Design of College Student Information Management System based on Big Data

Jiayi Lin

The Academic office of Jinan University, Guangzhou, China

linjy_1990@163.com

Abstract. The study aims to present an integrated overview of one holistic college students' affairs management system, which is designed to tap into big data with a view to enhancing educational administration. It adopts the B/S architecture, allowing real-time interactive operation of the data through web-based interfaces. It mainly includes basic information management of students, academic affairs, and student affairs supported by a solid data processing layer based on Hadoop components. It integrates various data sources, such as the tracking of student locations using RFID technology, in an attempt to monitor academic performance, campus activities, and even mental health. By employing advanced data mining techniques, the system aims to provide timely insights and early warnings to both students and counselors, thereby fostering a supportive educational environment. This innovative approach not only smooths the administrative processes but enriches the in-school experience due to personalized support and data-driven decisions.

Keywords: Big Data, Student Management System, Browser/Server Architecture

1 Introduction

Big data refers to data sets that cannot be acquired and processed within the expected time using traditional data processing tools [1, 2]. Compared with traditional data sets, big data has the essential characteristics of fast data production speed, large data volume, complex data structure, and low data value density [3]. However, compared with traditional data sets, the fundamental advantage of big data is not the vast data set. Still, effective information is obtained after repeated processing, thereby improving data use and economic value. The big data processing process includes data collection, storage, management, analysis and visualization [4]. A single computer is complex to meet the requirements of massive, high-speed and parallel processing of big data, and distributed data processing is often used to solve this problem [5]. The Hadoop ecosystem is currently the primary tool for system development based on big data, including MapReduce, HDFS, HBase, Hive, Sqoop, etc., which can meet the needs of big data processing [6-9].

S. Yang et al. (eds.), *Proceedings of the 2024 Seminar on Educational Technology and Management Information Systems (ETMIS 2024)*, Advances in Social Science, Education and Humanities Research 911,
https://doi.org/10.2991/978-94-6463-674-1_19

The Internet of Things (IoT) refers to the interconnection of devices or individuals through wireless radio frequency technology, sensor technology, embedded technology, network technology, and more [10]. Each device or person in the IoT has a unique radio frequency identification (RFID) code, which enables identification, positioning, tracking, communication, and other functions [11, 12]. It is an essential means of data collection and forms the basis for the development of big data processing technologies. Colleges and universities widely use student cards to facilitate student attendance, book borrowing, campus consumption, and more [13, 14]. Electronic tags are preinstalled on the card, with basic information such as the student's name, student number, and class prewritten. The student card serves as the student's identity in the student card management system (shown in Figure 1). Radio frequency identification mobile phone cards combine student cards with mobile phone cards, making them easy to carry [15, 16]. Students can recharge, make purchases, and query information by linking their student card account to their mobile phone [17]. This improves the security and utilization of student cards and accelerates the formation of intelligent digital campuses. It serves as the primary data collection source for student management systems based on big data.

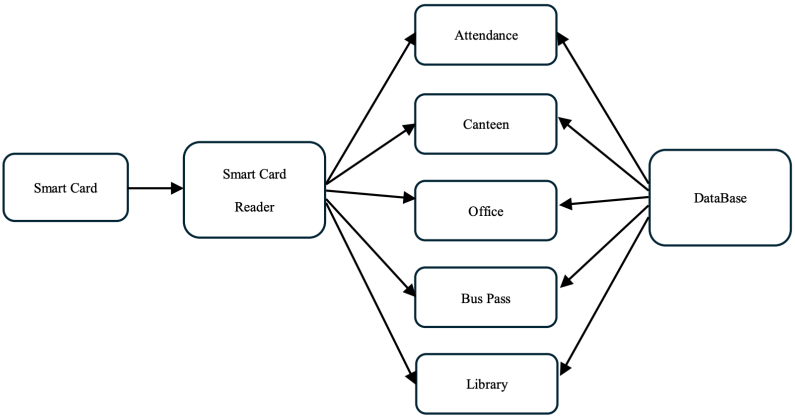

Fig. 1. Architecture of Student Card System [18]

With the advancement of comprehensive reforms in colleges and universities, the number of departments and individuals responsible for student management has gradually increased. Accurate, thorough, and real-time access to student information is essential for effective student management. Big data is crucial for managing student information. This study aims to design a data source, data storage, data processing, and application layers to develop a student management system. It enables student management information sharing to improve the efficiency and quality of student management.

