SUPPORTING INFORMATION SEEKING: COLLECTIVE INTELLIGENCE APPROACH

Yuichi Osato, Hiroshi Sugimura and Kazunori Matsumoto
Course of Information and Computer Sciences, Graduate School of Kanagawa Institute of Technology
1030 Shimo-ogino, Atsugi-shi, Kanagawa 243-0292, JAPAN

ABSTRACT

This paper proposes an intelligent method that supports information seeking in libraries. Libraries have long history of reference services, and already provide search tools and databases. We first discuss limitations of the current technologies and identify their disadvantages. We then show a general model of information seeking process that consists of several tasks having different purposes and goals. The tasks cover diverse kinds of roles so that a supporting tool need to have wide functions. In early steps of the process, we show a use of mind map is useful to facilitate understanding the seeking goal. We also show an importance of sharing experience and knowledge in a community. We compare conventional social intelligence techniques and identify the disadvantages of them. We propose a tool with new approaches.

KEYWORDS

Information seeking, information retrieval, library service, collective intelligence,

1. INTRODUCTION

In these few decades, explosive amount of information is digitalized and is provided through the Internet. Our processing ability is frequently overflowing with obtainable information, we sometimes fail to reach necessary information. This situation is recently called information explosion, it causes crucial problems and challenges a study of new technologies. We in this paper focus on university libraries, which have long histories and experiences as knowledge centers, are commonly confronted with these problems.

Libraries already provide services and instructions, including use of computer tools and reference databases. In particular, the open public access catalog system [5,7], OPAC for short, is widely used tool, which carries out information retrieval depending on partially specified data such as keywords. The OPAC operates efficiently if the user would have a clearly defined goal of the seeking, and could express the goal in clearly defined expression. In many practical situations, these conditions cannot be satisfied [5]. The category search [2] runs different manner. Information is hierarchically classified according to predefined categories, and then the search is carried out by going down or up the hierarchy. The classification relies on skills of expert persons, thus the cost of maintenance becomes large. We need further studies on tools

Recent studies introduce information sharing practices, which realize mass users participation. This idea is called collective intelligence or social intelligence. The social bookmark and the folksonomy [6] are the typical examples of the idea. By collecting practices, we aim at improving quality of information content and organization. We compare several social intelligence techniques, and conclude they are inadequate for our purpose. Base on the investigation, we propose a tool involving new social approaches.

2. INFORMATION SEEKING PROCESS

Information seeking is an interactive and iterative process, in which many decisions being made by the user. Many studies propose [5,7] own models based on different points of view. In the cognitive approach, the process becomes a mental model with which a user behaves. Another interesting approach is the berry-picking model. This focuses on a dynamic behaviors of the user, which change unpredictably depending on a

result of previous steps. We here propose a relatively simple model, which involves four main tasks, acquisition, interpretation, selection and organization. In Figure 1, the bold lined rectangles show the tasks. The tasks serially proceed from left to right, but backtrack, shown by dense dotted line, to any previous tasks if necessary. These tasks are repeated until the user obtain a satisfactory result or abandon the process. Since the process start with an ambiguous goal, we need to collect relevant information to confirm a shape of the goal. We necessarily make decisions in each tasks to proceed. Outlines of the tasks are listed below.

1. ***Acquisition task:*** by exploratory collecting information the user gradually builds up concrete shape of the goal. Browsing around relevant information is also necessary. In early stages of the iteration, information search and browsing tools which can run with uncertain partial request, is desirable. It is useful to collect knowledge or experience from other persons who have encountered similar seeking. Since the goal is uncertain, a mental state of the user is in confusion and have some frustration.
2. ***Interpretation task:*** for each part of the collected information, interpreting its meaning and deciding whether it relates to the current goal. Associations of ideas are useful to enrich current information, and to create a new possible direction of the seeking. Conversely, we decide a narrowing of the search if information becomes too diverse. Referring to other person's knowledge is also useful. If further information is necessary in interpretation, we backtrack to the previous task.
3. ***Refinement task:*** refining current goal if necessary based on results of the interpretation task, or proceed to the organization task. With the new goal, we back to the acquisition or interpretation task.
4. ***Organization task:*** investigating and organizing obtained information. The process is closed if we have well organized result, otherwise back to adequate previous task and iterate similar procedures.

