Personalized Knowledge Acquisition through Interactive Data Analysis in E-learning System

Zhongying Zhao
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
Graduate School of Chinese Academy of Sciences, Beijing, China
Email: zy.zhao@sub.siat.ac.cn

Shengzhong Feng[1], Qingtian Zeng[2], Jianping Fan[1], and Xiaohong Zhang[1]
[1]Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
[2]College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, China
Email: sz.feng@siat.ac.cn, qtzeng@sdust.edu.cn ,{jp.fan, xh.zhang}@siat.ac.cn

***Abstract*—Personalized knowledge acquisition is very important for promoting learning efficiency within E-learning system. To achieve this, two key problems involved are acquiring user's knowledge requirements and discovering the people that can meet the requirements. In this paper, we present two approaches to realize personalized knowledge acquisition. The first approach aims to mine what knowledge the student requires and to what degree. All the interactive logs, accumulated during question answering process, are taken into account to compute each student's knowledge requirement. The second approach is to construct and analyze user network based on the interactive data, which aims to find potential contributors list. Each student's potential contributors may satisfy his/her requirement timely and accurately. Then we design an experiment to implement the two approaches. In order to evaluate the performance of our approaches, we make an evaluation with the percentage of satisfying recommendations. The evaluation results show that our approaches can help each student acquire the knowledge that he/she requires efficiently.**

***Index Terms*—E-learning, knowledge acquisition, knowledge requirement, potential contributor**

I. INTRODUCTION

Personalized support becomes even more important, when e-Learning takes place in open and dynamic learning and information networks [1, 2, 3]. Personalized knowledge acquisition, is one of the most important phases for realizing user-adaptive or personalized e-learning systems. It involves several aspects. The first is to acquire what knowledge the student requires. The second problem is to find who can provide the related knowledge to satisfy student's requirement. Others include how to offer the knowledge, in what forms and what time. In this paper, we aim to solve the first two problems.

Interactive Question Answering (QA) system, which can be seen as virtual seminar, has been embedded into an e-learning system to improve learning performance [4, 5, 6]. During this system, students communicate their knowledge in the form of posing questions, selecting questions to answer and browsing others' questions and answers (Q&A). The e-learning system can store all these interactive logs into the data base in the form of question table, answer table and user table. All of these historical data contain a tremendous amount of information about the users' requirements and relations.

In this paper, we propose two approaches to achieve personalized knowledge acquisition. The first approach aims to mine what topics (what kind of knowledge) the user requires and to what degree. All his/her interactive logs, including posing question, answering questions and browsing answers, are taken into account to compute the knowledge requirement. This approach, however, does not imply whom a student can turn to when he/she has knowledge requirements. And the tightness of relations between students is also not reflected although it is very important for users to acquire knowledge. Thus, our second approach is to construct a user network for all the users in e-learning system. The user network describes each student's potential contributors list and the relation strength, which can improve the personalized knowledge acquisition.

Compared with others' excellent research results, the work in this paper is a supplement to achieve personalized E-learning, especially in personalized knowledge acquisition.

The remainder of the paper is organized as follows. In section 2, we discuss the related work. Section 3 describes the framework for personalized knowledge acquisition. Section 4 presents the approach for mining user's knowledge requirement. The construction and analysis of the user network is addressed in section 5, which aims to find the potential contributors. Section 6 combines our two approaches to achieve personalized knowledge acquisition. To evaluate our approaches, we design the experiment and evaluation in section 7. Section 8 concludes the whole paper and discusses the future work.

II. RELATED WORK

doi:10.4304/jcp.5.5.709-716

To realize a user-adaptive or personalized e-learning system, user model and modeling are two of the key problems [7]. User model can be built based on user's behavior, the contents of a web page or both. A human behavior based user model can be learned by observing the user's actions such as web log file, path, clicking and downloads frequency [8-10]. A divisive hierarchical clustering (DHC) algorithm to group terms is proposed by Kim H.R.& Chan P.K.[11]. Dwi H. Widyantoro et al. propose a three-descriptor model to represent a user's interest categories and an adaptive algorithm to learn the dynamics of the user's interests through positive and negative relevance feedback [12].Trajkova and Gauch build user profiles automatically from the web pages and focuses on improving the accuracy of the user profile based on concepts from a predefined ontology [13]. Liu lu and Wu lihua model user's interest and value related characteristics in recommender systems [14]. A vision-based approach to detect user's interests on web pages is proposed in literature [15]. Considering that users' interests may be inferred from what they read and how they interact with documents, Rajiv Badi et al present models for detecting user interest based on reading activity [16].

