Weak Information Structures for Community Information Sharing

Harumi Maeda*, Masao Kajihara*, Hidekazu Adachi*, Atsushi Sawada†, Hideaki Takeda*, and Toyoaki Nishida*

* Graduate School of Information Science, Nara Institute of Science and Technology
  8916-5 Takayama, Ikoma, Nara, 630-01 Japan
  Phone: +81-743-72-5265, Fax: +81-743-72-5269
  E-mail: {harumi-m, masao-k, hideka-a, takeda, nishida}@is.aist-nara.ac.jp

† Department of Information Science, Graduate School of Engineering, Kyoto University
  Yoshida-Honmachi, Sakyo, Kyoto, 606-01 Japan
  Phone: +81-75-753-5992, Fax: +81-75-753-5977
  E-mail: sawada@kuis.kyoto-u.ac.jp

Abstract: We discuss the requirements for community information sharing. We then propose the weak information structures to integrate heterogeneous information such as static information (e.g., local sites information) and dynamic information created in word-of-mouth communication. The weak information structures connect various information media without defining the semantics rigorously. By leaving the interpretation of the semantics to tacit human background knowledge, it becomes compact and robust. We have developed an information sharing system for community called InfoCommon which provides people with intelligent assistance for exchanging and sharing knowledge and ideas. We have evaluated InfoCommon at the ICMA'S96 Mobile Assistant Project.

1 Introduction

The World Wide Web has become popular for human information sharing on the Internet. As large-scale information resources on the Internet are increasing rapidly, it becomes more and more difficult to obtain information we need. Although a number of search tools are available, there are few intelligent systems which help us reorganize vast information obtained from the Internet from our point of view.

On the other hand, programming languages or knowledge representations such as first-order logic or frames have been used for computer information sharing. Unfortunately, these are so logically rigid that one must spend tremendous amount of efforts on information acquisition. This defect forces too much on human effort and hence significantly hinders accumulation of a large amount of useful information.

We investigate an intermediate information representation called the weak information structures, which are weaker than well-defined knowledge representations for human information sharing.

In this paper, we focus on a kind of communities in which a group of people meet together and are united by shared interests as a range of human information sharing. We attempt to build a system to support community information sharing using the weak information structures.

We discuss the requirements for community information sharing. We then propose the weak information structures to integrate heterogeneous information such as static information (e.g., local
sites information) and dynamic information created in word-of-mouth communication.

We have developed an information sharing system called InfoCommon which provides people with intelligent assistance for exchanging and sharing knowledge and ideas. We have evaluated InfoCommon at the ICMAS'96 Mobile Assistant Project.

In what follows, first we analyze requirements for community information sharing and describe the weak information structures. We then present InfoCommon and experimental results at the ICMA'S96 Mobile Assistant Project[1] and make discussion.

2 Issues in Community Information Sharing

In this section, issues in community information sharing will be discussed.

2.1 Important Information in Community

Community is a “group of people living together and/or united by shared interests, religion, nationality.” In this paper, we focus on a kind of communities in which people meet together and are united by shared interests.

By the above definition, local sites information where people meet together and information that people share interests is important. Personal information is also important to activate human-human interaction.

In addition, it is known that informal information created in word-of-mouth communication is essential to support community information sharing.

2.2 Hypothesis of Information Activity in Network Communities

As computer network technologies progress, virtual communities in which people do not live together and which are supported by computer networks have been formed. Mailing-lists and Forums are such systems which support virtual communities. We call such communities “network communities.”

We analyze how network communities are formed on mailing-lists.

Firstly, friends or companions who have the same/similar interests start a mailing-list. In the beginning, acquaintances of founders participate in. Newcomers then take part in when they happen to know the mailing-list by word-of-mouth or watching publicity.

First Messages created by newcomers are mainly questions except for self-introduction. This is because most of newcomers participate in the mailing-list to get information they need.

