Active Support for Using Academic Information Resource in Distributed Environment

Baoning Li and Tetsuo Kinoshita

Graduate School of Information Science, Tohoku University, Sendai 980-8577, Japan

Summary
This research proposes a novel approach to actively support using distributed academic information resources. With this approach, passive academic information resources are extended into autonomous and active entities. The extended academic information resources can actively perform tasks to support use of them, so that the burden for users can be reduced. Moreover, multiple extended academic information resources can organize in a decentralized way to autonomously cooperate with each other, and this will enable a more flexible support for using them in the distributed environment. In order to realize the extension of academic information resources, the Active Information Resource (AIR) scheme is adopted in the proposed approach, and an agent-oriented design method for AIR is developed in this research. Furthermore, the proposed approach is verified in two typical tasks of support for using distributed academic information resources: to search and to share. Firstly, this research designs a knowledge enhanced search method for personal academic information collections. Secondly, this research develops an active sharing mechanism for distributed personal academic information collections. The results of experiments on two prototype systems verified that the proposed approach has better performance than conventional approaches do. This confirms that the proposed approach is appropriate for actively support using distributed academic information resources.

Key words:
Active Information Resource, Information Retrieval, Information Sharing, Agent, Agent-based System

1. Introduction

With the explosion of the amount of information available in digital form, as well as the rapid growth of computer networks, particularly the Internet, numerous electronic academic materials have been made available from various information sources. These electronic academic materials can be considered as valuable academic information resources that are continuously being generated, accumulated and updated in the distributed environment.

Some characteristics of academic information resources include: well organized metadata is available, meaningful relations between academic information resources exist, researchers' accumulated knowledge about academic information resources can be reused, etc.

Due to the large amount of electronic academic information resources, the complex relations among them, and the distributed nature of them, it is becoming increasingly difficult for researchers to efficiently and effectively use these information resources. Therefore, an effective support for using academic information resources is required. Such a support approach should effectively use the characteristics of academic information resources to provide mechanisms and functions such as managing, searching, showing and sharing. Furthermore, it should be adaptive in the distributed environment.

There are many conventional technologies can be applied to support the use of academic information resources. Typically, academic information resources in a conventional support system is modeled as passive entities, which can be accessed and manipulated by some functions required by the tasks of support[1]. In such circumstances, there are usually gaps among the characteristics of academic information resources, the manipulation of these resources, and the support for using them. Usually the characteristics of academic information resources are not effectively used, and the support is less effective in the distributed environment.

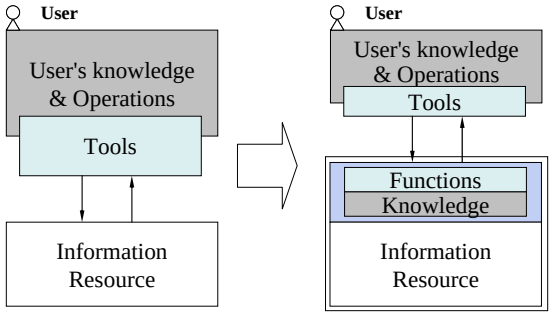


Fig.1 Extending an Information Resource

Manuscript received June 20, 2007
Manuscript revised June 27, 2007

In order to realize an active support for using academic information resources in the distributed environment, this research proposes a novel approach, as depicted in Fig.1.

With this approach, passive academic information resources are extended into autonomous and active entities. Pertinent knowledge and functions are combined with academic information resources, and the characteristics of these information resources can be effectively used. The extended academic information resources can actively perform tasks to support use of them, so that the burden for users can be reduced. Moreover, multiple extended academic information resources can organize in a decentralized way to autonomously cooperate with each other, and this will enable a more flexible support for using them in the distributed environment.

This research aims on the following two main challenges:

Challenge-1. To develop a design method for extending academic information resources into autonomous and active entities.
This research adopts the Active Information Resource (AIR) scheme[2] to realize the extension of academic information resources, and develops an agent-oriented design method for AIR.

Challenge-2. To study how to use the extended electronic academic information resources to enable the active support.
The proposed approach is verified in two typical tasks of searching and sharing the distributed academic information resources.

2. Agent-Oriented Design Method for Active Information Resource

An AIR is an information resource that is enhanced and extended with information resource specific knowledge and functions for actively and flexibly facilitating use/reuse of it.

An agent-oriented design method for AIR is developed in this research. This design method consists of the design model of AIR, and two agent-based design models of AIR.