With the development and application of Web technology, interactive question answering system has also become an active research area. An interactive question answering system named BuyAns is developed for personalized E-learning implementation [4]. HITIQA is designed to allow intelligence analysts and other users of information systems to pose questions in natural language and obtain relevant answers or the assistance [17]. With fully-implemented interactive QA system named FERRET, Sanda Harabagiu et al achieved a surprising performance by integrating predictive questions into the context of a QA dialogue [18]. Chun-Chia Wang et al proposes a repository-based Question Answering system for collaborative E-learning [19].

Although many Web-based learning techniques have been proposed to assist adaptive/personalized Web-based learning, few researches have attempted to acquire user's knowledge requirement and potential contributors. Personalized knowledge acquisition, including the acquisition of knowledge requirement and potential knowledge contributors, plays very important role in promoting personalized learning efficiency. We have done a lot of work in user modeling, including interest mining [20, 21], knowledge requirement acquisition [22]. This paper is an expanded version of the conference paper [21]. Some modifications of our QA system are described in this paper. Another approach for capturing the potential knowledge contributors is proposed, which can accelerate the personalized knowledge acquisition combining with our previous works.

III. FRAMEWORK

As illustrated in Fig.1 is the framework to describe the whole process for personalized knowledge acquisition within e-learning systems. This framework can be embedded into any e-learning system as an assistant component to realize user-adaptive or personalized learning or instruction. The main components contained in the framework include the following:

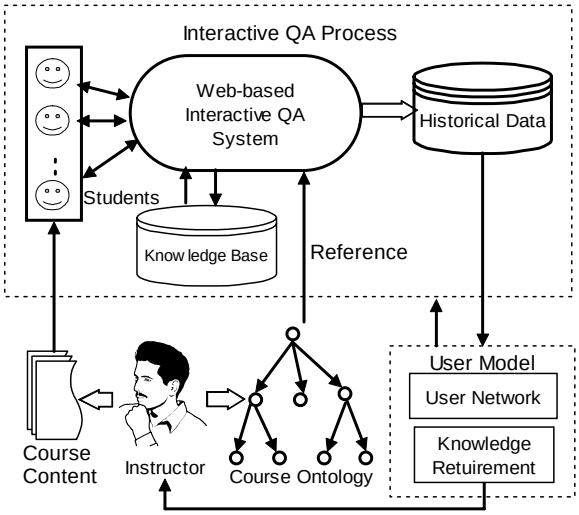

Figure 1. The framework for personalized knowledge acquisition within e-learning systems

(1) Course ontology (or concept hierarchy): The course ontology is predefined by the instructor with suitable granularity and scale. It presents the structure of the course content and provides a reference for QA board structure. An example for Artificial Intelligence (AI) course ontology is described in Fig.2 and [20]. In this course ontology, there are 8 leaf terms (T1-T8), which corresponds to the QA boards.

(2) Interactive QA process: Based on the predefined course ontology, the corresponding board structure of the user-interactive QA system can be generated. Within their favorite boards, users can post their most urgent questions, browse their favorite answers, and select others' questions to answer. All these interactive QA data can be recorded and accumulated to historical data base to acquire student's knowledge requirement and potential contributors. For the questions posed by students, some of them are answered by knowledge base. And others require students or instructor to give answers, which will be discussed in section IV.

(3) User modeling: It is a critical part for personalized knowledge acquisition. This process includes acquiring each student's knowledge requirement and constructing user network. According to the historical data, we can compute each student's knowledge requirement in every question. Then the mapping relation between question and topic is used to compute his/her knowledge requirement in each topic. The historical interactive data are also used to construct user network. The user network describes the relations between students during the QA process, which can be used to find the potential contributors for each student.

User modeling is done once a week. Then each student's knowledge requirement and user network can be used to guide the next interactive QA process and achieve personalized knowledge acquisition. Furthermore, based on student's knowledge requirement model, the

instructor can adjust her/his teaching materials and offer personalized helps to each student.

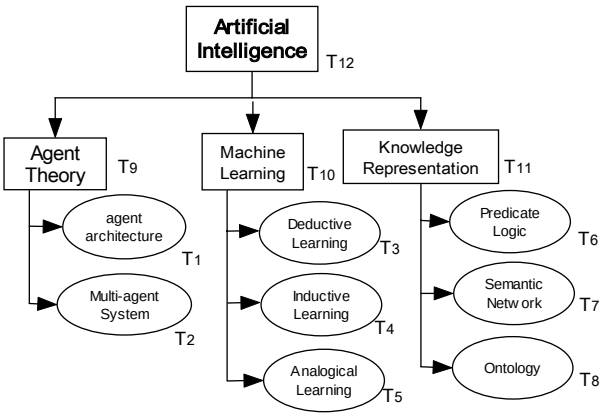

Figure 2. The concept hierarchy for AI course ontology

IV. MINING USER'S KNOWLEDGE REQUIREMENT

During the interactive QA process, students can propose questions, browse others' answers and select their favorite questions to answer. While browsing the answers given by others, the student is asked to submit his/her rate through clicking the button "satisfying" or "not satisfying". The button "satisfying" corresponds to 1 while the button "not satisfying" corresponds to 0.

