SECURITY OF HEALTHCARE DATA AT REST IN CLOUD WITH MOBILE HEALTH MONITORING SYSTEM

Nikhita N Nerkar¹, Prof. Vina M. Lomte²

1Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune-58, India
2Assistant Professor, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune-58, India

ABSTRACT

Mobile health (mHealth) monitoring using Cloud, applies the common mobile communications and cloud computing technologies to provide feedback decision support, which has been considered as a new approach to improve the quality of healthcare service while lowering the healthcare cost. Unfortunately it is also at a position of serious risk on both clients/ mobile user’s privacy and intellectual property of monitoring service providers, which could prevent the wide adoption of mHealth technology. Mobile health monitoring system provides an ease of efficient information flow and security to information from user to the company (our system) and vice versa. The system is used to monitor the health parameters of the user and stratify the user queries. The client, trusted authority, the company and server are unaware of parameters passed in the network. For security purpose we implemented Identity Based Encryption. Further in this paper we will see the improvised usage of algorithm which we have implemented i.e. Hierarchical Identity Based Encryption. The system promises authentication, ease of use and cost effective. The goal of the proposed system to make it more users friendly and provide salvation of the data.

Keywords: Mobile Health Monitoring, Security, Identity Based Encryption, Hierarchical Identity Based Encryption, Authentication

I. INTRODUCTION

Personalization can be defined as the modification of the nature of an activity where a user is busy in or wants to get busy in the unique characteristics of the user and the context within which the user currently resides. Personalization in telemedicine is important since every patient's demands are different. Personalization healthcare provides an opportunity to achieve more ardent treatment where issues can be superscripted and hopefully be protected at earliest possible wink.

In the previous years, the mobile devices were hindering multi-purpose machines which needed mounting, installing and few other steps. These devices failed in real world healthcare
scenarios, due to their less compatibility. But now days the mobile device allow us quicker access to any information and has made our means of access very easy. The mobiles are playing vital role when it comes to healthcare business. The facilities support number of users/patients to store their health data, personal data and access it as and when needed. As the healthcare companies are empowering the use of mobile computing devices whereas the logical step is apply an emphatic method to nail down the security of data which being accessed onto these devices. This security method involves implementing strong access controls at both the ends i.e. at device level and at network levels.

Conventionally, privacy protection apparatus were applied by simply removing client’s personal identity information or by using anonymization technique which generally failed to serve as an effective way to deal with privacy of mobile Health systems due to the increasing amount and assortment of personal ascribable information. Personal identifiable information (PII) is basically any information, which is recorded relating to an identifiable individual. Whereas any information, which if linked to an identifiable individual, becomes personal in nature, be it biological, biographical, genealogical, reputational. As we know, the scope of PII might not necessarily be restricted to SSN (such as names), name and address, which are generally considered as PII traditionally. Undeniably, the art re-identification techniques [5], [6] have shown that any attribute could become personal identifiable information in practice [3].

In this paper, we design an Intelligent Mobile Health Monitoring System. We first identified the design problems on privacy preservation and provided our solutions. To affluence the understanding, in our paper we start with the basic scheme, so that we can easily identify the possible privacy rupture. We then provide an improvised scheme by addressing the identified privacy problems.

II. LITERATURE SURVEY

Anterior years there was healthcare crisis for the tribal communities as there were very less doctors and more number of patients which made situation difficult for the patients as they had to wait for days to get a single appointment. According to the paper [5], there was a need to have a contemporary approach to manage this healthcare crisis. There was one such approach discussed was using telemedicine, where the health of the patient was monitored primitively to improve the durability and quality of life of the patients.