Newcomers try to find information by themselves in vain, and ask other members. They get chances to talk with and know others by asking questions, and they are recognized by others as well. Discussions seldom occur unless people do not know the people with whom they want to talk.

We set up a hypothesis that there are a series of processes for newcomers in information activity in network communities: “search → asking → knowing people → discussion.” We think that helping these processes facilitates community information sharing.

In addition, we claim that the existence of special persons who answer others’ questions and encourage others to send messages is important to activate network communities.

2.3 Problems in Information Sharing on the WWW

As we pointed out in section 1, The WWW has become popular for human information sharing on the Internet. Here we focus on some problems in information sharing on the WWW. (1) It is difficult for users to get information they need when the menu structures of hypertexts are different from user knowledge about topics. (2) If search results are too many when using search engines, it is difficult for users to find useful information from them. We need more intelligent systems to reorganize vast information from our point of view.

2.4 Information Sharing Environment with Mobile Terminals

Most of groupware and information sharing systems are used on workstations or personal computers at work and at home. Matsushita[2] pointed out that building information sharing environment with mobile terminals will be one of the most anticipated issues in this field.

We must consider some limitations in current technology of mobile terminals, for example: (1) slowness of communication, (2) expensiveness of communication cost, (3) poorness of security, and (4) smallness of screen.
2.5 Requirements Analysis for Community Information Sharing

From considerations described in the previous section, we analyze the requirements to build a successful system for community information sharing.

1. Contents of information are important. For example, local sites, personal, word-of-mouth information and information that people share interests should be stored.

2. Special members assist general users to use the system.

3. The system supports information activity in network communities: “search → asking → knowing people → discussion.”

4. Users can access heterogeneous information from the users’ point of view with/without menus.

5. Users can integrate and reorganize personal and public information.

In addition, in order to use mobile terminals, we must consider the following issues to overcome the limitations of them.

6. The system should have easy and simple user-interface.

7. Interaction between mobile terminals and the server should be reduced.

3 Weak Information Structures

The weak information structures connect a wide variety of information media such as natural language texts, hypertexts and images without defining the semantics rigorously (Figure 1). By leaving the interpretation of the semantics to tacit human background knowledge inherently shared with people, they become compact and robust. Moreover, the weak information structures are easy to generate from raw data for both of humans and computers, and therefore can reduce the cost of information acquisition and integration.

We call the information structures “weak” in the sense that we need not define semantics rigorously, as compared with knowledge representations such as semantic networks, frames or first-order logics.

In community information sharing, the weak information structures are used to integrate heterogeneous information such as static information (e.g. local sites information) and dynamic information created in word-of-mouth communication. We think that background knowledge shared by members can be utilized to understand the meaning of the relations.

4 InfoCommon

We have developed a shared-card information sharing system called InfoCommon which allows seamless keyword-based access to a variety of information cards to support community information sharing.

4.1 Architecture

We employed several design principles to facilitate community information sharing. First, we try not to enforce anyone particular concept. Instead, we allow much freedom in the usage of terms and the structure of shared information and to incorporate useful information from various viewpoints. Second, we make the information space a single seamless space. This releases users from working with a rigid menu. Third, we enable users to build a personal information space where they can organize relatively small amounts of information as desired.

InfoCommon is composed of (a) PDAs (Personal Digital Assistant) possessed by users, (b) InfoCommon information server which handles user requests and (c) News server which stores user messages and (d) InfoCommon information base. The InfoCommon information base consists of (d1) a knowledge base which links keywords and information cards using the weak information structures and (d2) an information card base which
stores static information (Figure 2).

In addition, InfoCommon information volunteers (a) answer to user questions as the help desk to activate communications and (b) add information cards and adjust the weak information structures to improve search results.

4.2 Functions

InfoCommon supports the following functions.

Content-based Information Retrieval Given a set of keywords, InfoCommon will respond with the set of information cards connected to the keywords. The result of retrieval is stored in the user’s local information base where the user can re-arrange the collection of information cards, and add/remove nodes/links as desired.