For the question posed by a student, he/she has to wait for others to answer. In order to shorten the time spent in waiting for answers, we classify the questions into $PlainQ$ and $Probe\text{-}intoQ$. $PlainQ$ is answered by knowledge base, while $Probe\text{-}intoQ$ is answered by other students or their instructors.

A. Classification of questions

(1) $PlainQ$ refers to the question that has an exact or accurate answer, such as querying a fact or a definition. Based on the questions proposed by students, we extracted the corresponding pattern. The pattern set for $PlainQ$ is given as follows.

$PSplain = \{$ *What is* (*the definition of*) [*concept*]?
Which | *What* [*areas* | *fields*] *does* [*a concept*] *cover* | *include* | *contain*?
When does the [*concept*] *be proposed* | *presented* | *given* (*in*) (*some fields*)?
Who proposes | *presents* | *gives the* (*concept of*) [*concept*]?...$\}$,
where
[] *represents compulsory field*, () *represents optional field*.

If a question satisfies one pattern, it will be submitted to knowledge base. And the answer will be given by knowledge base immediately. With the enlarging of knowledge base, new patterns are being added to the pattern set in order to satisfy user's requirement.

(2) $Probe\text{-}intoQ$ refers to the question that can be analyzed and answered from many different perspectives. E.g. What conclusions can be drawn ……? How to understand……? What are the reasons for……?

For this kind of question, it is often dealt with as the following. (1) It is answered by some helpful students. For this circumstance, the questioner has to spent long time waiting for answers. (2) The instructor is charge of every question. Since the instructor is the expert for this course or domain, his/her answers are often accurate and make questioner satisfied. A disadvantage is that, however, with students and questions increasing, the instructor will become the bottleneck. (3) All the others are encouraged to answer the proposed questions. But only some/few of them can satisfy the questioner's knowledge requirement. That is, a lot of junk information/answers are made during this process. To improve the interactive QA efficiency and achieve personalized knowledge acquisition, we should find the potential contributors for each student. This will be discussed in section V.

B. Historical data

In the interactive QA system, students interactively exchange their knowledge by questioning, answering, and browsing. All these historical data can be stored in the form of user table, question table, and answer table. The follows are the attributes of each table.

TABLE 1. User table

UID	*OperType*	*QID/AnsID*	*Judge*

TABLE 2. Question table

QID	*QTopic*	*WhoPro*	*QType*	*NumAns*

TABLE 3. Answer table

AnsID	*QID*	*WhoAns*	*NBrow*	*NSatisfy*	*NNoSatisfy*

For the user table, UID is the primary key. $OperType$ represents the operation the student makes, and $OperType \in \{\text{proposing, browsing, answering}\}$. $QID / AnsID$ refers to the question or answer given by this student. The selection of QID or $AnsID$ is determined by $OperType$. $Judge$ means whether the student is satisfied with the given answer. $Judge = NULL,\ if\ \ OperType = \text{answering}$. Otherwise, $Judge = \begin{cases} 1, \text{if the students is satisfied} \\ 0, \text{if the students is not satisfied} \end{cases}$.

For the question table, QID is the primary key. $QTopic$ refers to the board that the question belongs to, and $Qtopic \in \{T_i \mid i = 1, 2, ..., 8\}$, T_i is the leaf term of course ontology. $WhoPro$ is the UID that proposes this question. $QType$ denotes the type of the question, and $QType \in \{PlainQ, Probe\text{-}intoQ\}$. $NumAns$ denotes the number of the answers given to this question.

For the answer table, $AnsID$ is the primary key. QID denotes the question that the answer belongs to. $WhoAns$ represents the UID that gives this answer. $NBrow$ means the number that the answer is browsed. $NSatisfy$ is the total number of $Judge = 1$ for this answer. Similar to the above, $NNoSatisfy$ is the total number of $Judge = 0$ for this answer. And according to the description of interactive process in this section, we can derive that $NBrow = NSatisfy + NNoSatisfy$.

C. Acquiring user's knowledge requirement

From the historical data, we can get each student's interactive logs through operations on the table. In order to compute each student's knowledge requirement from the interactive logs, we firstly define $QASet_i$ of student i.

$$QASet_i = \{(Q_j, AnS_j, p_{i,j}, f(B)_{i,j}, f(A)_{i,j}) \mid j = 1,2,...,n\},$$

where

(1) Q_j is a question.

