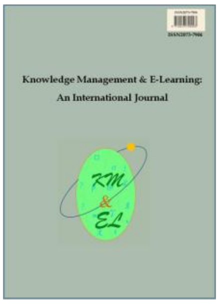
Knowledge Management & E-Learning, Vol.5, No.3. Sep 2013

Knowledge Management & E-Learning



ISSN 2073-7904

A study on online learner profile for supporting personalized learning

Jie Yang
Central China Normal University, China

Recommended citation:
Yang, J. (2013). A study on online learner profile for supporting personalized learning. *Knowledge Management & E-Learning*, *5*(3), 315–322.

Knowledge Management & E-Learning, 5(3), 315–322

A study on online learner profile for supporting personalized learning

Jie Yang*

Faculty of Education Technology
Central China Normal University, China
E-mail: stone2139313@gmail.com

*Corresponding author

Abstract: Digital learning as a popular learning approach has received increasing attention in modern education. The learner profile in online learning plays a critical role in supporting personalized learning. This article uses an information flow-based approach to build the learner profile for supporting personalized learning. The learner profile includes the individual profile to capture the personal features and the community profile to capture the social features in online learning environment.

Keywords: Online learning; Information flow; Personalized fearning; Learner profile; Community profile

Biographical notes: Jie Yang is a master student in Department of Educational Technology, Central China Normal University. Her research interests include the design and development of learning resources and technology-enhanced learning.

1. Introduction

With the rapid development of information technology and lifelong learning, online learning has been turning into important complement of the traditional teaching gradually, and personalized learning support services has become one of the main focuses in educational research area. Learning Support Service System, including service targets, service environments and service contents, is an efficient tool to provide an important quality assurance to modern online education. Among them service targets is the essential prerequisite. The modern distance education pilot program originally included only four universities including Tsinghua University in 1994. More recently the number of universities participating in this program had reached up to 68 till 2004, and the number of students registered for this program has increased to more than 2.3 millions. The learner profiles can be established for online learners to help them learn more efficiently.

A learner profile based on information flow is presented in this paper. From the information flow between online learners, the profile captures learners' interests and classifies the learners, which contribute to offering personalized service to learners.

2. Related work

A learner profile is used to present the learners' interests and preferences. Then how to capture these learners' information? This draws great attention of scholars. Learner interests can be captured either explicitly or implicitly (Agichtein, Brill, Dumais, & Ragno, 2006; Shen, Tan, & Zhai, 2005; Teevan, Dumais, & Horvitz, 2005, 2010). By explicit approach, learners proactively communicate useful information to the system, e.g., by registering information or compiling questionnaires. Instead, by implicit approach, learners' interests are automatically captured based on the actions they have taken in a specific context. For example, how learners navigate from one webpage to another, or which documents they have downloaded etc. The ways to define the types of information to be captured has become an important issue, as it would affect the effectiveness of personalization. Learner profiles are generally represented as sets of weighted keywords, semantic networks or hierarchies of concepts (Gauch, Speretta, Chandramouli, & Micarelli, 2007).

Previous studies describe few methods to extract learner's features in which the following characteristics can be seen. First, they lack of standard sources of learner profile information. The learner profile is formed by acquiring learners' preference from the registration information and the imprint online (Banga, Shah, Patel, & Patel, 2006). The effectiveness of this method highly depends on the level of learner's cooperation. Moreover, this method does not consider the learners' social attributes. For learners with the common interests, communications and cooperation can further be suggested. Last but not least, most learners profile cannot update adaptively. This paper is motivated to tackle these shortfalls. This paper aims to present a new method to build the learner profile.

3. Building the learner profile

A high effective and complete learning support service system, including service targets, service environments and service contents, is required to establish for online students to learn individually. Therefore, this paper constructs a personalized learning support services based on information flow. Fig. 1. shows a graphic representation of this personalized learning support services implementation procedure. The following subsections explain the individual components shown in the figure.

Information space stores all messages, which comes from online students' exchange through kinds of communication equipment.

In this paper, there are two types of learner profile: individual profile and community profile. Individual profile reflects the differences between learners' interest and the individual learner's preference. Each online learner has individual profile only. Community profile covers the entire learning community common interest and reflects the same preference of community learners.

Interest extractor process can be divided into the following steps. First, through the analysis of the message content in the information space, we can filter out those unwanted messages to us. Only useful information is kept. Second, we characterize each message with a vector and calculate the message vector. Finally, we can obtain the learners' interest by classifying the message vector.

Learning resources space stores all kinds of learning resources. These resources are organized by the form of Area-to-Topic.

