RESEARCH ON SMART LEARNING SYSTEM FOR UNIVERSITIES IN EPIDEMIC ENVIRONMENT

Hua Wang1 and Jian Yu
School of Information and Electronic Engineering, Zhejiang University of Science and Technology, China
Email: {1wanghua96@126.com}

Abstract
The outbreak of the COVID-19 pneumonia epidemic has posed challenges to online teaching and cooperation in colleges and universities. In order to address these challenges, a smart learning system based on big data has been designed and developed by employing an path recommendation algorithm to implement the learning recommendation strategy. Several models are constructed to represent question domain, such as learning object model, question model and relationship model. We show how it can be applied to an actual prototype platform of smart learning system under the epidemic environment to demonstrate the usefulness of the proposed model.

Key Words: Smart learning, learning model, Learning path recommendation, Epidemic environment.

1. Introduction
The outbreak of the COVID-19 pneumonia epidemic has posed challenges to all walks of life, and the education industry is no exception. In response to this sudden war without gunpowder, the Education Ministry of China has put forward the requirements of “Suspension of classes without stopping teaching, suspension of classes without stopping school” [1], and issued the “Regarding the Organization of Online Teaching in Regular Colleges and Universities during the Period of Epidemic Prevention and Control” [2]. Major colleges and universities are prepared to carry out comprehensive online teaching according to their own conditions. However, it is the first time for this kind of comprehensive, continuous and systematic online teaching activities to be carried out. There is no experience to learn from, and everything must be explored and advanced. Under such circumstances, our department relies on the online and offline hybrid first-class courses and the complete set of recorded MOOC [3] resources as the basis, and uses the online teaching platform developed by our university to supervise students’ sign-in, preview, evaluation, discussion, etc.

When students study at home, the individual learning environment has changed from centralized to distributed, and the strong connection between teachers and students has changed to the weak one [4]. The teaching management method also undergoes essential changes. It is difficult for teachers to understand the learning status of students in real time, the restraint and supervision power on students is reduced, and the control of the classroom becomes more difficult. Online teaching requires more autonomy for learners [5]. In order to improve students’ interest and autonomy in learning at home, teachers need to obtain timely feedback from students and master their learning dynamics [6]. Therefore, accurate analysis of academic conditions and teaching by learning are the key factors affecting the effect of online teaching.

Since online teaching and cooperation may not be natural for teachers and students [7-9], there may be some challenges when implementing "suspension of classes without stopping teaching". Decision makers need to consider two main points:

1. Balance numbers and no-screen-activity. Simply substituting online lectures and discussions for school class time may harm the health of students. At this time, lectures can be shortened and combined with non-digital learning activities.

2. Always pay attention to the emotional health of students. Viruses and school closures can make students feel uneasy and confused. Technical means can be used to provide relevant links, interaction and support while learning.

In particular, learners have almost no assistance in efficient learning content delivery in the presence of multiple learning resources. Endowing an online learning system with a learning path recommendation property could push individual learning contents for learners [10]. Saito and Watanobe put forward a learning path recommendation based on neural networks applied in programming education [11]. There have been past attempts at creating recommendation models such as [12-14]. However, in our opinion, most of the current proposals suffer from the limitation of learning contents when dived in complete online teaching-learning scenarios under epidemic environment. As of now, most studies cloud not touch upon learning objects and their relationship modeling in the context. We are prone to use a practical modeling method for this point.
2. The proposed learning system

Design ideas
A smart learning system based on big data should be able to perceive and collect the learning process and result data; it can diagnose learners’ knowledge level through big data analysis, also, it can provide wisdom for students based on several factors: students’ knowledge, ability and interest preferences, to name a few. The target system should be developed to provide personalized learning content and paths; to provide individual and group learning process and result data analysis; to provide information assistance for teachers to guide learning; to evaluate the effectiveness of learning resources and tools based on the learning process; to promote resource iteration and update, so that the system has the ability to evolve.

When supporting students' learning after class, it is mainly conducted in the form of self-directed learning. The teacher selects resources from the resource library, or makes and uploads them by themselves, and the system pushes them to students according to the push strategy or chooses them independently. Here, teachers should provide students with task-driven resources, personalized listening resources, and various support personalized exploration resources. The manifestation of specific resources can be video, audio, picture, animation, construction tools, etc. During the push, resources are arranged in different ways according to specific subject teaching contents, subject characteristics, academic conditions and learning methods, forming guidance resources of different paths to suit the learning of students at different levels, and realize personalized teaching and hierarchical teaching.