(2) AnS_j is the set of answers to Q_j.

$$|AnS_j| = \begin{cases} 0, \ if \ Q_j \text{ is a question without answer.} \\ 1, \ if \ Q_j \in PlainQS. \\ m\,(m>1), if \ Q_j \text{is a } Probe\text{-}intoQ. \end{cases}$$

(3) $p_{i,j} = \begin{cases} 0, if \ Q_j \text{ is proposed by student } i. \\ 1, if \ Q_j \text{ is not proposed by student } i. \end{cases}$

(4) $f(B)_{i,j}$ denotes the number of answers browsed by student i to Q_j.

(5) $f(A)_{i,j}$ denotes the number of answers given by student i to Q_j.

By common sense, the student who posts question Q_j usually requires the corresponding knowledge more urgently than those only browsing. For question Q_j, the knowledge requirement of student i is defined as follows:

$$KR_i(Q_j) = \frac{1}{\pi}\arctan(\alpha p_{ij} + \beta(f(B)_{ij} - Meanf_{B_i}) + \gamma f(A)_{i,j}) + \frac{1}{\pi}\arctan(\beta Meanf_{B_i}) \quad (1)$$

where

(1) α, β, γ are parameters and $\alpha > \beta > \gamma$,

(2) $Mean\ f_{B_i}$ is the average number of browsing of student i

Since the structure board of the QA system is generated according to the course ontology and all the questions and answers are distributed in each board. From the question table (Tab.2), we can clearly know the topic that each question belongs to. Then we can compute each student's knowledge requirement about each leaf term of course ontology, i.e. student's knowledge requirement about each sections of course content.

$$KR_i(T_k) = \frac{1}{m}\sum_{j=1}^{m} KR_i(Q_j), \text{ where } (Q_j, T_k) \in MR, \ m = |\{Q_j \mid (Q_j, T_k) \in MR\}| \quad (2)$$

V. CAPTURING POTENTIAL CONTRIBUTORS

A. User Network Construction

a. All questions without users involved.

From the data base described in section Ⅳ, we can construct the mapping relations between all the questions and answers (Fig.3).

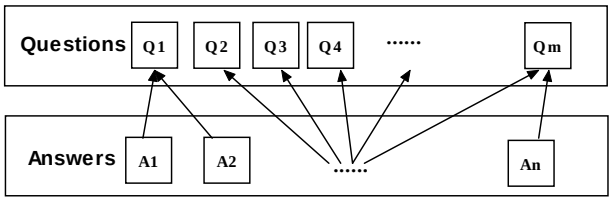

Figure 3. Mapping relations between questions and answers

b. One question with users involved.

In order to construct user network, we firstly consider the situation for one question Q_j. From the database, we can select all the users related to this question, including those who propose/answer this question and those who browse the answers to this question. To intuitively represent the relations between users, questions, answers, we have defined the following rules:

(1) Circles are employed to denote users while squares are used to denote questions or answers;

(2) There is a direct edge from U_i to Q_j if user U_i proposes the question Q_j;

(3) There is a direct edge from U_i to A_j if U_i gives the answer A_j;

(4) There is a direct edge from A_j to U_i if U_i browses the answer A_j and considers it good;

According to the database and the rules defined above, we can construct a graph for the question Q_j. Two examples are shown in Fig.4 and Fig.5. Fig.4 represents two special situations. One is that no one has given answer to Q_j proposed by U_p. The other is that the Q_j posed by U_p is a *PlainQ* and it has been answered by data base. In Fig.5, U_p proposes Q_j and then U_1, U_2, and U_3 give 4 answers in total. U_2 and U_4 browse U_1's answer and judge it good. U_6 browses the answers given by U_2 and U_3 and consider them valuable. U_5 also browses U_3's answer and gives a good estimate.

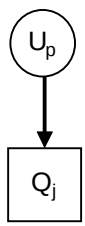

Figure 4. The question Q_j without answer

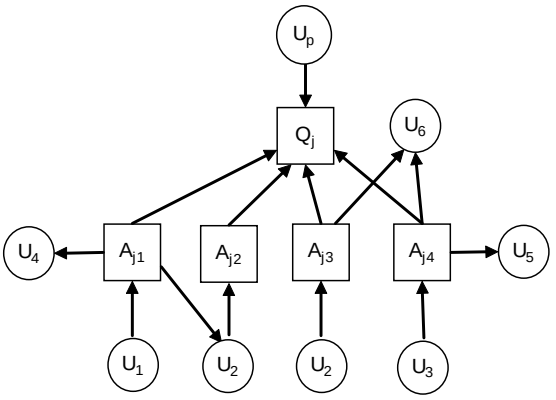

Figure 5. The question Q_j with many answers

To clearly describe the relations between users involved in one question, it is necessary for us to delete the question(s) and answers (Q&A). We present the following steps for Q&A elimination.