Personalized recommendation is the connection between online students profile and learning resources. The system can recommend appropriate learning resources to online students.

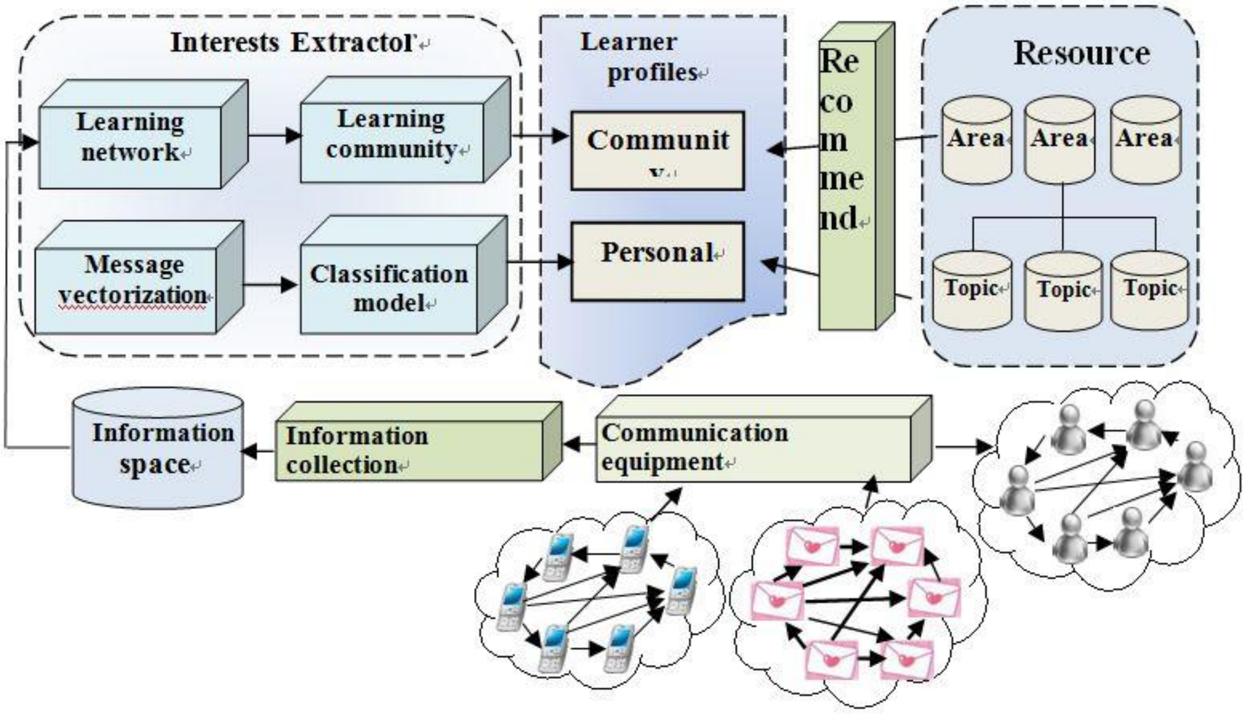


Fig. 1. Implementation of personalized learning support

3.1. Extracting the learner profile

Online learners communicate with each other by email, QQ, mobile phones and so on. As a result, an information flow network consisted of various messages is established. These messages provide us lots of useful information. For individual, they imply that who you contact and when you communicate with others; for community, they indicate that how often you contact somebody and what topics people online have in common. The information network can reflect learners' social activities and the relationship between online learners (Russo & Koesten, 2005). Each learner has his own personalized features and his social attributes, so learner profile in this paper consist of individual profile and community profile.

Community profile includes the entire learning community's common interests. Individual learners of the same preference can be seen in the community profile. Fig. 2. shows the extraction of community profile process. The process can be divided into three steps.

Firstly, through the analysis of the message content in the information space, we can build a learning network, which can illustrate the relationships among learners. Among the network, each node represents an online learner and each connection between two nodes indicates the connection messages between two learners. The connection's weight means the number of contacts between learners (Burges, 1998). Equation 1 in Table 1 describes the method to compute each connection's length between two nodes. In this way, the smaller of connection's length in the network, the closer the relationship between learners.