In the learning process, students can choose a path that suits them according to their own learning situation. For example, when they are doing autonomous learning, they first choose an exploratory path and choose an inquiry tool for learning, but they cannot solve the problem due to difficulties encountered in the inquiry process. Someone can choose the path of watching the micro-class video for learning, no matter which path can achieve the purpose of learning knowledge. After students conduct self-study, they need to take a self-test to test their self-study mastery. Students can choose testing questions of different difficulty level according to their own learning situation to test. When students choose the difficulty level, the system automatically pushes the testing questions of the corresponding difficulty. It will be automatically reviewed and the test result will be given after answering. At the same time, according to the students' answering situation and the analysis of the testing questions, the system could diagnose the weak part of the student's learning, find the corresponding resources to push, help them strengthen and consolidate their learning, in order to achieve the purpose of automatic diagnosis and push. Such a complete guidance process can help students to learn and evaluate in a targeted and directional manner without the guidance of a teacher. And in this process, the system can record data to understand the student’s study habits and learning level, to more accurately push the learning resources and diagnose the learning situation in future learning. They both serve the classroom teaching because the preview and homework are carried out in conjunction with the synchronous learning in the classroom. The preview is the pre-learning of classroom teaching, and the homework is the testing of classroom teaching. Therefore, when supporting the learning process of pre-class preparation and after-class homework, the application of the smart learning ecosystem is more inclined to teacher-led behavior. In the preview process, the teacher pushes the prepared preview resources. The resources pushed here can include the preview study guide, courseware, and various material resources that could support the student's preview. After the students preview, the teacher formulates corresponding exercises to test the students and evaluate the students' preview situation. In the process of assigning homework, the teacher pushes homework resources. After students complete the homework, the system automatically reviews and gives homework evaluation.

In the above three processes, the system can obtain feedback data from the learning process and evaluation results, such as the number of views and collections of resources and tools. These data can reflect the teaching resources provided in the system from multiple dimensions. With the continuous accumulation and analysis of data, resources and tools are evaluated, and teaching resources and tools are continuously improved to achieve better application effects. Teachers will continue to track and obtain the feedback information of the system during the learning process of students, and provide targeted guidance.

1. The overall architecture of the smart learning system
Through investigation and analysis of college students' learning status, questions and solutions based on big data analysis technology and sensing technology employed in Internet of Things and other information technologies, the overall architecture of the smart learning system is designed by using smart learning concepts and learning processes. We divide the smart learning system into five levels, i.e., the sensor layer, the data foundation platform layer, the service cloud layer, the application layer, and the self-adaptive presentation layer.

The intelligent learning system model proposed in this paper adds a sensor layer and a self-adaptive presentation layer on the basis of the IBM big data system architecture [15]. In the smart learning environment, data sources are more diverse. In addition to the data generated by the system itself, various hardware perception devices and other systems are the basis of system data. Therefore, under the data foundation layer, a sensor layer is provided. The smart learning system emphasizes personalized, adaptable, and intelligent learning, so a self-adaptive presentation is needed in the display of learning content. The basic data is cleaned, processed, and incrementally processed to form meaningful information. We call it the data foundation, which corresponds to the data processing and storage layer in the
IBM big data system architecture. The service cloud layer corresponds to the data analysis layer of the IBM big data system architecture, which is also the core of our system. Here we mainly refer to the general model of the adaptive learning system, including the learner model, the domain knowledge model and the adaptive engine based on the following data foundation platform to provide personalized learning services through the engine which has self-evolving learning capabilities as shown in Fig. 1.

![Fig. 1 The overall architecture of the smart learning system](image)

(1) The sensor layer
The sensor layer is composed of various application terminals such as personal computers, mobile phones, tablets, cameras, microphones, other collection hardware, a wide variety of smart devices and external application systems, and is mainly responsible for the perception and acquisition of various learning data of learners. With the support of big data and cloud computing, the systems can be connected. The smart learning system can be associated with other systems to obtain more extensive data and serve as the analysis of the smart learning system. The output of one system can be the input of another system. For example, the scores of the students in the previous examinations in the performance management system can be imported into the smart learning system through data connection, which becomes the most direct means of evaluating students' learning level.

(2) The data foundation platform layer
Through the data interface, the data foundation platform obtains the original data collected through the sensor layer to form images, sounds, images, and learners' application process data in other learning systems. The data resource also include records of offline learning activities, and the use of system web logs, data streams, click stream and other information to form a rich structured and semi-structured original data pool.

