Methods of Teachers' Personal Privacy Security Protection Based on Big Data Analysis

Haichao Zhang
Center of Education Technology and network, Jilin Engineering Normal University, Changchun, China

*Corresponding author e-mail: zhanghc@jlenu.edu.cn

Abstract. With the advent of the era of big data, these diverse, maritime and high-speed terminal data will be calculated through cloud-based storage, sharing, and mining of information, cheaply, efficiently, and quickly. The creation and driving of new industries, new products, new services, and big data are constantly emerging, which has profoundly changed people's daily lives. The value of data in the era of big data is constantly being discovered and developed to provide a convenient and comfortable life. It brings unprecedented challenges to privacy protection. The collection of personal information is throughout life. The preservation of information is vulnerable to online attacks and leads to large-scale disclosure of personal information. During the operation of big data, problems such as illegal collection of privacy information, excessive analysis and illegal transactions occurred, which constantly threatened the stability of personal life and the security of personal property. The purpose of this article is to study teachers' privacy protection methods based on big data analysis. This paper uses the fusion technology of ORAM and group signatures to study the privacy protection of teachers. It aims to better protect the privacy of teachers in a big data environment. It also provides reference suggestions for other researchers and contributes to the protection of teachers' privacy.

Keywords: Big Data, Teachers’ Personal Privacy, ORAM Technology, Group Signature Technology

1. Introduction
At present, it is in the era of big data. Big data is through the compatibility and storage of a large amount of data, accurately find the required information, and achieve purposeful retrieval through artificial intelligence technology[1]. Big data is divided according to the type of data, and the capacity is above 10T. The real-time update speed of big data information is very fast, which is convenient for users to check the information in a timely manner. Through the update of data information, the system can maintain in a relatively active state to ensure the stability of the system operation[2]. At the same time, accurate classification of installation information is guaranteed. Constructing relevant models in the field of education privacy, investigating the correlation between variables, providing effective support for privacy protection, and using data mining technology and data analysis technology are
inevitable trends in the future education of modern information technology.

Information technology has provided brand-new opportunities for the development of teachers' professional abilities, especially the establishment of specific goals and common interests through the Internet, and has gradually become an important carrier for the development of teachers' professional abilities. In addition, the communication and data sharing models (QQ groups, cloud computing, etc.) formed through various social platforms combined with cloud storage have gradually become the main method for informal learning in the community. However, information technology has brought convenience to teachers' study life, and also brought new privacy issues to teachers' study life, leaking people's actual information, intellectual property rights, and confidential data [3-4]. At present, how to protect the privacy of educational networks is mainly a legal and technical issue. Protecting personal privacy and security is an important issue in the current construction of network regulation. The existing laws and regulations do not reflect the in-depth integration of information technology and education, do not make legal requirements, from formulating to publishing to regularly respond to real-world problems and solve the network environment. Specific education issues are not possible. On the contrary, technical means cannot completely resist any infringement, but they have been highly targeted, and specific technologies can fully protect the privacy of users in specific stages and situations. For this reason, this article studies how to protect teachers' personal privacy based on the analysis of big data.

Information technology has brought convenience to teachers' learning life, and also brought new problems of privacy and security. To this end, this article will analyze the significance of privacy protection for users in the cloud environment and traditional privacy protection technologies. In order to protect teachers' learning and social privacy and security issues in a big data environment, this paper proposes a fusion technology of ORAM and group signatures to protect the privacy of teachers' learning and living communities in a cloud environment. This technology not only solves the problem that traditional encryption schemes cannot prevent cloud data access mode sequences from leaking, but also solves the problem that traditional cloud data access mode controls cannot protect cloud data access mode sequences. In addition, this paper also uses a questionnaire survey method to conduct teacher privacy security surveys to make the research more realistic, thereby effectively protecting the privacy of the teacher learning community in a big data environment, and promoting the continuous high-quality development of teacher careers in the information age.

