Development and Testing of the Information Security Protocol in the Medical Cloud Platform

Liudmila Babenko[1], Evgenia Ishchukova[1], Dmitriy Alekseev[1] and Alexander Shumilin[1]

[1] *Southern Federal University, 2, Chekhov Street, Taganrog, 347922, Russia*

Abstract

The suggested protocol of ensuring information security for data, inside medical information system will be considered in the paper. Development and implementation an architecture of the medical cloud system for storing, systematization, and processing of surveys results (EEG samples) jointly with a protocol for ensuring protection of confidential data based on a fully homomorphic cryptosystem are also demonstrated. During the development such medical system to ensure the security for confidential medical data inside an infrastructure is a main and the most important purpose. Hierarchical division of data streams into layers, standardization of data transmission protocols, and formats of their storage ensure the creation of a universal, flexible and reliable medical information system. Regarding scientific novelty, the designed cloud architecture of a medical system differs from analogs in easy scalability, flexibility, versatility (integration with most existing systems from various manufacturers), and ensuring a high level of security of stored and processed data through the use of the developed protocol.

Keywords

privacy, information security, cloud computing, data processing, big data, distributed systems, encryption, medical cloud information system, homomorphic encryption, personal data

1. Introduction

In the century of global informatization and active development in information technology spheres, medical institutions within performing diagnostic studies process and systematize significant amounts of data for the subsequent rehabilitation and treatment of patients [3]. The effectiveness of the provided medical services is directly proportional to the efficiency and ease of use of this information by specialists of medical organizations [9-12]. The presence of tasks related to the storage, systematization, and processing of increasing amounts of data determines the relevance of the development and integration of medical information systems (MIS) into medical institutions [5-8]. The ability to operate with the data in digital format ensures promptness of the medical doctors to obtain essential information about the selected patient. This way increases the speed of decision-making on the diagnosis and treatment methods [2, 4].

Medical organizations under the law are the owners of the personal data of their patients (clients). They are directly involved in the collection, systematization, accumulation, storage, clarification, updating, modification, distribution, and destruction of such information. One of the problems in the design of medical information systems is the need to integrate mechanisms for protecting confidential information [3]. The category of personal medical data that requires non-traditional approaches to their protection includes dynamically changing indicators of the results of medical examinations of patients (for example, EEG indicators). Because the requirements of the legislation establish the need to protect

AISMA-2021: International Workshop on Advanced in Information Security Management and Applications, October 1, 2021, Stavropol, Krasnoyarsk, Russia

EMAIL: lkbabenko@sfedu.ru (Liudmila Babenko); jekky82@mail.ru (Evgenia Ishchukova); dalekseev@sfedu.ru (Dmitriy Alekseev); ashumilin@sfedu.ru (Alexander Shumilin)

ORCID: 0000-0003-2353-7911 (Liudmila Babenko); 0000-0002-6818-1608 (Evgenia Ishchukova); 0000-0001-6578-4158 (Dmitriy Alekseev); 0000-0002-4237-4380 (Alexander Shumilin)

personal data, the main task in the implementation of a cloud storage system, systematization, and processing of medical data is to ensure the security of stored information.

The purpose of this research is the development and implementation of the architecture of a cloud storage system for systematization and processing of survey results (for example, EEG), as well as an algorithm for ensuring the protection of confidential data based on fully homomorphic encryption [1].

2. Cloud platform structure

To solve the problem of collecting, storing and processing data, it is necessary to create an information structure that would allow automating business processes, on the one hand, and effectively protect systematized data using fully homomorphic encryption methods, on the other hand. This structure must meet high security requirements, be scalable in the face of growing data volume, and be flexible. The authors have developed a cloud platform that consists of 4 general layers. The structure of this medical platform is presented in fig. 1.