Information Sending (Posting News) InfoCommon is built on a conventional News service. Users can send messages to the News server. These messages are viewed as message cards in InfoCommon.

Personalizing Information Users can edit and reorganize gathered information and message cards and personal memoranda.

4.3 User Interface

The information unit in InfoCommon is called “card.” There are three kinds of cards: (a) an information card is a piece of static information stored in the information card base, (b) a message card is a piece of dynamic information created by users, which is stored in the News server, and (c) a memo card is a piece of users’ personal memoranda, which is stored in PDAs.

InfoCommon provides visual interface for retrieving and sending cards. A relation between cards is denoted by a link.

Figure 3(a) shows an example result when a user inputs “daibutsu, temple” in information retrieval. An icon of the card “Todaiji Temple” which is related to both of “daibutsu (great image of Buddha)” and “temple” is shown in the middle of the screen. Some card icons which are related to “Todaiji Temple” such as “Nandaimon Gate” and “Daibutsuden Hall” (these are buildings of Todaiji Temple) are linked with the card icon of “Todaiji Temple”. A card icon of “Model Course around Narı Park” is shown but not linked with that of “Todaiji Temple”, because the former and the latter are not related directly.

4.4 Search Algorithm

This section describes the algorithm which finds a set of cards when the server receives user input in information retrieval.

- **step 1:** to remove unnecessary words/symbols from the input string and to expand keywords referring synonyms
- **step 2:** to select card candidates by full-text “AND” search
- **step 3:** to select card candidates by full-text “OR” search when card candidates are not selected by **step 2**
- **step 4:** to add card candidates by path-finding of the weak information structures and generate links
- **step 5:** to define maximum 10¹ cards from card candidates using predefined weights

The weak information structures which are referred in **step 4** are defined by (1) predefined relations between cards, (2) predefined relations between concepts or (3) relations of cards defined by users or information volunteers.

We describe the path finding below. The path finding is based on the idea of “spreading activation” on semantic networks [3]. Path-finding with distance n searches a set of units which are

---

1. screen size of PDA and usability are concerned
connected by links with distance \( n \). Distance 1 denotes an extent between keys and values. Figure 4 illustrates how the algorithm works to an input: “temple, Prince Shotoku?” and answers “Horyuji-Temple” and “Chogosonji-Temple.”

5 Social Experiment at the ICMAS’96 Mobile Assistant Project

5.1 ICMAS’96 Mobile Assistant Project

ICMAS’96 is the Second International Conference on Multiagent Systems, which was held in December 9 - 13 of 1996 in Kyoto, Japan.

Nippon Telegraph and Telephone Corporation (NTT), Kyoto University, and Nara Institute of Science and Technology jointly planned the ICMAS’96 Mobile Assistant Project to exhibit a Telescript-based system which provides (1) e-mail, e-forum and internet access services, (2) conference and tourist information for local sites, and (3) social match making based on participants’ profiles and schedules to arrange meetings, teas, dinners and so on.

The project is the world first experiment in applying mobile computing systems to community support. 100 personal intelligent communicators with handy phones were loaned to conference participants to actually try out the system.

People could access various information about the participants through their PDAs (Figure 5). One server machine (HP9000 model 800160) and 30 telephone lines were set up at the conference site[1].

5.2 Overview of Experimental Result of InfoCommon

We evaluated the usefulness of InfoCommon at the ICMAS’96 Mobile Assistant Project.

Figure 6 is a photo taken in the Nara Park where an excursion was held, which shows how a user actually used InfoCommon.

The InfoCommon information base stored static information such as abstracts of papers, session, local sites information and participants’ profiles.

The number of information retrievals was 351 and information sending was 32 over the 5 day period.

In what follows, We first show some examples of how users actually used InfoCommon at the conference and then analyze how InfoCommon was used by examining the log file and questionnaires in details.
5.3 Examples

5.3.1 Example 1: “fipa”

Table 1 describes message exchanges about a topic “FIPA (The Foundation for Intelligent Physical Agents).”

