

# Dynamic Learning Style Prediction Method Based on a Pattern Recognition Technique

Juan Yang, Zhi X. Huang, Yue X. Gao, and Hong T. Liu,

**Abstract**—During the past decade, personalized e-learning systems and adaptive educational hypermedia systems have attracted much attention from researchers in the fields of computer science and education. The integration of learning styles into an intelligent system is a possible solution to the problems of “learning deviation” and “cognitive overload.” In this study, we propose a learning style prediction method based on a pattern recognition technique. The main contributions of this method are: (1) it is a form of middleware that can be applied to other intelligent tutoring systems, and (2) it can process topic-dependent data to make predictions and update learning style profiles in a recursive manner. Experimental evaluations demonstrated the effectiveness of this prediction method.

**Index Terms**—Learning behavior, Learning style, Pattern recognition.

## I. INTRODUCTION

DURING the past decade, personalized E-learning systems and adaptive educational hypermedia systems (AEHSSs) have attracted much attention from researchers in the fields of computer science and education. Many intelligent learning systems have been proposed and implemented in the past few years based on research into human-computer interactions and intelligent systems in artificial intelligence. Most of these original intelligent systems have focused on addressing superficial requests by learners in the same manner as standard intelligent systems in other application areas. However, these user request-driven systems cannot perform to the desired level if they are components of web-based learning procedures. Educators have investigated this problem and found that the learning materials recommended by these systems do not satisfy the requests of learners sufficiently and they fail to meet the potential needs of learners in some cases. In particular, problems known as “learning deviation” and “cognitive overload” [1] are of great importance, and it has been proposed that most learners are unsure of their actual needs during the learning process, which may lead to inaccurate requests. To address these problems, it might be beneficial if intelligent learning systems were designed to analyze the actual needs of learners, thereby helping to improve their learning performance.

Many educational theories suggest that the integration of learning styles into learning activities may improve learning

J. Yang is with the Department of Computer Science, Sichuan Normal University, China, e-mail: jkxy\_yjuan@sicnu.edu.cn.

Z. X. Huang is with South West University, China, e-mail: zhixinghuang@swu.edu.cn.

Y. X. Gao is with the Department of Computer Science, Sichuan Normal University, China, e-mail: gyx@sicnu.edu.cn.

H. T. Liu is with the College of Computer Science, Chongqing University of Posts and Telecommunications, China, e-mail: liuh@cqupt.edu.cn.

performance [2], [3], [4]. Thus, the integration of learning styles in an intelligent system may be a possible solution to the “learning deviation” and “cognition overload” problems. The integration of user profiles with learning style data could enhance intelligent learning systems by providing a learner with personalized learning behavior guidance that is appropriate to their potential needs.

In this study, we propose a learning style prediction method based on a pattern recognition technique. This method functions as a form of middleware that can be applied to other intelligent tutoring systems, while it can process topic-dependent data to make predictions and update learning style profiles in a recursive manner. Experimental evaluations demonstrated the effectiveness of our novel prediction method.

This paper is organized as follows. Section 2 introduces related research about learning styles and user profiles with integrated learning styles. Section 3 describes the graphical user interface of our benchmark prototype system. Section 4 provides formal descriptions of the learning style theories and section 5 presents a pattern recognition model for the dynamic inference of the learning style of a new learner. The experimental evaluations and analyses are presented in section 6. Section 7 gives our conclusions and identifies some future research areas.

## II. RELATED WORK

Various learning style classification methods have been used to construct user profiles in AEHSSs, including approaches based on the Myers-Briggs [5], Kolb [6], Herrmann [7], and Felder-Silverman learning style families [8]. The adaptive educational system [9] uses the field-independent/-dependent cognitive style to construct a learner profile and this allows the learning process to be adapted to different learners. The INSPIRE system [10] uses the Honey-Mumford learning style to construct learner profiles, while the iWeaver system [11] uses the Dunn-Dunn learning style to facilitate navigation and content representation. Most other intelligent learning systems use the Felder-Silverman learning style family to support their personalized learning services [12], [13], [14], [15], [16], [17], [18], [19], [20], [23], [24], [27], [28].

AEHSSs usually employ two main methods for determining learning style preferences [21], i.e., questionnaires and automatic detection. Compared with questionnaires, automatic detection can obtain more accurate information because the data are acquired dynamically in the current learning environment. Many useful techniques are available for the automatic detection of learning styles, such as neural networks

[13], Bayesian networks [14], rule-based reasoning [15], [27], classifiers [28], and data mining [16]. These techniques abstract the typical learning behaviors from human-computer interaction processes, which can represent the characteristics of different learning style preferences. The browser plug-in tool iLesson [22] uses the data mining agent “PolyAnalyst” to monitor learning behaviors. The data collected by the agent are used to infer the learning style preferences, which are categorized into iLesson groups based on different learning style clusters. The Protus system [16] uses the sequential pattern mining algorithm “AprioriAll” to predict learning styles by grouping learning behaviors into 16 different clusters based on combinations of four Felder–Silverman learning style dimensions. Garcia et al. [14] proposed the use of Bayesian networks to predict learning style preferences by constructing conditional probability tables for the learning behaviors and different learning style dimensions. Zatarain-Cabada et al. [13] used neural networks to predict learning style preferences by training a seven-neuron network using sample questionnaire data obtained from learners and their specific learning behaviors. AHA! is a well-known adaptive hypermedia system that contains several authoring tools [25], [26] and Stash et al. [23] exploited the arbitrary relationships available in this system to predict learning styles.

Unfortunately, most of the user profiles generated by these methods are highly dependent on the tutoring systems. Thus, few of these methods can be reused in other learning systems without the technical support of the original platforms. Stash et al. extended AHA! by using an adaptive engine to apply adaptation strategies that correspond to the learning styles of a chosen domain model, which makes AHA! more intelligent when evaluating learners [23]. They also proposed an XML-based language called LAG-XLS to design learning style strategies [24]. User model updating and adaptation are implemented using this rule-based engine. The major advantage of this approach is that it separates the learning style instructional strategies from the tutoring system. Thus, various predesigned instructional patterns for different learning style preferences can be controlled completely by the authors instead of the system developers. Ozpolat and Akar [28] developed a generic prediction module that categorizes learning style preferences based on NBTree classification combined with a binary relevance classifier. The module used to predict learning styles is totally independent of the tutoring system and the learning content. However, the main drawback of the prediction module is that it places additional burdens on learners, such as asking them to complete the prediction procedure before learning. Latham et al. proposed a conversational intelligent tutoring system for predicting learning styles based on natural language interactions with the system [27]. However, the most important factor that determines the efficiency and effectiveness of this method is the performance of natural language processing, which is still one of the most time-consuming research problems in computer science.

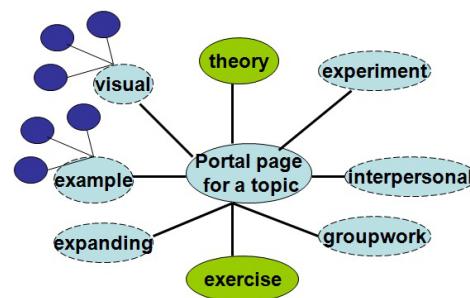


Fig. 1. Star topology for a learning topic

### III. INTRODUCTION TO THE BENCHMARK PROTOTYPE SYSTEM

This prediction mechanism is middleware but it still needs a benchmark to indicate its functionality and effectiveness. In this section, we introduce the benchmark system, which is a prototype system called “Programming in Java” (PIJ). The main function of PIJ is to collect data on learning behaviors using log files. To improve the computational efficiency and to decrease the complexity, the organization of learning objects (LOs) in PIJ follows three distinct rules.

(1) The overall content comprises a series of topics in a given order, which is consistent with their prior/subsequent constraints.

(2) Each topic must contain a sufficient number of LOs to provide the various types of LOs demanded by learners.

(3) The construction of the LOs within a single topic follows the star topology.

Figure 1 illustrates the star topology for a topic and Figures 2–5 show the PIJ interfaces. Figure 2 shows the main PIJ interface where topics are listed. Figure 3 illustrates the portal page of the topic “Introduction & Java Platform.” Figure 4 illustrates the content layout of the LO “Theory” about the current topic. The content is provided in a linear order and “Go to Portal” buttons are scattered in the middle of the paragraphs, which allows some learners to learn in jumps. Figure 5 shows the exercise interface for the same topic. Learners provide their self-evaluations by submitting their exercise attachments. Regardless of the results, the learners have completed this phase of the learning task if they submit their exercise attachments. The learning path is defined as closed if and only if the learner submits their exercise attachment and clicks the “Submit” button. Each learning path used to construct the learning information pattern must be a closed path within the same topic.