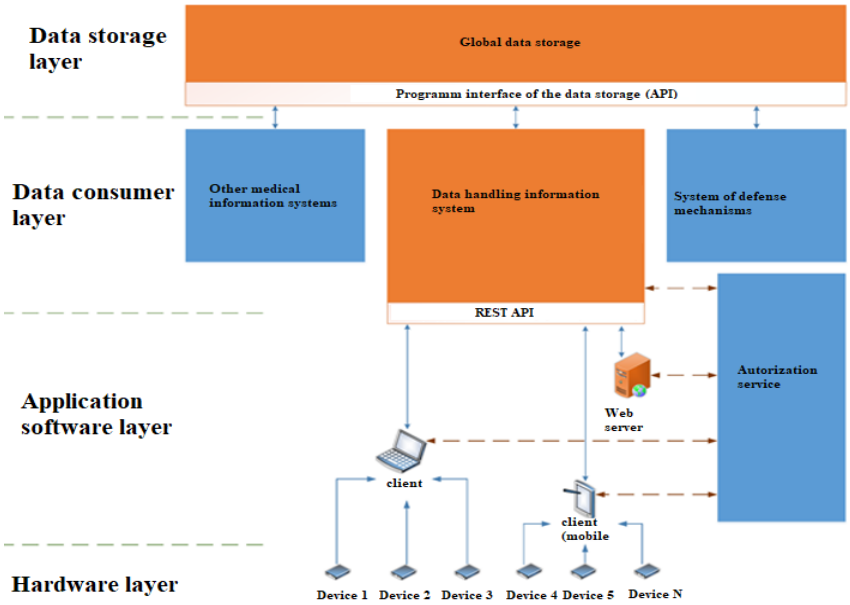

Figure 1: The structure of the developed medical platform

Data storage layer: a global data warehouse, that includes a database for storing initial data of examinations and reports, as well as anthropometric, diagnostic, and demographic information about patients. **Data consumer layer:** it includes systems that receive and process data from the global warehouse (the previous layer) or transfer new data to them. This layer is linked to the data storage layer through a standardized programming interface (Storage API). **Application software layer:** this layer contains software for customers, where medical data are generated and/or displayed (examinations in the form of signals, reporting, and personal data of a patient). **Hardware layer**: contains physical devices for conducting surveys. In general, there can be various types: electroencephalographs, cardiographs, biofeedback systems, wearable fitness trackers, etc. As a result, the algorithm has been developed to ensure the protection of confidential data based on a fully homomorphic system.

This approach differs from the known ones by using the principles of separation and personal data protection mechanisms, as well as the ability to integrate with most existing software and hardware systems and information systems. The key feature of the medical information system is the subsystem of protection mechanisms, which is an algorithm for ensuring the protection of confidential data based on a fully homomorphic cryptosystem, proposed, implemented and researched by the author.

Being the core of the platform, the data processing information system interacts with application applications using REST API requests. Authorization of users is carried out using the authorization module, which receives an encrypted token from an external authorization service and compares it with

the token received from the user. This ensures a high degree of system security. In the general case, a data processing information system can consist of parallel virtual nodes that provide work with their own pool of users. The central element of the Data Processing Information System is the control manager, which fulfills user requests and interface module requests. The management manager also handles the preparation of data for storage in the global data warehouse by managing the format converter. The data obtained can be used by both medical and research organizations, and the platform, accordingly, can perform the task of intellectualizing business processes by simplifying the access of specialists to information (telemedicine services, professional consulting SaaS, automated decision support systems) and a research task (study of algorithms for ensuring the protection of information, correlates and markets of diseases, pharmacological studies).

The cloud platform implements the following functions: providing convenient tools for transferring data between system users; creation of an interface and an extensive database for a research analysis system (including using machine learning algorithms); creation of interfaces for integration into existing medical information systems (MIS); creation of a cloud service (SaaS) for storing, classifying and processing data created using various equipment with support for many popular data formats; creation of a subsystem for ensuring the protection of examination results using the algorithm proposed by the author (on the example of EEG).

To confirm the effectiveness and high level of data security provided, within the framework of the proposed algorithm, a cloud-based medical information system for storing, processing and systematizing confidential data was implemented. The server part of the application has been developed using the Incoding Framework, which is responsible for the business logic and work with the database, a set of REST API requests has been implemented for various scenarios of interaction between users and groups of the platform. The OAuth2.0 service was used to authorize users. The database was a MySQL instance on Amazon Relational Database Service (RDS). The web management interface was developed using the ASP.NET MVC Framework. The platform is deployed on a cluster of instances from Amazon Elastic Compute Cloud (EC2).

