

Received: January 4, 2018

Revision received: March 21, 2018

Accepted: March 28, 2018

Copyright © 2018 EDAM

[www.estp.com.tr](http://www.estp.com.tr)

DOI 10.12738/estp.2018.5.031 • October 2018 • 18(5) • 1323-1331

*Research Article*

# A Study on Computer Distance Individualized Instruction based on Web Mining Algorithm

Lelai Shi<sup>1</sup>

Wuhan Textile University

Dingqiang Duan<sup>2</sup>

Wuhan Textile University

## Abstract

To study the computer distance individualized instruction, this paper applies literature investigation method, summarization method and case study to mine the learning data. Based on the analysis on the current distance teaching system, it is found that the system can't satisfy the learner's requirement of individualized instruction. To solve this problem, this paper proposes distance individualized instruction model based on data mining technology, and analyzes the model and its implementation model. It also modifies Apriori algorithm with the aim to mine learner logs to find out their frequent transversal path in the distance instruction system through Web mining. The result was then solved out to enable the individualized recommendation of its interface, which validated the feasibility and effectiveness of the teaching model and provided a new reference direction for distance individualized instruction.

## Keywords

Computer Distance Instruction • Data Mining Technology • Individualized Instruction • Web Mining

\*This work was supported by the Doctoral Research Start-up Fund of Wuhan Textile University (No. 175018).

<sup>1</sup>Center for Industrial Economic Studies, School of Economics, Wuhan Textile University, Wuhan 430200, China. Email: shilelai@hotmail.com

<sup>2</sup>Correspondence to: Dingqiang Duan, Centre for Industrial Economic Studies, School of Economics, Wuhan Textile University, Wuhan 430200, China. Email: ubrer@live.cn

**Citation:** Shi, L. L., Duan, D. Q. (2018). Study on Computer Distance Individualized Instruction based on Web Mining Algorithm. *Educational Sciences: Theory & Practice*, 18(5), 1323-1331. <http://dx.doi.org/10.12738/estp.2018.5.031>

The rapid development of information technology and network technology facilitates the computer distance teaching becoming a new method to learn. It enables people enjoy teaching service offered by distance instruction system at home, and will become an important method for lifelong learning.

Relevant literature home and abroad shows that although the study and application on computer distance instruction is relatively mature, the learning resources in the distance teaching system are all sharing resources, which can't satisfy the individualized needs of learners and differentiate their diversity (Özbek and Eker, 2015). This greatly contradicts the individualized instruction advocated by education reform, and traditional distance teaching platform can't settle this problem from the root. The rise of data mining technology makes individualized service of computer distance instruction possible, as it can extract valuable patterns from numerous data (Springer and Pear, 2008). If those patterns are effectively employed, the learners can enjoy individualized instruction service.

To settle the problems existing in the traditional distance instruction system, this paper proposes distance individualized instruction model and its implementation model based on data mining technology. It also transforms Apriori algorithm (Cakir and Simsek, 2010) with the aim to mine learner logs of the actual users to find out their frequent transversal path in the distance instruction system through Web mining, which enables the individualized recommendation of its interface. This validated the feasibility and effectiveness of the teaching model and provided a new reference direction for distance individualized instruction.

This paper is grouped into four chapters. The first chapter is the introduction that introduces the background and present problems of computer distance instruction system and proposes the research content—the computer distance individualized teaching model based on data mining technology. The second chapter illustrates the distance instruction system and data mining technology, thus laying a theory basis for the following passage. Based on the first two chapters, it also presents the computer distance individualized instruction model and its implementation model. After a deep analysis of the model, it figures out the individualized learning interface via Web mining, which validated the feasibility and effectiveness of the model. The last chapter concludes the paper.

## **Basic theory and relevant technology**

### **Distance teaching system**

With the help of computer and network technology, computer distance instruction system enables the sharing of quality teaching resources and breaks the limit time and space for teachers and learners, thus, learners can choose lessons needed following their requirements at any time and any place and communicate with their teachers (Zhang *et al.*, 2010). This greatly facilitates the self-directed learning.

Figure 1 presents a distance teaching system model commonly used at this period (Allen *et al.*, 2013). The system consists of five parts -- courseware learning, Q&A system, communication system, homework system and examination system. Below the five parts, there are four subordinate databases – courseware database, Q&A database, homework database and examination database.