2 Current Status of Information Management of College Students

Big data processing is an effective technology that can quickly and accurately collect social media data from different users or enterprises [14]. These data are diverse, rich, and variable, and can provide us with valuable information to better meet the needs of social development. Data analysis plays a vital role in student management and has excellent value. Tulasi [19] emphasized that applying big data in higher education has opened a new chapter. It can explore the correlation between data and provide managers with an effective decision-making basis to better guide the education process. Richardson and Lenarcic [20] proposed a new method, that is, by integrating information systems such as mobile SMS platform management system, school broadcast and network terminal, to quickly evaluate and reflect on the situation of students, to find problems in time and reduce management costs. This method can help college administrators better manage students and improve the efficiency and quality of school management. Richardson and Lenarcic [20] applied valuable and actionable information to the field of education by analyzing a large number of data sets, to deeply explore the behavior and learning situation of learners and the learning effects of their use of different learning resources. This method can help us better understand students' learning status and provide valuable suggestions for educational decision-making. Yang and Zhang [21] mainly introduced the application of data mining in the education management information system of college students, provided some methods and ideas for better management of student information, and proposed the application research method of data mining in the education management information system of college students including educational data. [22] believed that big data can be regarded as an algorithm for analyzing and processing data, so it can be used to create student information management strategies. Experimental results show that big data is essential in developing and implementing student information management systems.

The current research reflects the critical value of big data in student management and the development and design of specific student management systems. However, the content of the management system involved in the research is relatively simple, mainly for student performance management. This paper applies big data thinking and involves comprehensive content on student system management, covering student safety, academics, and psychology. It monitors students' academic performance, campus activities, daily attendance, and mental health. Through data analysis, it sends early warning notifications to counselors and students themselves about upcoming problems. Counselors can help and guide students to solve problems and regulate student behavior. It transforms the previous post-processing management method into preventive management before and during the event.

3 Analysis of College Student Management Information System based on Big Data

3.1 Target Analysis

The college student information management system, based on big data, provides robust software and hardware access capabilities for establishing a unified student information management database for the entire university. It identifies and connects multiple types of terminal devices to meet the personalized needs of users accessing the system. Through big data, student data is collected, stored, cleaned, and mined to create information files for all students at the university. Data analysis and visualization tools provide graphical displays of campus safety status, campus epidemic trends, student employment distribution, and student mental health status. Users log in to the system through authorized accounts, with different user identities having different operational permissions. Users can upload, download, query, and modify information on relevant pages, making it easier for student management personnel to fully grasp student dynamics. At the same time, this system also creates convenient conditions for students to access campus information.

3.2 Functional Requirements Analysis

The college student management system based on big data is divided into four functional modules: student basic information management, academic affairs information management, student affairs information management, and logistics information management. The responsibility for ensuring data authenticity, real-time accuracy, and validity lies with the student's department, the Academic Affairs Office, the Student Affairs Office, and the Logistics Office. Students, department (college) administrators, and school administrators are the three types of users of the system. The student's academic information is imported into the system by the Academic Affairs Office. After logging into the system for the first time, the student can access other functional modules after modifying the password and completing their personal information. The school administrator assigns the department (college) administrator account. After logging in, the department administrator can upload, download, modify, delete, publish, lock, encrypt, and manage the information for their department (college), as well as view and download information from other departments (colleges). After logging into the system, the school administrator can view all system data and open, add, delete, or lock the functional modules of the department (college) according to business needs. The system functional module diagram is shown in Figure 2.

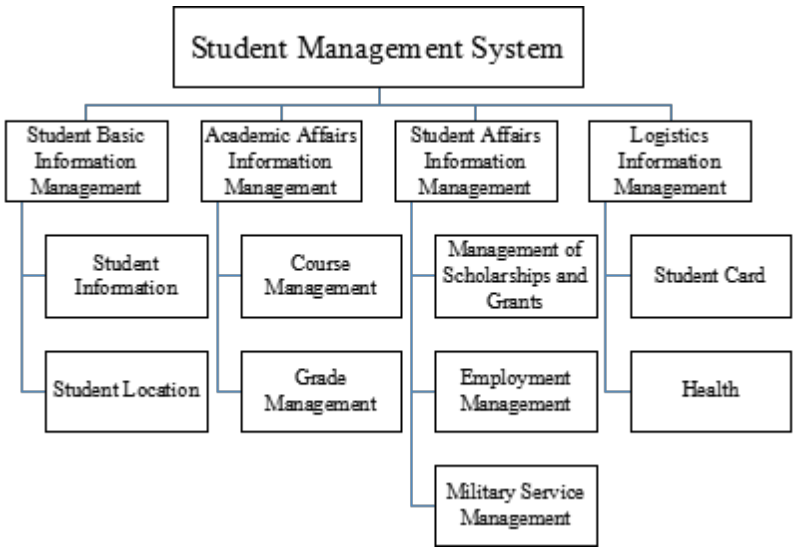

Fig. 2. System Function Module Diagram

Student Basic Information Management.