Furthermore, the data in the data layer is integrated, converted and cleaned through ETL [16] and other technologies, and stored in the learner information database, learner behavior database, domain model database, learning resource database, learning tool library, guided strategy library, rule database and other databases by using continuous data collection and conversion.

(3) Service cloud layer
This layer is the core of the smart learning system. The adaptive engine provides personalized learning services based on the learner model and domain model. According to the design principle of "ecological perspective", it constantly evaluates the service results and updates engine rules during the service process. Realize the self-evolution of the engine. The core content of the adaptive engine includes the diagnosis of learners' knowledge level, the
recommendation of learning paths, the recommendation of learning resources, and the evaluation of the effectiveness of resources and tools.

(4) Application layer
The application layer is a collection of various business services for users, including learning content presentation, learning tool support, learning context creation, learning activities, teacher guidance activities, learning strategies, etc. The application layer provides modular services. In addition, the application layer also provides external data access services, which can import or enter external data such as offline learning results.

(5) Self-adaptive presentation layer
The adaptive presentation layer controls which application is displayed and how it is displayed based on user information and personal preferences. Each person can have their own personalized software interface. Self-adaptive rendering mainly includes:

- Self-adaptive presentation of the device according to the device performance, system, screen size, etc.
- Elastic computing, that is, the calculation subject is automatically assigned according to the computing power of the device used. For example, if the user uses a mobile terminal, the logical calculation subject is the mobile terminal, which uses the form of a thin client.
- Self-adaptation of network status, that is, to identify the network status, wired network, wireless network or mobile network, different network status uses different calculation and display methods.
- Self-adaptive interface style, that is, comprehensively considering the user's usage habits and user-defined interface forms, presents different styles of display methods according to users' different preferences. This kind of self-adaptation also includes adaptive application navigation, that is, giving different application navigation according to different user roles to meet the needs of users.

2.1 Learning object model
Learning objects include learning resources and tools, which are the basic elements that support the learning of knowledge points. The types of learning objects are diversified, such as text, video, animation, pictures, games, personalized learning tools, etc. Diversified learning objects can provide learners with different learning preferences with more suitable learning options. Because the type of learning object may be a composite type, for example, there may be both text, pictures, and video in a tutorial case. In order to record the type of learning object more scientifically, we use a six-dimensional vector to represent it where each component takes a value. The range is [0,1]. For example, if the type of a learning object is (0.2,0.0,4,0.8,0,0), it indicates that the learning object contains both text, animation and pictures. Because the content of the learning object is different in the degree of abstraction and the way of knowledge expression, its applicability to learners of different learning styles is also different. Therefore, it is necessary to record the applicable learning style of the learning object. The applicable learning style is a four-dimensional vector. The learning object has a direct influence on the training and development of the learner's core literacy of knowledge, so we incorporate the core literacy of knowledge into the attributes of the learning object. The core literacy is not only integrated and comprehensive, but also hierarchical and continuous. Based on the above analysis, the key attributes of learning objects in Table 1 are determined.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Type</th>
<th>Limit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Object Title</td>
<td>varchar</td>
<td>1024</td>
<td>A title for a learning object</td>
</tr>
<tr>
<td>Applicable Learning Style</td>
<td>vector4</td>
<td></td>
<td>Custom four-dimensional vector</td>
</tr>
<tr>
<td>Media Type</td>
<td>vector4</td>
<td></td>
<td>Media type identification</td>
</tr>
<tr>
<td>Preset Learning Time</td>
<td>int</td>
<td>12</td>
<td>Domain experts estimate the time it takes to complete the learning object</td>
</tr>
<tr>
<td>Average Learning Time</td>
<td>int</td>
<td>12</td>
<td>The average time it takes for multiple learners to learn the object</td>
</tr>
<tr>
<td>Number of Learning</td>
<td>int</td>
<td>11</td>
<td>Total number of learning times for multiple learners</td>
</tr>
<tr>
<td>Total Learning Time</td>
<td>int</td>
<td>11</td>
<td>Total learning time of multiple learners</td>
</tr>
<tr>
<td>Evaluation of Learning Objects</td>
<td>float</td>
<td>(0,1)</td>
<td>The average value of multiple learners' evaluations</td>
</tr>
<tr>
<td>Supporting Knowledge Points</td>
<td>string</td>
<td>1024</td>
<td>Knowledge points related to the content of the learning object</td>
</tr>
<tr>
<td>Associated Content Path</td>
<td>string</td>
<td>1024</td>
<td>Thumbnails, physical files, end-to-end adaptable files</td>
</tr>
<tr>
<td>Core Literacy of Knowledge</td>
<td>vector1</td>
<td>2</td>
<td>Support for the development of core knowledge</td>
</tr>
</tbody>
</table>

2.2 Question model
According to IRT (Item Response Theory) [17], after learners have learned a target knowledge, the learner's mastery of the target knowledge can be diagnosed through question push and learner's feedback information. The purpose of the question is to test the knowledge level of the learner, that is, to select the question through a certain
strategy, to evaluate the knowledge level of the learner, and to update the data into the learner model. The solution of questions has a direct relationship or requirements to the core literacy of knowledge, so the question model also includes the core literacy of knowledge. The question attribute items, value types and attribute constraints determined are shown in Table 2.