The PIJ system is used as a benchmark to indicate the functionality and usability of the learning style prediction mechanism. If the intelligent learning system is different from the PIJ, the data collected by the original tutoring system are transformed into formal descriptions using the method introduced in the next section.

### IV. FORMAL DESCRIPTION OF THE UNDERLYING LEARNING STYLE THEORIES

In this section, we explain the underlying learning style theories and transform them into formal descriptions.



Fig. 2. Main Programming in Java (PIJ) interface



Fig. 3. Portal page for the topic "Introduction & Java Platform"

The solution to most problems in object-oriented design seems flippant: You create another type of object. The new type of object that solves this particular problem holds references to other objects. This new object, generally called a container.

9. Exception handling: dealing with errors

Ever since the beginning of programming languages, error handling has been a particularly difficult issue. Because it's so hard to design a good error-handling scheme, Java's exception handling stands out among programming languages, because in Java, exception handling was wired in from the beginning and you're forced to use it. It is the single acceptable way to report errors. If you don't write your code to properly handle exceptions, you'll get a compile-time error message. This guaranteed consistency can sometimes make error handling much easier.

10. Java and the Internet

If Java is, in fact, yet another computer programming language, you may question why it is so important and why it is being promoted as a revolutionary step in computer programming. The answer isn't immediately obvious if you're coming from a traditional programming perspective. Although Java is very useful for solving traditional standalone programming problems, it is also important because it solves programming problems for the World Wide Web.

Scattered in the middle → Go to portal

Introduction and deploying of java platform

To compile and run a program, you must first have a Java programming environment. There are a number of third-party development environments, but in this lesson you are using the Java Developer's Kit (JDK) from Sun, which is free. If you are using another development system, you will need to look in the documentation for that system to determine how to compile and run programs.

Get on the Internet and go to <http://java.sun.com>. There you will find information and links that will lead you through the process of downloading and installing the JDK for your particular platform.

Once the JDK is installed, and you've set up your computer's path information so that it will find `javac` and `java`, go to command window, and change your directory into the jdk installing directory, go to the sub directory `bin`, then type:

Fig. 4. Content layout of the learning object "Theory," which is related to the current topic

#### A. Formal descriptions of the psychological foundations of Felder & Silverman learning styles

The learning style used is that proposed by Felder & Silverman for engineering students, which distinguishes between learning preferences based on the following five dimensions.

- The way learners perceive information: sensing or intuitive
- The way learners acquire information: visual or verbal

Exercise 6: Modify the previous exercise so that the values of the data in DataOnly are assigned to and printed in main().

Exercise 6: Write a program that includes and calls the storage() method defined as a code fragment in this chapter.

Exercise 7: Turn the Incrementable code fragments into a working program.

Exercise 8: Write a program that demonstrates that, no matter how many objects you create of a particular class, there is only one instance of a particular static field in that class.

Exercise 9: Write a program that demonstrates that autoboxing works for all the primitive types and their wrappers.

Exercise 10: Write a program that prints three arguments taken from the command line. To do this, you'll need to index into the command-line array of Strings.

Exercise 11: Turn the AllTheColorsOfTheRainbow example into a program that compiles and runs.

Exercise 12: Find the code for the second version of HelloDate.java, which is the simple comment documentation example. Execute Javadoc on the file and view the results with your Web browser.

Exercise 13: Run Documentation1.java, Documentation2.java, and Documentation3.java through Javadoc. Verify the resulting documentation with your Web browser.

Exercise 14: Add an HTML list of items to the documentation in the previous exercise.

Exercise 15: Take the program in Exercise 2 and add comment documentation to it. Extract this comment documentation into an HTML file using Javadoc and view it with your Web browser.

Exercise 16: In the Initialization & Cleanup chapter, locate the Overloading.java example and add Javadoc documentation. Extract this comment documentation into an HTML file using Javadoc and view it with your Web browser.

Browse button → Submit button → Go to portal

Fig. 5. Exercise interface for the topic "Introduction & Java Platform"

- The way learners organize information: inductive or deductive
- The way learners process information: active or reflective
- The way learners understand information: sequential or global

Further details of these five learning style preference categories are provided in Table I. One of the premises of this study is that the same preference-biased learners would have similar behavioral tendencies in the same environmental conditions. According to study [8], if a learner prefers a certain learning style, they will exhibit a specific learning behavior, whereas the reverse may not be the case, e.g., if a learner prefers practical exercises, they cannot be classified as an active user because we do not know whether other behaviors might affect their identification as an active or reflective user. Thus, the learning style preference is a sufficient condition for the corresponding learning behaviors, whereas the reverse may not be the case. Therefore, sensitive words are used as indicators to capture learning information rather than as the basis for inferring learning style preferences. Sensitive words selected from literacy expressions are transformed into formal descriptions, where the methods used to select the sensitive words were reported by [14], [16], [23], [27] and [28]. In addition to sensitive words, other LO options are included in the PIJ to make the learning information vector more effective.

We synthesized the most commonly used and readily monitored behaviors from previous studies [14], [16], [23], [27] and [28], i.e., tools for monitoring interpersonal communication were reported by [14], [16], [27] and [28], and quizzes/exercises/questions for monitoring behaviors were reported by [14], [16] and [27]. However, the types of behavior that are most useful for recognizing learning style preferences are unknown. Therefore, we selected the most commonly and readily monitored behaviors to construct the behavior information vector. Whether a learner conducted sequential learning was monitored by [14], [16] and [28]. By contrast, [14] and [28] monitored the learning sequence, and [16] monitored the extra hyperlinks visited by users. In our method, we use the learning sequence to identify whether a user conducts sequential learning, but we also monitor any extra hyperlinks visited by users as supplementary information. Both types of information are considered when constructing a learning information vector. Thus, the learning information

vector may be more meaningful compared with a specific learning behavior. Of course, other learning behaviors can also be considered, such as the mouse click time or the time spent staring at the screen. The prediction results will be more accurate if more meaningful subtle learning behaviors are considered. However, given the computational performance of this method, dimensionality reduction is necessary for the learning behavior vector.

Eight abstract labels are used to identify the different kinds of LOs that cover the sensitive words extracted from the Felder & Silverman learning style model. The necessary constraints of the LOs are also listed in Table II. PIJ should conform to these constraints.

### B. Propositions and prerequisites for the establishment of learning information

We connect the detailed behaviors with the learning information in a logical and quantitative manner. Five variables are used as prerequisites to describe and quantify the detailed behaviors. The specific descriptions of these variables are provided in Table III. The propositions and prerequisites used to establish the learning information based on these five variables are described in Table IV. A set containing a subject's learning-related information could be used as a substitute if we need to construct a learning path. If a learner has visited more than two LOs with the label "example," the learner's "learn from example" proposition is true. By contrast, if a learner did not visit any LOs labeled as "example" or they visited only one LO labeled as "example," the learner's "learn from example" proposition is false. After a learner completes a learning topic, a set containing their learning information is constructed from their original log file.

## V. DYNAMIC PREDICTION OF LEARNING STYLES BY RECOGNIZING INFORMATION PATTERNS

In this section, we propose a mechanism for predicting learning styles based on recognizing the information patterns of subjects. The architecture of the prediction model is shown in Figure 6. Initially, sample learners are required to complete two components: the Index of Learning Style (ILS) questionnaire and the PIJ benchmark. The results obtained from these users are employed as labeled samples to implement the supervised training procedure. The learning information obtained from these users is mapped onto their learning style preferences using mutual similarity pattern recognition.

### A. Construction of patterns

Three steps are involved in this process. First, the pretested learning style preference patterns of the sample learners are used to construct a multidimensional space. Second, their mutual similarity patterns are constructed based on their learning information related to the same learning topic. The final step compares the mutual similarity pattern with the pretested learning style patterns to identify the key learning style dimension.