The implemented platform supports the ability to integrate with medical information systems using two-way data transfer via the HL7 protocol. Integration with a number of software and hardware systems for recording EEG signals was carried out.

3. Confidential data protection protocol development

Protocol preconditions:

- The users of the medical information system are medical doctors and patients. Each of the users who works with the system is registered in it. Access is granted for use by a unique identifier (login) and password.
- A medical doctor informs patients about the need for a series of monitoring examinations (for example, daily monitoring of brain activity during a fixed period for the patient's rehabilitation) when they visit a medical institution. The dynamics of changes in indicators and their average values over a long time are important for the correct diagnosis and choice of treatment methods for patients, At the same time, the medical doctor provides his unique identifier to the patient, which is necessary for his initial registration in the medical information system. During registering in the system, the patient shares his data and the unique identifier of the medical doctor who ordered the series of examinations.
- Next, a list of available medical doctors who are collaborating with the concrete patient is displayed in the user's account (in a mobile phone or a PC). In the personal account, the function of adding new medical doctors is available (for potential scaling of the system and adding new specialists of a similar or different profile). When an attempt is made to bind the medical doctor's unique identifier to the patient's account, the medical doctor receives a notification request to confirm this action. After confirming the request, the medical doctor has access to view the results of medical examinations of a particular patient.
- Server (let’s call them “S”) that is used for storing and organizing all data, contains a database with access rights for each registered medical doctor of the system to patient results.

Protocol steps:

Step 1. On the side of each medical doctor in the medical information system, a pair of keys (public and private) is generated and stored for use by each of his patients. Keys are generated immediately after confirmation of the request to join the medical doctor to the personal account of a particular patient.

$$public\ key\ D_1\ for\ P_1$$
$$private\ key\ D_1\ for\ P_1$$

Step 2. After the medical doctor belongs to the patient (in his account), the medical doctor's public key (unique for each patient) is sent to the patient.

$$public\ key\ D_1\ \rightarrow P_1$$

Step 3. During the examination, a patient records the parameters of the biorhythms of the brain (alpha, beta, and theta rhythms). The changes in wave activity (of each type of rhythm) are a changing series of numerical data (updated once per second).

$$alpha\_1 \dots alpha\_N$$
$$beta\_1 \dots beta\_N$$
$$theta\_1 \dots theta\_N$$

Step 4. The obtained data (indicators of the rhythms of brain activity) from a medical device (for example, a wireless encephalography) are encrypted using a public key that was received from the medical doctor on the patient's side (mobile application. The encrypted data is transmitted and stored on the server. In the regards to cipher texts for alpha_1, beta_1, theta_1, they might be obtained the following way:

$$Cipher\_alpha\_1 = Encrypt(alpha)\ public\ key\ D_1\ for\ P_1$$
$$Cipher\_beta\_1 = Encrypt(beta)\ public\ key\ D_1\ for\ P_1$$
$$Cipher\_theta\_1 = Encrypt(theta)\ public\ key\ D_1\ for\ P_1$$

Step 5. The data that is sent to server is stored there without further decryption. Data on the values of the rhythms of brain activity are statistically accumulated over a long time (rehabilitation course). In parallel, the value of the number of studies performed for each patient is taken into account.

The use of the homomorphic encryption algorithm allows operation adding these data without preliminary decryption, to calculate the average value for each of the rhythms (based on the results of a series of examinations, for example, during a rehabilitation course. The calculation of the average value is conducted because of multiplying the total value of each rhythm (for a series of examinations) by the multiplicative inverse for the value of the number of examinations performed.

$$Cipher\_alpha = Cipher\_alpha\_1 + Cipher\_alpha\_2 + \cdots + Cipher\_alpha\_N$$
$$Cipher\_beta = Cipher\_beta\_1 + Cipher\_beta\_2 + \cdots + Cipher\_beta\_N$$
$$Cipher\_theta = Cipher\_theta\_1 + Cipher\_theta\_2 + \cdots + Cipher\_theta\_N$$

where N – quantity of all sessions.