U_i walks to the Q_j along the direction of the edge and replace the Q_j if there is a direct edge from U_i to Q_j;

(2) U_i walks to A_j along the direction of the edge and replace the A_j if there is a direct edge from U_i to A_j;

(3) If there are N(N>1) direct edges from U_i to U_j, all the edges should be merged and marked a weight N.

(4) For other edges without weight, we assign 1 as the weight.

Following the 4 steps, we can eliminate all the questions and answers. Fig.4 is transformed to a isolated point named U_p. Fig.5 is converted to a weighted graph as illustrated in Fig.6. That is a simple user network since only one question is taken into account.

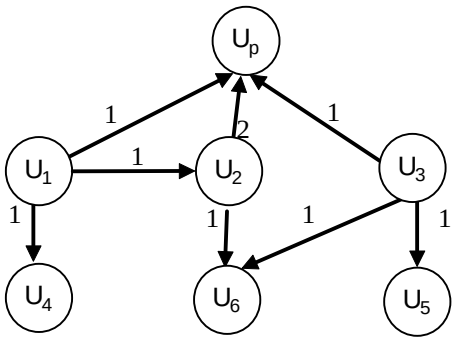

Figure 6. The user network for one question Q_j

To facilitate the computer processing, we adopt matrix to denote the graph above. The user network for Q_j is defined as $UM_{N*N}(T_k)^j$, where Q_j belongs to the topic T_k.

$$UM_{N*N}(T_k)^j = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{21} & & w_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{bmatrix}, \text{ where}$$

(1) N is the total number of users in the interactive QA system.

(2) w_{ij} is the weight marked on the directed edge from U_i to U_j.

c. One topic with users involved

The user network for all questions can be considered to be a merge problem, since we have described the construction of simple user network for one question above. For the topic T_k, we can easily construct the matrix, denoted by $UM_{N*N}(T_k)$.

$$UM_{N*N}(T_k) = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{21} & & w_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{bmatrix}$$

$UM_{N*N}(T_k)$ can be computed from the following formula:

$$UM_{N*N}(T_k) = \sum_{j=1}^{j=M} UM_{N*N}(T_k)^j, \quad (3)$$

where M is the number of questions on board T_k in QA system.

The following algorithm presents the whole process for constructing the user network.

Algorithm. Constructing user network for the interactive QA system

INPUT: historical data base (user table, question table, answer table)

OUTPUT: user network for each topic $UM_{N*N}(T_k)$

Step 1: define and initialize a N*N matrix $UM_{N*N}(T_k)$, $UM_{ij}(T_k) = 0$ (i,j=1,2, … ,N);

Step 2: for every question in topic T_k, do

If U_i give an answer to U_j, $UM_{ij}(T_k) += 1$;

If U_j browses U_i's answer and judge it good, $UM_{ij}(T_k) += 1$;

Step 3: output $UM_{N*N}(T_k)$

B. User network analysis

According to the construction and definition of user network, we can see that the $UM_{N*N}(T_k)$ is essentially a directed graph. As to $UM_{N*N}(T_k)$, w_{ij} refers to the weigh marks on the directed edge from U_i to U_j. It represents how many contributions U_i makes to U_j. And,

(1) If $w_{ij} \geq 1$, we say U_i is a potential contributor to U_j, and U_j is a potential beneficiary.

(2) $\sum_{j=1}^{j=N} w_{ij}$ represents how many contributions U_i makes to others.

(3) $\sum_{i=1}^{i=N} w_{ij}$ means the total gains that U_j obtains from others.

In (1) above, we consider U_i a potential contributor because it has made contributions to U_j in the past. It may help U_j in the future but not always. For the same reason, U_j is a potential beneficiary. Therefore, the user

network $UM_{N*N}(T_k)$ describes and quantifies the strength of relations between users under the topic T_k. Here, the relation refers to contributor – beneficiary.

From the *j*th column of $UM_{N*N}(T_k)$, we can get each U_j's potential contributors list by sorting U_i according to the w_{ij} decreasing.

VI. Personalized Knowledge Acquisition

From the first approach, we can obtain each student's knowledge requirement, including which topics they require, to what degree, who has the same or the similar knowledge requirement with each other. The second approach aims to find each student's potential contributors, which can help students acquire the knowledge that they require quickly and accurately. In this section, we combine the two approaches to achieve personalized knowledge acquisition.