Table 2. List of key attributes of the question

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Type</th>
<th>Limit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Type</td>
<td>int</td>
<td>12</td>
<td>The type of question, such as multiple choice, true or false</td>
</tr>
<tr>
<td>Question Content</td>
<td>varchar</td>
<td>2048</td>
<td>Question description file</td>
</tr>
<tr>
<td>Reference Answer</td>
<td>varchar</td>
<td>2048</td>
<td>Question answer description file</td>
</tr>
<tr>
<td>Degree of Difficulty</td>
<td>float</td>
<td>(0,1)</td>
<td>Indicates the difficulty of the question</td>
</tr>
<tr>
<td>Distinction</td>
<td>float</td>
<td>(0,1)</td>
<td>Indicates the degree of topic discrimination</td>
</tr>
<tr>
<td>Preset Answer Time</td>
<td>int</td>
<td>12</td>
<td>Domain experts’ valuation of the time it takes to solve the question</td>
</tr>
<tr>
<td>Average Answer Time</td>
<td>int</td>
<td>12</td>
<td>The average time it takes for multiple learners to answer the question</td>
</tr>
<tr>
<td>Number of Questions</td>
<td>int</td>
<td>12</td>
<td>The number of times the learner asked the question</td>
</tr>
<tr>
<td>Problem Evaluation</td>
<td>double</td>
<td>(0,1)</td>
<td>The average value of multiple learners' evaluations</td>
</tr>
<tr>
<td>Knowledge Points</td>
<td>string</td>
<td>1024</td>
<td>Knowledge points needed to solve the problem</td>
</tr>
<tr>
<td>Core Literacy of Knowledge</td>
<td>vector12</td>
<td></td>
<td>Support for the development of core knowledge</td>
</tr>
</tbody>
</table>

2.3 Learner model

The learner model reflects the learner’s individual characteristics, including personal information, knowledge level, interest preferences, learning habits and other learner characteristics. Based on the goals of the smart learning system and the characteristics of college students, this we gives the following learner model:

\[
\text{Student Model}=(\text{bi, kl, ip, lh, lt})
\]

- bi is the basic information representing the learner’s name, ID, password, gender, age, class, school, hobbies, contact information, technical operation level and physiological characteristics, etc.
- kl is the knowledge level indicating the course number, knowledge point number and mastery level.
- ip is the interest preference denoting types of learning objects that learners tend to choose when learning, including text, video, animation, pictures, games, personalized learning tools, and so on. The preference of the learning object indirectly reflects the preference of the learner's learning style. The learning methods of the smart learning mainly include inquiry learning, listening learning and reading learning. For example, if learners like to choose personalized tools for learning, it reflects learners' preference for inquiry learning. Therefore, the interest preference proposed in this paper does not express the learning style preference as an independent element. Because learners may have different preferences for learning objects of different media types, the same method is used to represent the types of learning objects. The learner’s interest preferences are described by a six-dimensional vector as: \(\text{ip}=(\text{text, video, animation, pictures, games, personalized learning tools and others})\), where each component has a value range of \([0,1]\).
- lh is the learning history recording the course number, number of logins, login time, learning time, learning activity records, learning behaviors in the learning process and other information. Among them, learning activity records could be homework records, resource clicks, access times, residence time of a single resource, self-study records, practice records, testing records, and activity sequences such as participation topic records and discussion records. It can also include information about changes in emotional states and facial expressions during activities, forming big data and storing them in the database.
- It is the learning terminal recording the type of learning terminal, the number of using, equipment performance information, system version, screen resolution and network operation status, etc. Some resources can be used well on the PC side, but may not be presented on the mobile side, because different learning terminals have different levels of support for the presentation of learning objects. In order to enable the learning object to display adaptive presentation on the terminal device, the system will record the information of the learning terminal used by the learner, so as to better recommend learning resources to the learner, and also provide reference for the development and construction of the learning object to ensure learning The individualized learning effect of the reader.

