

Research on Data-driven Feedback Teaching Service

Jiangbo Shu^{1, a}, Li Wang^{1, b}, Xu Wang¹, Min Zhi¹, Taihe Cao¹ and Hai Liu¹

¹National Engineering Center for E-Learning, Central China Normal University, Wuhan, China

^ashujiangbo@mail.ccnu.edu.cn, ^bwinlily@foxmail.com

Key words: Data-driven, Feedback, Big data, Teaching services

Abstract. In recent years, most of the teaching service models in universities are linear in the exploration of big data in education, and people who use data are mainly researchers and policy makers, the process of data generation is static. Aiming at this problem, this paper proposes a data-driven feedback teaching service model. The model is characterized by: 1) the service mode is closed and nonlinearity; 2) the process of data generation is cyclic and dynamic; 3) the results of the data analysis are fed back to the participants in the teaching activities, intervening the process of generating data effectively. Finally, two examples are given to illustrate the effectiveness of the proposed model.

Introduction

With the continuous accumulation and depth development of big data in education, big data will be propitious to our individualized education and has a profound impact on teaching and management [1]. The application of big data in the field of education, mainly refers to the three elements of on-line decision, learning analysis, data mining. Its main function is to carry on the application and research of predictive analysis, behavior analysis, and academic analysis and so on [2, 3].

In China the biggest problem in the field of education is that there is no awareness of retained data and data collection. At present, colleges and universities that in better use of data are mainly concentrated in data analysis for network system, support for teaching and learning management, analysis for particular students, data analysis support for scientific research and so on. Such as Tsinghua University, Zhejiang University, Fudan University, etc. [4-6].

At present, most of the research on big data in education is still in theory in domestic universities. There is no real landing and the data analysis is not feedback to the teachers, students or learning system. Aiming at this problem, this paper attempts to research the data driven feedback teaching service. Through the analysis and processing of the data generated on the teaching platform, on the one hand, the effective data features are implicit feedback to the teaching platform, in order to optimize the service of the teaching platform, on the other hand, the specific data characteristics are explicit feedback to the participants in teaching activities, such as teachers, students and administrator, to provide support for the timely adjustment of teaching strategies and the timely optimization of teaching plan.

Data-driven feedback teaching service model

Most of the teaching service models are linear at present, which appears as the teaching support platform (teacher, student) → big data in education → data processing and visualization → education researchers and policy makers. The characteristics of the linear model is that data generated in the progress of using teaching support platform by teachers and students, then it had the corresponding processing of data including noise reduction, clean, analysis, mining and visualization, finally presented to education researchers and policy makers.

If the results of big data in education processing are fed back to platform maintainers, teachers,

students and the other directly related people, they can obtain the data with the minimum cost. In this way it will be able to make the education data results benefit more users, it can effectively help platform maintainers to optimize the design of the system, teachers and students to adjust teaching strategies or learning plan.

Based on the above ideas, it proposes data driven feedback teaching service model as shown in Fig. 1.