These tasks are arbitrarily iterated until the user is satisfied with the obtained information. We remark here decisions which are made in the process depends on experience and knowledge of the user. In libraries, advices from librarians, who are experts in library science, are useful. We discuss in the next chapter the tools which are required for each task.

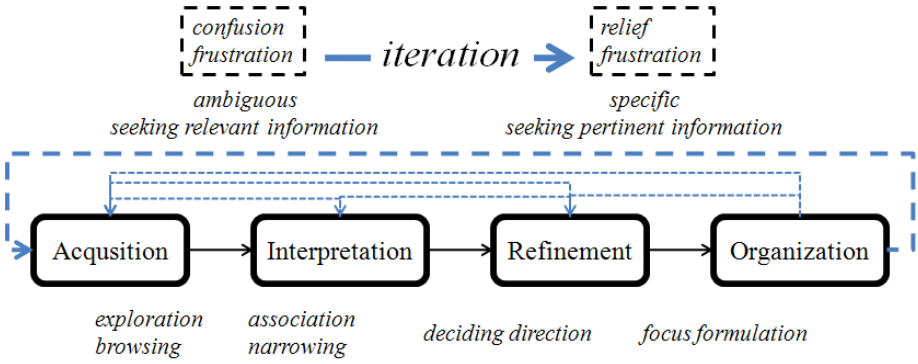

Figure 1. An Outline of Seeking Model

3. DESIGN OF INTELLIGENT TOOL

As we stated above, the tasks in the information seeking process cover diverse kinds of roles so that a supporting tool need to have wide intelligent functions. An outline of the overall structure of the proposed tool is shown in Figure 2, where the left parts implement collective intelligence technology. Explanation of the right parts, shown with the dotted lines, is omitted in this paper. In the following sections, we explain three major functions of the tool.

3.1 Scope Browser

Scope browsing aims to facilitate understanding an outline of a specific topic. This functions is especially useful in the Acquisition and Interpretation tasks. Typically a scope note in a pathfinder [5,7,9], a brief

textual explanation is the most common and a useful practice for this purpose. In general, there are complex relations among a set of topics. We often focus on some familiar topics, and extend our understanding to other unknown topics by investigating relations among them. Ontologies [3,4] in the semantic Web would be used similarly. The standard language for ontology description [3] is almost a subset of the first-order calculus, thus it is useful when we deal with logical aspects of topics. Since the seeking process begins with ambiguous and unclear state, the standard logical ontology is unsuitable in early stages of the process. We then need a method that can deal with rough image of our thinking and structures of relating concepts. Based on these observations, we use the mind map [2] technique to describe topic structures. Mind map writing has few restriction rules so that most persons can use it without special instructions. Relating images, concepts, topics, etc. spread from a specified topic which is put in the center. Mind maps are created by librarians and volunteers, and are stored in the knowledge base. Browsing the maps helps quick understanding of topics.

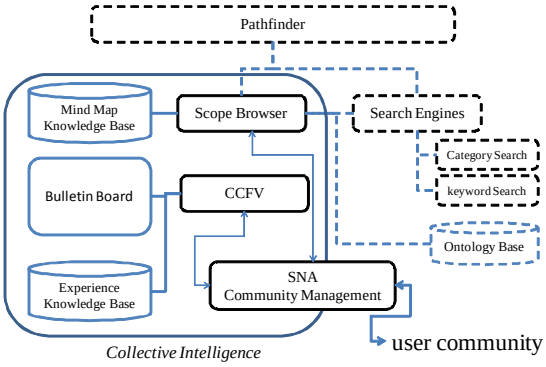

Figure 2. An Outline of Tool Structure

3.2 Managing Collected Knowledge

Cooperation is important in human community. It has synergistic effect to give and to take opinions with other persons. This work is considerably relying on implicit individual skills. Thus social sharing of knowledge is necessary and is a challenge to information technologies. Many libraries maintain reference databases, which collect questions from users with advices from librarians. Internet forums often have similar Q&A databases. Ratings of products are frequently found anywhere in the Internet. The social bookmarks and the folksonomy are mechanism to share knowledge on the Web through bookmarks or tags. In Table 1, we summarize the comparison of these techniques. From this comparison, we conclude the following. First, it is hard to access to simple Q&A databases without ratings or priorities. Second, dependence to expert persons, such as librarians and moderators, increases the cost of maintenance, and at the same time brings a delay of update intervals. Third, to give numerical ratings often lacks rational justification.