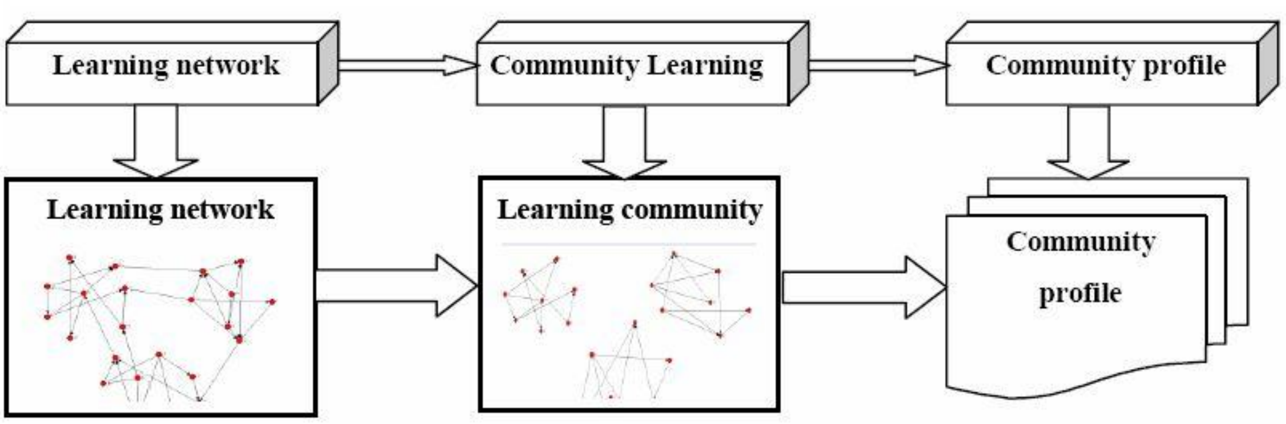


Fig. 2. Extraction of the community profile

Table 1
Connection's length computation

$$\mathrm{Length}_{ij} = \frac{t}{\mathrm{Num}_{ij}} \qquad (1)$$

where

Length_{ij} = The connection's length between Node i and Node j;

t = The minimum number of exchange messages between learners;

Num_{ij} = The number of messages sent by learner i to learner j and sent by learner j to learner i.

Secondly, we discover the learning community from the learning network. Learners in the same community means they are all interested in one certain topic. According to the definition of the connection's length (Equation 1 in Table 1), we can look for the strongest coupling in the network constantly and delete it until the network do not meet the condition of N-1 Betweenness rule. The reason is that the longer the connection's length, the smaller probability of belonging to the same learning community.

Finally, we obtain the community profile by analyzing community messages' content and learners' preference in this community. This extraction process is semi-automated and the extraction method is to mine useful features of the learners in community network.

Individual profile reflects the differences between learners' interests and represents the individual learner's preferences. Each online learner has one individual profile only. Fig. 3. shows the extraction of individual profile process. The process can be divided into three steps.

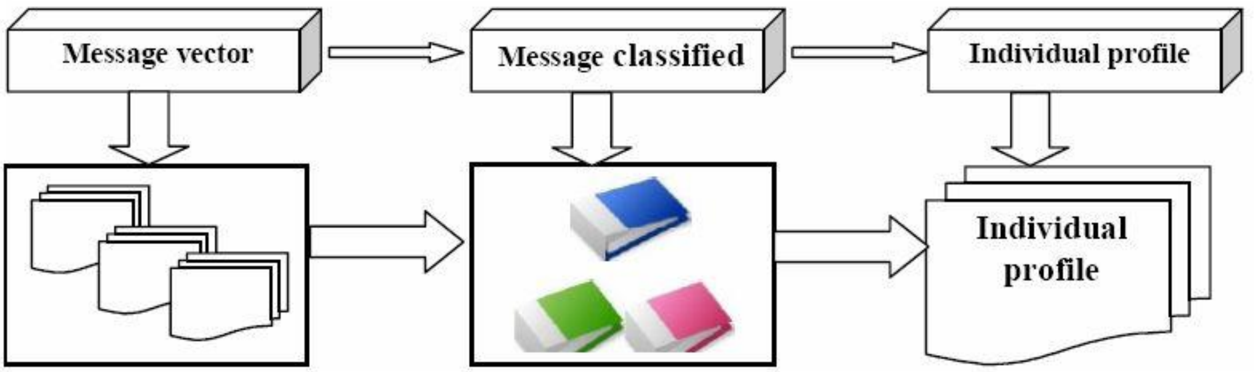


Fig. 3. Extraction of the individual profile

Firstly, individual messages are modeled with a term vector in the following form:

$$Mi = (w_1, w_2, \ldots, w_n)$$

Where: n is the total number of terms in the message;