The current distance instruction system already enables resource sharing and breaks the limit of time and space, thus realizing rapid information transmission and combining self-directed learning and assistant learning. In this way, it makes education more social and helps learning blend into daily life (Abler and Wells, 2005). However, the model didn't organically integrate databases of different systems. It only provides single system module function instead of analyzing learning characteristics and behaviors of various learners. The learning resources it offers for learners are the same, which neglects the individual discrepancy (Almaoui et al., 2007). Therefore, it can't offer individualized instruction service following the learner's trait, violating learner-centered constructivism learning theory and modern education development trend.

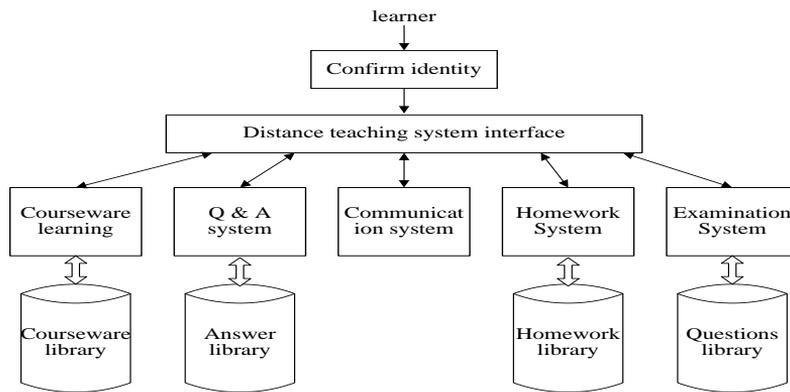


Figure 1. Traditional distance learning model.

**Data mining technology**

In the era of big data, numerous data are produced at every moment. Except for simple data query and retrieval, the data mining technology makes it possible to extract valuable patterns through data acquisition and analysis that serves as a reference for decision making.

At present, the most commonly-used data mining algorithms are cluster analysis, association analysis, decision tree, neural network and genetic algorithm, etc. after establishing the data mining model, any algorithm can be applied the mine data, and the mining process can be realized through computer science and statistics. Most importantly, an optimal algorithm should be figured out. Figure 2 are major procedures of data mining (Dai and Zhang, 2008).

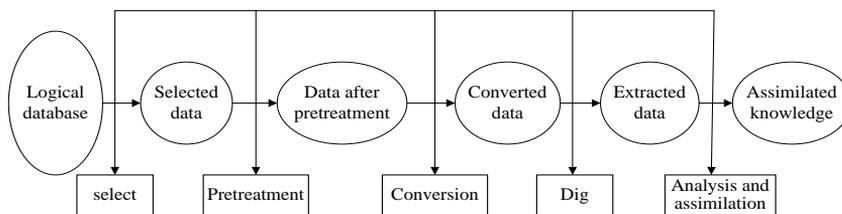


Figure 2. Data mining process and steps.

## Computer distance individualized instruction system based on data mining technology

### System model

This paper proposes a model of computer distance individualized instruction system based on data mining technology to settle the problem in the current computer distance instruction system and offer learners a learning interface catering to their interests, needs and learning status (Salinas and Zhang, 2015). The functions of its different modules are shown in Figure 3.

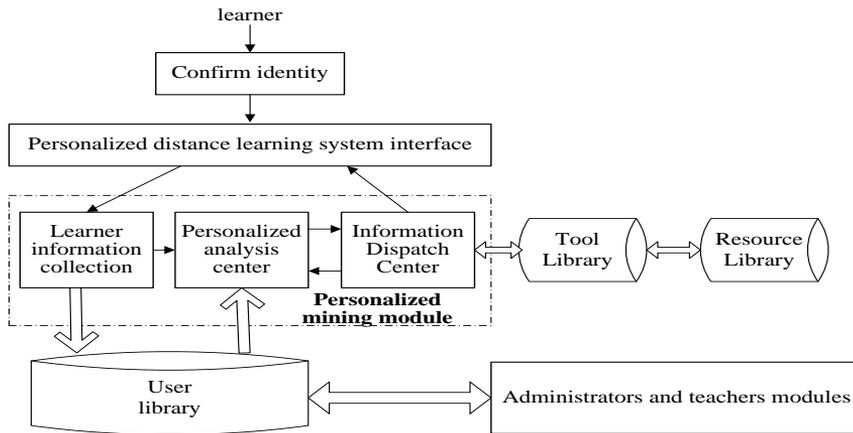


Figure 3. Personalized distance learning system model based on data mining.

User database. The database involves in database of learning behavior, learning record and learning characteristic. The learning behavior database composes of the learning time, order and content, etc.; the learning record database is the learner’s archival information like history grade and progress; the learner characteristic database is made up of the characteristics of learner’s knowledge structure, emotion etc. The sorting of those characteristics playing a significant role in offering individualized service.