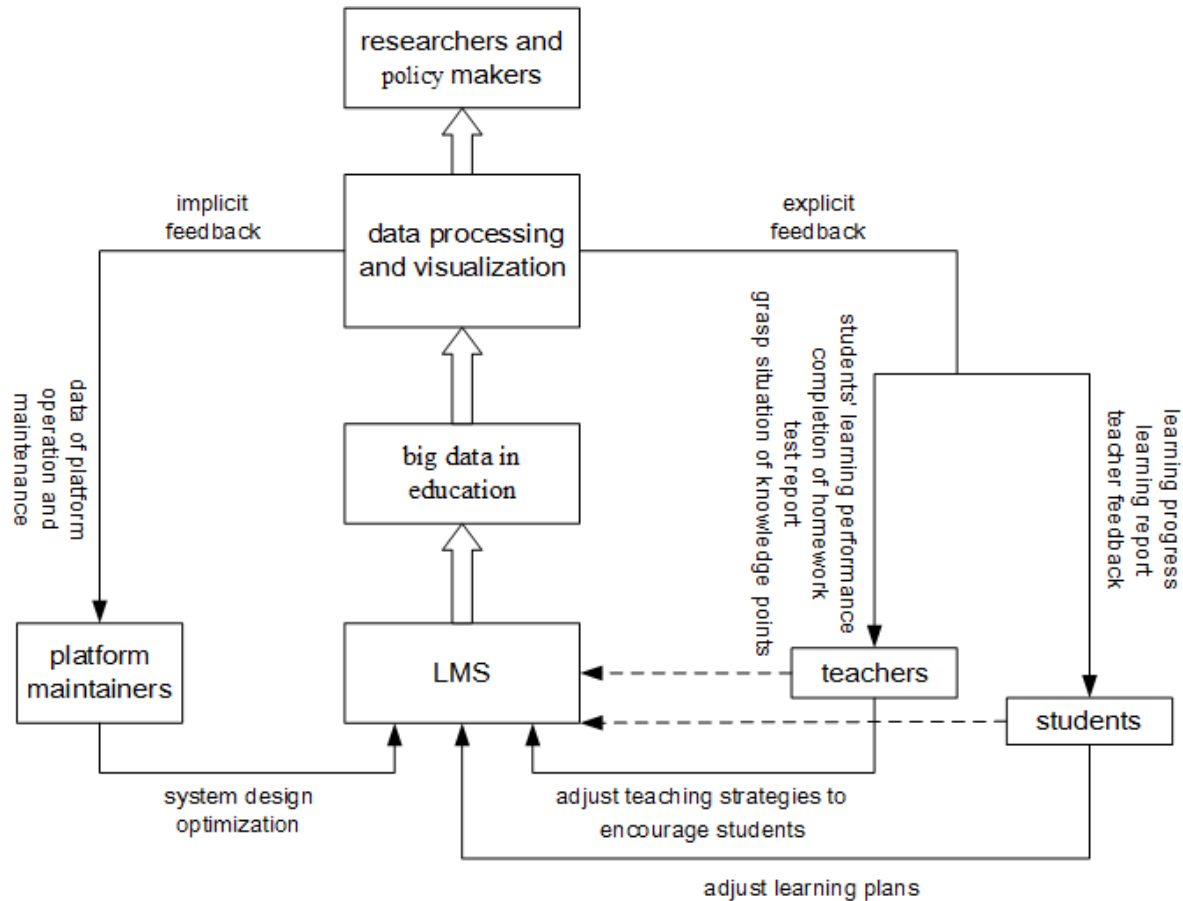


Fig. 1 Data driven feedback teaching service model.

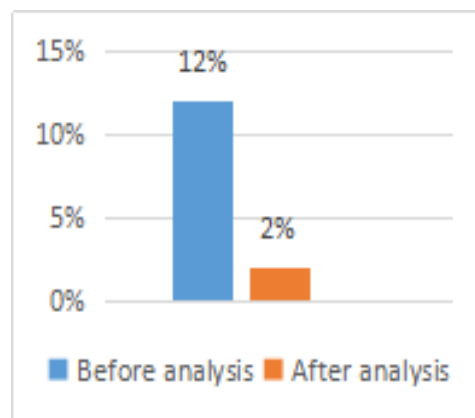


Fig. 2 Comparison of specific gravity.

The main characteristics of feedback teaching service model are the big data in education not only provide services to researchers and policy makers by processing and visualizing, but also have two directions. On the one hand, implicit feedback data to platform maintainers, On the other hand,

explicit feedback data to teachers and students. After getting feedback, platform maintainers can optimize the design of the system, teachers can adjust teaching strategies in a timely manner, and take timely incentives for students to promote the virtuous circle of teaching, and students can adjust their study plan, and promote the virtuous cycle of learning.