w_i is the weight of term i

Table 2
TF-IDF feature weight computation

$$W_i = TF_{im} \times \log\left(\frac{|D|}{DF(i)} + 0.01\right) \quad (2)$$

where

W_i = TF-IDF weight for term i in message m

TF_{im} = frequency of term i in message m

IDF_i = inverse document frequency of term i in the message

$= \log(|D|/DF(i) + 0.01);$

|D| = the number of messages in training set

DF(i)= the number of term i occurring in message m

The message vectors are used in calculating the similarity scores between messages, such as similarity of a new message and a message in the information space. After term vectors are created for all messages, the weights for terms are computed. In computing term weights, we use TF-IDF. Messages with the same terms are more likely semantically related. This information is captured in term frequency (TF). TF tends to weight the commonly occurring term more than the rarely occurring terms. Inverse document frequency (IDF) fixes it by introducing a general importance of the term. Equation 2 in Table 2 describes the method to compute individual TF-IDF weight values for each term.

Secondly, two classifiers are created. one is for useful messages and the other is for useless messages. We then filter out those unwanted messages to us and retain the useful ones. Equation 3 in Table 3 describes the method to compute useful messages classifier.

Table 3
Useful messages classifier computation

$$M_x = \alpha \frac{1}{|D_x|} \sum \frac{M_x}{|M|} \qquad (3)$$

where
Mx = useful messages classifier
Dx = the number of messages in training set
|M| = the length of message m

After all messages have been filtered by the useful messages classifier, we can create topic classifiers for each learning topic, which are the sub-topics of the community profile. Equation 4 in Table 4 describes the method to compute the topic classifiers.

Table 4
Topic classifier computation

$$T_k = \frac{1}{|D_x|} \sum \frac{M}{|M|} \qquad (4)$$

where
Tk = topic classifier
Dx = the number of messages in training set
|M| = the length of message m

Table 5
Individual profile computation

$$W_i = \frac{\alpha * \sum(m \in from_j) \cap (m \in T_i) Sim(M,T_i) + \beta * \sum(m \in to_j) \cap (m \in T_i) Sim(M,T_i)}{\alpha * \sum(m \in from_j) Sim(M,T_i) + \beta * \sum(m \in to_j) Sim(M,T_i)} \qquad (5)$$

where
Wi = individual profile
fromi = the number of messages sent by learner i
toj = the number of messages sent to learner j
Sim(M,Ti) = the similarity between message m and topic t

Finally, we can compute the individual profile. Individual profile reflects the differences between learners' preference. Equation 5 in Table 5 describes the method to compute individual profile.

3.2. Updating the learner profile

Learners' preference will change with time, so learner profile should be updated from time to time. We track the change in learners' interests using a method of information flow tracking. The time attribute in the message is used for the tracking purpose. We add a time factor like following form in the individual profile computation.

$$2^{\frac{-\mathrm{age(m)}}{\mathrm{hl}}}$$

Where: age(m) is the age of the message m;

hl is the message's half-life

In this way, learner profile will be update adaptively without learners' cooperation.

4. Summary and future work

This paper presents a few approaches for mining information automatically and extracting learners' preference to build learner profile. Our work focuses on a method based on information flow especially for online learner to build learner profile. This method helps solve several key problems, such as extracting learner personalized features implicitly，describing learner personalized features and social features, and the updating learner profile adaptively. In addition, the learner profile content supports the personalized recommendations of resources and personalized learning services.

In order to further improve the effectiveness of the learner profile, we plan to include more message sources and additional learner features. We also will apply few statistical models to infer and extract semantic information and relations in the information space. Another interesting study that we wish to perform in the future is to recommend personal learning resource for online learners according to the learner profile.

References

Agichtein, E., Brill, E., Dumais, S., & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 3–10). New York, NY, USA: ACM

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*, 121–167.

Banga, J., Shah, N., Patel, B., & Patel, A. (2006). *Systems and methods of network operation and information processing, including data acquisition, processing and provision and/or interoperability features*. US Patent Application 12/088,236.

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User profiles for personalized information access. *Lecture Notes in Computer Science, 4321*, 54–89.

Russo, T. C., & Koesten, J. (2005). Prestige, centrality, and learning: A social network analysis of an online class. *Communication Education, 54*(3), 254–261.

Shen, X., Tan, B., & Zhai, C. (2005). Context-sensitive information retrieval using implicit feedback. *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 43–50). New York, NY, USA: ACM.

Teevan, J., Dumais, S. T., & Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 449–456). New York, NY, USA: ACM Press.

Teevan, J., Dumais, S. T., & Horvitz, E. (2010). Potential for personalization. *ACM Transactions on Computer–Human Interaction, 17*(1), article 4.