Resource database. It is a database related to teaching resource such as courseware, test questions and homework.

Tool database. It transforms the four separated systems – homework system, Q&A system, communication system and homework system in the traditional system into subordinate applications that can be invoked.

Individualized mining module. It is the key to realize individualized teaching, and can be divided into three parts.

(1) learner information collection. This part is the foundation of individualized instruction system. The quality and quantity of data collection will have directly impact on the analysis result. It uses intelligent Agent to acquire data like learner’s basic information, interested questions and learning progress, and deliver the result to the individualized analysis center. (2) Individualized analysis center. It is made up of user information

database and information process module. The user information database stores user's data. The information process module conducts individualized analysis via data mining technology (including methods like statistical analysis, association rule, clustering and clarifying) and sends the instruction to information dispatch center. (3) Information dispatch center. According to the analysis result from individualized analysis center, the information dispatch center dispatches information interface that can satisfy the specific requirement of learners, meanwhile, learning information feedback will be sent back to individualized analysis center to update the user's information database. (4) Manager and teacher module. These two modules are respectively in charge of the management of distance instruction system as well as the management of learning resources and users. (5) The system interface of individualized distance instruction. Learners use the interface to learn, and the interface send the learner's information and learning content to the individualized analysis center to form an individualized interface catering to the learner's characteristics. This is a dynamic circulation.

### System implementation model

To better set up a computer distance individualized instruction model, this paper has designed a system implementation model (Ai and Zhang, 2010) shown in Figure 4.

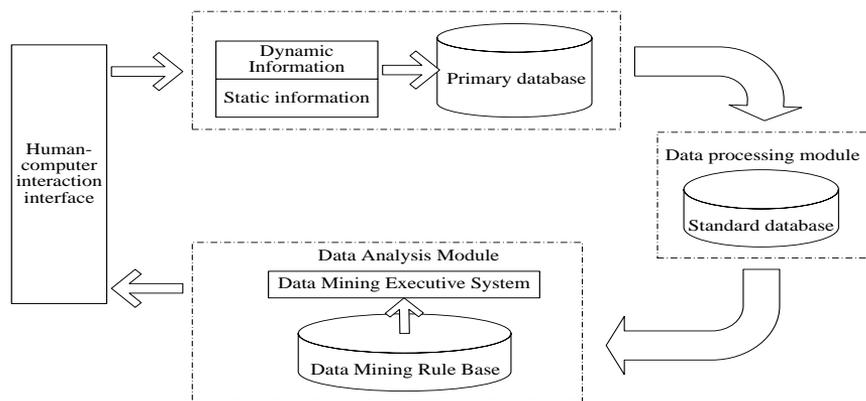


Figure 4. Implementation Model of Personalized Distance Learning System Based on Data Mining.

The model is made up of four modules, of which the functions are explained as follows: (1) Human-computer interaction interface. It is a module enabling human-computer interaction through natural language and semantic query. The individualized interface is a result of data mining. (2) Data collection module. This module collects information like commonly-visited websites of learners, learning time and hobbies, and transforms it to data processing module. The quality and quantity of the information will directly impact the effectiveness of system analysis; thus, it is a fundamental module of the system. (3) Data processing module. To speed up data processing, this module needs to extract data related to learners' behaviors from numerous historic data in the database and the updated data. (4) Data analysis module. This module is the core of the system. Analyzing the extracted data of learners, it chooses suitable model and data mining algorithm to find out the learner's interest.

**Offering learners individualized learning interface with Web mining**

**The overview of Web mining.** The Web mining is the application of data mining technology in Web, a process mining the potential pattern and valuable information in the data produced while browsing the Web with data mining technology. It is often grouped into Web content mining, Web structure mining and Web usage mining (Preidys & Sakalauskas, 2010). This paper applies Web usage mining, and its process shown in Figure 5 mainly consists of three steps: data pretreatment, pattern discovery and pattern analysis.

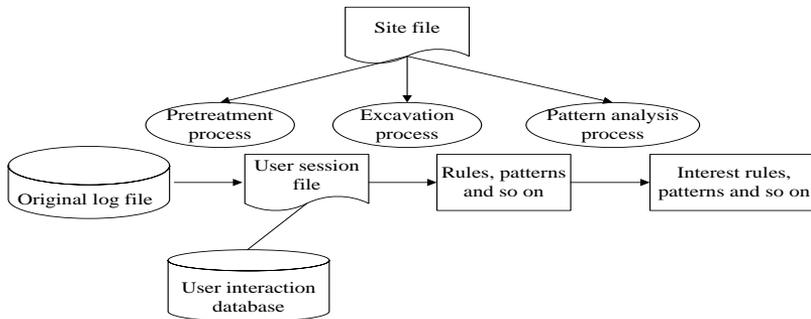


Figure 5. Web log mining process.