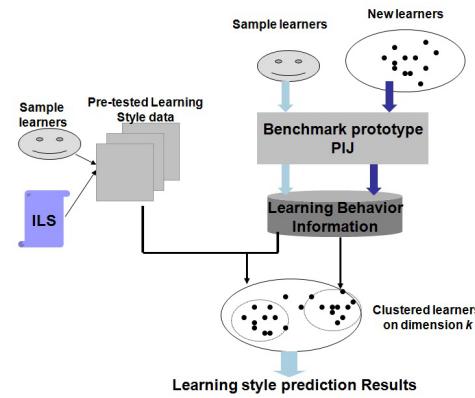


Fig. 6. Architecture of the prediction model

#### 1) Constructing the learning style preference patterns:

Suppose we have already collected the normalized learning style preference data for a sample of  $n$  students using the ILS instrument (normalization is a technique that standardizes a range of values within specific limits to ensure the accuracy of subsequent analyses). The learning style preference similarity with respect to dimension  $t$  for any two sample students can be computed using their pretested data  $p_i^t$  and  $p_j^t$ :

$$\alpha_{ij}^t = \frac{e - e\sqrt{|p_i^t - p_j^t|}}{e - 1}. \quad (1)$$

In Formula (1),  $p_i^t \in [0, 1]$  represents the quantified learning preference value on dimension  $t$  for the sample student  $i$ . In Formulae (1),  $e$  is a mathematical constant that is approximately equal to 2.71. Formula (1) is also called a "squashing function" because  $\alpha_{ij}^t$  approaches 1 if the sample learners  $i$  and  $j$  have more similar learning style preferences with respect to dimension  $k$ . Otherwise,  $\alpha_{ij}^t$  equals 0 if their preferences are totally opposite. Thus, if the sample learner  $i$  is an extremely sequential learner and sample learner  $j$  is an extremely global learner, such that  $p_i^{\text{sequential/global}} = 0, p_j^{\text{sequential/global}} = 1$ , and  $\alpha_{ij}^{\text{sequential/global}} = 0$ , learner  $i$  and learner  $j$  share nothing in common with respect to learning style dimension  $t$ . By contrast, if  $i$  and  $j$  are both extremely sequential learners, their learning style similarity level will have a high value, such as  $p_i^{\text{sequential/global}} = p_j^{\text{sequential/global}} = 0$  and  $\alpha_{ij}^{\text{sequential/global}} = 1$ . The similarity values between sample learners can be stored in an  $n \times n$  matrix  $M_t$ , where  $M_t$  is obviously a symmetric matrix.

#### 2) Constructing the mutual similarity patterns:

First, we need to construct the learning behavior similarity relationships between the sample learners based on their learning information related to the same topic. A vector  $[B_1^i, B_2^i, \dots, B_K^i]$  is used to describe the learning information of sample learner  $A_i$  with respect to a topic.  $K$  is the dimensionality of the vector and an element  $B_j^i$  represents a learning behavior described in Table IV. The learning behavior similarity relationship among any sample learners  $A_i$  and  $A_j$  can be computed using Formulae (2) and (3), as follows:

$$D(A_i, A_j) = \sqrt{\sum_{l \leq K} (B_l^i - B_l^j)^2}, \quad (2)$$

TABLE I  
SENSITIVE WORDS EXTRACTED FROM LITERACY EXPRESSIONS

Literacy descriptions [8]	Sensitive words
Active learners tend to retain and understand information best by <b>doing something active, trying things out, discussing, or performing group work</b> , whereas reflective learners are the opposite.	Experiment; Interpersonal communication tool; Group work assignment
Sensing learners tend to prefer <b>facts</b> and solving problems using <b>well-established methods</b> . They are patient with respect to details, such as performing <b>hands-on work</b> , and they like <b>knowledge that has apparent connections with the real word</b> . Intuitive learners prefer innovation, but dislike repetition, and they like <b>abstraction</b> and <b>mathematical formulation</b> .	Example; experiment; concepts; visual
Sequential learners tend to understand in <b>linear</b> steps whereas global learners tend to learn during repeated bursts of activity.	In order
Visual learners prefer visual representations, such as <b>pictures, diagrams, and flow charts</b> , whereas verbal learners prefer <b>written</b> and spoken <b>explanations</b> .	Visual; concept; theory

TABLE II  
FULL TYPES OF THE NECESSARY LEARNING OBJECTS AND THEIR CONSTRAINTS

LO labels	Explanations	Quantitative constraints in the LO topology refer to a single topic
Theory	An LO that contains conceptual material about a specific topic, such as <b>definitions and formulations</b>	number(LO.theory)=1
Example	An LO that contains instances of a given topic	number(LO.example)≥2
Visual	An LO that contains supplementary visual resources about a specific topic, such as <b>pictures, graphs, and charts</b>	number(LO.visual)≥2
Experiment	An LO that contains hands-on experiments about a specific topic	number(LO.experiment)≥1
Expanding	An LO that contains an expanding resource about a specific topic, such as <b>explanations of reference information</b>	number(LO.expanding)≥1
Interpersonal	Any communication tools can be abstracted as an LO node labeled as “interpersonal,” such as real-time communication tools (MSN and QQ), discussion tools (forums and chat rooms), and asynchronous communication tools (email)	number(LO.interpersonal)≥1
Groupwork	An LO containing a previously designed assignment that targets a specific topic and requires cooperation	number(LO.groupwork)≥1
Exercise	An LO containing an exercise/exam that is designed to reflect the learning state of a learner with respect to a specific topic	number(LO.exercise)≥1

TABLE III  
VARIABLES USED AS PREREQUISITES

Variables	descriptions
participate(LO.a)	If a learner has participated in an LO labeled “a” in a single topic (i.e., they have read the LO and clicked the “ok” button on the pushed window), the variable <b>participate=true</b> or else <b>participate=false</b>
visit(LO.a)	If a learner has visited an LO labeled “a” in a single topic (they have only read the LO), the variable <b>visit=true</b> or else <b>visit=false</b>
num(LO.a)	The number of different LOs labeled “a” visited in a single topic
time(LO.a)	The number of times the same LO labeled “a” was visited in a single topic
inorder(LO.a, LO.a...LO.a)	If a learner uses the “go to portal” button more than twice within the LO “a,” <b>inorder=false</b> or else <b>inorder=true</b>

TABLE IV  
PROPOSITIONS AND THE PREREQUISITES FOR THE ESTABLISHMENT OF LEARNING INFORMATION

Propositions	Prerequisites for the establishment of learning information	
	True (1)	False (0)
Group work assignment	$\exists \text{participate}(\text{LO.groupwork})$	$\neg \exists \text{participate}(\text{LO.groupwork})$
Interpersonal communication	$\exists \text{participate}(\text{LO.interpersonal})$	$\neg \exists \text{participate}(\text{LO.interpersonal})$
Learning by example	$\exists \text{visit}(\text{LO.example}) \wedge (\text{num}(\text{LO.example}) > 1)$	$\neg \exists \text{visit}(\text{LO.example}) \vee (\exists \text{visit}(\text{LO.example}) \wedge (\text{num}(\text{LO.example}) = 1))$
Reading theory (concept)	$\exists \text{visit}(\text{LO.theory}) \wedge (\text{time}(\text{LO.theory}) > 1)$	$\neg \exists \text{visit}(\text{LO.theory}) \vee (\exists \text{visit}(\text{LO.theory}) \wedge (\text{time}(\text{LO.theory}) = 1))$
Hands-on experiments	$\exists \text{participate}(\text{LO.experiment})$	$\neg \exists \text{participate}(\text{LO.experiment})$
Expanding reading	$\exists \text{visit}(\text{LO.expanding})$	$\neg \exists \text{visit}(\text{LO.expanding})$
Continuous learning	$\text{inorder}(\text{topics})$	$\neg \text{inorder}(\text{topics})$
Visual	$\exists \text{visit}(\text{LO.visual}) \wedge (\text{num}(\text{LO.visual}) > 1)$	$\neg \exists \text{visit}(\text{LO.visual}) \vee (\exists \text{visit}(\text{LO.visual}) \wedge (\text{num}(\text{LO.theory}) = 1))$

$$d(A_i, A_j) = (e^{\sqrt{K}} - e^{D(A_i, A_j)}) / (e^{\sqrt{K}} - 1). \quad (3)$$

Formula (2) is the formulation used to compute the Euclidean distance, which is one of the most commonly used

techniques for comparing the similarity of two vectors in a multidimensional space. Formula (3) is used to normalize the Euclidean distance in area  $[0,1]$ , so  $d$  approaches 1 if the learning information of the two learners is more similar. In the opposite case,  $d$  is approximately equal to 0. This normalization formula is the same as that used to normalize the similarity relationship between the pretested learning style data for the sample learners. For instance, we consider the learning information for four hypothetical learners and their learning style preferences with respect to the different dimensions of a specific topic, “Introduction & Java Platform,” where the hypothetical data are listed in Table V.