In our interactive QA system, each student registered is assigned an agent as the assistant. The agent plays very important role in the process of personalized knowledge acquisition. It is in charge of the call of our two approaches once a week. And it maintains the user's knowledge requirement computed from the first approach. The potential contributors list is also maintained by the agent after sorting based on the decrease of w_{ij} (i=1,2, … ,N). Each agent can browse the board and recommend others' new questions and answers to its own user.

To achieve personalized knowledge acquisition, all the results from the first and second approach are taken into account. The higher student's knowledge requirement degree, the more we select from the potential contributors list. In this paper, we adopt the 5-3-2 rule. Take the topic T_j for example, U_i's agent selects K contributors from U_i's potential contributors list. And,

$$K = \begin{cases} 5, \text{ if } \delta_1 \le KR_i(T_j) \le 1 \\ 3, \text{ if } \delta_2 \le KR_i(T_j) < \delta_1 \\ 2, \text{ if } 0 < KR_i(T_j) < \delta_2 \end{cases}$$

That means, if $\delta_1 \le KR_i(T_j) \le 1$, the agent selects 5 contributors with the highest weight from U_i's potential contributors list. When U_i proposes a question on T_j, its agent will recommend it to the 5 users and invite them to give answers. Once a new answer appears, the agent notifies its user to browse. Meanwhile, the agent forward the corresponding link to U_i's potential beneficiaries in order to satisfy their implicit knowledge requirement. Then users browse the questions and answers and submit their evaluations. Al l these behaviors are recorded in the database, which will be used to compute or update each user's knowledge requirement and user network.

VII. Experiment and Evaluation

A. Experiment design and result

To evaluate the performance of our approaches, we conduct an experiment with AI course ontology. As to AI, we only select 3 chapters and 8 sections to simplify the experiment. The concept hierarchy is shown in Fig.2.

10 students majoring in the Artificial Intelligence are invited to attend this experiment. During the first four weeks of the experiment, we encourage every student to answer others' questions actively and helpfully. And all the students are encouraged to post their urgent questions on their favorite boards. During the interactive process, each student pays more attention to his/her favorite questions. Once the answer that a student browses satisfies his/her requirement, he/she clicks the "satisfy" button to submit the judge. Otherwise, "not satisfy" button is clicked.

After four weeks, we have collected and stored abundant Q/A historical data in the form of user table, question table and answer table. Then each student's agent calls our two approaches. The first proposed approach is applied to compute the knowledge requirement of each student shown in Fig.7. The second approach is invoked to construct user network for each topic. An example of user network for T_8 is shown as follows:

$$UM_{N*N}(T_8) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 5 & 1 & 0 & 8 & 0 & 0 \\ 0 & 4 & 0 & 0 & 1 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 1 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 5 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

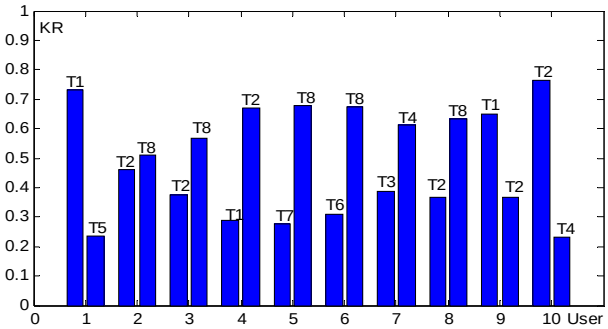

Figure 7. The knowledge requirement of each student

B. Evaluation

In this subsection, we make an evaluation to examine the performance of our two approaches. The performance is measured in terms of percentage of satisfying recommendations (Perc of SatRec). A satisfying recommendation is defined as a positive recommendation such that, after receiving and browsing it, student is satisfied with the recommendation and click "satisfy" to

submit his/her evaluation. Thus, the percentage of satisfying recommendations is the total number of recommendations divided by the number of satisfying recommendations.

We set $\delta_1 = 0.6$, $\delta_2 = 0.3$ in this evaluation and the potential contributors are selected based on the rule described in section Ⅵ. All the results computed from the first and second approach are taken into account to achieve personalized knowledge acquisition. According to the student's knowledge requirement $KR_i(T_j)$ computed from the first approach, the agent selects K contributors on top of the potential contributors list. When U_i proposes a question on T_j, its agent will recommend it to the K students and invite them to give answers. Once a new answer appears, the agent notifies its user to browse. Meanwhile, the agent forward the corresponding link to U_i 's potential beneficiaries in order to satisfy their implicit knowledge requirement. Then users browse the questions and answers and submit their evaluations.