Data-driven teaching service applications

The Cloud Classroom (short for CC), an integrated teaching support service platform which is developed by CCNU, has provided teaching support services for a majority of teachers and students. This paper takes the teaching practice of CC as an example to illustrate the application of data-driven teaching service. And optimize the teaching services through data processing and analysis.

User login service optimization. CC is a platform providing teaching support services for students and teachers in CCNU. It differs from other open platforms that the new users don't need to register by themselves. Namely, they cannot retrieve the passwords via the mailbox. Permissions of users' are set by the platform administrator. That is, when users forget their password, they can only turn to platform maintainers for help. In the process of teaching, users can get services of the platform, only if they log in. Many users forget their passwords, in the process of logging in, so that they cannot access the system, making it hard to carry on with the teaching activities.

To address this issue, 37336 CC users were selected as the research object. The analysis of users' password characteristics and related factors acts as the background who doesn't bind the mailbox and need to retrieve the password from system administrator. According to the risk value that the user forgets the password, it asks high-risk users to bind mailbox recommendations to reduce the probability of retrieving the password.

There are three aspects which will affect the user password risk value.

- 1) Changing Interval: That is the time interval for changing the password. The longer the interval, the higher the risk of forgetting.
- 2) Password Complexity: The complexity of the new password is calculated by applying the password complexity calculation method. And the rules are as follows:
 - a. Characters are divided into four categories: numbers, lowercase letters, uppercase letters, and special characters (@ #! and the like).
 - b. Each type of character is given a complexity and denoted as P_1 : number is defined as 1, lowercase letters is defined as 2, uppercase letters is defined as 3 and special characters is defined as 4.
 - c. Giving the transformation between characters a complexity and denoted as P_2 : the conversion complexity between same characters type is 0, numbers to lowercase characters is 1, numbers to uppercase characters is 2, numbers to special characters is 2, lowercase characters to uppercase characters is 1, lowercase characters to special characters is 2 and uppercase characters to special characters is 2.

Consequently, the complexity of the password is converted to characters' complexity and transformation's complexity. And it can be formalized as follows:

It sets the password as $S = S_1 S_2 S_3 \cdots S_n$ and the complexity of S is denoted by $P(S)$. $P(S)$ is calculated as follows.

$$P(S) = \sum_{i=1}^n P_1(S_i) + \sum_{i=1}^{n-1} P_2(S_i, S_{i+1}) \quad (1)$$

For instance, the complexity of 'Gre72' is calculated as follows.

$$P(Gre72)=P_1(G)+P_1(r)+P_1(e)+P_1(7)+P_1(2)+P_2(G,r)+P_2(r,e)+P_2(e,7)+P_2(7,2) \\ =3+2+2+1+1+1+0+1+0=11 \quad (2)$$

- 3) Password Characteristics: Users are divided into the following categories on the summary of 37,336 users' passwords. Related to birthday, telephone number, name and some meaningful words and numbers. The weight of password is set to 0 for users who do not change the password. And according to different categories to give 5-20 weight range for users who has changed.

Then, giving different weight value for different factors after comprehensive analysis of the above three factors. And according to the experience, the value of changing interval accounts for 20%, password complexity accounted for 40% and password characteristics accounted for 40%.

Let *UN* denotes as student number and faculty number, *PW* denotes as password, *PWS* denotes as password classification, *PWSW* represents the proportion of password classification, *PWC* represents the password complexity, *PCCW* represents the weight of the password complexity, *TI* represents the time interval, *TIW* represents the weight of the time interval and *PWIC* represents the overall complexity of the password. Then, it has the following formulas.

$$PWSW = ROUND(PWS * 40\%,0) \quad (3)$$

$$PWCW = ROUND(PWC * 40\%,0) \quad (4)$$

$$TIW = ROUND(TI / 100 * 20\%,0) \quad (5)$$

$$PWIC = PWSW + PWCW + TIW \quad (6)$$

Eq. (3) calculates the values of the user's passwords for different feature categories. Eq. (4) calculates the value of the password complexity. Eq. (5) calculates the time interval value. And Eq. (6) calculates the total value of users' password complexity. According to the above algorithm, users' integrated password risk value can be calculated. And part of the user password risk value is shown in Table 1.

