

Research on Digital Library Personalized Information Service Model Based on Agent Model

Xu Yan
Xi'an Physical Education University
Xi'an, Shaanxi, China, 710068
guilinxuyan@yahoo.com.cn



ABSTRACT: Taking the personalized information needs of users as the research object, this article introduced the individualized service process of digital library, established the Agent model, built multivariate function by combining with the user preference model, and expounded the process of personalized information acquisition, analysis and push. We could conclude that digital library personalized information service system operated coordinately between each module, and had a complete chain of relationships. It could discovery and track information resources independently, effectively solve the problem of information overload, and realize the optimization of user information needs.

Categories and Subject Descriptors:

I.2.11 [Distributed Artificial Intelligence]: Multiagent Systems; **H.3.7 [Digital Libraries]**

General Terms:

Information Service

Keywords: Digital Library, Information Service

Received: 21 July 2013, Revised 25 August 2013, Accepted 2 September 2013

1. Introduction

Along with the development of the virtual information system, the user can easily get information, but is difficult to obtain accurate information, especially for scientific research clients and academic users. They urgently need the information service with personality, high quality which needs the library to provide specialized and personalized service according to the different requirements of the users. Personalized information service arises at the historic moment. It can actively provide users with the personalized information based on the user's interests

and hobbies, let the reader quickly and accurately find needed information from the vast amounts of information, and learn autonomously by using library resources [1]. Under the conditions of personalized information service, the user can make the corresponding information according to their requirements, and digital library can also actively service the users and improve the service quality according to user's needs. However, personalized information service needs a robust technical support. Agent is used to design personalized service model of the library. Taking the user as the center, it researches the user's behavior and interest and searches, organizes, selects and establishes a targeted, user-oriented personalized service mechanism. The library actively searches knowledge for the reader instead of providing information passively, which has improved the utilization of library resources [2].

There are many researches on personalized information services of digital library in the world. On the personalized push service and push system of the digital library, for example, My Library system put forward by Cornell university includes three service contents: Linksmen, My updates and Contentents [3]. "My library" system developed by zhejiang university library includes such functions as bookmarks, custom library digital resources, the latest information, links, search engine, and custom WEB page style. For the research of personalized recommendation service, the typical research projects at abroad are: ihrnina, Cite. The Seer, Fab [4]. Domestic research of this area is "digital library personalized recommendation system" developed by School of information, renmin university of China and library. Adopting the Agent, this article designs personalized service model, and takes the user as center to research the user's behavior, interests and habits, and researches the personalized information service model of digital library based on Agent.

2. The Process of Personalized Service System in Digital Library

The contents of personalized information service mainly includes three modules: the customization and retrieval of personalized information, and personalized service interface. Therefore, personalized information service system is a structure which constantly circulates, and mainly has three modules: the acquisition of user information, processing of personalized information and personalized collection management and scheduling which is shown in figure 1:

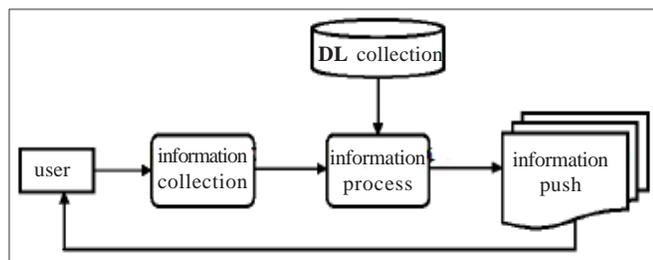


Figure 1. The process of personalized service system

Digital library obtains the user's personalized information via the custom, and understands the needs of users, and then provides users with more accurate information service. That is the process of information acquisition; analyze the user's habits, interests and the information demand, and establish accurate profiles. At the same time, keep track of users' condition, adjust the content and view of the digital library service in time, realize synchronous updating between user requirements and information service. This is the process of personalized information; Personalized collection includes information push, mining and prediction. Information push usually adapts the modeling method of user preference. In view of the new users who have set the needed information, it sends the online information to the user regularly or irregularly by push technology, and provides convenient conditions for the user. Information mining technology deeply analyzes the information that the user needs, and then classifies the information based on the characteristics of information content to improve the automation level of information retrieval. Information prediction analyzes the previous information, then predicts the information point and direction of future information. And then sends the predicted information to the user.