If our approaches can achieve personalized knowledge acquisition, that means expressing student's knowledge requirement and helping him/her acquire knowledge precisely and quickly, the percentage of satisfying recommendations should be keep higher. The percentage of satisfying recommendations of the 10 students is calculated every day. After 60 days for the content recommendation, the percentages of satisfying recommendations are shown in Fig.8. We can see that all the percentages are stable and always higher than 0.7. Further more, the changing spectrum is narrowing down with time passing by. Thus, the experimental results show that the two proposed approaches realize personalized knowledge acquisition to some degree.

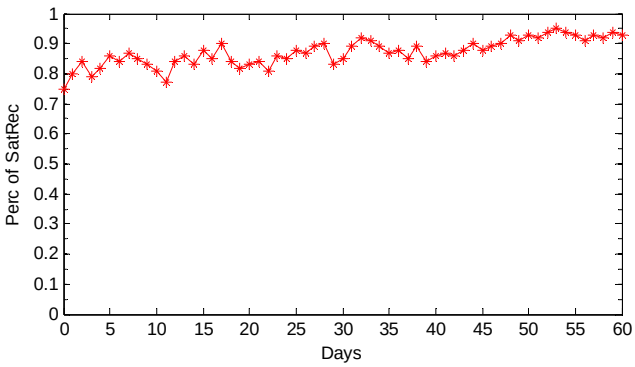

Figure 8. The Students' percentage of satisfying recommendations within 60 days

VIII. CONCLUSION

Personalized knowledge acquisition is very important for learning efficiency. In this paper, we propose two approaches to achieve personalized knowledge acquisition in interactive QA system. The first approach is to acquire the knowledge requirement of the users from their historical QA logs. It aims to mine what kind of knowledge or what topic is required by students. And it also quantifies the urgency of each student's knowledge requirement. The second approach presents the construction of user network based on students' interactive history. It aims to find the potential contributors and improve the personalized knowledge acquisition quickly and accurately.

An experiment is conducted to implement our proposed approaches, obtaining the knowledge requirement and potential contributors list of each student. The evaluation combined with the experiment results indicate that our approaches realize personalized knowledge acquisition in interactive QA system to some degree.

In our future work, we will extend or modify our interactive QA system, so that users can upload their resources for knowledge sharing. Then the potential beneficiary can learn from his/her potential contributors and acquire more knowledge that he/she implicitly requires.

ACKNOWLEDGMENT

This research is partly supported by National High-tech Research and Development Program of China (863 Program, No.2006AA01A114), National Science Foundation of China (No.60603090, 90718011 and 50875158), the Research Foundation of Shandong Educational Committee of China (No.J08LJ77), and the Taishan Scholar Program of Shandong Province. The authors would like to thank all the students for their efforts in our experiment.

REFERENCES

[1] Chen, C., Lee, H., and Chen, Y. Personalized e-learning system using item response theory. Computers & Education. 2005(44), pp.237-255.

[2] Dolog, P., Henze, N., Nejdl, W., and Sintek, M. Personalization in distributed e-learning environments. In Proceedings of the 13th international World Wide Web Conference on Alternate Track, 2004, pp.170-179.

[3] Thyagharajan, K. & N., R. Adaptive content creation for personalized e-learning using web services. Journal of Applied Sciences Research, 2007, 3(9), pp.828-836.

[4] Dawei Hu, Wei Chen and Qingtian Zeng et al. Using a user-interactive QA system for personalized E-learning. International Journal of Distance Education Technologies, 2008(6), pp.1-22.

[5] J. C. Hung, C.-S. Wang, C.-Y. Yang, M.-S. Chiu, and G. Yee. Applying word sense disambiguation to question answering system for e-learning. In Proceedings of the 19th international conference on advanced information networking and applications, 2005(1), pp.157-162.

[6] Wang, Y., Wang, W., and Huang, C. Enhanced semantic question answering system for e-learning environment. In Proceedings of the 21st international conference on Advanced Information Networking and Applications Workshops. 2007(2), pp.1023-1028.

[7] Brusilovsky, P., Sosnovsky, S., and Shcherbinina, O. User Modeling in a Distributed E-Learning Architecture. User Modeling, 2005(3538), pp.387-391.

[8] Kazunari Sugiyama, Kenji Hatano, Masatoshi Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. In Proceedings of the 13th international conference on World Wide Web, 2004, pp.675-684.