As the advancement in the telemedicine approach, [2] made various modified approaches to improvise the healthcare services. In this, the components that accomplish four distinct functions i.e. data assemblage and depot, data acumen, alert proclamation, feedback, and newscasting. The BP and glucose sensors were used to measure the patient’s current details; these sensors used custom built software to transfer data from sensors to the mobile device of the patients. The Data which is stored on the mobile was further transferred to the web server for further process. But if the connection is low the data there itself is stored for the longer time waiting for stronger connection to retransmit it. According to [4], Parkinson’s disease (PD) is the second most common neurodegenerative disorder after Alzheimer. The telemonitoring technique was also used for PD, which heavily depends upon the design of simple tests that can be self-conduction quickly and fortuitously. The data was collected using the Intel At-Home Testing Device (AHTD), where telemonitoring system designed to aid remotely, and Internet-enabled measurement of a variety of PD-related motor deterioration symptoms. Then the data is collected at the patient’s home, disseminated over the internet, and handled appropriately in the clinic to predict the UPDRS (Unified Parkinson’s Disease Rating Scale) score.
In [10], the pervasive computing concept was considered. In this system, it collected the patient’s physiological data through the bio-sensors. Then this data was assembled in the sensor network and a curt of the collected data is diffused to the patient’s personal computer. The devices forwarded the respective data to the medical servers for the further analysis. After the data was analyzed, the medical server provided the feedback to the patient’s personal computer. The patients could take necessary actions depending upon the feedback. The system contains three components: Wearable Body Sensor Network, Intelligent Medical Server and Patients Personal Home Server.

All the telemonitoring system started making homes of the patient’s emplacement of adding various flavor to the healthcare innovations. The increasing advancements in the bio sensors, computing power and development of complexity algorithms are simultaneously increasing the privacy problems of the patients. In [7] the author discussed about the increasing of the patient’s emplacement of adding various flavor to the healthcare innovations. The Intel Health Guide had integrated a considerable number of privacy protections into the technology, providing security to both the patient and health care provider. The concept Privacy protections begin with patient consent to use the device. The security was developed into the health guide where patient must enter 4 digit pin (personal identification number) before he/she uses the cell phone/device. But the drawback was when the user received the device it was booted by the provider of the device so the PIN was no more private. And even the data shared between the user and the service provider was transmitted on the internet which was encrypted using Secure Sockets Layer (SSL/TLS) technology.

Advancement to all this [1], the author discussed about storing the data on to the cloud. The data stored on the cloud was in encrypted format and user was recognized with the help of token generated for per user and supplied the respective tokens to the user. The user could not access the data healthcare system until he enters the correct token before every stride.

The above discussed systems had common problem where security was breached, and the user details were not safe to be maintained, which might negatively influence the user. These systems are overcome with the few more innovation in the concept of healthcare and preserving data related to the user and the health related data in our paper.

In this paper we design a mHealth system referring the concepts of described in [1]. To ease our understanding, we should start with basic scheme, so we can predict the privacy breaches and provide an improved scheme to avoid these breach. The improved scheme allows the company (service provider) to work in offline or online mode after the setup stage and supplies data to the cloud securely. We applied further scheme by applying encryption technique once at the setup stage and then the company plays no role in encryption and decryption, as it shifts the further task to the cloud and the client side. This reduces the communicational and computational burden.

III. SYSTEM MODEL, PRELIMINARIES AND SECURITY

A. System Model

To add in our previous discussion, we elaborate our intelligent mHealth system. It consists of four parties: the company who acts as the health care service provider, a semi trusted authority (TA), the individual client, and the cloud server (we can simply say cloud). The service provider stores its data or we can say program in an encrypted format in cloud server. Here the cloud just acts a storage device. Individual client collects its data while getting registered to the system. While storing this data it gets transformed into attribute vectors. A semi trusted authority is responsible for authenticating the user to use the system and distributing the private keys and the tokens to the respective user. We consider a cloud server a neutral server as it neither combines with TA nor with the service provider.
Our system can be broadly introduced in basic 4 steps: setup, store, tokengen, and query. At very first time when system is initialized, the setup phase runs and thus system parameters are published. Then the service provider initiates the mHealth monitoring program also known as branching program (see part B.). This branching program is further encrypted and then resulting cipher text is delivered to cloud for storage, which corresponds to store algorithm.