Next, the average for each cipher text will be computed:

$$Cipher\_average\_alpha = \frac{Cipher\_alpha}{N}$$
$$Cipher\_average\_beta = \frac{Cipher\_beta}{N}$$
$$Cipher\_average\_theta = \frac{Cipher\_theta}{N}$$

Step 6. If necessary to obtain calculated data for each rhythm (alpha, beta, and theta), the medical doctor sends a request to the server in the form of a patient ID to receive data.

$$GET\ (ID\ P1) \rightarrow Server$$

Step 7. After receiving the request, the server checks whether the medical doctor has access to the data of a particular patient. The check is performed by the medical doctor's ID.

Step 8. If you have access rights, the server sends encrypted data (average values of the brain activity rhythms) to the medical doctor, who made the request.

$$Cipher\_for\_P1 = (Cipher\_average\_alpha,$$
$$Cipher\_average\_beta, Cipher\_average\_thera)$$

Step 9. The medical doctor decrypts the received data using his private key (generated for a specific patient).

$$Average\_alpha = Decrypt(Cipher\_average\_alpha)\ private\ key\ D1\ for\ P1$$

$$Average\_beta = Decrypt(Cipher\_average\_beta)\ private\ key\ D1\ for\ P1$$
$$Average\_theta = Decrypt(Cipher\_average\_theta) private\ key\ D1\ for\ P1$$

The sequence of steps (from the server-side, patients, and medical doctors) of the algorithm is demonstrated at fig. 2.

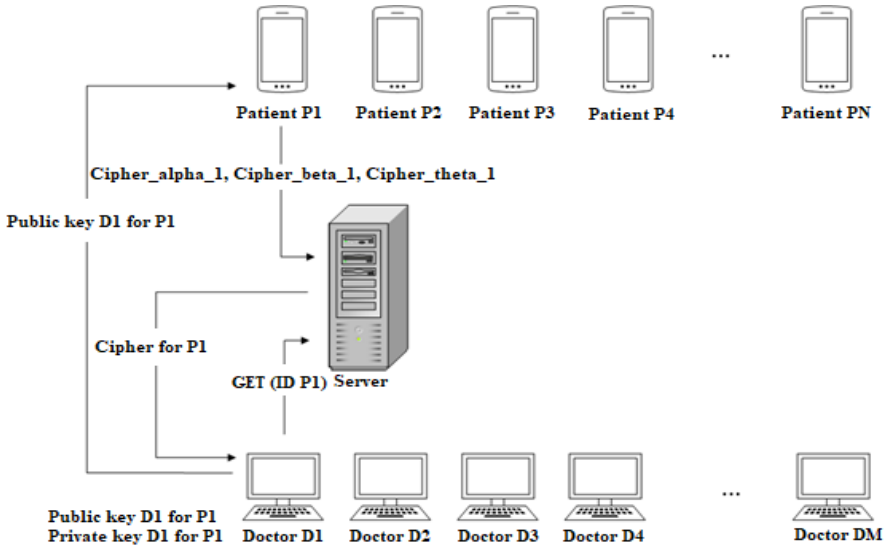

Figure 2: Protocol for ensuring medical data protection

4. Experimental part

As part of the work, the authors proposed a protocol for ensuring the protection of confidential data based on a completely homomorphic cryptosystem for brain activity data. The BGV scheme was chosen as the fully homomorphic encryption algorithm under study. To test the properties of the proposed information security algorithm, the HELib cryptographic library was chosen. To obtain a comprehensive assessment of the efficiency of the algorithm proposed by the authors, it is necessary to analyze the following parameters: the execution time of encryption, decryption, addition, multiplication, the signal-to-noise ratio of the ciphertext to the plaintext.

During testing, measurement results were obtained demonstrating the high efficiency of the protocol developed by the authors. The average execution time of the encryption function was 115463817 ns, the average execution time of the decryption function was 8644635 ns, the average execution time of the homomorphic addition function was 391472 ns, the average operation time of the multiplication function was 5987452 ns. It is known that the ciphertext obtained with the help of fully homomorphic encryption schemes has redundancy. The degree of this redundancy is an important parameter of fully homomorphic encryption schemes. This redundancy can be determined by examining the signal-to-noise ratio. During testing of the protocol, it was found that the signal-to-noise ratio was 1.94567.