Table 1. User password risk value

UN	PW	PWS	PWSW	PWC	PWCW	TI	TIW	PWIC
2012980049	199612	5	2	4	2	1114	2	6
2007980090	19640730	5	2	8	3	1113	2	7
...
2015111025	2015111025	0	0	6	2	16	0	2

According to system maintenance statistics, about 12% of users had forgotten the password. Thus, it defines users who are the lowest 12% integrated password risk value as High-risk groups. And the system will urge high-risk people to bind mailboxes. Thus, when the user forgets the password, they can retrieve the password through the mailbox self-help.

Figure 2 is the proportion of users who had forgotten password. It can be seen that the password to retrieve significantly decreased according to the data. On the one hand, the study from three aspects of the password risk analysis is valid. On the other hand, it can optimize the system's quality of service through this method.

Course Forum Service Optimization. At present, online course forums have become important components of various education websites and online courses [7]. They also give students good online learning circumstances. Forum discussion is an important way for the interaction of online teaching, which can improve the social existence and cognitive existence of students, deepen

students' understanding of the knowledge they have learned, and enhance students' sense of belonging to the learning community [8]. Course forum help teachers and students learn from each other, enhance mutual feelings, and create a collaborative learning atmosphere [9]. But in the practical teaching process, the effectiveness of the course forum needs further exploring.

In this study, it chooses a well-developed course "Pedagogy Basis" in CC. The time range of analysis is from Feb-15-2016 to Apr-15-2016, which is called Phase I. The course has 1 instructor and 75 students. Statistics show that teachers posted a total of 11 times, the total number of students replying to teachers is 309; students posted a total of 2 times, number of replies among students is 2. Through data sorting, you can get Table 2, which lists partial interactions between teachers and students.

Table 2. Teachers and students' posting and replies relationship table in Phase I.

Poster	Replies	Number
2010980110	2014210474	15
2010980110	2013212394	13
...
2014211255	2014212369	1

According to the data of teachers and students' posting and replies relationship table, the social network analysis is carried out. The teachers' teaching staff number and student's student number are set as nodes, teacher and student roles are set to T and S respectively. NetDraw is used to draw the social network relationship diagram, which is shown in Fig. 3.

In Fig. 3, according to the analysis of central degree, there are four shapes, namely Circle, Box, up Triangle, Square, with their central values being 0, 1, 2 and 31. The central value of the shape indicates the number of people who have posts and replies interaction with the person represented with the shape. From Figure 3, it can be seen that the teacher with the shape of Square, is surrounded by a number of Boxes and up Triangles, and there is a lot of isolated circles on the left. This shows that some students and teachers have a small amount of replies, the other part of the students do not post, and they also have no reply interaction with teachers and other students, finally, it illustrates that students do not actively participate in forum posting interaction.

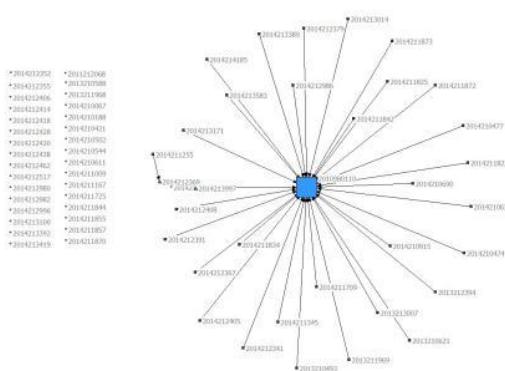


Fig. 3 Reply interactions between students and teachers in Phase I.