(1) Data pretreatment. This step transforms learners’ original visiting logs into database used in the stage of log mining. What it needs to process concludes data cleansing, user identification, conversation identification and visiting path optimization. Its result plays significant role in the data mining. (2) Pattern discovery. This step works out the maximal forward reference set by the maximal forward reference algorithm. Based on the set, it figures out the maximal forward reference sequence, from which the frequent visiting sequence can be found out. The process is supported by multiple technologies like data mining technology, statistic theory and artificial intelligence. (3) Pattern analysis. At this step, it removes valueless result from the process of pattern mining to gain users’ frequent visiting path and pattern. Therefore, it can restructure the website and offer the learner an individualized interface.

**The application of Web mining technology in distance instruction.**

**The mining process:** Most of the computer distance instruction systems are based on Web, thus, a Web database can be formed through mining Web log to record its frequent traversal path. Based on this, a Web interface suitable for individual requirement and the individualized distance instruction can be realized.

Table 1  
Instance Data

TID	Page reference sequence
15	ABCD
25	ABC
35	BCDE
45	BEFG
55	BCDF
65	EFG

To figure out user’s frequent traversal path, this paper modifies Apriori algorithm to form a sequential frequent-path interface (Xia *et al.*, 2013). To illustrate the implementation process of frequent algorithm, it also selects database containing six transactional records which are chosen from a user visit table of a distance instruction system. As shown in Table 1, the capitals are page labels preset.

Assume that the minimal support threshold is 3; D is the database;  $C_K$  is a candidate item set of frequent K-reference;  $L_K$  is set of all frequent K- reference. Figure 6 shows the diagram of frequent path mining algorithm (Joksimović *et al.*, 2015)

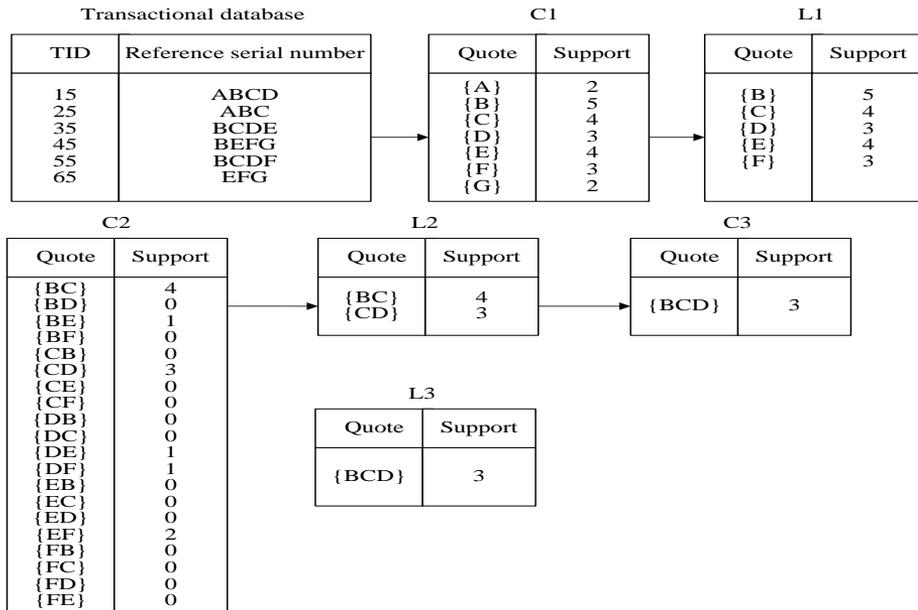


Figure 6. Frequent path mining algorithm diagram

**The application of mining result.** Once the mining result solved out by the algorithm, the computer distance instruction system will automatically adjust the learning interface to comply with learners’ habit next time when they login. Thus, learners can know more about their learning condition and help themselves with their self-directed learning. The result can also assist teachers to know the progress and ability of learners, hence adjusting their teaching content and progress. In this way, they can provide learners practical learning content and advice, thus achieving the real individualized instruction.