Table 1. Typical Collective Intelligence Techniques

Technique	Example	Description, Merits/Demerits
Q&A database	Library reference database	Reliable answers from librarian; Difficult to find helpful cases; Expensive cost of maintenance; Update speed is relatively slow;
	Internet Forum Wiki Answers	Moderator reject dreadful answers, but reliability depends on situations; Necessary information may not obtainable; Difficult to find helpful cases; Update speed depends on situations;
Ratings or Recommendations	Many commercial sites on Web	Numerical ratings are widely used; Necessary information may not obtainable; Reliability depends on situations;

Social Tags	Social bookmark sites Folksonomy sites	Reliability depends on community; Necessary information may not obtainable;

Our proposal is named here Comment and Call for Vote approach, hereafter CCFV for short. The CCFV is developed to solve the problems in the above investigations. As the key point of the CCFV, we restrictedly admit comments on information which can be voted by yes/agree or no/not agree. By simple calculation, comments that have large yes/agree votes are treated as reliable and important comments, and then they become social knowledge. In general, a large community often separates several groups having different opinions. Social activities, such as the vote for a comment, must consider a structure of groups. For this reason, an analysis method explained below is applied in our tool.

3.3 Social Network Analysis

Social network analysis [8], SNA for short, purposes to model behaviors of actors based on their connections and relations to others. The SNA module in Figure 2, we carry out community management functions. If we find information exchanges among actors pass through particular actors, then they plays a central role in the community. By investigating properties of communications, we can separate actors into clusters. In university libraries, we increase the accuracy of a group separation by using background personal information about users, when obtainable.

4. CONCLUSION

We define the information seeking model as an interactive and iterative process involving the tasks, in which many decisions being made by the user. As the tasks cover diverse kinds of roles so that a supporting tool having wide intelligent functions is desirable. For this purpose, we show the collective intelligence approach is useful. We propose a new tool based on the investigation of the current methods. In the proposal, we extract social knowledge from prioritized comments with votes. By applying community management with SNA, we identify groups of the community and manage the reliability of the social operations. The proposed is partly implemented in our library, some execution examples will be shown in the presentation.

REFERENCES

[1] Baeza-Yates, R. and Ribeiro-Neto, B., 1999. Modern Information Retrieval, Addison Wesley.

[2] Buzan,T. ,1995. The mind map book, BBC Books.

[3] Cantador, I. et al, 2008. News@hand: A Semantic Web Approach to Recommending News. *Proceedings of the 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems,* Hannover, Germany..

[4] Hoser, B., 2008. Information Retrieval vs Knowledge Retrieval: A Social Network Perspective, *Proceegings of ESWC Workshop on Collective Semantics,* Tenerife, Spain.

[5] Kuhlthau, C.C., 1991. Inside the search process: information seeking from the user's perspective, *Journal of the Americal Society for Information Science, 42(5).*

[6] Millen, D., et al, 2006. Dogear: Social bookmarking in the enterprise, *Proceedings of International Conference on Human Factors in Computing systems,* Montréal, Canada.

[7] Rubin, R., 2004. Foundations of Library and Information Science, Neal-Schuman Publisher, Inc.

[8] Wasserman, S. and Faust, K., 1999. Social Network Analysis, Structural Analysis in the Social Sciences, Vol. 8, Cambridge Unversity Press.

[9] Wilson, M., 2009. The Tetris model of the information seeking process, *Technical Report of Swansea University, http://eprints.ecs.soton.ac.uk/.*

[10] http://wiki.answers.com/

[11] http://www.ndl.go.jp/jp/library/collabo-ref.html