agent contractor and a SE agent contractor respectively, and normalizes them within [0, 1].

$$E_p = w_r * Relevance + w_p * Presence + w_s * Social\_relation - w_c * Cognitive\_load$$

2. A manager selects n contractors having the highest evaluations, and requests them to send their result messages.

$$E_s = w_r * Relevance - w_n * Network\_load$$

At present, the weights w_r, w_p, w_s, w_c, w_r, w_n, n and the threshold τ in the procedure of contractor mode are experimentally set by a designer. In future work, we have a plan to apply a simple learning method to adjust them depending a user.

After receiving the result messages, the manager integrates *Task results* in the messages and indicates them like Fig.3. The Web page of the hit lists from *Task results* in the result messages are sorted depending on the number of search engines which support the Web pages[6].

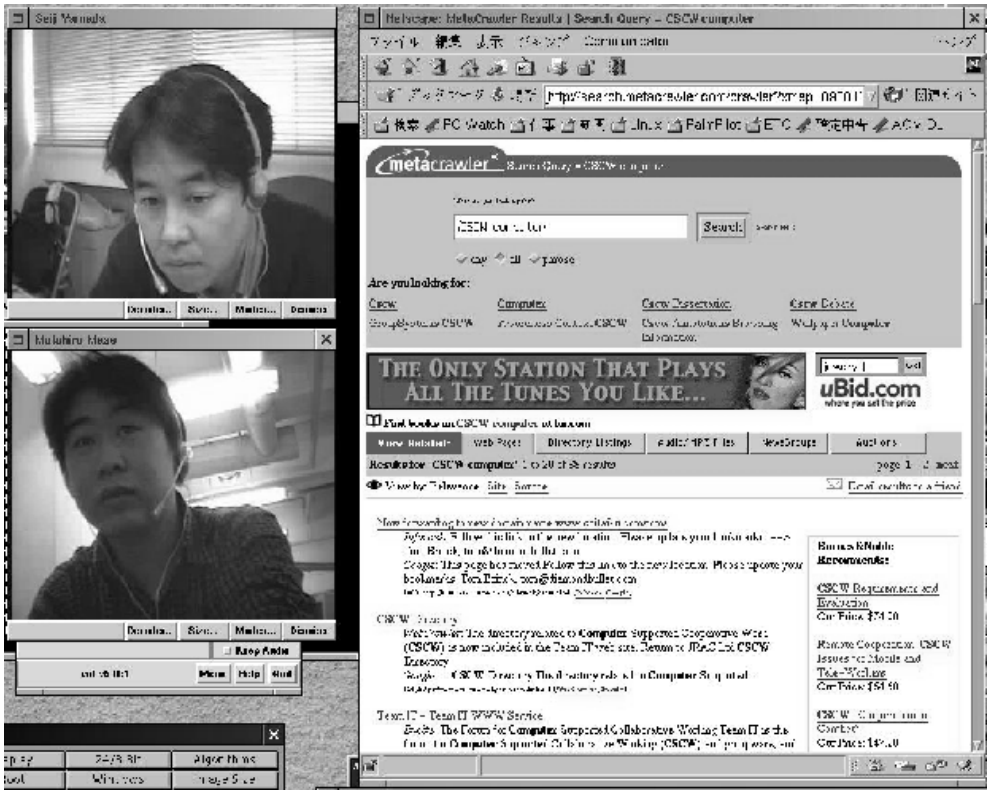

Figure 3 Indication of results.

5 Implementation

We are currently implementing the whole system using Perl, C and GTK+. The vic[1][2] and vat[2] applications are also introduced to audio/video communication between persons.

6 Conclusion

We proposed an information search framework which searches open information and closed information through both the WWW and a human group. We integrated the information search in the WWW with that in a human group, and developed the information search system as multi-agent system. Knowledge profiles for each person and each search engine are build, and the system selectively searches information resources consisting the WWW and personal knowledge. As a result, a user acquires integrated results of hit lists from the WWW and direct audio/video connections to a person who can answer the query.

Unfortunately the implementation is not fully done yet. The implementation and the development of experimental evaluation method for our system are our future work.

REFERENCES

1 A. E. Howe and D. Dreilinger. Savvy search: a metasearch engine that learns which search engines to query. *AI Magazine*, 18(2):19–25, 1997.

[1] http://www-nrg.ee.lbl.gov/vic/
[2] http://www-nrg.ee.lbl.gov/vat/

2 S. McCanne and V. Jacobson. vic: a flexible framework for packet video. In *Proceedings of the third international conference on Multimedia*, pages 511–522, 1995.

3 M. Mori and S. Yamada. Bookmark-agent: Information sharing of urls. In *Poster Proceedings of The 8th International World Wide Web Conference (WWW-8)*, pages 70–71, 1999.

4 H. Nakanishi, C. Yoshida, T. Nishimura, and Toru Ishida. FreeWalk: Supporting casual meetings in a network. In *International Conference on Computer Supported Cooperative Work (CSCW-96)*, pages 308–314, 1996.

5 J. Rucker and M. J. Polanco. Siteseer: Personalized navigation for the Web. *Communications on ACM*, 40(3):73–75, 1997.

6 E. Selberg and O. Etzioni. The metacrawler architecture for resource aggregation on the Web. In *IEEE Expert*, volume January-February, pages 11–14. IEEE, 1997.