3. Establishment of Agent Model

3.1 The Conception of Agent Technology

Agent is a concept that develops in the field of artificial intelligence (AI) [5]. It can continue a role of a software entity which has the capability of perception, problem solving and communicating with the outside world. It can automatically search the information that the user may be interest in according to the user-defined criteria, and passes it to the specified place according to the specified time. It becomes the mediation of access to resources. Agent technology is a branch of distributed computing,

and is widely used in business, manufacturing, finance, e-commerce, etc.

3.2 The Function of Agent Model

Agent technology provides the technical support for the personalized information service of the library. It has such characteristics as autonomy, sociality, activity, learning, and intelligent. It can search the network information resources according to user's request and preferences, and then filter, analyze and mine the information, finally send to the user according to the priority of the correlation. because of its autonomy, it can substitute the user to get access to the information according to user personalized information demand without user intervention, and then actively push the information to the user which realize the fundamental change of service manner.

3.3 The Principle of Operation

Firstly, the user submits their information needs to intelligent retrieval Agent.

Secondly, Agent starts the retrieval function, analyzes and understands the user's information demand, automatically goes into the digital library, talks to library server, and retrieves, analyzes and processes the user interface.

Thirdly, optimize the results according to the needs and the way of thinking of users.

Finally, send the optimized results to the user.

Intelligent Agent is the latest software design pattern which is beyond the technology of client/server oriented system. According to the design pattern of Agent intelligent retrieval Agent designs the information access system which can automatically retrieve the user interface, database of digital library and the information resources according to the characteristics and the way of thinking of information users to adapt to the environment of the library.

3.4 The Establishment of Agent Model

The abstract model of Agent can be expressed as: $Ag: RE \rightarrow Ac$, of which the set of the external environment condition is: $E = \{e, e', \dots\}$, the action set of Agent is $Ac = \{a, a', \dots\}$, R is the performance set of state change of E and Ac . According to the change of environmental condition, the model maps a state to action, and makes the motor output act on the environment, to achieve a closed loop. The specific process is shown in figure 2.

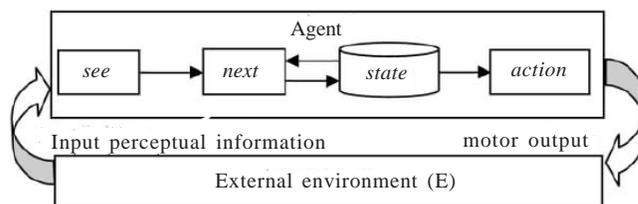


Figure 2. The abstract model of Agent

Starting from the initial internal state i_0 , Agent observes the external environment state e , generates a perception