- Design Model of AIR

There are 6 main parts in the proposed design model of AIR, which is shown in Fig.2: *Information Resource*, *Domain Knowledge-base*, *Knowledge about Information Contents*, *Information Extraction Unit*, *Information Processing Unit*, and *Contact & Cooperation Unit*. With the interaction between these 6 parts, an AIR can autonomously perform its tasks [4].

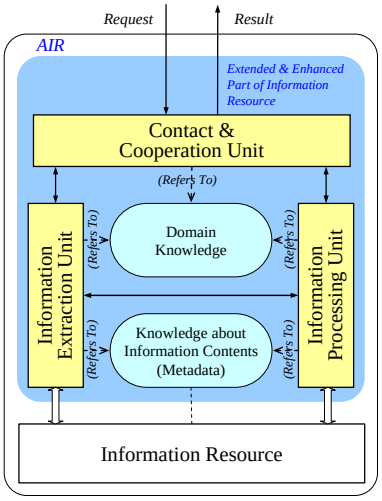


Fig.2 Design Model of AIR

- Agent-Based Design Models of AIR

When an AIR is designed with the agent-oriented approach, the knowledge and functions of the AIR should be mapped into knowledge and functions of an agent or multiple agents. Thus an AIR can be designed using one of the following design models:

− Single-Agent-Based Design Model
In a single-agent-based AIR, the knowledge and functions of an AIR is mapped into knowledge and functions of one agent. Generally, the amount of information in a single-agent-based AIR is relatively small. Since the structure of such an AIR is relatively simple, its design, implementation and maintenance would be accordingly easy.

− Multiagent-Based Design Model
For an AIR with fairly large quantity of information, complex functionality of information processing and corresponding knowledge about information use/reuse, it should be more suitable to deploy multi-agents. In a multiagent-based AIR, each agent is in charge of only a certain part of the AIR's functionality and corresponding knowledge. In such an AIR, the knowledge and ability of a certain agent could be relatively simple, which means the development and maintenance of each agent could be easy.

3. Knowledge Enhanced Search Method for Personal Academic Information Collections

When a personal academic information collection is searched for a researcher's personal use, not only the literal

contents, but also the characteristics of them should be considered[3]. However, when a conventional search method is used to search such a collection, usually the relations among the academic information resources are not properly considered, and the researcher's personal knowledge about the academic information resources are often neglected. Consequently, the recall and efficiency of search could be low [5].

This research uses two types of knowledge about these information resources to enhance the process of a conventional search method:

- **K_R**: Knowledge on Relations among the personally collected academic information resources
- **K_U**: User's Knowledge about the personally collected academic information resources

The single-agent-based design model of AIR is used to realize the proposed knowledge enhanced search method. Each piece of personally collected academic information resource is extended into an AIR (PerC-AIR: Personally Collected academic information resource AIR). PerC-AIRs are designed to automatically discover and update their K_R through autonomous cooperation. At the meantime, K_U can be automatically maintained by each PerC-AIR which keeps track of its user's operation on a piece of academic information resource. Moreover, a knowledge enhanced search in distributed PerC-AIRs is performed through autonomous cooperation between them.

In a K_R enhanced search, relevance spreads from relevant PerC-AIRs to non-relevant ones, and relevance is accumulated in non-relevant PerC-AIRs. The influence of a relevant PerC-AIR on a related one is determined by the relevance value of the relevant PerC-AIR, the strength of their relation, and the depth of iteration.

In a K_U enhanced search, the researcher's evaluation or opinion on each item in the collection is reused. Relevant or non-relevant items can be efficiently determined when there is proper K_U of PerC-AIR indicating so.

Fig.3 depicts a combination of K_R and K_U enhanced search, in which basically the protocol and the algorithm for the K_R enhanced search are used, except that K_U is used in each step of the K_R enhanced search.

To verify the proposed search method, some experiments were performed on a prototype system which consists of 110 distributed PerC-AIRs. As demonstrated by the average results of multiple searches shown in Fig.4, the proposed search method has better performance compared to the conventional method. Particularly, the use of a combination of both K_R and K_U can lead to a better performance and a higher efficiency.