2. Method

2.1 ORAM technology

The so-called ORAM technology refers to a technology in which a user executes a protocol through ORAM and starts making a data access request to the cloud. Each access request fetches multiple data blocks (including read and write processes) from the cloud once [5]. This prevents the cloud and attackers from using the user's access request technology to obtain the actual access mode of the cloud data series [6]. In other words, if ORAM technology is used, the cloud and the attacker cannot know the following information: The data that the user actually needs to access. During the period from the last access to the data currently accessed, the data accessed this time is the latest data. The data is the same. The current access is a read or write operation, and it is impossible to distinguish between two user access request sequences of the same length in calculation. The specific mathematical format used to calculate unrecognizability is defined as follows: \( Y = (O_{PM}, aM, datam) \ldots (O_{Pl}, a1, datal) \) is the actual time that the client must send to the server to access the request sequence. \( O_{Pl} \) indicates read operation, and write operation \( a1 \) indicates read or write data [7]. Identifying \( A \) (listed as an access request sequence processed by ORAM technology, the cloud only accepts the \( A \) (code request) sequence. For two real request sequences, it is not possible to calculate the corresponding \( A \) (Y) and \( A (+1) \).
2.2 Probabilistic Encryption Strategy Based on ORAM in Big Data Environment
The big data environment is the medium for teachers' learning life. The interaction between each user and the cloud follows the interaction protocol of the latest Oram solution [8-9]. Big data cannot infer the true sequence of data access patterns based on users' access requests, because when users obtain shared content in the teacher's learning life through the big data environment, all content is protected in the cloud through probabilistic encryption. Probabilistic encryption means that each encryption key is the same, but the random number is different, so the same plaintext will generate a different ciphertext each time. That is, after the user requests the data, the data will be re-encrypted. The ciphertext is completely different from the previous ciphertext [10]. The key belongs only to the corresponding user. The cloud cannot tell if the data read here is the same as the data read previously. In the big data environment, the Oram-based teacher life community cannot only protect the confidentiality of data, but also effectively ensure that the access pattern sequence of the teacher life community will not be leaked [11].

2.3 Group signature technology
Group signature, as a special digital signature technology, can not only protect data integrity, but also have anonymity and traceability. Each group consists of a group administrator and several members [12]. The group administrator generates a private key and a group public key for each group, and publishes each signed private key to the group members. Members of each group can generate arbitrary message signatures, while other members of the group can verify the validity of the signature with the group public key. Group signature technology not only provides anonymity protection, but also uses a private key to track the identity of the signer as needed.

3. Experiment
Step1: This article's investigation on the protection of personal data privacy is mainly combined with personal data privacy protection awareness, protection content, protection methods, and protection effects. Specifically, the personal data privacy protection questionnaire survey aims to understand the personal data privacy protection aspects. The actual situation of personal data, to understand the specific conditions of personal data privacy protection awareness, protection content, protection methods, protection effects, etc., I hope to understand the main problems in personal data privacy protection through this questionnaire.

Step2: The main survey object of this questionnaire is the teachers in the school. The reason is that the teacher group is also an important component of Internet users. It has a strong representation in the protection of personal data privacy, and the actual questionnaire is easier to carry out. A total of 100 questionnaires were distributed. 90 valid questionnaires were actually recovered, and the effective recovery rate was 90%.

Step3: Comparison and evaluation of data. The obtained data is compared by a computer, and the obtained data results are represented by a chart to make the effect more prominent. Through careful analysis with data and charts, and then make assessments and provide feedback, ensure the objectivity and accuracy of the feedback.
4. Discussion

Table 1. Personal data privacy

| Privacy awareness                | percentage | Reasons for privacy leaks        | percentage |
|----------------------------------|------------|----------------------------------|------------|
| Personal qualities               | 86%        | personal reasons                 | 100%       |
| Personal particulars             | 100%       | Business or organizational reasons| 60%        |
| Personal credit                  | 100%       | Technical reasons                | 57.2%      |
| Network trace                    | 79.56%     | Benefit-driven                   | 58.56%     |
| Personal Basic Information       | 100%       | Legal defect                     | 71.27%     |

As shown in Table 1. Regarding the hazards of personal data privacy disclosure, all respondents believe that the leakage of personal data privacy will threaten their personal safety, their study, work and life will be troubled, the questionnaire of the security questionnaire of online applications will be questioned, and social information ethics. The bottom line is facing challenges and the social credit system is harmed. The proportion of those who believe that the process of social informatization is affected is 48.07%, and the proportion of which is provocative of laws and regulations is 40.88%. In terms of the main reasons for the privacy of personal data, all respondents believe that they are mainly individuals. The reasons for weak protection awareness, interest-driven, and lack of regulations are 58.56% due to industry self-discipline, 54.69% due to technical reasons, and 71.27% due to lack of morality, indicating that the public has leaked personal data privacy. The hazard awareness is higher. With regard to the recognition of personal data privacy content, all respondents believe that personal basic information, personal special information, and personal credit information belong to the scope of personal data privacy, while only 86.19% of the recognition of personal characteristics information, network traces. The awareness of information is only 79.56%. It shows that most of the public have a high level of privacy awareness in terms of name, gender, age, ID number, archives, property status, bank card and other account passwords, credit records and repayments, and business reputation. Privacy in terms of height, weight, IP address, recycle bin files, history, etc. is low. With regard to the protection of personal data privacy, all respondents believe that individuals need to strengthen their awareness of data privacy protection. E-commerce operators also need to strengthen the industry's self-discipline behavior. The government should also establish and improve a legal and regulatory system to protect the privacy of personal data. Good network information protection order.