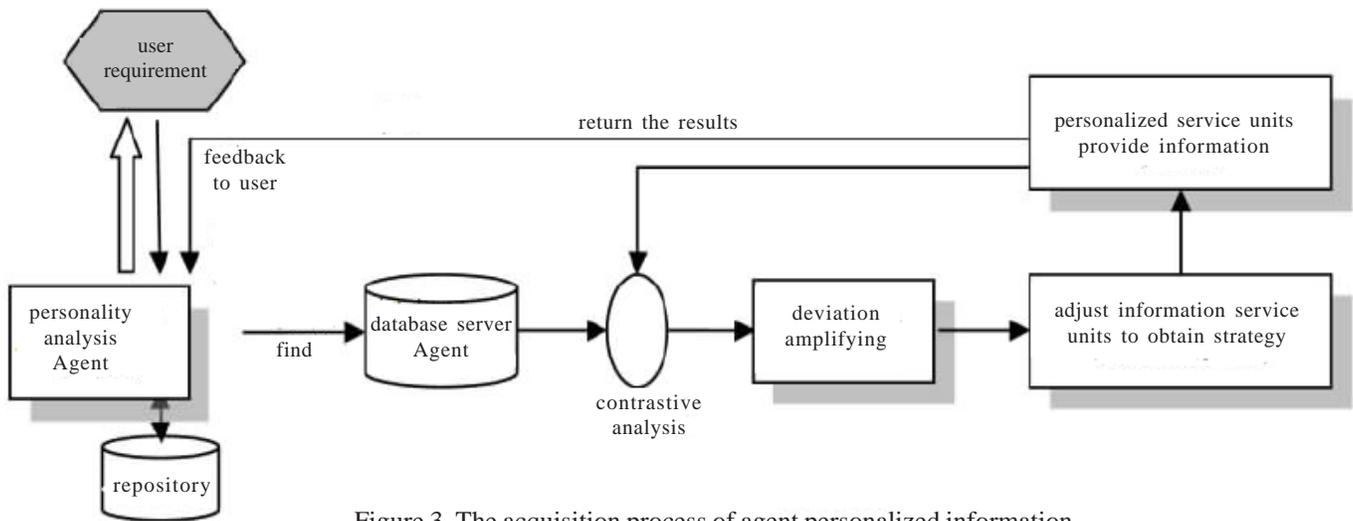


Figure 3. The acquisition process of agent personalized information

function $see(e)$, and updates the internal state of Agent through the function $next$, and then changes to $next[i0, see(e)]$, finally compares with state library. Agent selects an action r through the action $\{next[i0, see(e)]\}$, where r is alternating sequence in the environmental condition e and $a: r: e0 \xrightarrow{a_0} e1 \xrightarrow{a_1} e2 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} eu$, where a is the body of an action. After the action, Agent continues to percept the outside world by see , update the state through $next$, and select the action performed through action.

4. The Analysis of Agent Model Function

4.1 The Acquisition Process of Agent Personalized Information

The key of personalized service is the acquisition of the user's dynamic information. There are two ways to obtain the information for Intelligent Agent: one is to regularly poll Web log, and regularly collect user's information; the other is to track the Agent in real time, finds the changes and acts timely to analyze the user's personalized information. The specific process is shown in figure 3.

According to user requirements, personality analysis Agent connects with the database Agent that can store user information, entrusts them to find the information which matches the model, judges through the knowledge base related rules to extract the personalized information service, and sends results to personality analysis Agent,

and then adjusts the weights of the algorithm, finally feedback the results to the user[6]. When the user submits the application of "digital library", for example, "retrieve class Agent" in personality analysis Agent will add the "digital library", and gives a certain weight, and constructs a structure of personalized information by reasoning algorithm. Then retrieve the information around the personalized information structure, extracts the personalized information and then feedbacks the analysis to personality analysis Agent. The result includes "digital library", and system forecast information, such as "digital library", "distributed library", "digital library construction", etc.. After the retrieval, adjust the corresponding weights of knowledge reasoning algorithm.

4.2 The Analysis Process of Personalized Information

Personalized processing module is the core of the whole system, which is responsible for the normal operation of the whole system. It mainly processes user demand, extracts the content and translates into the information that can identify by the machines. Its structure is as follows:

User information collection Agent mainly collects data from the subscriber interface module, and classifies simply, and then sends the collecting information to the personality analysis Agent which is a key part of the personalized processing module. According to the corresponding rules of knowledge base, it uses reasoning algorithms to process the collecting information, separates out the

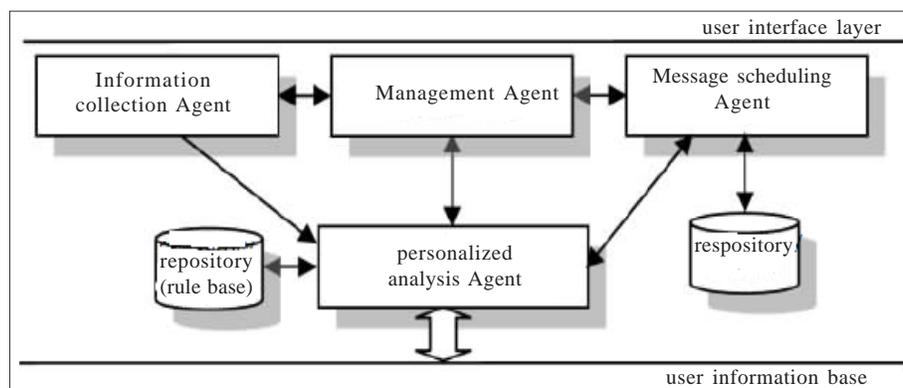


Figure 4. Analysis process of personalized information Agent

personalized content, such as the user's personal information, journal website that user often visits, the keyword used, the user's IP address and frequency of access, etc. When the processing is completed, put the results into the information database or pass it to the information presentation module, and change the knowledge base. Knowledge base is a repository which mainly saves the database Agent needed, processing results, and the data send back to the user after processing [7]. At the same time, the knowledge base is also a rule base and Agent can analyzes, judges according to the rules. Information scheduling Agent schedules the database according to the character parameter through precess of personality analysis, and selects the appropriate information for the user, meanwhile, feedback the user's activity information to the personality analysis Agent, and reprocess by the personality analysis Agent. Management Agent is responsible to coordinate and manage other Agents.