The client accesses the mHealth system, by raising query to the cloud. When client registers to the system for the very first time, TA approves the client and runs the tokengen algorithm. The next time when client wishes to raise the query (which is the attribute vector represented by collected health data), he initially sends the key to the TA and then inputs the query and TA inputs the master secret algorithm. The client receives the token for the corresponding query whereas TA gets no useful information regarding the query.

In the last phase, the client sends the token in order to receive the perception from the cloud which comes under the query phase. The query is resolved and perception is sent to client in semi decrypted cipher text. The client then performs the decryption task on the partially decrypted cipher text and obtains the final decisions provided by the monitoring program on the cloud’s input query. The cloud also obtains no useful information from the query raised by the client.

B. Preliminaries: Branching Program

Here, we will formally describe the branching programs [16], which includes decision tree. For our ease we have only considered the binary branching program. A private query protocol can be easily derived using this decision tree [1]. Let \( v \) be set of attribute vectors of the client, \( v = \{v_1, v_2, v_3, v_4, \ldots v_n\} \). Branching program is a set of triple \( \langle\{p_1, p_2, \ldots, p_k\}, L, R\rangle \) where \( p \) is set of nodes in the branching program. Each decision node is again a set of pair \( (a_i, t_i) \), where \( a_i \) is attribute index and \( t_i \) is the threshold value. The \( t_i \) is compared with \( v(a_i) \) is compared at the node. For each decision node \( i \), if \( v(a_i) \leq t_i \), set \( h = L(i) \), else \( h = R(i) \). Repeat the process recursively for \( p_h \), and so on, until one of the leaf nodes is reached with decision information [1].

To construct branching program introduced in MediNet project [17] [18]. For instance let us consider, “[Systolic BP: 150, Missed one medication=0 (indicating he did miss the medication), Energy Expenditure: 900 kcal, salt intake: 1000 milligrams]” and the respective threshold is “\( t_1 = 130 \), \( t_2 = 0 \), \( t_3 = 700 \) kcal, \( t_4 = 1500 \)”. The recommendation returned from the monitoring program (Fig. 2) would be “D4, D5, D6” (by following the path through comparing each attribute element with the respective threshold at each node), which indicates the clients need to “notify next kin, modify daily diet, and take regular medication”. The health data related to the input attribute vector can be sampled by the client.
C. Security: Encryption and Decryption technique

Encryption is a technique which is widely used to generate the cipher text. Homomorphic encryption technique does the calculation on this encrypted data without decrypting any of the cipher text. The Homomorphic encryption scheme was proposed by Paillier cryptosystem [19], [20]. This encryption technique allows the client to gain the token in response of the input of attribute vectors provided by TA. Homomorphic encryption is generally used in cloud data storage as the data more secure.

Decryption technique to MDRQ by using BF-IBE as shown in [22]. Using this technique neither the TA nor the cloud gets any valuable information on client’s decryption results.

D. Multi-dimensional range Query

We are using multi dimensional range query concept in our system, so that we can access the data even if it is in encrypted format. The concept was firstly introduced by Elaine Shi [21]. In this system, the sender encrypts the message under range \([r_1, r_2]\) or even under the set of vector \(v\), and a receiver can decrypt the message with the privacy key corresponding to the range \([r_1, r_2]\). The cipher text generated assures the privacy of both the message content and the data under which the encryption was performed.

In our system, we are using this concept similar to CAM [1], where MDRQ plays a very important role. In this system we will be using BF-IBE [13] to develop MDRQ. But MDRQ can also be build by using Anonymous Identity Based Encryption scheme [14]. If we compare normal IBE where only normal message is encrypted and made secure, where as in A-IBE scheme it can hide the data as well as the sender receiver information also. All the comparisons between the input vectors and the threshold at an intermediate node are implemented by MDRQs. At every decision node attribute vector \(a_i\), and their respective thresholds represented in two minimum root sets: \([0, t_1]\) and \([t_1, max]\) as shown in Figure 1. Every index of next decision node is encrypted under the respective range.