[9] Feng Qiu, Junghoo Cho. Automatic identification of user interest for personalized search. In Proceedings of the 15th international conference on World Wide Web. 2006, pp.727-736.
[10] Himanshu Sharma, Bernard J. Jansen. Automated evaluation of search engine performance via implicit user feedback. The 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2005, pp.649-650.
[11] Hyoung R. Kim and Philip K. Chan. Learning implicit user interest hierarchy for context in personalization. In Proceedings of Intelligent User Interfaces. 2003, pp.101-108.
[12] D.H. Widyantoro, T.R. Ioerger, and J. Yen. Learning user interest dynamics with a three-descriptor representation. Journal of the American Society of Information Science and Technology, 2001, 52(3), pp.212-225.
[13] J. Trajkova and S. Gauch. Improving ontology-based user profiles. In Proceedings of the Recherched' Information Assiste par Ordinateur (RIAO), 2004, pp.380-389.
[14] Liu Lu, Wu lihua. User modeling for personalized recommender system. Tsinghua science and technology. 2005(10), pp.772-777.
[15] Yin Liu, Wenyin Liu and Changjun Jiang. User interest detection on web pages for building personalized information agent. Advances in web-age information management. 2004 (3129), pp. 280-290.
[16] Rajiv Badi, Soonil Bae, J. Michael Moore et al. Recognizing user interest and document value from reading and organizing activities in document triage. In Proceedings of the 11th international conference on intelligent user interfaces. 2006, pp.218-225.
[17] S. Small, T. Liu, N. Shimizu, T. Strzalkowski. HITIQA: an interactive question answering system a preliminary report. In Proceedings of the ACL workshop on multilingual summarization and question answering. 2003(12), pp.46-53.
[18] Sanda Harabagiu, Andrew Hickl, John Lehmann and Dan Moldovan. Experiments with interactive question-answering. In Proceedings of the 43rd annual meeting on Association for Computational Linguistics. 2005, pp.205-214.
[19] Wang, C. C., Huang, J. C., Shih, T. K., & Lin, H. W. (2006). A repository-based question answering system for collaborative e-learning. Journal of Computers. 2006, 17(3), pp. 55-68.
[20] Yongquan Liang, Zhongying Zhao, Qingtian Zeng. Mining User's Interest from Reading Behavior in E-learning System. The 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD), 2007(2), pp. 417-422.
[21] Zhongying Zhao, Shengzhong Feng, Yongquan Liang, Qingtian Zeng, Jianping Fan. Mining user's interest from interactive behaviors in QA System, First International Workshop on Education Technology and Computer Science. 2009(12), pp.1025-1029.
[22] Qingtian Zeng, Zhongying Zhao, Yongquan Liang. Course Ontology-based User's Knowledge Requirement Acquisition from Behaviors within E-Learning Systems, Computers & Education, In Press, Corrected Proof, Available online 28 May 2009, doi:10.1016/j.compedu.2009.04.019.

Zhongying Zhao is born in Shandong Province, 1983. She is a PhD candidate both in Shenzhen Institute of Advanced Technology and Institute of Computing Technology, Chinese Academy of Sciences, China. Her research interests include data mining, parallel and distributed computing, high performance computing, and personalized e-learning.

Shengzhong Feng is a professor at the Shenzhen Institute of Advanced Technology, CAS. His research focuses on parallel algorithms, grid computing and bioinformatics. Specially, now his interests are in developing novel and effective methods for digital city modeling and application. Before came to SIAT, CAS, he had been working in the Institute of Computing Technology, CAS, and participated in the Dawning supercomputer research and development. He graduated from the University of Science and Technology of China in 1991, and received his PhD from Beijing Institute of Technology in 1997.

Qingtian Zeng is a professor at the College of Information Science and Technology, Shandong University of Science and Technology. His research interests are Workflow Mining, Domain Ontology and Petri net. He has published papers in leading international journals such as IEEE MASC. He received the support from Natural Science Foundation of China and the Excellent Young Scientist Foundation of Shandong Province of China. He is serving on the editorial board of the Journal of Software Engineering and Asian Journal of Information Management, and was the Program Co-Chair of KSEM07, GCC07, WaGe2007, and KSEM09.

Jianping Fan is a professor at the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. His research interests include high performance computing, Grid computing, and computer architecture. He took part in and developed Dawning I, Dawning 1000, Dawning 3000, Dawning 4000 and other series of Dawning supercomputers. He has published more than 70 papers and 1 book. He also acquired 11 pending or issued patents. Professor Fan has received many awards, such as outstanding award of CAS science and technology progress, first price of national science and technology progress , first price of Beijing science and technology progress, and the outstanding young scientist of CAS. He has actively participated in various professional activities. He served as editor of journal of computer research and development, Advisor of 11th five years science and technology development plan of ministry of information industry, and general chair of HPC China2007 and GCC2008.

Xiaohong Zhang is a PhD candidate in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China. Her research interests include massive data management, data access pattern analysis and data mining.