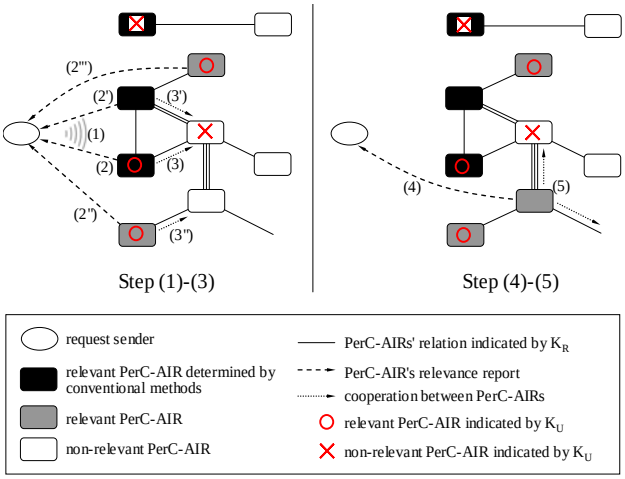
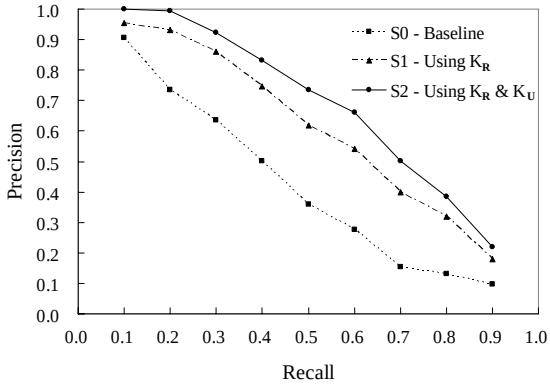
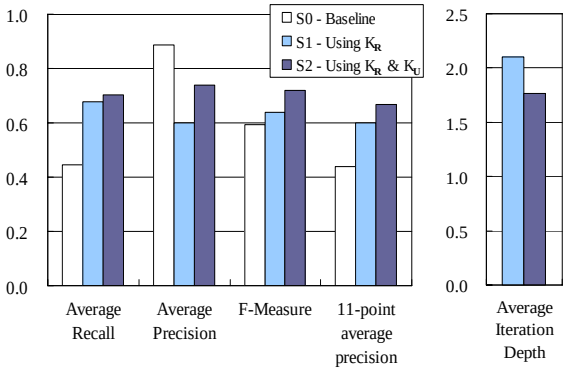


Fig.3 Combining K_R and K_U Enhanced Search



(a) Evaluation by precision and recall



(b) Evaluation by different kinds of measures

Fig.4 Performance of Knowledge Enhanced Search

It can be concluded that the proposed knowledge enhanced search method can be used to search personal academic information collections effectively.

4. Active Mechanism for Sharing Distributed Personal Academic Information Collections

Since the information resources in a researcher's personal academic information collection are also potentially valuable to some other researchers, making use of shared personal academic information collections could be an efficient way for researchers to obtain academic information resources [4].

However, in a conventional mechanism for sharing such collections, there is usually a lack of efficient functions for discovery of valuable collections, as well as an effective search method for them.

This research proposes an active mechanism for sharing distributed personal academic information collections. The proposed mechanism has the following two main functions:

(F1) Autonomous Discovery of Valuable Collections
Through this function, collections that are relatively valuable for a researcher can be discovered in an autonomous way, so that a search can be efficiently limited in the valuable collections. Col-AIRs (Collection AIRs) are used to realize this function. There is one Col-AIR in each collection, which automatically gathers information about the collection from all PerC-AIRs in the collection. Using this information, Col-AIRs of different collections can autonomously cooperate to discover valuable collections for any researcher that owns one of these collections.

(F2) Knowledge Enhanced Search in Shared Collections
The K_R enhanced search method proposed in Chapter 3 is used, with some modification according to the consideration of a collection's domain relevance for a researcher: if the main part of a collection's academic domain is overlapped with a researcher's research interest domain, the collection's domain relevance for the researcher will be a high value.
When a researcher submits a request to search a collection shared by another researcher, the items in the collection that are strongly related to some initial relevant items should also be considered as relevant ones to the request and be given some predicted relevance values. However, if the collection's domain relevance for the search requester is low, these predicted relevance values should decrease, reflecting that the items are of relatively lower possibility to be relevant.
PerC-AIRs are used to realize this function, with each collection's domain relevance being automatically estimated by cooperative Col-AIRs.

To verify the proposed mechanism, some experiments were conducted on its two main functions. The prototype system used in the experiments consists of three distributed collections, each of which has 1 Col-AIR and around 200 PerC-AIRs. The single-agent-based design model of AIR is used to build Col-AIRs and PerC-AIRs.