Figure 1. Personal data privacy theft and harassment

As shown in Figure 1. Many people do not develop a good habit in terms of personal data privacy protection. Most network users do not have the habit of clearing the browser history on a regular basis. The proportion of cleanups that they only think of is as high as 86.19%, and the proportion of cleanup
once a month is 10.49%. The proportion of cleaning up once a week is only 3.32%. In addition, survey respondents have little knowledge about the technical aspects of personal data privacy such as questionnaire ORAM, digital signatures, and digital certificates. Most people only heard about it, and did not understand its true role and meaning. Less specific gravity, very clear that the specific gravity of the principle is even less. Moreover, most people are not aware of the collection and use of personal data privacy by commercial companies. In terms of website information collection, most people do not allow personal information to be collected by the website, and they do not mind very much in terms of access records, browser models, search records, and login conditions. And if the commercial company is mainly to provide more high-quality and efficient services, 85.64% of the respondents allow them to collect personal data privacy. In terms of whether the privacy of personal data has been stolen, 47.51% of the respondents' personal data privacy has been stolen, and there are often harassing calls or harassing emails to promote merchandise. 49.17% of the respondents even though their personal data privacy has been stolen. It has been taken, but there will not always be harassing calls or harassing emails to promote products. Only 3.37% of the respondents have almost no harassing phone calls or harassing emails due to better protection of personal data.

5. Conclusion
With the rapid development of the Internet, computer networks have become an important part of people's daily lives, and computer networks have gradually changed the way people live and work. The convenience of this Internet poses a certain threat to personal privacy and a free and open space for people. It also threatens people's personal property security and personal data privacy on the Internet platform. Major issues that need to be addressed to protect today's social development network platforms have gradually come into being. Teacher learning platforms have gradually become an important form of teacher learning in a big data environment, but existing privacy and security issues have seriously hindered the progress of teacher information technology and promoted the development of teacher capabilities. To this end, this research aims to analyze the privacy security issues faced by teachers in the existing big data environment from a technical level, and propose specific protection countermeasures for the fusion technology of ORAM and group signatures based on the shortcomings of existing privacy protection technologies. strategy. However, in order to further ensure the development of teachers' privacy and security capabilities through big data technology, it is not enough to rely solely on technical means. To further improve laws and regulations and restrict ethics, teachers' careers in the information age must continue to play complementary roles and jointly promote high-quality development.

References
[1] Yi Ai, Zongping Li, Mi Gan. A solution to measure traveler’s transfer tolerance for walking mode and dockless bike-sharing mode[J]. Journal of Supercomputing, 2018, 75(1):1-18.
[2] Yahya Al-Dhuraibi, Fawaz Paraiso, Nabil Djarallah. Elasticity in Cloud Computing: State of the Art and Research Challenges[J]. IEEE Transactions on Services Computing, 2017, PP(99):1-1.
[3] Kenneth Guang-Lih Huang. Uncertain intellectual property conditions and knowledge appropriation strategies: Evidence from the genomics industry[J]. Industrial & Corporate Change, 2017, 26(1):41-71.
[4] K. RazaviPour, M. Yousefi. Iranian english language teachers' job satisfaction and organisational climate in public and private schools[J]. Issues in Educational Research, 2017, 27(4):842-858.
[5] de Carvalho, Elisabeth, Björnson, Emil, Sørensen, Jesper H. Random Pilot and Data Access in Massive MIMO for Machine-type Communications[J]. IEEE Transactions on Wireless Communications, 2017, PP(99):1-1.
[6] Chen Zhao, Chao-Fu Wang. Characteristic Mode Design of Wide Band Circularly Polarized Patch Antenna Consisting of H-Shaped Unit Cells[J]. IEEE Access, 2018, PP(99):1-1.
[7] Dima Kozakov, David R Hall, Bing Xia. The ClusPro web server for protein–protein docking[J]. Nature Protocols, 2017, 12(2):255-278.

[8] Jianguo Chen, Kenli Li, Senior Member. A Parallel Random Forest Algorithm for Big Data in a Spark Cloud Computing Environment[J]. IEEE Transactions on Parallel & Distributed Systems, 2017, 28(4):919-933.

[9] Joanna Michalik, Andrzej Cacko, Jakub Poliński. An interactive assistant for patients with cardiac implantable electronic devices: A study protocol of the LUCY trial[J]. Medicine, 2018, 97(39):e12556.

[10] Siti Nuradilah Azman, Iskandar Ishak, Nurfadhлина Mohd Sharef. Towards an Enhanced Aspect-based Contradiction Detection Approach for Online Review Content[J]. Journal of Physics Conference Series, 2017, 892(1):012006.

[11] Filip Debelva, Irma Mosquera. Privacy and Confidentiality in Exchange of Information Procedures: Some Uncertainties, Many Issues, but Few Solutions[J]. Intertax, 2017, 45(5):362-381.

[12] Chong Shangguan, Jingxue Ma, Gennian Ge. New upper bounds for parent-identifying codes and traceability codes[J]. Designs Codes & Cryptography, 2017(2):1-11.