User A tried to find information about “FIPA” using information retrieval in vain. He asked about “FIPA” by information sending. The Information volunteer and users answered his question. Figure 7 shows a result of information retrieval when user A finally got information about FIPA (as item 12 of Table 1). Card icons whose messages are sent Via InfoCommon are linked together.

5.3.2 Example 2: “real world agent”

Figure 8 shows a result of information retrieval when a user input “real world agent.”

Card icons of session information, abstracts of papers and participants’ profiles concerning “real world agent” were found. In this case, card icons of a speaker’s profile, his abstract and session information are linked together.

5.3.3 Example 3: “nishimura”

Figure 9 shows a result of input “nishimura,” which is a name of a participant who used the system frequently. Icons of personal information of Nishimura, message cards created by him were displayed.

5.4 Analysis

5.4.1 Purpose

Answers to the question “For what did you use InfoCommon?” are shown in Figure 10. 59% used the system for information retrieval and 18% for information sending including News(11%), help desk(7%) and discussion.
5.4.2 Comparison with Traditional Newsreader

In this experiment, 3 services were prepared which access to 17 Newsgroups: InfoCommon, a traditional Newsreader system and an information navigation system[4]. Users could read and post messages in each newsgroup using the Newsreader. In contrast, users could access messages regardless of newsgroups based on keywords in InfoCommon.

32 out of 48 messages (67%) were posted via InfoCommon.

The answer to the question “Which service did you mainly use for reading News ?” is as follows; Newsreader 33 persons (64%), InfoCommon 14 persons (27%), other 4 persons (9%).

About “Which service did you mainly use for posting News ?”: Newsreader 9 persons (52%), InfoCommon 7 persons (41%), other 1 person (1%).

The reasons as to why InfoCommon was used for News are described as follows: “Because keyword search was easy and useful (14 persons)”, “Because I found interesting topics in InfoCommon(5 persons)” and “Because I had a question (4 persons)”. We determined that InfoCommon added new facility to the conventional Newsreader.

Major reasons of the choice “didn’t use” were “slow information retrieval (8 persons)” and “I couldn’t connect to the server (4 persons)”. These problems involve server response and so can be easily fixed.

5.4.3 Topic

Answers to the question “For what topic did you use InfoCommon” are displayed in Figure 11: research(33%), presentation(18%), people(16%), sightseeing(15%) and restaurant(11%) and others (7%).

5.4.4 Statistics

Table 2 shows the ranking of frequently asked keywords in information retrieval. The most frequent input word was “icstat,” a special keyword to display statistics concerning frequently asked keywords and frequently discussed subjects. This result suggests that users used InfoCommon to know what other people are interested in and/or what are hot before information retrieval or sending. We found that such statistics are useful for sharing information among participants.

“Nara” and “keihan” are names of area where the conference was held. “Fipa” and “agent” are research keywords concerning the conference topic. “Nishimura” is a name of a participant. “Food”, “lunch”, “banquet”, “kamameshi” and “restaurant” can be considered as food information. The above result almost meets with that of questionnaires described in the previous section. The difference is that keywords about research, presentation and people are less found in Table 2 and this is because these words vary individually.