The current learning information patterns of the four hypothetical sample learners can be depicted as a  $4 \times 4$  matrix  $M$ , where each entry  $m_{ij}$  in matrix  $M$  is the normalized Euclidean distance  $d(A_i, A_j)$  between two sample students  $i$  and  $j$ , e.g.,  $m_{1,2} = d(A, B) = (e^{\sqrt{8}} - e^{\sqrt{6}})/(e^{\sqrt{8}} - 1) = 0.33$ .

### B. Identifying the key learning style dimension based on pattern recognition

The method described here can be viewed as a form of “problem transformation method” for multilabel classification, where the objective is to identify the learning style preferences based on the current learning topics. Another assumption of the present study is that most learners only favor a specific learning style in one dimension, whereas their preferences are ambiguous or unknown in other dimensions of a specific learning environment [22]. For example, if learners take a course such as “graph theory,” even a naturally strong verbally biased learner might not exhibit his/her learning preference. Thus, the aim of identifying the key learning style dimension is to transform this multilabel classification problem into a single-label classification problem. The advantage of this approach is that it can allow several different learning style families to be used together because their mutual similarity-based learning style preference patterns are independent of each other.

Suppose that there are  $t$  dimensions in the learning styles of the Felder & Silverman model, the target function  $f(X)$  used for pattern recognition is set as

$$\begin{aligned} \min f(X), \quad & \sum_{i \leq t} x_i^2 = 1 \\ f(X) = \text{var}(M, \sum_{i \leq t} x_i \cdot M^t) \\ &= \sqrt{\sum_{i \leq n} \sum_{j \leq n} (m_{ij} - \sum_{h \leq t} x_h \cdot \alpha_{ij}^t)^2} \quad \sum_{i \leq t} x_i^2 = 1. \end{aligned} \quad (4)$$

In Formula (4),  $X$  is a vector,  $X = (x_1, x_2, \dots, x_t)$ , and  $t$  is the dimensionality of the learning styles. Formula (4) is used to identify the key learning style similarity matrix that represents the current collective learning behaviors. We construct the goal function as a nonlinear programming problem [29], which allows us to determine a possible solution by minimizing the differences of the matrices.  $\sum_{i \leq t} x_i^2 = 1$  is a constrained condition used to normalize the vector (to ensure that  $\forall x_i \in [0, 1]$ ). Since the matrices  $M$  and  $M^t$

are symmetric, the function  $\text{var}()$  can be transformed further to compute the Euclidean distance between two  $\frac{n(n-1)}{2}$  dimensional vectors, as shown in Formula (4). In the optimal situation, we have

$$M = \sum_{i \leq t} x_i \cdot M^i. \quad (5)$$

If Formula (5) has a solution, it must be the only solution. Otherwise, we need to use a heuristic search algorithm to find the approximate optimal solutions. We use the simulated annealing (SA) algorithm to find the approximate optimal solutions for the goal function because this algorithm can always converge with a probability of 1. The convergence time is satisfied because the solution space of this problem is relatively small compared with other problems. In the SA algorithm,  $f(X)$  is referred to as the evaluation function. If the new solution is not better than the original one, the new solution will be accepted with a probability of  $e^{-df/T}$ . The detailed process of the SA algorithm is described in the Appendix.

### C. Predicting learning style preferences

After identifying the key learning style dimension, where the learning information of the subject has the highest projected value in the multidimensional space, the learning style dimension  $k$  should be used as a scale tool to cluster the new learners according to their learning behaviors. This clustering process is simple compared with the clustering processes used in previous studies because the multilabel classification problem has been transformed into a single-label classification problem.

The first algorithm is called the *Initialization Cluster Core Construction Algorithm*, which is a variation on the 2-means clustering algorithm. This algorithm produces nodes that represent the original sample learners who had extreme learning style preferences with respect to dimension  $k$ , which are used as the initial core sets for the different clusters. The pseudo-code of the algorithm is shown in Table VI.

In Table VI, the set  $P_c$  and  $N_c$  contains the core nodes for dimension  $k$ , which will change constantly as more testers are integrated into the learning process. We apply polarization processing to the learning style preference to satisfy the requirements of the computational model. Thus, if a learner is an active learner, they would be defined as positive in the “active/reflective” dimension. By contrast, if a learner is negative with respect to the same dimension, they would be defined as a reflective learner. The parameters  $d_n$  and  $d_p$  are used to calculate the different distances from the current node to the negative cluster and positive cluster, using Formulae (2) and (3). For the nodes that are already in the positive cluster, if  $d_n > d_p$ , the node is removed from its original cluster because it is closer to the negative cluster than the positive cluster, and vice versa.

After updating the core sets of the clusters that represent the cognitive preference center with respect to learning style dimension  $k$ , the new users are clustered according to their

TABLE V  
HYPOTHETICAL DATA

Sample learners	Learning style data					Learning information							
	Active	Sensing	Global	Visual	Inductive	Group work assignment	Interpersonal communication	Learning by example	Reading theory	Hands-on experiments	Expanding reading	Continuous learning	Visual
A	0	1	1	0	1	F	F	T	T	T	T	F	
B	0	0	1	1	0.1	T	F	F	F	F	T	T	
C	1	1	0	0	0.9	T	F	T	T	T	F	F	
D	0.6	0.6	0.1	0.9	1	F	T	T	T	T	T	T	

TABLE VI  
INITIALIZATION CLUSTER CORE CONSTRUCTION ALGORITHM

**Initialization Cluster Core Construction Algorithm**

- Input:** Set A contains the sample nodes where the learning style preferences with respect to dimension  $k$  are no more than 0.1 or no less than 0.9
- Output:** Set  $P_c$  and  $N_c$
- $P_c = \emptyset$ ,  $N_c = \emptyset$
- Place a sample node into set  $P_c$  if its learning style value with respect to dimension  $k$  is no less than 0.9, but place it into set  $N_c$  if its learning style value with respect to dimension  $k$  is no more than 0.1;
  - Compute the average normalized Euclidean distance  $\bar{d}_p$  between the members of set  $P_c$  and  $\bar{d}_n$  between the members of set  $N_c$ ;
  - For each member in set  $P_c$ 
    - Compute its average normalized Euclidean distance from the other members of set  $P_c$  and from the members of set  $N_c$ ; If  $d_n > d_p$  Delete this node from set  $P_c$  //sample is inaccurate;
    - For each member in set  $N_c$  Compute its average normalized Euclidean distance  $d_n$  from the other members of set  $N_c$  and  $d_p$  from the members of set  $P_c$ ;
    - If  $d_n < d_p$  Delete this node from set  $N_c$  //sample is inaccurate;
  - Go back to step 2 and repeat until  $\bar{d}_p$  and  $\bar{d}_n$  converge.

TABLE VII  
CLUSTERING ALGORITHM

**Clustering Algorithm**

- Input:** Unknown nodes set L, sets  $P_c$ ,  $N_c$ , P, N, U
- Output:** Sets P, N, U
- P**= $P_c$ , **N**= $N_c$ , **U**= $\emptyset$ ;
- For each unknown node l in L**
- Compute the average normalized Euclidean distance  $\bar{d}(l, p)$  between  $l$  and the members of set  $P_c$ ;
  - Compute the average normalized Euclidean distance  $\bar{d}(l, n)$  between  $l$  and the members of set  $N_c$ ;
  - Case: $|d(l, p) - d(l, n)| \leq \varepsilon$  add  $l$  in Cluster U;  
Case: $d(l, p) - d(l, n) > \varepsilon$ , add  $l$  in Cluster P;  
Case: $d(l, n) - d(l, p) > \varepsilon$ , add  $l$  in Cluster N;

similarity with the nodes in the core sets. The second algorithm, i.e., “Clustering Algorithm” (Table VII), classifies the new nodes into three different clusters. “Positive” is defined as learners with preferences on the following poles: active, sensing, global, and visual. “Negative” is defined as learners with opposite learning preferences: reflective, intuitive, sequential, and verbal. “Unknown” is defined as learners with no obvious tendencies with respect to the learning style dimension  $k$ .