2.4 Relational model

The relational model describes the relationship between knowledge elements. When modeling the relationship between knowledge elements, we establish three relationships, namely, the relationship between knowledge points, the one between learning objects and knowledge points, and the one between knowledge points and problems.
The relationship model between elements is shown in Fig. 2. Here, the establishment of the relationship between knowledge points is to facilitate the comprehensive diagnosis of knowledge level. The establishment of the relationship between learning objects and knowledge points is to push learning resources and provide reference information, and the establishment of the relationship between knowledge points and questions is to test learners. This provides a basis when selecting questions for learning effects.

In Fig. 2, the line segment between the points represents the relationship between the knowledge points. The line segment between the knowledge points has two forms. One is the solid line representing the predecessor/successor relationship, and the other is the dashed line representing the correlation relationship. The relationship between the learning object and the knowledge point represents the degree of support of the learning resource to the knowledge point. The line segment between the knowledge point and the question represents the degree of support for solving the problem. The line segment between the question and the core literacy represents question solution requiring the core literacy. And the line segment between the learning object and core literacy indicates the degree of support of the learning object to the core literacy development.

### 2.5 Intelligent learning path recommendation algorithm based on ant colony algorithm

#### Algorithm: Learning path recommendation

- The learner selects the target knowledge point \( K_n \);
- Generate the fully connected network graph \( G \) of the learning object according to \( K_n \);
- Initialize the road section evaluation information;
- Build Referential Learner Population (RLP);
- //according to the evaluation information of the learners in the RLP
- Update the road section evaluation information;
- Create ant colony \( C_0 = \{c_1, c_2, \ldots, c_l\} \);
- //initialize ant colony algorithm
- Set learning path length \( len \);
- Determining the initial learning object;
- //Perform optimization calculations for \( len \) times in a loop
- For each \( c \in C \):
  - a) Select the road section according to the probability selection function;
  - b) Record the route;
  - c) Modify the taboo list;
- Update road section information table;
- Output the learning path according to the path evaluation value.

If the learner takes an assessment after learning, and the knowledge level is improved, the assessment is adopted, otherwise it is not adopted. The reason is that the learning object may not be very effective, but the possibility of error and misleading is extremely low. Therefore, it can be ignored. After the learning, the decline of the knowledge level is mainly due to the learner’s personal reasons (such as forgetting). The global path evaluation is specifically expressed as follows. Suppose that user \( u \) has completed the learning task, let \( P \) represents the learning path, \( \theta_m \) is the evaluation of the learning path by user \( u \), \( K_u(t+1) \) represents the level of knowledge after learning, and \( K_u(t) \) represents the level of knowledge before learning, then the evaluation update formula for each section on path \( P \) is:

\[
\theta_m(t+1) = \begin{cases} 
(1 - \mu) \cdot \theta_m(t) + \theta_m^p, & (p_m \rightarrow p_n) \in P \\
(1 - \mu) \cdot \theta_m(t), & (p_m \rightarrow p_n) \notin P 
\end{cases}
\]

We now show how to recommend the learning path according to global path evaluation as shown in Fig. 3.

### 3. The implementation of the proposed system

To demonstrate the usefulness of the proposed model, we show how it can be applied to an actual prototype platform of smart learning system under the epidemic environment. Some snapshots are demonstrated as follows. Fig. 4 shows how to model the learning object and Fig. 5 demonstrates...
how to build the question model. Personalized learning function model for learners is explained in Fig. 6.

The personalized learning function module recommends suitable learning resources and tools to learners based on the learners' interest preferences, knowledge level, learning history and learning style. Learners can also choose learning resources and tools for learning. Fig.7 shows the student learning scenario on the mobile terminal.

4. Conclusion
We established a smart learning system for college and university students under the epidemic environment by focusing on the main hub of the smart learning system and the system engine to conduct key research. The principles, processes, and algorithms of various services of the smart learning ecosystem are explained from three aspects, that is diagnosis of learners' knowledge level, recommendation of smart learning paths, acquisition of auxiliary information for learning guidance.

Aiming at the discontinuity of the learning system used by college students, the planning of teaching, and the modularization of learning, the learning path recommendation strategy analysis was carried out, and how to select the recommendation algorithm to implement the recommendation strategy was further discussed. Finally, based on the improved ant colony algorithm, the relevant parameters of the system are optimized, and a complete description of the algorithm is given. This research expands the application scope of ant colony algorithm and realizes intelligent learning path recommendation based on learner model and domain knowledge model.

The smart learning system is implemented and the practical application experiments are carried out.
References