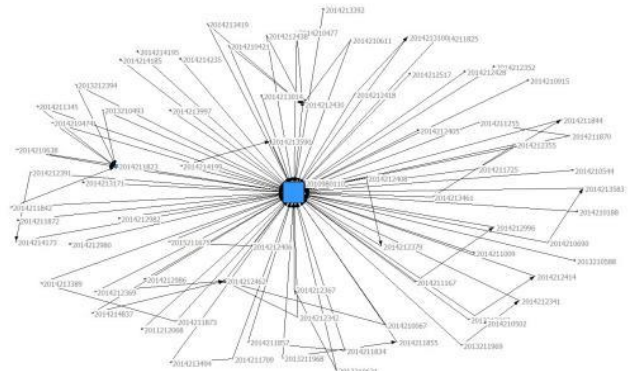


Fig. 4 Reply interactions between students and teachers in Phase II.

The analyst returns the results to the instructor. After taking these data characteristics, teachers take some measures to mobilize the students, to encourage students to post, participate in the forum topic interaction, make full use of curriculum forum resources. Through the follow-up study on the forum of the "pedagogy basis" course, it selects Apr-15-2016 to Jue-15-2016 as a phase, which

called Phase II, and then analyzes the forum posting and replies data again. The statistics shows that teachers posted a total of 13 times, the total number of students replying to teachers' posts is 781; students posted a total of 20 times, 46 students replied to other students' posts. NetDraw gave us the social network diagram in Fig. 4.

In Fig. 4, it shows that there is no isolated point, that is, all students are involved in the discussions of the course forum, and the interactions among students increased, which shows that providing feedback data to teachers, can effectively improve the activity degree of curriculum forum, and optimize the quality of the course forum.

Summary

In this paper, in the exploration of big data in education, it puts forward a data driven feedback teaching service model. And the validity of the model is illustrated by two application case. With the popularity of national data awareness, the author suggests that teaching institutions should try to consider the role of teaching activities participants in the service system, and the resulting data transfer mode in the top-level design of teaching services, in order that data can be used to achieve maximum efficiency in the data service system.

For the record, the feedback teaching service model is only a guiding model. Different teaching service platform needs to be practiced and optimized according to its special features.

Acknowledge.

This research was partially funded by the self-determined research funds of CCNU from the colleges' basic research and operation of MOE (CCNU16JYKX027), the National Natural Science Foundation of China under Grant (No.61505064), Hong Kong Scholars Programs (Grant Nos. XJ2016063), and the Project of the Program for National Key Technology Research and Development Program (2014BAH22F01).

References

- [1] Changhong Yu and Yunwu Wang, "The Goal, Content and Strategy of Digital Campus Construction in the Background of Large Data," *China Educational Technology*, 10(2013), 30-35.
- [2] Lihai Zheng, "The Mode of Educational Management in the Big Data," *China Educational Technology*, 07(2015), 32-36.
- [3] Qiao Chen, Zhi Shi and Yeqin Shao, "Research on the Construction of Individual Cloud Resources Cabinet in Big Data," *Modern Educational Technology*, 23(2013), 111-114.
- [4] Qingchao Ke, "Big Data and Wisdom Education," *The Chinese Journal of ICT in Education*, 24(2013), 8-11.
- [5] Zhuoxuan Jiang, Yan Zhang and Xiaoming Li, "Analysis and Prediction of Learning Behavior Based on MOOC Data," *Journal of Computer Research and Development*, 52(2015), 614-628.
- [6] Yanfang Yu, "Research on the Design Idea and Technology Framework of the Credit Information Management Platform of the Credit Bank under the Background of big data in education," *Distance Education in China*, 06(2015), 53-59.
- [7] Dandan Gao, "The Evaluation of Network Learning," *E-education Research*, 11(2009), 25-29.
- [8] Sujuan Yang, "Research on Cognitive Existence of Web - based Course," *Distance Education in China*, 03(2011), 41-45.
- [9] Xiao Liu, "Analysis on the Application of Curriculum Forum," *China Educational Technology*, 01(2011), 77-79.