Now, we can infer the learning style preference of a new subject with respect to dimension  $k$  based on the following rule.

If  $l \in P$ , then  $l$  is  $k$  “positive” learner.

Else if  $l \in N$ , then  $l$  is  $k$  “negative” learner.

Else  $l$  is a neutral or “unknown” learner with respect to dimension  $k$ .

*D. Dynamic changes in the learning information patterns*

When new testers are added to the prediction method, they are treated as new reliable training data that replace the most

similar original samples with respect to the specific learning style dimension. The clustering procedure that considers newly added examples is actually a 2-means clustering procedure (which corresponds to steps 2–5 in Table VI). The re-computation of the centers of the clusters in dimension  $k$  means that the typical learning information pattern associated with this type of learning style preference is updated gradually. Finally, the learning information patterns will converge on a stable state when sufficient homogeneous testers have been integrated in the learning process. During the dynamic pattern recognition process, the samples used to predict learning style preferences should be updated by adding new testers, or modified if the learning circumstances change, such as lesson replacement (e.g., changing an experimental-based lesson into a pure theory-based lesson, or changing a language lesson into a political lesson) or a change in a student’s cognitive level.

For example, the original samples A, B, and C have the same learning information pattern with respect to a specific topic (e.g.,  $<00000011>$ ) so their mutual relationship value is 1, and they are all core nodes in the positive cluster of the “active/reflective” dimension after initialization. When a new learner E is added to the process, who has a similar learning information pattern (e.g.,  $<00000111>$ ) to the original samples A, B, and C, E is classified as an active-biased learner. E is then treated as new training data that randomly replaces one of the most similar original samples in the “active/reflective” dimension. If we suppose that E replaces A, the mutual similarity relationship may be changed to 0.89, 0.89, 1. After comparing the new similarity relationship with

TABLE VIII  
PARTIAL RECORD OF A CLOSED LEARNING PATH IN THE TOPIC  
“INTRODUCTION & JAVA PLATFORM”

In the “Introduction & Java Platform” topic	Information
Visit(theory)	Groupwork assignment=0
Visit(example)	Interpersonal communication=1
Take(exercise)	Learning by example=1
Visit(example)	Reading theory=0
Participate(experiment)	Hands-on experiment=1
Participate(interpersonal)	Expanding reading=0
Visit(visual)	Visual=0
Take(exercise)	
Submit(exercise)	

the old one, it is clear that they have the same nature. Suppose that another two new learners F and G are added to the prediction process, which have the same learning information patterns as E ( $< 00000111 >$ ), and that they replace the original samples B and C. The new instances E, F, and G construct exactly the same collective learning information similarity pattern as the original three samples but the learning information pattern of the active learner is amended from  $< 00000011 >$  to  $< 00000111 >$ .

## VI. EXPERIMENTS AND RELATED ANALYSIS

### A. Pattern recognition experiments

PIJ is a typical programming lesson in computer science, so it was used as a prototype benchmark to collect the learning behaviors of a sample of students. PIJ is used as an example of topic formation in Figures 2–5. The learning behavior of each student sample is transformed into a learning information vector using the reasoning rules described in Table IV. Table VIII shows a partial learning record within a topic for a specific subject.

As mentioned earlier, the learning information pattern evolves dynamically in different learning contents. Thus, learning information pattern recognition should be processed recursively for different learning topics to obtain a more accurate learning style prediction result.

We used the ILS to survey 50 sample students who were majoring in computer science. These were second year students from three different schools: Sichuan Normal University, Southwest University, and Chongqing University of Posts and Telecommunications. The students helped us to construct their learning style preference patterns and learning information patterns. The pretested learning style preferences of the sample learners conformed to four different distributions based on the normalized ILS results, as shown in Figure 7. The ILS results were transformed into normalized decimals between 0 and 1. We mapped the positive learning style preference scores (the scores on the active/sensing/global/visual scale were 5–11) onto the area [0.7–1], the negative learning style preference scores (the scores on the reflective/intuitive/sequential/verbal scale were 5–11) were mapped onto the area [0–0.3], and the balanced scores were mapped onto the area [0.4–0.6].

The similarity matrices of the students with respect to different learning style dimensions are shown in Figure 8. Figures 8 and 9 show graphical representations of the matrices, which demonstrate the effectiveness of our pattern recognition process. As each element  $m_{ij}$  in a matrix approaches 1, the corresponding spot in position  $(i, j)$  is lighter. By contrast, if  $m_{ij}$  approaches 0, the corresponding spot is darker. Thus, the darkest spots in the figures represent 0 and the lightest spots represent 1. In Figure 8, matrix M1 represents the similarity relationship between sample learners with respect to the “sequential/global” dimension, matrix M2 represents the similarity relationship with respect to the “sensor/intuitive” dimension, matrix M3 represents the similarity relationship with respect to the “active/reflective” dimension, matrix M4 represents the similarity relationship with respect to the “visual/verbal” dimension, and the last matrix M (Figure 9) represents the learning information similarity relationships of the sample learners for the topic “Introduction & Java Platform.” Figure 8 shows that the multidimensional space comprised four different types of mutual similarity relationships that corresponded to the learning style dimensions of the sample learners. The multidimensional space can be expanded by adding new types of mutual similarity relationships with respect to the new learning style dimensions, even if these learning style dimensions come from different learning style families. The sample learning style preference patterns were generated using Formula (1). The learning information patterns of the sample learners were generated using Formulae (2) and (3) and are shown in Figure 9.

In this experiment, the solution was  $X=[0,0,1,0]$ , which means that the highest projected value (approximately equal to 1) of the mutual similarity of the learning behaviors of the sample learners was in the “active/reflective” dimension. This solution was computed using the SA algorithm. The SA algorithm converged within 0.5 s, which is an adequate level of performance for a web-based tutoring system. To achieve this satisfactory computing performance, the original temperature in the SA algorithm was set to 5,000 degrees, the minimum temperature was 0 degrees, and the attenuation coefficient was 0.05. The maximum number of consecutive bad solutions rejected was 500. This pattern recognition performance is acceptable compared with other techniques, such as sequential pattern mining and Bayesian network training. After the SA algorithm completed the pattern recognition process, matrix M3 was distinguished from the candidate learning style preference patterns: M1, M2, M3, and M4. Compared with the other three patterns, M3 was the most similar to M. The results of this experiment facilitate two meaningful conclusions, as follows.

(1) The learning behaviors of sample learners can be used as a basis for making comparisons to predict the learning styles of new learners with respect to the dimension “active/reflective.” Thus, the current learning behaviors of sample learners cannot be used to predict the learning styles with respect to other dimensions because they do not reflect the preferences for alternative learning style dimensions.

(2) The unobserved dimensions of learning styles can also be recognized based on their learning behaviors in the context of appropriate learning topics.