Student information is imported from the backend by the Academic Affairs Office, including the student's name, student number, gender, date of birth, ID number, college, major, class, time of admission, and more. After logging into the system, students can view this information and complete their personal contact details, dormitory information, home address, hobbies, family economic status, and more. Student location data is collected through the radio frequency identification (RFID) student card, which tracks the time students enter and leave the campus, dormitory, library, and other locations.

Academic Affairs Information Management.

The Academic Affairs Office is responsible for managing academic affairs information, publishing course details, creating competition notices, promoting competition progress, and more. Students can query the course offerings for their major and all courses available at the university this semester, query and print their academic transcripts and timetables, receive various competition notices, and check competition results. The school administrator can import students' grades for the college and issue academic warning reports to students who fail.

Student Affairs Information Management.

The Student Affairs Office is responsible for managing student affairs information, including publishing assessment policies for outstanding student cadres, scholarships, and grants at all levels; opening application channels for various awards and scholarships regularly; and publishing employment and military service information. Stu-

dents can query the above information, apply for awards and subsidies online, and complete employment and military service registration. In this module, departments can allocate dormitories to college students, import class teacher information, and view, summarize, review, and export the results of various subsidy applications submitted by students.

Logistics Information Management.

The Logistics Department issues student cards, manages online binding, loss reporting, and recharging of student cards, publishes dormitory information for departments and colleges, provides health information, arranges mental health surveys for the entire university, and sets up online counseling for mental health. After students bind their student cards, they can use them for medical treatment or to purchase medicines at the school hospital, shop at the school supermarket, and purchase water and electricity. They can also conduct online mental health consultations and make offline appointments through the client interface. Information such as book borrowing and entering or leaving the campus will be uploaded to the student card management system in real-time and further integrated into the student information management system. Departments and colleges can view and modify relevant student information within this module.

4 System Architecture of a College Student Management System based on Big Data

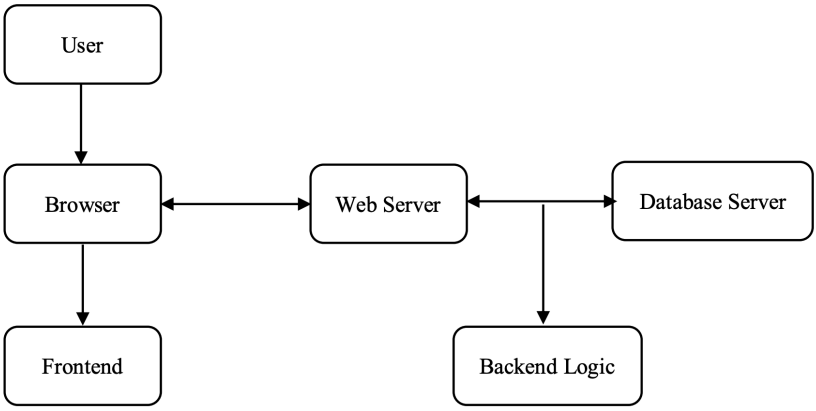

Fig. 3. B/S development mode

The college student management system based on big data adopts the B/S (Browser/Server) development mode, i.e., browser requests and server responses (shown in Figure 3). The access terminal must only log in to the system through a Web browser

for identity authentication. After the Web server accepts the access request, the database server stores the data information and provides the access result. The Web server then feeds back the access result to the user, thereby realizing data interaction. To meet the real-time processing of large-scale, multi-source heterogeneous data and to perform model analysis, this system mainly uses Hadoop open-source components, and the program is written in Java. It adopts a layered design, divided into a data source layer, a data storage layer, a data processing layer, and an application layer. The upper layer can access the lower layer, and the lower layer provides services for the upper layer. The system architecture diagram is shown in Figure 4.