<table>
<thead>
<tr>
<th>no.</th>
<th>user</th>
<th>system</th>
<th>manipulation</th>
<th>date</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>user A</td>
<td>InfoCommon</td>
<td>retrieval: keyword “fipa”</td>
<td>Dec 9 23:12</td>
<td>User A searched information about FIPA.</td>
</tr>
<tr>
<td>2</td>
<td>user A</td>
<td>InfoCommon</td>
<td>posting</td>
<td>Dec 9 23:17</td>
<td>User A asked a question about FIPA.</td>
</tr>
<tr>
<td>3</td>
<td>information volunteer</td>
<td>InfoCommon</td>
<td>posting</td>
<td>Dec 10 00:35</td>
<td>The information volunteer answered to user A’s question.</td>
</tr>
<tr>
<td>4</td>
<td>user B</td>
<td>InfoCommon</td>
<td>retrieval: keyword “icstat”</td>
<td>Dec 10 00:51</td>
<td>User B searched statistics and found a subject about FIPA.</td>
</tr>
<tr>
<td>5</td>
<td>user B</td>
<td>InfoCommon</td>
<td>retrieval: keyword “fipa”</td>
<td>Dec 10 00:56</td>
<td>User B searched information about FIPA.</td>
</tr>
<tr>
<td>6</td>
<td>user B</td>
<td>InfoCommon</td>
<td>posting</td>
<td>Dec 10 00:58</td>
<td>User B answered to user A’s question.</td>
</tr>
<tr>
<td>7</td>
<td>user C</td>
<td>Newsreader</td>
<td>read “misc” newsgroup</td>
<td>-</td>
<td>User C found information about FIPA.</td>
</tr>
<tr>
<td>8</td>
<td>user C</td>
<td>Newsreader</td>
<td>post “misc” newsgroup</td>
<td>-</td>
<td>User C answered to user A’s question.</td>
</tr>
<tr>
<td>9</td>
<td>user C</td>
<td>Newsreader</td>
<td>“misc” newsgroup</td>
<td>Dec 11 23:56</td>
<td>User C answered to user A’s question again.</td>
</tr>
<tr>
<td>10</td>
<td>user D</td>
<td>InfoCommon</td>
<td>retrieval: keyword “icstat”</td>
<td>Dec 12 09:01</td>
<td>User D searched statistics and found a subject about FIPA.</td>
</tr>
<tr>
<td>11</td>
<td>user D</td>
<td>InfoCommon</td>
<td>retrieval: keyword “fipa”</td>
<td>Dec 12 09:28</td>
<td>User D searched information about FIPA.</td>
</tr>
<tr>
<td>12</td>
<td>user A</td>
<td>InfoCommon</td>
<td>retrieval: keyword “fipa”</td>
<td>Dec 12 10:01</td>
<td>User A searched information about FIPA and found his own question and answers by others.</td>
</tr>
</tbody>
</table>
5.4.5 Change of Search

Figure 12 shows that InfoCommon was used from evening to midnight. This result meets with a report of Kawaura et al.[8] which says “the computer communication is often done during night.”

On the other hand, it does not meet with the general investigation about user activities of PDAs which says the communication in mobile computing is done irrespective of time. We analyze that the reasons are perhaps (1) the system was slow, (2) people were busy during the daytime, and (3) information retrieval was not a very urgent task.

5.4.6 Search Results

Figure 13 shows answers to the question “How did you feel the search results of InfoCommon?” 81% felt that the search results were fine (very good + good + moderate) as shown in Figure 13.

Figure 14 shows different answers to the same question as before. 44% answered that the search results were interesting.

5.4.7 Satisfaction

51% answered that they were satisfied with InfoCommon(Figure 15).

5.4.8 Usefulness

55% answered that InfoCommon was useful for getting the information they need (Figure 16(a)). On the other hand, 48% answered that it was not useful for knowing people (Figure 16(b)), 60% for discussion (Figure 16(c)). The results indicate that the system is useful for information retrieval but not for knowing people nor discussion.

We feel that the period (5 days) was too short to form the kind of community in which people create many active discussions. We need to conduct a longer-term experiment to evaluate the usefulness of the system for discussion in the future.
6 Discussion

6.1 Contents of Information in Community Information Sharing

The result of questionnaires shows that (a) local sites information such as restaurants and sightseeing, (b) personal information and (c) information about research and presentations were searched. In addition, it is found that statics are useful by analyzing the log file.

These result support the consideration in section 2 and suggest that people are interested in what others are interested in.

6.2 InfoCommon as an Information Retrieval Tool

We analyze that the following results support that InfoCommon is useful as an information retrieval tool for community; (a) 81% felt that search results were fine, (b) 51% were satisfied the system, and (c) 55% answered that the system was useful for information retrieval.