![Figure 1: Branching Program at the company level](image-url)
E. Key private proxy re-encryption

We are using proxy re-encryption, which allows the proxy server to re-encryption key to transform the cipher text (first level cipher text) encrypted for cloud storage into second level cipher text, that could be decrypted by user without allowing proxy obtain any useful information on the underlying message. In our system, the service provider delivers the monitoring program to cloud in an encrypted format using MDRQ scheme and then this cipher text is stored in an untrusted cloud. Than rekey algorithm is run by TA instead of service provider and allows TA to encrypt the cipher text stored in cloud.

IV. PROPOSED WORK

A. Introduction

The MHM System tries to overcome the scenario which are discussed above in literature survey where telemonitoring is done, the data is collected on the basis of sensors, stored onto the device, transmitted to the service provider and accordingly the feedback is reported to the user. To describe the fundamental idea we start with the basic scheme and then show the flow of the system and final graphical representation of the results and screenshots. The elite notion that buried the performance of the system is the privacy which is very difficult to be breached in this system as Hierarchical Identity based encryption is used to encrypt and decrypt the data. Using this convoluted algorithm is one of the elite tasks to be performed. The system is divided into multiple time slots generally known as “slots”. There can be N maximum number of users, who are requesting to access the system. When a client attempts to access the system, he is assigned an index, $i \in [1,N]$ by Trusted authority.

B. System Architecture

1) Trusted Authority

The Trusted Authority (TA) is a web application where user is authenticated and token is generated for every other user. Let us consider that user A is performing login for the first time. The user won’t be able to access our monitoring system until he/she is not authenticated. We can say it acts as an arbiter between the user and the data stored in cloud.

2) Service Provider

Service Provider creates branching program and stores the monitoring data (branching program) in the cloud server in an encrypted format.

3) Mobile Users

An Android application is developed for the user which is will be operated as per all other play store applications. Through this application a GUI interface will be developed which will help user to access our health monitoring system and efficiently resolve the user query.

4) Cloud Server

The data stored on the cloud is the data which is supplied by the user while creating an account, token generated for respective user, the branching program done by the service provider. The data collected and stored will always be stored in an encrypted format. The below steps are referred from [1].

Setup: This step is performed by TA.

Store: This is performed by the service provider also known as company. For every node Kj, whose child nodes are not at all leaf nodes, the service provider runs $CL(j) = \text{AnonEnc}(id,PP,L(j))$ and $CR(j) = \text{AnonEnc}(id,PP,R(j))$, to encrypt the left and the right child nodes indices under id with
either id ∈ S[0,tj] or id ∈ S[tj+1,Max], respectively (where PP is known as System parameters). And when the child nodes of Kj have leaf nodes, then the service provider generates the cipher text CL(j)=AnonEnc(id,PP,mL(j)) and CR(j)=AnonEnc(id,PP,mR(j)), where mL(j) and mR(j) denote the attached information at the two leaf nodes. All the generated cipher texts are delivered and stored in the cloud.

Tokengen: To generate the private for the attribute vectors v= (v1,v2,v3…vn). The user inputs every attribute which is represented in set v, and delivers all set to TA. Than TA runs AnonExtract (id,msk) function, through which all identity id ∈ Svi in the identity set and delivers all the clients their respective private keys.

Query: A client raises the query, which is further sent to cloud to resolve. To run the AnonDecryption algorithm on the ciphertext generated in the Store algorithm client needs to send the private key set generated by TA to the cloud. The decryption result decides which cipher text should be next decrypted to get proper results. The decryption result decides whether next right node should be considered or the left one. This process can be continued till it reaches the leaf node and gives the client attached information.

C. Flow Diagram

The below diagram Figure 3, explains us the flow of our proposed system. The flow explains us the flow of the data from user to TA to Cloud Server.

![Figure 3: System Flow](image)
1) **User Flow**

The user registers himself through the GUI interface login page. If the user is authenticated user he can easily raise the query in the system. But if the user is not authenticated by the TA, he has to wait for it, till the TA logins and approves him as a authenticated user.