4.3 The Push of Personalized Information of Agent

The main function of push dule is to recommend information to users after the process of recommending module according to the previous information. The user preference modeling in personalized recommendation system of digital library describes the integration of information requirements from submission to cancellation, including dynamic renewal to accurately reflect the user's preference. User preference model can be formalized as a triple:

$$P: P = (M, F, \psi)$$

Of which $M = \{M^{(0)}, M^{(1)}, \dots, M^{(M)}, \dots, M^{(n)}\}$,

$$N = \{F^{(1)}, F^{(2)}, F^{(M)}, \dots, F^{(n)}\},$$

$\psi: M \times F \rightarrow M, \{M^{(M+1)} = (\psi^{M(i)}, F^{(i+1)}), i = 0, 1, \dots, n\}$

Here, M is the state set produced in the process of use of the user preference model; F is the feedback information set collected for the user preference model, and its role is to drive the updating process; ψ represents the renewal function of the user preference model.

Where, $M^{(0)}$ is the initial state of the user model; $M^{(i)}$ means user preference mode updated i times; $F^{(i)}$ is the information of updating $M^{(i+1)}$ after feedback the i th times. The formalized description reflects that the user preference model is updating constantly, and synchronously falls the user's preference [8].

In view of the characteristics of user in digital library recommender system, this paper raises the question that combines the short-term and long-term interests. Set the long-term preference got based on their age, professional background, education level, etc; set the analyzed preferences as the short-term preferences, such as the classified preference, the sample documents and query, and so on. In the final consideration of user preferences, we determine the preference model $\alpha M^L + (1 - \alpha) M^S$, where M^L is the eigenvector of long-term preference, M^S

M^S is the eigenvector for short-term preference, α is the parameter between (0 ~ 1), and adjusts the influence of the short-term and long-term interests .

Describe classification system T as a tree structure: $T = (C, R)$, where $C = \{c1, c2, \dots, ci, \dots, cn\}$, and it is the set that belongs to all categories in the classification system, ci is the i category; $R = R = \{\langle ci, cj \rangle \in r \mid 1 < i < j < n\}$, which means hyponymy among the classifications in C and $\langle ci, cj \rangle \in r$ means ci is the classification of cj .

Based on the rule-based, we conclude that the user may be interest in some category, such as ci, cj, ck , etc, and $\{ct1, ct2, \dots, cti, \dots\}$ can be applied to ci . Thus M^L can be shown as

$$\begin{aligned} M^L &= ci \cup cj, \dots \cup ck, \\ &= \{ct1, ct2, \dots, cti\} \cup \{ct'1, ct'2, \dots, ct'i\} \cup \dots \\ &= \{ct1, ct2, \dots, cti, \dots, ct'1, ct'2, \dots, ct'i\} \end{aligned}$$

As to the condition that the user selects multiple categories, we can consider according to whether the preference category is child-parent or sibling when we establish the user preference model in the user preference model. In the condition that the user selects multiple preference classification, the initial preference model M^S is expressed as:

$$\begin{aligned} M^L &= \alpha ci \cup \beta cj \\ &= \alpha \{ct1, ct2, \dots, cti\} \cup \beta \{ct'1, ct'2, \dots, ct'i, \dots\} \\ &= \{\alpha ct1, \alpha ct2, \dots, \alpha cti, \dots, \beta ct'1, \beta ct'2, \dots, \beta ct'i, \dots\} \end{aligned}$$

Where α and β are the weights of preference of sibling classification.