Table 1 summarizes the ranks of the collections' values estimated by Col-AIRs (EstVal), comparing to another list of reference values (in ApprVal). ApprVal is calculated based on the researchers' evaluation of all academic information resources in three collections, and therefore relatively ideal for ranking the values of the collections. The minute difference between the two ranks indicates that the proposed function (F1) is effective for autonomously discovering valuable collections for a researcher.

Table 1 Rank of Collections' Values

	$Val_a(B)$	$Val_a(C)$	$Val_b(A)$	$Val_b(C)$	$Val_c(A)$	$Val_c(B)$
# in EstVal	1	2	4	6	3	5
# in ApprVal	1	2	5	6	4	3
(Difference)	(0)	(0)	(1)	(0)	(1)	(2)

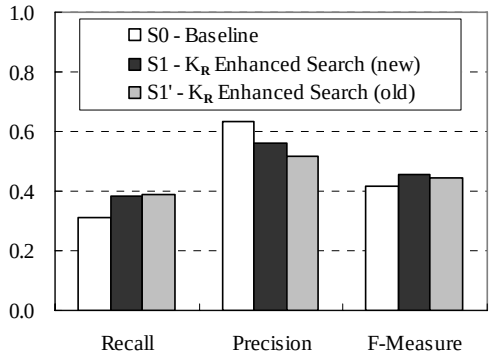


Fig.5 Performance of multiple search

Table 2 Comparison of S1 vs. S0 and S1'

%	*Recall*	*Precision*	*F-Measure*
S1 vs. S0	+ 23.4	- 11.5	+ 9.1
S1 vs. S1'	- 1.5	+ 8.9	+ 2.7

Furthermore, Fig.5 shows the average results of multiple searches conducted in the three collections. The proposed search method S1 showed a better performance than the

conventional method S0 did. Moreover, the consideration of a collection's domain relevance proved to be appropriate (S1 is better than S1') as shown in Table 2. This confirms that the proposed function (F2) is effective for searching shared collections in the distributed environment.

It can be concluded that the proposed mechanism can be used to support academic researchers to efficiently share the academic information resources in the distributed environment.

5. Conclusion

This research established a novel approach to actively support using distributed academic information resources. The proposed approach is to extend passive academic information resources into autonomous and active entities to realize an active support.

The achievements of this research include:

- An agent-oriented design method for AIR was developed.
- The proposed approach was applied and verified in two typical tasks to search and share the academic information resources based on the following functions.
  - Knowledge Enhanced Search Method for Personal Academic Information Collections,
  - Active Mechanism for Sharing Distributed Personal Academic Information Collections.

With the proposed approach used, it can be expected that researchers' burden of using distributed academic information resources will be reduced.

References

[1] A. Chervenak, I. Foster, C. Kesselman, C. Salis- bury and S. Tuecke, "The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets," Journal of Network and Computer Applications, Vol.23, No.3, pp.187-200, 2000.

[2] Tetsuo Kinoshita, "A Method for Utilizing Distributed Information Resources Effectively: Design of Active Information Resource," IEICE Technical Report, AI99-54, pp.13-19, 1999 (in Japanese).

[3] H. Bruce, "Personal, Anticipated Information Need," Information Research, Vol.10, No.3, 2005.

[4] C. Bailey, and M. Clarke, "Managing Knowledge for Personal and Organizational Benefit," Journal of Knowledge Management, Vol.5, No.1, pp.58-67, 2001.

[5] Baoning Li, Toru Abe, Tetsuo Kinoshita, “Design of Agent-based Active Information Resource,” Proc. 1st Int. Conf. Agent-Based Technologies and Systems (ATS'03), pp.233-244, Univ. Calgary, 2003.

[6] Baoning Li, Toru Abe, Yukio Iwaya, Tetsuo Kinoshita, "On a knowledge enhanced search method for personal academic information collections," Proc. the 2006 IEICE General Conference, DAS-1-1, 2006.

Baoning Li received his B.Eng. and M.Eng. degrees from Beihang University, Beijing, China, in 1994 and 1997, respectively. He also received a Ph.D degree in applied information science from Tohoku University in 2006.

Tetsuo Kinoshita is a professor of Information Synergy Center and Graduate School of Information Science, Tohoku University, Japan. He received his Dr.Eng. degree in information engineering from Tohoku University in 1993. His research interests include agent engineering, knowledge engineering, knowledge-based systems and knowledge-based systems. He received the IPSJ Research Award, the IPSJ Best Paper Award and the IEICE Achievement Award in 1989, 1997 and 2001, respectively. Dr. Kinoshita is a member of IEEE, ACM, AAAI, IEICE, IPSJ, JSAI, and Society for Cognitive Science of Japan.