5. Conclusion

Hierarchical division of data streams into layers, standardization of data transmission protocols, and formats of their storage ensure the creation of a universal, flexible and reliable medical information system. Regarding scientific novelty, the designed cloud architecture of a medical system differs from analogs in easy scalability, flexibility, versatility (integration with most existing systems from various manufacturers), and ensuring a high level of security of stored and processed data through the use of the developed algorithm.

The novelty of the developed algorithm lies in the correct operation with different types of data (the mechanism of the algorithm is the same for different types of studies). It is also possible to use the algorithm in an untrusted environment, which will increase the efficiency of security systems (there is no need to decrypt data to calculate for example average values of patient indicators over a long period).

The theoretical significance of the research is to expand knowledge about the process of integrating encryption mechanisms into medical information systems, based on cloud architecture. The findings can serve as a basis for research in various areas of information security: security analysis of cryptographic protocols, cryptographic information security tools, information security, and cloud computing. The developed architecture allows performing a quick integration into existing medical systems. A single space for storing data makes it possible to learn a significant array of classified medical information using machine learning methods and algorithms.

We also would like to add the following: the provided platform has passed preclinical tests and the procedure of validation. This is a universal ecosystem for the implementation of telemedicine services and new scientific research in the field of information security, based on the use of fully homomorphic encryption methods. As for the practical implementation, it is possible to replicate the developed architecture, for example, to the infrastructure of another provider whose data center is located in the Russian Federation.

6. Acknowledgement

The current research was funded by RFBR, project number 20-37-90138.

7. References

[1] Mitkina P.A. Features of storing medical information // Modern scientific research and innovations. 2017. No. 5 [Electronic resource]. URL: http://web.snauka.ru/issues/2017/05/82546 (date accessed: 10/07/2019).

[2] Health Insurance Portability and Accountability Act // [Electronic resource]. URL: https://en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act (date accessed: 10/08/2019).

[3] DICOM // [Electronic resource]. URL: https://ru.wikipedia.org/wiki/DICOM (date of treatment 10/08/2019)

[4] L.-Y. T. a. M.-S. H. Li-Chin Huangc, "A reversible data hiding method by histogram shifting in high-quality medical images," The Journals of systems and software, vol. 86, pp. 716-727, 2013

[5] M. G. a. R. D. Jessica Fridrich, "Detecting LSB Steganography in Color and Gray-Scale Images," Binghamton.

[6] Logistic map // [Electronic resource]. URL: https://en.wikipedia.org/wiki/Logistic_map (date accessed 10/08/2019)

[7] Abdulrahman Alsalmany // Cloud System for Encryption and Authentication Medical Images // IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 20, Issue 1, Ver. II (Jan.-Feb. 2018), PP 65-75 [electronic resource] https://www.researchgate.net/publication/332571801_Cloud_System_For_Encryption_And_Authentication_Medical_Images (date accessed: 29.09.2019)

[8] Plotnikov A.V., Prilutskiy D.A., Selishchev S.V. // "DICOM standard in computer medical technologies", [Electronic resource]. URL: https://mks.ru/library/article/1997/dicom.html (date of treatment 10/08/2019)

[9] Kotyashichev I. A. Information protection in "Cloud technologies" as a subject of national security / I. A. Kotyashichev, E. A. Byrylova. - Text: direct // Young scientist. - 2015. - No. 6.4 (86.4). - S. 30-34. - URL: https://moluch.ru/archive/86/16357/ (date of access: 09.06.2020).

[10] Kereytova M.R., Malysh V.N. Information security in medical information systems // NIK. 2012. no. URL: https://cyberleninka.ru/article/n/informatsionnaya-bezopasnost-v-meditsinskih-informatsionnyh-sistemah (date accessed: 06/11/2020).

[11] Boychenko I. V. Building IT infrastructure for healthcare based on the paradigm of cloud computing // Medical doctor and information technologies. 2011. No. 3. URL: https://cyberleninka.ru/article/n/postroenie-it-infrastruktury-zdravoohraneniya-na-osnove-paradigmy-oblachnyh-vychisleniy (date accessed: 09.06.2020).

[12] Rohan Jathanna. Int. Journal of Engineering Research and Application www.ijera.com ISSN: 2248-9622, Vol. 7, Issue 6, (Part - 5) June 2017, pp. 31-38 (date of access: 10.06.2020).