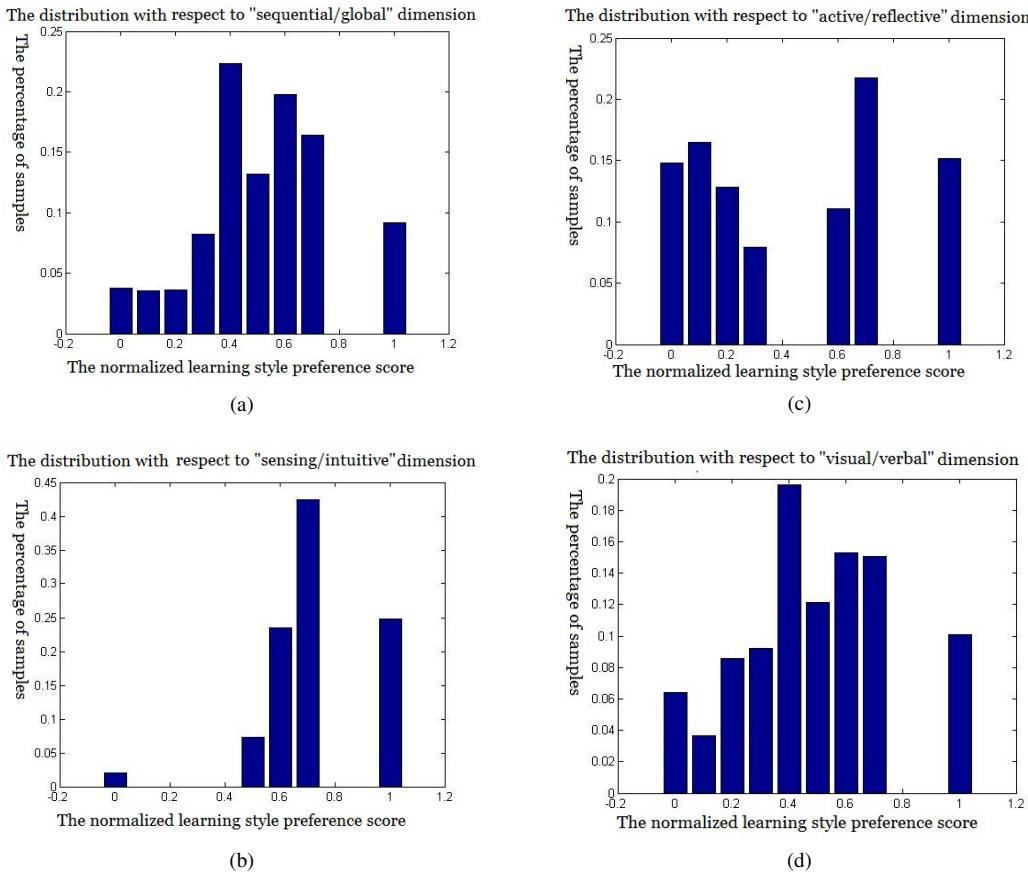
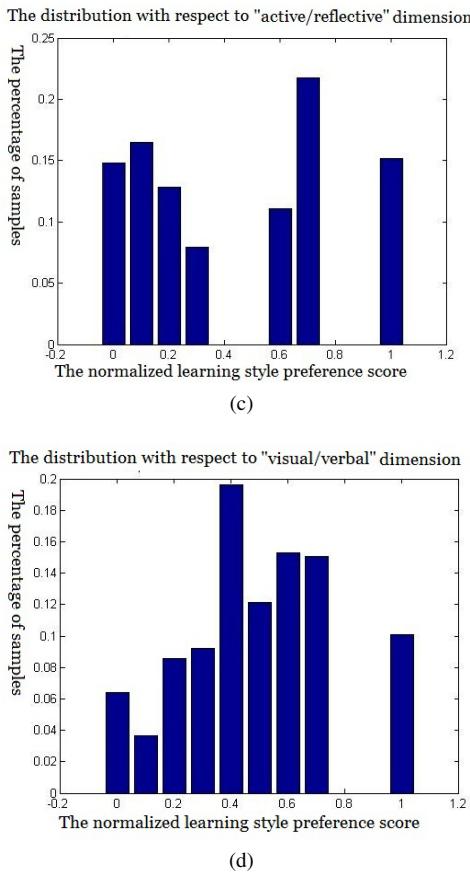


Fig. 7. Distributions with respect to different learning style dimensions

The conclusions of this experiment are very useful for predicting the learning styles of new learners. Conclusion (1) indicates that the prediction of a learning style is more accurate for a specific topic, while Conclusion (2) solves the problem of predicting the unobserved dimensions of the learning style. In this experiment, the learning information related to sample learners, which was transformed from their learning behaviors, was constructed based on their current learning topics. After identifying the learning style dimension in multidimensional space, this learning information can be used to predict the learning styles of new students. Thus, we can cluster the new students by comparing their learning information with that of sample learners in the context of the same learning topic. The new learners should have the same learning background as the samples (e.g., in our experiment, the sample learners and new learners were second year students who were majoring in computer science). If the backgrounds of the learners change, new samples need to be collected to construct new learning style patterns and learning information patterns.

#### B. Clustering experiments with new learners

We classified the new learners into their corresponding clusters. First, we built the initial cores of the clusters using the Initialization Cluster Core Construction Algorithm. Seven sample learners were selected automatically to form the cores of the clusters, who were scattered in different positions, as shown in Figure 10. The x-axis in Figure 10 represents the



normalized learning style preference scores with respect to the “active/reflective” dimension and the y-axis is the normalized Euclidean distance of the learning information vectors for the sample learners. Clearly, the initial core nodes of the “active” cluster are distributed in the upper right whereas the core nodes of the “reflective” cluster are distributed in the bottom left. Only one node did not belong to the “active” cluster or the “reflective” cluster, which we labeled as an unreliable sample node and it was deleted from the set.

Thirty additional test students were invited to participate in this experiment to verify the efficiency of the learning style prediction method. As mentioned above, these learners had the same learning background as our original samples. Thus, they could be treated as reliable sample data for clustering the new learners. Figure 11 shows the clustering results obtained for these 30 students in the context of the topic “Introduction & Java Platform.” In Figure 11, the x-axis represents the normalized learning preference scores of the test students, which were obtained using a questionnaire, and the y-axis is the normalized Euclidean distance of the testers’ learning information vectors. These 30 volunteers were asked to complete the questionnaire carefully and the learning style preference scores for the “active/reflective” dimension were used to verify whether they were clustered into the appropriate corresponding clusters. The spots distributed in the top right and circled by a dashed black line represent the test students who were clustered in the “active” cluster and

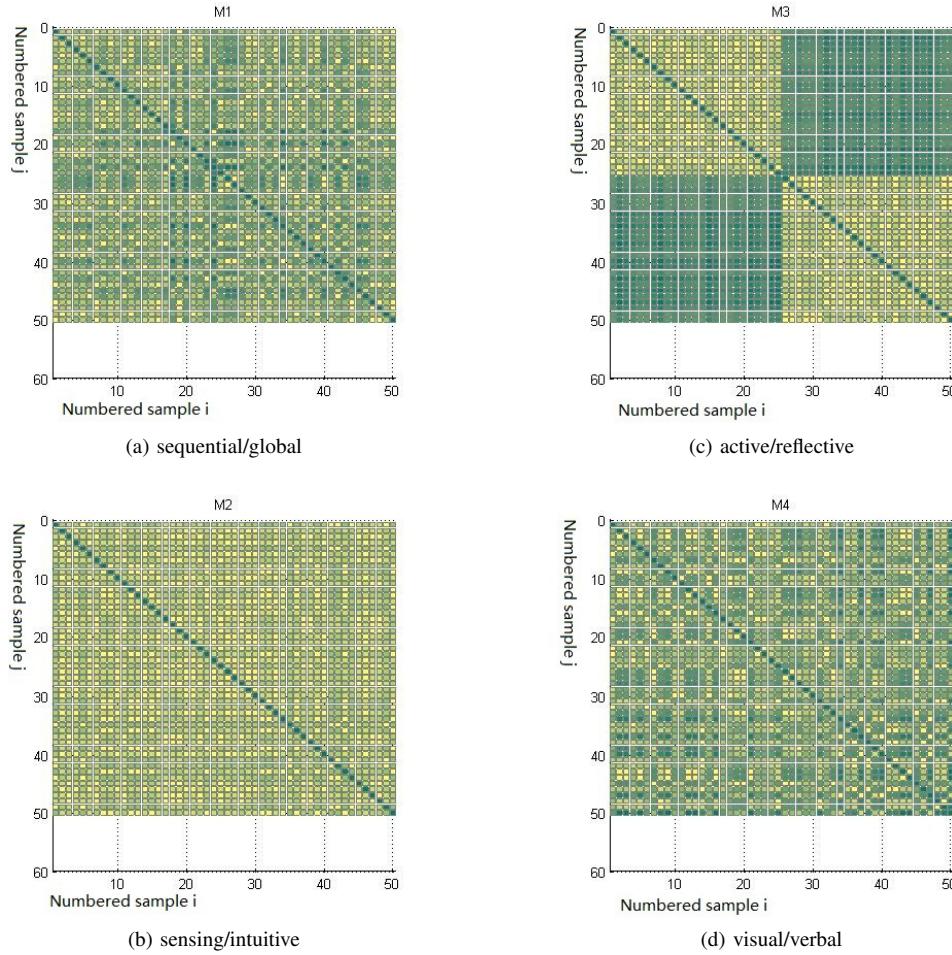


Fig. 8. Similarity relationships among different learning style dimensions

who had normalized learning style scores of greater than 0.6 ( $>0.6$ ).

The spots distributed in the bottom left and circled by a dashed black line represent the test students who were clustered in the “reflective” cluster and who had normalized learning style scores of less than 0.5 ( $<0.5$ ). Their learning behaviors were consistent with their pretested learning style preferences with respect to the “active/reflective” dimension. The training data update process had not commenced at this stage of the experiment because the students were highly likely to behave in an unpredictable and random manner during the first topic learning period.