6.3 InfoCommon as an Information Exchange Tool

We could not prove the usefulness of InfoCommon as an information exchange tool for community because (a) the number of information sending was only 32 and (b) only 21% answered that it was useful for discussion.

However, the result that 32 messages out of 48 (67%) were via InfoCommon suggests the InfoCommon can be used as an information exchange tool. We analyze that this is because the simplicity of its user-interface and existence of information volunteers.

In addition, the comparison with the traditional Newsreader suggests that InfoCommon does not replace the traditional Newsreader but add new functions to it.

6.4 Hypothesis of Information Activity in Network Communities

We attempt to evaluate whether the system is useful about the hypothesis of information activity in communities: “search → asking → knowing people → discussion” that we set up in section 2.

(1) We evaluate that the system supported the process “search” from the results of questionnaires.

(2) The process of “search → ask” is partially supported by the system because the log file shows that some users asked questions after information retrieval (as example 1 in section 5).

(3) Half of users answered that the system was not useful for “knowing people.” However, we consider that this is because there were other systems in which users can access participants information without connecting the server[6] and users did not need to use InfoCommon for the purpose. We expect that InfoCommon is useful for knowing people and try to verify it in the future.

(4) The process of “discussion” cannot be supported by InfoCommon. We feel that the period was too short to form the kind of community in which people create many active discussions. We need to conduct a longer-term experiment to evaluate the usefulness of the system for discussion in the future.

6.5 Weak Information Structures in Community Information Sharing

The following results indicate the system’s usefulness for community information sharing: (a) 51% of users were satisfied with the system, (b) 81% felt search results were fine and (c) 55% answered that the system was useful for getting information they need. We consider the results are supporting evidence of that people who share interests in community can have common background knowledge and they can understand the meaning of the semantics of the weak information structures.

7 Related Work

This research is a part of the Knowledgeable Community Project [7] which aims to develop distributed knowledge bases.

Ishida et al[8][9] investigate to facilitate conversation on wide-area information networks to support community activities.

CYC[10] and ARPA Knowledge Sharing Effort [11][12][13] have made a significant contribution in the sense they shed light on the importance of knowledge and information sharing and that they have presented a self-completed computational model. Their approach orients computer information sharing while ours is for human information sharing.

Gaines uses semantic networks as information representation for group knowledge sharing[14]. Our approach is based on much weaker information representation than semantic networks.

Kautz studied the use of agents in assisting and simplifying person-to-person communication
for information gathering tasks[15]. They focus on the use of a software agent. We concentrate on the process of how humans create knowledge and information.

8 Conclusions

We discussed requirements for community information sharing. We then proposed the weak information structures to integrate heterogeneous information such as static information (e.g., local sites information) and dynamic information created in word-of-mouth communication. The weak information structures connect various information media without defining the semantics rigorously. By leaving the interpretation of the semantics to tacit human background knowledge, it becomes compact and robust.

We developed an information sharing system for community called InfoCommon which provides people with intelligent assistance for exchanging and sharing knowledge and ideas. We evaluated InfoCommon at the ICMAS’96 Mobile Assistant Project.

As a future research, we plan to conduct a longer-term experiment to evaluate the usefulness of the system for discussion.

Acknowledgements

The ICMA’96 Mobile Assistant Project was conducted collaboratively by NTT, Kyoto University, Nara Institute of Science and Technology, and Kobe University and in cooperation with NTT Kansai Mobile Communications Network Inc., General Magic Inc., NTT Multimedia Service Promotion Headquarters and Keihinna Interaction Plaza Inc. We wish to thank Mr. Fumio Hat-tori, Mr. Ichiro Morihara, Mr. Yoshiyasu Nishibe, Prof. Toru Ishida, Prof. Katsumi Tanaka, Mr. Shiro Takata and Mr. Hisao Nojima for their help in this project. We would like to thank the participants of this project for their cooperation.

References