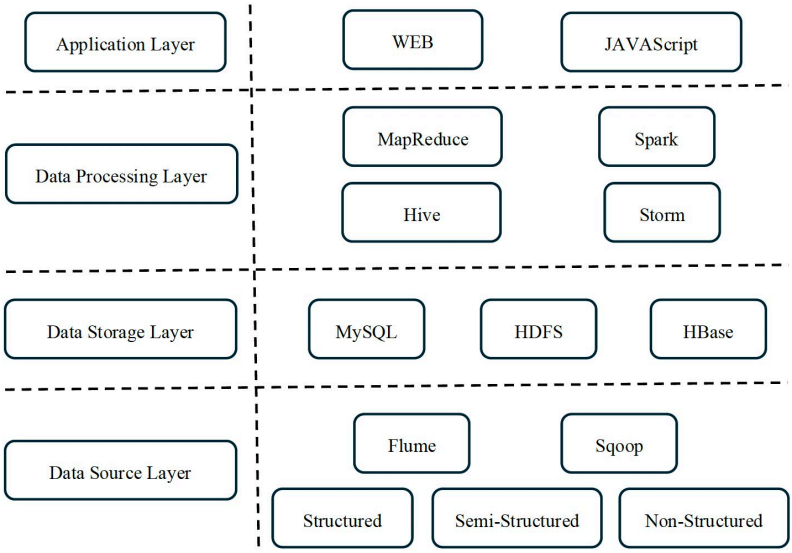

Fig. 4. System Architecture Diagram

4.1 Data Source Layer

The main task of the system data source layer is to collect data for the system. The data related to students mainly comes from three aspects. First is students' access data. The campus LAN and wireless network coverage of campus provides a convenient and secure way to collect data. Students need to log in to the campus network and enter their student ID and password for verification. Network access logs can be collected using Hadoop's Flum component. This data type has diverse formats, large data volumes, and free generation time. The second is student card data. The card records students' book borrowing, restaurant dining, access control management, recharge and other data, which can be obtained through the student card management system. Third is the data exported from various management systems or generated in daily management work by relevant departments involving student enrollment, grades, funding, rewards and penalties. This data type can be imported into the student management

information database through the system interface. The system data source presents the coexistence of structured, semi-structured and unstructured data. Flum is the data source layer's most important big data information collection tool. It can collect log data, social media data and other data sources distributed on multiple servers and move them to HDFS storage.

4.2 Data Storage Layer

The data storage layer is mainly used to store data on student information, pre-processing, system user and user interaction results. Structured data can be stored in the MySQL database, and semi-structured and unstructured data can be stored in the non-relational database HBsase. HBsase data tables are divided into two levels to build data attributes: column families and column qualifiers. Column qualifiers can be added dynamically later, which is convenient for system administrators to set new table structures according to actual needs. HDFS supports high-throughput access and can store data on multiple campus servers. It is easy to expand and can effectively avoid downtime problems caused by data concentration on a single server. Sqoop can easily realize the exchange between MySQL database and Hadoop, and Flume completes the multi-channel collection of data sources. Therefore, the data is imported into HDFS in batches through Sqoop and Flume components to prepare for data processing and back up the data to other directories. Since student information contains personal privacy, the original data will be desensitized and cleaned at this level. Then the data in various formats will be converted, processed, integrated and pre-processed to establish a unified data processing unit, extract data features, and provide a reliable data processing set for the system.

4.3 Data Processing Layer

The data processing layer not only transmits data for users to access the system, but also analyzes student academic performance, activity trajectory, campus consumption, book borrowing and other data in combination with data calculation models. Clustering algorithms and association rule mining algorithms are used to classify students' characteristics. Based on historical data and personal characteristics, students' mental health monitoring, academic warning, career planning, etc. are carried out. This layer uses the external API interface of each Hadoop component to connect and combine HBase and Hive. SQL or Hive_SQL language can directly access the combination of the two database tables, including data reading and writing, query, statistics, and analysis, which is convenient for batch management of related data in HDFS and improves data processing efficiency and accuracy.

4.4 Application Layer

The application layer is the interface for users to access the system and the window for the system to provide data to users. It provides access results to this layer by calling the lower layer data. Hadoop's file access interface is usually in the form of Java

API, using Java language to write the GUI interface, which interacts well with Java. In this system, in addition to the standard system access interface, the query results are displayed to users in the form of graphics. The JavaScript open-source component Echart is installed to realize data visualization functions, including straight line graphs, curve graphs, bar graphs, etc., to meet the user's interactive graphic analysis needs. Through visual data reports, staff in charge of student management can understand students' learning difficulties, behavioral tendencies, interests and hobbies, to take timely intervention measures and formulate training strategies. Students can view their academic, health, book borrowing, self-study, behavioral habits, etc., and understand their level among the students around them, providing a certain value reference for students to understand themselves and strive for excellence.