It is interesting that few of the test students were classified as ambiguous or unknown in this experiment. The spots circled by solid red lines indicate unknown classifications because the learning behaviors of these subjects differed from their pretested learning style preference results. The spots circled by solid blue lines indicate neutral or ambiguous learners because their learning behaviors were consistent with their neutral pretested learning style preference with respect to the “active/reflective” dimension. More than four ambiguous or unknown test students are overlapping and indistinguishable in Figure 11. These results show that the highest projected values of the learning behaviors of sample learners were found in only one dimension of multidimensional space in the context

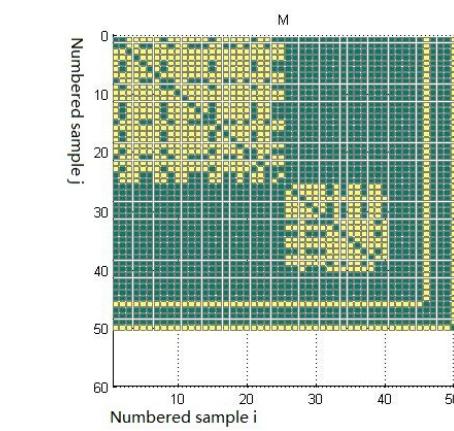


Fig. 9. M (learning behaviors of the sample learners)

of the current learning topic. Thus, the pattern recognition process can be applied iteratively to other topics to identify the dominant learning style dimension, thereby generating more accurate data to complement the learning style profiles of students.

After six runs of learning pattern recognition, the learning style predictions of the overall learning procedure for the

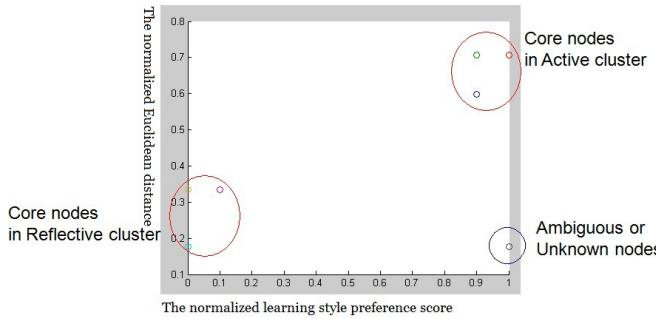


Fig. 10. Initial cores of the clusters

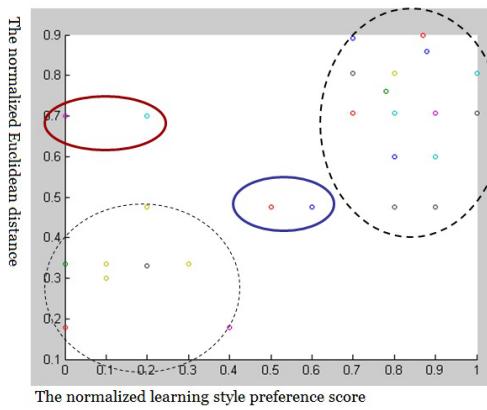


Fig. 11. Clustering of learners in the “Active/reflective” dimension

PIJ lesson are shown in Table IX. We updated the learning information patterns and recomputed the centers of the corresponding clusters in topic runs 4, 5, and 6, after 5–10 testers were integrated in the prediction process each time. There were no distortions in the collective learning information patterns or in the gradual convergence of the learning information patterns in our experiment. Most of the newly added test students had slight random variations in their behaviors compared with the original test subjects, which caused minor movements in the centers of the corresponding clusters. However, there were no obvious changes in the learning information patterns in our practical experiments. In Table IX, “unknown learners” indicate test students with learning behaviors that differed from their pretested learning style preference results, while “ambiguous learners” indicate those with ambiguous pretested learning style preferences where their actual learning behaviors supported the results.

### C. Prediction accuracy and related discussion

Table IX shows that over 30% of the test students who participated in the first two runs of the experiment were classified as unknown or ambiguous because the learning information of the sample learners was inaccurate in the initial stage. The learning processes of the sample learners and test learners changed rapidly, where they reduced the number of unnecessary browsing links, avoided inherent bias errors, etc. The learning behaviors of the sample and test

TABLE IX  
CHANGES IN THE LEARNING STYLE PREDICTIONS THROUGHOUT THE PROCESS

Topic run	<i>k</i> dimension	Unknown(%)	Ambiguous(%)
1	Active/reflective	20%	10%
2	Sensing/intuitive	22%	15%
3	Sequential/global	18%	5%
4	Visual/verbal	5%	10%
5	Sensor/intuitive	4%	12%
6	Active/reflective	5%	11%

TABLE X  
RELIABILITY OF THE PREDICTED RESULTS

Learning style dimension	Mean	Sample standard deviation(s)	Confidence interval	Confidence levels
Active/reflective	0.84	0.38	(70.5–97.5%)	85.3%
Sensing/intuitive	0.84	0.38	(70.5–97.5%)	85.3%
Sequential/global	0.77	0.43	(61.5–92.5%)	79.6%
Visual/verbal	0.85	0.35	(72.5–97.5%)	88.1%

learners gradually appeared more ordered as they learned topics in the same nature. This is the main reason why the prediction process converged rapidly in topic runs 5 and 6. The rapid convergence in topic run 4 may be attributable to the amendment of the learning information pattern associated with the “visual/verbal” learning style dimension, where eight samples were replaced three times. Table X illustrates the reliability of the predicted results by presenting the 95% confidence intervals for the final predicted results and the confidence levels associated with a sampling error of 10%.

The predictions and accuracy of our method were compared with those of other studies [16], [27], [28]. Unfortunately, only limited reference values were available for comparison with the results obtained in the present study because of differences in the experimental environments and datasets used. Compared with a previous study by [16], however, our method obtained more accurate predictions. Klasnja-Milicevic et al. used the “mean absolute error (MAE)” to evaluate the accuracy of prediction. If we define inaccurate results as clusters where the MAE values exceeded 0.4 (mean value of  $|p_i - p_j| > 0.4$ ), only topics 1 and 4 obtained relatively accurate results. In these topics, 75% of the learning styles were predicted accurately, whereas the other topics only had 40–70% accuracy. The average accuracy in the previous experiment was about 65%. In our study, the worst prediction accuracy was 63% for Topic 2, whereas the best prediction accuracy was 85% for Topic 4. The average accuracy of the predictions in our study was about 75%.

Another study [28] used 30 graduates in an experiment with 10 sample students. They achieved accuracies of 73.3% for the “sensitive/intuitive” dimension, 73.3% for the “global/sequential” dimension, 70% for the “active/reflective” dimension, and 53.3% for the “visual/verbal” dimension. The overall average accuracy for these dimensions was about 70%.

In another study [27], 75 learners participated in a learning

style prediction experiment and the researchers achieved a higher prediction accuracy than other studies. Their experimental results showed that the prediction accuracies for each dimension were: sensing = 70%, intuitive = 80%, active = 100%, reflective = 73%, sequential= 82%, global = 61%, visual = 68%, and verbal = 71%. The average overall accuracy for these dimensions was 75%, which matched our average accuracy. Thus, the method proposed in the present study delivers a relatively high prediction accuracy compared with previous methods.

## VII. CONCLUSIONS AND FUTURE WORK

In this study, we developed a learning style prediction method based on a pattern recognition technique. The prediction process is divided into two steps. The first step identifies the key learning style dimension, which is used to transform a multilabel classification problem into a single-label classification problem. The second step uses the continuously updated learning information to classify the new learners into three clusters based on the selected learning style dimension.

The concepts and methodology used in the development of the mathematical model can also be applied to other learning style approaches and other intelligent learning systems after some modifications. Our method predicts learning styles by observing critical learning behaviors and the prediction accuracy of this approach is higher than that of some previous studies.

To facilitate the widespread use of this learning style prediction method, we plan to encapsulate the main components into open-source web services, which can be invoked by any intelligent learning system. The organization of the learning resources will also be wrapped in a standard manner that complies with open standards, such as IEEE LOM.

In future research, we aim to use deep neural networks to construct an autoencoder/decoder to filter the feature learning information from over 1,000 records to improve the efficiency when recognizing the information patterns.