2) **Trusted Authority**

When the setup is ready, and user registers for the very first time, the TA has to approve the user and generate token for the respective user at least once. When the user is authenticated user as we can see in the user flow that he is ready to access our health monitoring system by raising the queries and other application details. The data whichever is shared on the system is always stored on the Cloud Server in an encrypted format. Also, whenever the user raises the query, with the help of branching program stored on to cloud is responded back with the query results.

### D. Algorithm

**Algorithm:** Proposed Algorithm

- **Setup (1^\*):**
  - Output: PP(G,G_T,q,g,y,H_i,i=1,2,3,4)
  - PP -> System Parameters
  - G,G_T -> public groups
  - g -> random primitive root from G
  - H_i -> Cryptographic hash function

- **Extract (id, msk):**
  - Output: sk_{id}=H_1(id)^{msk}

- **Encrypt (id, PP, m):**
  - m ∈ M
  - Output: C= (c_1,c_2,c_3)
  - With r=H_3(m||σ)
  - C_1=g^r
  - C_2= σ ⊕ H_2(e(H_1(id),y)^r)
  - C_3=m ⊕ H_4(σ)
  - σ -> random element from M

- **Decrypt (C,sk_{id}):**
  - σ= c_2 ⊕ H_2(e(sk_{id}, c_1))
  - m= c_3 ⊕ H_4(σ).

### V. SECURITY AND EVALUATION ANALYSIS

#### A. Security

Here, we evaluate our proposed MHM system. We observed that no information is gained by the company from either of the clients query and in addition to that cloud also does not obtain any information related to the branching program generated by the company as proxy re encryption technique is been used.

When the company loads the branching program on the cloud, the key privacy makes sure that the cloud is all together unknown about the node, leaf nodes, threshold values and the nodes. This mask privacy defeats the attacks on cloud data.

In addition to this, only the client knows the decisions to his raised query as the results are directly supplied to the client. Moreover, TA and company have no clue regarding the raised query and cannot get any information from it as it is masked with the keys. Even the attackers attack on TA and company to fetch the information related to the query will fail.
B. Evaluation

To implement our MHM system, we used a laptop with 2.27 GHz processor with a 3GB of RAM to simulate the cloud server and the company, and a mobile phone (Android platform) with 2GHz processor with 512MB of RAM to simulate a client. Figure 4 explains us that the company and cloud requires only 6.23 ms for encryption and 10.076 ms for decryption process respectively when number of nodes in branching program is 15.

Figure 4: Comparison of two algorithms used for encryption & decryption our design
Figure 5: Tokengen algorithm time cost for individual client

Figure 6.1 Creation of Branching program

Figure 6.2 Uploading of Branching program in encrypted format
VI. CONCLUSION

In this paper, we proposed a improvised version of Mobile Health monitoring System, which enables client to have an interactive application by providing various decisions related to the disease symptoms raised in form of query. The system along with it provides privacy to the user details and medical data. To apply privacy to the client data we have used IBE and also the evaluated the security performance by applying HIBE instead of IBE. We have highlighted what kind of encryption approach and methods are reasonable for maintaining between the data security and performance tradeoff. For both cloud user and provider’ perspective, our proposed integrated data encryption architecture is practically helpful in reducing data security and privacy risk in cloud environment.

REFERENCES

5. Shirley Ann Becker, Ramprasad Sugumaran, and Kiran Pannu Northern ,” The Use of Mobile Technology for Proactive Healthcare in Tribal Communities”
6. Kinshu K, Maiga Chang, sabine graf and guangbing Yang,”Adaptivity and Personalization in Mobile Learning” © 2010 Old City Publishing, Inc. Published by license under the OCP Science imprint, a member of the Old City Publishing Group Tech., Inst., Cognition and Learning, Vol. 8, pp. 163–174
9. Independently conducted by Ponemon Institute LLC,” Data Loss Risks During Downsizing As Employees Exit, so does Corporate Data” Publication Date: February 23, 2009 Sponsored by Symantec Corporation