5 Conclusion

Big data plays a crucial role in student data collection, storage, and processing. Data visualization provides administrators with more concise data analysis results, making student information management more scientific and intuitive. Applying big data to the development of student management information systems helps migrate and integrate existing student information from various departments. By collecting real-time data generated by students, it becomes easier to establish a unified database for the entire university. This not only improves data utilization and transforms massive amounts of data into valuable information in a relatively short time, but also provides students with a reliable and timely channel for acquiring information, thereby effectively improving the efficiency of student management work.

Reference

1. CASADO R, YOUNAS M. Emerging trends and technologies in big data processing. Concurrency and Computation: Practice and Experience, 2015, 27(8): 2078-2091.
2. WU W-T, LI Y-J, FENG A-Z, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research, 2021, 81-12.
3. YAQOOB I, HASHEM I A T, GANI A, et al. Big data: From beginning to future. International Journal of Information Management, 2016, 36(6): 1231-1247.
4. HUANG F. Data processing [M]. Encyclopedia of big data. Springer. 2022: 312-316.
5. SHAKHOVSKA N, BOYKO N, ZASOBA Y, et al. Big data processing technologies in distributed information systems. Procedia Computer Science, 2019, 160561-566.
6. LAMPI J. Large-scale distributed data management and processing using R, Hadoop and MapReduce [D]; J. Lampi, 2014.
7. KUMAR N. Big data using Hadoop and Hive [M]. Mercury Learning and Information, 2021.
8. MUTASHER W G, ALJUBOORI A F. New and Existing Approaches Reviewing of Big Data Analysis with Hadoop Tools. Baghdad Science Journal, 2022, 19(4): 0887-0887.
9. DEMCHENKO Y, CUADRADO-GALLEGO J J, CHERTOV O, et al. Big Data Algorithms, MapReduce and Hadoop ecosystem [M]. Big Data Infrastructure Technologies for

Data Analytics: Scaling Data Science Applications for Continuous Growth. Springer. 2024: 145-198.
10. ABDUL-QAWY A S, PRAMOD P, MAGESH E, et al. The internet of things (iot): An overview. International Journal of Engineering Research and Applications, 2015, 5(12): 71-82.
11. KHELIFI F, BRADAI A, BENSLIMANE A, et al. A survey of localization systems in internet of things. Mobile Networks and Applications, 2019, 24761-785.
12. DUROC Y. From identification to sensing: RFID is one of the key technologies in the IoT field. Sensors, 2022, 22(19): 7523.
13. SONG Y, ZHU K. Experience and Thinking on the Construction, Management and Operation of Campus Card System in Colleges and Universities; proceedings of the 2023 3rd International Conference on Business Administration and Data Science (BADS 2023), F, 2023 [C]. Atlantis Press.
14. ZHANG Y. Challenges and Strategies of Student Management in Universities in the Context of Big Data. Mobile Information Systems, 2022, 2022(1): 3537468.
15. CHIAGOZIE O G, NWAJI O G. Radio frequency identification (RFID) based attendance system with automatic door unit. Academic Research International, 2012, 2(2): 168.
16. WIJAYA R, KRISTIANTO S, HASIBUAN Y B, et al. Contactless Student Attendance System Using BLE Technology, QR-Code, and Android; proceedings of the Conference on Innovative Technologies in Intelligent Systems and Industrial Applications, F, 2022 [C]. Springer.
17. ZHOU X. The Application and Research of Intelligent All-in-One Card in the Smart Campus of Colleges and Universities. Journal of Computer and Communications, 2018, 6(08): 82.
18. ALIAS H, AZMI A A, SALIM S. Student matric card payment system using RFID technology. Southeast Asian Journal of Technology and Science, 2020, 1(1): 21-27.
19. TULASI B. Significance of big data and analytics in higher education. International Journal of Computer Applications, 2013, 68(14):
20. RICHARDSON J, LENARCIC J. The blended discourse of SMS communication in a mobile student administration system. Proceedings Ascilite Auckland, 2009, 8429.
21. YANG H, ZHANG W. Data mining in college student education management information system. International Journal of Embedded Systems, 2022, 15(3): 279-287.
22. FENG S, ZHAN L, ZHAO H. Construction and realization of university student information management system based on big data technology; proceedings of the Journal of Physics: Conference Series, F, 2021 [C]. IOP Publishing.