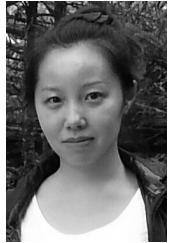
## ACKNOWLEDGMENTS

This research was supported by the following foundations or programs: the Research Foundation of the Sichuan Department of Education (14SA0008), the Fundamental Research Funds for the Central Universities (No. XDKJ2014B012), the National Society Science Foundation of China (13CGL146), and the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ130518).

## REFERENCES

- [1] M. J. Eppler and J. Mwnfia, "Concept of information overload: a review of literature from organization science, accounting, marketing, MIS, and related disciplines," *The Information Society*, vol. 20, no. 5, pp. 325-344, 2004.
- [2] N. Barjaktarevic, W. Hall, and P. Fullick, "Incorporating Learning Styles in Hypermedia Environment: Empirical Evaluation," *Proc. of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems (AH '03)*, pp. 41-52, 2003.
- [3] S. Graf and Kinshuk, "Providing Adaptive Courses in Learning Management Systems with Respect to Learning Styles," *Proc. of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (AACE '07)*, pp. 2576-2583, 2007.
- [4] H. Pashler, M. McDaniel, D. Rohrer, and R. Bjork, "Learning styles: concepts and evidence," *Journal of Psychological Science in the Public Interest*, vol. 9, no. 3, pp. 105-119, 2009.
- [5] M. H. McCaulley, "The MBTI and Individual Pathways in Engineering Design," *Journal of Engineering Education*, vol. 80, no. 5, pp. 537-542, 1990.
- [6] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [7] N. Herrmann, *The Creative Brain*, Lake Lure, NC, Brain Books, 1990.
- [8] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Journal of Engineering Education*, vol. 78, no. 7, pp. 674-681, 1988.
- [9] E. Triantafillou, A. Pomportsis, and E. Georgiadou, "AES-CS: Adaptive educational system base on cognitive styles," *Proc. of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems (AH '02)*, pp. 10-20, 2002.
- [10] K. A. Papanikolaou, M. Grigoriadou, H. Kornilakis, and G. D. Magoulas, "Personalizing the interaction in a Web-based educational hypermedia system: the case of INSPIRE," *Journal of User Modeling and User Adapted Interaction*, vol. 13, no. 3, pp. 213-267, 2003.
- [11] C. Wolf, "iWeaver: Towards Learning Style-based e-Learning in Computer Science education," *Proc. of the 5th Australasian Computing Education Conference (ACE '03)*, pp. 273-279, 2003.
- [12] J. C. R. Tseng, H. C. Chu, G. J. Hwang, and C. C. Tsai, "Development of an adaptive learning system with two sources of personalization information," *Journal of Computers & Education*, vol. 51, pp. 776-786, 2008.
- [13] R. Zatarain-Cabada, M. L. Barron-Estrada, Viridiana Ponce Angulo, A. J. Garcia, and C. A. R. Garcia, "A framework for Creating, Training, and Testing Self-Organizing Maps for Recognizing Learning Styles," *Proc. of the 5th International Conference on E-learning and Games, (EDUTAINMENT '10)*, pp. 53-64, 2010.
- [14] P. Garcia, A. Amaldi, S. Schiaffino, and M. Campo, "Evaluating Bayesian networks' precision for detecting students' learning styles," *Journal of Computers & Education*, vol. 49, pp. 794-808, 2007.
- [15] Z. Jeremic and V. Devedzic, "Design patterns ITS: Student model implementation," *Proc. of the 4th IEEE International Conference on Advanced Learning Technologies (ICALT '04)*, pp. 864-865, 2004.
- [16] A. Klasnja-Milicevic, B. Vesin, M. Ivanovic, and Z. Budimac, "E-Learning personalization based on hybrid recommendation strategy and learning style identification," *Journal of Computer & Education*, vol. 56, pp. 885-899, 2011.
- [17] P. Paredes and P. Rodriguez, "Considering sensing-intuitive dimension to exposition exemplification in adaptive sequencing," *Proc. of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems (AH '02)*, pp. 556-559, 2002.
- [18] E. Brown, C. Stewart, and T. Brailsford, "Adapting for Visual and Verbal Learning Styles in AEH," *Proc. of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT '06)*, pp. 1145-1146, 2006.
- [19] C. A. Carver, R. A. Howard, and W. D. Lane, "Enhancing student learning through hypermedia courseware and incorporation of student learning styles," *IEEE Trans. Education*, vol. 42, no. 1, pp. 33-38, 1999.
- [20] J. Yang, H. T. Liu, and Z. X. Huang, "Smap: To Generate the Personalized Learning Paths for Different Learning Style Learners," *Proc. of the 5th International Conference on E-learning and Games (EDUTAINMENT '10)*, pp. 13-22, 2010.
- [21] P. Brusilovsky, "Methods and techniques of adaptive hypermedia," *Journal of User Modeling and User Adapted Interaction*, vol. 6, no. 2-3, pp. 87-129, 1996.
- [22] D. A. Sanders and J. Bergasa-Suso, "Inferring Learning Style From the Way Students Interact with a Computer User Interface and the WWW," *IEEE Trans. Education*, vol. 53, no. 4, pp. 613-620, 2010.
- [23] N. Stash, *Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System*, Eindhoven University of Technology, The Netherlands, 2007.
- [24] N. Stash, A. Cristea, and P. D. Bra, "Explicit Intelligence in Adaptive Hypermedia: Generic Adaptation Languages for Learning Preferences and Styles," *Proc. of the International Workshop on Combining Intelligent and Adaptive Hypermedia Methods/Techniques in Web-Based Education Systems (HT '05, CIAH Workshop)*, pp. 75-84, 2005.
- [25] P. D. Bra, A. Aerts, B. Berden, B. De Lange, B. Rousseau, T. Santic, D. Smits, and N. Stash, "AHA! The adaptive Hypermedia Architecture," *Proc. of the 14th ACM Conference on Hypertext and Hypermedia (HYPERTEXT '03)*, pp. 81-84, 2003.
- [26] P. D. Bra, D. Smits, and N. Stash, "The Design of AHA!," Available online <http://aha.win.tue.nl:18080/aha/ahadesign/>

- [27] A. Latham, K. Crockett, D. McLean, and B. Edmonds, "A conversational intelligent tutoring System to automatically predict learning style," *Journal of Computers & Education*, vol. 59, pp. 95-109, 2012.
- [28] E. Ozpolat and G. B. Akar, "Automatic detection of learning styles for an e-learning system," *Journal of Computers & Education*, vol. 53, pp. 355-367, 2009.
- [29] D. Cai, Z. Shao, X. F. He, X. F. Yan, and J. W. Han, "Mining hidden community in heterogeneous social networks," *Proc. of the 3rd International Workshop on Link Discovery: Issues, Approaches and Applications (LinkKDD '05)*, pp. 58-65, 2005.



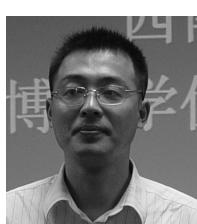
**Juan Yang** received her PhD in Artificial Intelligence from Southwest University, China, in 2007. She is a researcher in the Learning Technology Lab of the Department of Computer Science at Sichuan Normal University, China. She collaborates on research projects with the School of Computer and Information Science, Southwest University. Her research interests include intelligent learning systems, artificial intelligence in e-learning, and learning data analysis.



**Zhixing Huang** received his PhD in Artificial Intelligence from Southwest University, China, in 2006. He is an associate professor in the School of Computer and Information Science at Southwest University, China. He was a postdoctoral researcher in the Knowledge Grid Research Group, the Key Lab of Intelligent Information Processing of the Chinese Academy of Sciences, Beijing in 2010. From October 2006 to September 2007, he was a visiting researcher in the Ishida & Matsubara Lab in the Department of Social Informatics at Kyoto University, Japan. His research interests include machine learning and semantic computing.



**Yuexiang Gao** received his ME degree in Computer Application from Chongqing University, China, in 2003. He is a research fellow in the Learning Technology Lab of the Department of Computer Science at Sichuan Normal University. He is currently working toward his PhD degree at Southwest Jiaotong University, China. His research interests include web information analysis and data mining.



**Hongtao Liu** received his PhD in Artificial Intelligence from Southwest University, China, in 2007. He is an associate professor in the College of Computer Science at Chongqing University of Posts and Telecommunications, China. He is currently a visiting researcher at UCL, England. His research interests include machine learning and data mining.