As to the subclassing, we can give an attenuation factory ($0 < \gamma < 1$) shown as follows:

$$\begin{aligned} M^S &= \alpha \gamma \{cj \mid \langle ci, cj \rangle \in r\} \cup \beta ci \\ &= \{\alpha \gamma ct1, \alpha \gamma ct2, \dots, \alpha \gamma cti, \dots, \beta ct'1, \beta ct'2, \dots, \beta ct'i, \dots\} \end{aligned}$$

Recommended results can be recommended to the user by "push" system (push), or obtained through the "pull" (pull). In the traditional Internet services, information transmission was processed according to the Pull mode, and the serve provided by servicer is passive. The user himself finds information. The Push mode provided by Agent technology is active information service network, is the information for the user, the server for the user's information, that is information finds users. The server gets the user's demands and pushes the information retrieved. Agent technology service not only conveys information to users, but also can push users the updating and dynamic information according to user's preset information channels and requests to realize the real personalized information service[9]. The realization of the individualized information service is to use intelligent Agent technology to comprehensively mine the user's interests, and automatically recognize and extract according to the user specific information object template, and then pass

the filtered results to the user in accordance with user requirements. Finally realize the circulation of the whole service system.

5. Results and Discussion

From the above analysis, the whole personalized information system of digital library mainly composes of three modules, and the module is close related with each other, which has formed a complete chain of relations. The whole system model is based on the user, and takes the coordination between every link as the proof, which fully realized the personalized service of digital library. At present, the digital library personalized service is still in infancy, but as the digital library develops to networking, sharing, users doesn't just face with a library, but a huge net-library based on a number of libraries. It is difficult for user to obtain the accurate information. The library personalized service and resource discovery and tracking processed by the personalized service system independently, can effectively solve the problem of information overload, and realize the optimization of information needs. Design the library personalized service model with Agent, has high guidance in researching and realizing the library personalized service system, and is a good try is to realize the wide-area digital library personalized service.

However, in terms of theoretical research and technology, the Internet oriented digital library personalized active information service is a system engineering with multi-interdisciplinary, and needs the integrated application of kinds of theories and technologies, such as information retrieval, artificial intelligence, data mining and so on. Therefore, libraries of all types and at all levels should deeply researches the related problems on the basis of the existing service, especially the applied research of the technologies, such as semantic Web, ontology, net, ambient intelligence, etc [10]. It can solve the problems, such as how to make use of ontology technology to accurately obtain the user preference information, how to carry out precise retrieval of information resources, how to semantically match the user preferences and information resources, how to use semantic grid to solve contradictions of the scale and efficiency of the personalized active information service and how to effectively carry out individualized active information service

in the intelligence environment, etc. Obviously, only the effective technology integration, and good solution to these problems, can make the user as the organizer of information resource, meanwhile, uses the information efficiently and enjoys convenient service.

References

- [1] Lin, Ning. (2012). Research on the individual service model of the digital library on the basis of the intelligent Agent. *The Journal of the Library Science in Jiangxi*. (3) 77-81.
- [2] Zhenxing, Hou., Hongyan, Cui. (2013). Research on personalized active information service model of digital library. *Information Science*. (3) 35-39.
- [3] Weixin, Xu. (2013). Construction of personalized information service system of digital library. *Journal of Nanchang College of Education*. (2) 187-188.
- [4] Yufeng, Zhang ., Chuangye, Yan . (2010). Research on the model of personalized information service based on agent. *Information Science*. (5) 56-59.
- [5] Yongxian, Li., Pengwei, Zhu. (2011). Research on personalized information push service model based on Agent. *Information Technology*. (7) 55-58.
- [6] Xiquan, Huang. (2006). The Modeling of User Profile in the Digital Library Recommender System. *Information Science*. (1) 28-29.
- [7] Xiulian, Chen. (2008). Agent induviduation capacity information searches system be applied in figure library science flat roof. *Science and Technology of West China*. (1) 94-96.
- [8] Xiaoyong, Lin. (2006). Realization of personalized information service based on agent technology. *Journal of Library Science*. (4) 124-125.
- [9] Jun, Zhang. (2013). The application of ElGamal encryption technology to the information security of digital library. *IAES Journal of Electrical and Electronics Engineering*. 11 (12).
- [10] Genjian, Yu., kunpeng, Weng. (2013). Intrusion detection system and technology of layered wireless sensor network based on Agent. *IAES Journal of Electrical and Electronics Engineering*. 11 (8) 4238-4243.