## Conclusion

Has studied the distance individualized instruction, this paper has the following conclusions based on the research on the present computer distance instruction system and relevant literature home and abroad:

1. The analysis on the current computer distance instruction system finds out that the system can't satisfy individualized instruction, therefore, it needs to be further improved.
2. Based on the data mining theory, it proposes the model of computer distance individualized instruction system and analyze its implementation process in detail.
3. This paper mines the learner' behavior log in the instruction system with Web mining and modified Apriori algorithm. Consequently, it enables individualized instruction interface recommendation, thus validating the practicability and effectiveness of the model.

## References

- Abler, R. T., & Wells, I. G. (2005). Distributed engineering education: evolution of the telecollaboration stations for individualized distance learning. *IEEE Transactions on Education*, 48(3), 490-496. <http://dx.doi.org/10.1109/te.2005.849757>
- Ai, Y., & Zhang, J. (2010). The Application of data mining technology in distance learning evaluation. *International Forum on Information Technology and Applications*, 3, 145-148. <http://dx.doi.org/10.1109/ifita.2010.153>
- Allen, J., Gugerty, L., Muth, E., & Scisco, J. (2013). Remote technical support requires diagnosing the end user customer as well as the computer. *Human-Computer Interaction*, 28(5), 442-477. <http://dx.doi.org/10.1080/07370024.2013.770360>
- Almaoui, M., Kushki, A., & Plataniotis, K. N. (2007). Metadata-driven multimedia transcoding for distance learning. *Multimedia Systems*, 12(6), 505-520. <http://dx.doi.org/10.1007/s00530-006-0071-8>
- Cakir, O., & Simsek, N. (2010). A comparative analysis of the effects of computer and paper-based personalization on student achievement. *Computers & Education*, 55(4), 1524-1531. <http://dx.doi.org/10.1016/j.compedu.2010.06.018>
- Dai, S., & Zhang, P. (2008). A data mining algorithm in distance learning. *International Conference on Computer Supported Cooperative Work in Design, IEEE*, 1014-1017. <http://dx.doi.org/10.1109/cscwd.2008.4537118>
- Joksimović, S., Gašević, D., Loughin, T. M., Kovanović, V., & Hatala, M. (2015). Learning at distance: effects of interaction traces on academic achievement. *Computers & Education*, 87, 204-217. <http://dx.doi.org/10.1016/j.compedu.2015.07.002>
- Özbek, N. S., & Eker, İ. (2015). An interactive computer-aided instructional strategy and assessment methods for system identification and adaptive control laboratory. *IEEE Transactions on Education*, 58(4), 297-302. <http://dx.doi.org/10.1109/te.2015.2412512>
- Preidys, S., Sakalauskas, L. (2010). Analysis of students' study activities in virtual learning environments using data mining methods. *Ukio Technologinis Ir Ekonominis Vystymas*, 16(1), 94-108. <http://dx.doi.org/10.3846/tede.2010.06>
- Salinas, J. G. M., & Stephens, C. R. (2015). Applying Data Mining Techniques to Identify Success Factors in Students Enrolled in Distance Learning: A Case Study. *Mexican International Conference on Artificial*

- Intelligence*, Springer, Cham, 208-219. [http://dx.doi.org/10.1007/978-3-319-27101-9\\_15](http://dx.doi.org/10.1007/978-3-319-27101-9_15)
- Schiaffino, S., Garcia, P., & Amandi, A. (2008). Eteacher: providing personalized assistance to e-learning students. *Computers & Education*, 51(4), 1744-1754. <https://doi.org/10.1016/j.compedu.2008.05.008>
- Shen, R., Han, P., Yang, F., Yang, Q., & Huang, J. Z. (2003). Data mining and case-based reasoning for distance learning. *International Journal of Distance Education Technologies*, 1(3), 46-58. <http://dx.doi.org/10.4018/jdet.2003070104>
- Springer, C. R., & Pear, J. J. (2008). Performance measures in courses using computer-aided personalized system of instruction. *Computers & Education*, 51(2), 829-835. <http://dx.doi.org/10.1016/j.compedu.2007.09.002>
- Xia, H., Wu, P., & Hoi, S. C. H. (2013). Online multi-modal distance learning for scalable multimedia retrieval. *ACM International Conference on Web Search and Data Mining*, 455-464 <http://dx.doi.org/10.1145/2433396.2433453>
- Zhang, Y., Su, G., & Zheng, W. (2010). Converting legacy desktop applications into on-demand personalized software. *IEEE Transactions on Services Computing*, 3(4), 306-321. <http://dx.doi.org/10.1109/tsc.2010.32>