International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 2 | ISSN : 2456-3307

Hybrid Cryptographic Access Control for Cloud based Electronic Health Records Systems

Dhivya B, Siqqiue Ibrahim S. P., Kirubakaran R

Computer Science and Engineering, Kumaraguru College of Technology, Coimbotore, Tamilnadu, India

ABSTRACT

Cloud based Electronic Health Record (EHR) systems are next generation"big data systems" for facilitating a) efficient and scalable storage, and b) to foster collaborative care, clinical research and development. Mobility and use of multiple mobile devices in collaborative healthcare intrigue robust privacy preservation. Thus, large scale EHR systems require secure access to privacy sensitive EHR data, data storage and management. We provide a comprehensive solution with i) a cryptographic role based technique to distribute session keys to establish communications and information retrieval using Kerberos protocol, ii) location and biometrics based authentication method to authorize the users and iii) a wavelet based steganographic technique to embed EHR data securely using ECG biometric as the host in a trusted cloud storage. Based on a comprehensive security analysis, our model proves to be a scalable, secure and a reliable model to access and manage EHR data.

Keywords: Electronic Health Record(EHR), Access Control,Location Awareness, Steganography.

I. INTRODUCTION

Electronic health records (EHR) Systems offer more efficient means for delivering quality ensured healthcare services and to promote collaborative clinical research. EHRs consist of data pertaining to "all aspects of care" (e.g. genomic test results, diagnosis, medication, laboratory test results and imaging data). According to Australian Bureau of statistics over 23 millions of Australian population [3], if the average EHR data file is 01 Gigabytes,then, the overall EHR data amounts to petabyte scale. According to IBM big data is defined as any situation or event that generates data with any or all of the three properties:Volume, Variety and Velocity [1]. Thus, it is evident that we have a significant "big EHR data" management problem in terms of volume, veracity and velocity. Poorly implemented EHR based systems pose significant risks for patient safety and data misuses [10] . The main security concerns in big data systems include secure-storage, secure-access and secureretrieval [2]. In addition to robust access control mechanisms, location of data access is an important aspect of secure data usage. Recently reported incidents on illegal trade and stealing of patient data over mobile devices and technologies intrigue research on secure data usage based on location. Multi-factor user authentication using multiple features with location validations is a promising solution. Healthcare service facilities increasingly use mobile devices to improve the work flow dynamics and efficiency. However, user mobility and use of multiple mobile devices make the above mentioned security concerns more severe and demands robust privacy preserving measures. Based on the above discussion, the main research questions that we address here are: How to securely store and manage "big EHR data"? and How to ensure secure location-aware access to "big EHR data"? Cloud based utility services (e.g. storage) offer additional benefits to EHR systems such as being more cost effective with the ease of management and to collaborate with mobile technologies and devices to gather data [4][5]. In addressing the first research question we find ECG as an appropriate host signal which is a verifiable secure feature to store " big EHR data" in cloud. For the second research question we propose a cryptographic role based access control which is a scalable solution to facilitate access to large number of users. Our solutions

are expressed as three main contributions. Affiliation based multi-factor authentication with location verification the use multiple autentication features including location and biometrics. A similarity score computed based on the affiliations between the authentication attributes are used to declare the validity of the claimed authenticity. Securely establishing communications with EHR system - We use Kerberos protocol to generate the session keys based on the location and identity verifications to securely establish sessions based communications with the EHR system.Secure and efficient cloud based storage of EHR data.The use wavelets based steganographic technique to store EHR data securely in cloud by using the ECG of each patient as the host signal.

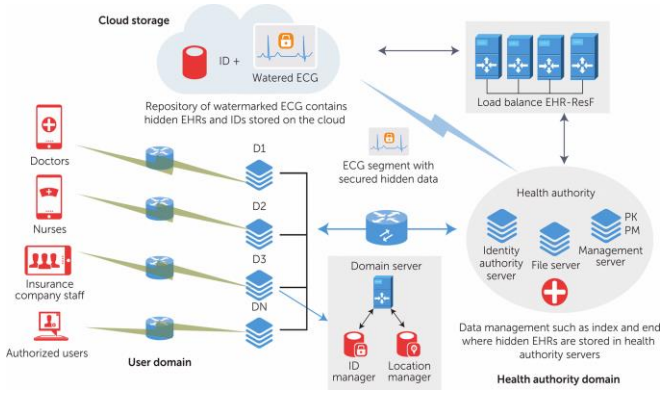


Figure 1. The main scenario of our proposed model where EHR hidden inside segments of ECG signals and Different users can retrieve different watermarked segments.

II. MODEL DESCRIPTION

The comprehensive cloud-based architecture (Figure 1) to facilitate secure utilizations of EHR data resources by enforcing cryptographic access control with context and location awareness. Based on the existing large scale ECG data collection based pilot studies [6], we assume the that ECG of sufficient length are available to store the EHR data for all the patients. To manage EHR efficiently and securely, The proposed novel design based on steganography. Steganography is used to hide confidential EHR data inside ECG host data. Strengths of steganography over traditional cryptography is evident due to its efficient and secure information concealment. Extraction of steganographically concealed EHR data is possible only for authorized users based on their security parameters [14]. Thus, guarantees the security of storage and retrieval of EHR by hiding inside ECG. Next, we will describe the main entities of the proposed architecture (see Figure 1). **Mobile users:** Health Authority (HA) validates the mobile users using multi-factor authentication. HA authenticate users based on fixed identity attributes and location by consulting the domain server. Based on this validation domain server (DS) informs the cloud service provider (CSP) to allow the requested formation to be transmitted. In a mobile service environment, location awareness is vital in distributing privacy sensitive data [9]. Face biometric is a widely accepted strong biometric feature that can be used in an unobtrusive manner to verify a user in a domain [8] using a mobile device equipped with a camera. Thus, we couple these two features to identify and validate the user as a legitimate user in a secure location known to the host network. Health authority (HA): The HA is a trusted third party, such as a government institution. For each user, HA provides unique identity credentials (username,password). User authentication and validations are processed by the Identity Authority (IU) residing in HA. Both users and the system admin share an offline process of generating users public/private keys by using certificateless public key cryptography (CL-PKC) [13] as it is not necessary to know all the private keys of the users by the key generating entity. These public/private keys are used to provide secure communication between users and HA. HA is also responsible for extracting the required information from the encrypted data. That select role based access control model to map the access requests to generate the session keys.The selected Kerberos protocol [12] to securely communicate the session keys to CSP and the user. Trusted CSP: It is the main data centre for holding and organzing patient EHR in a hierarchical manner to provide an efficient and secure mechanism for distributing the session keys among users. EHR resource facilitator (EHR-ResF): EHRResF performs Haar wavelets based EHR data hiding in ECG host data using the patient key, extract the encrypted data and transmit to using the session key.

III. MAIN OPERATIONS

A. Authentication and Secure Session Establishment

Kerberos protocol [12] has two main parts: i) Authentication server (AS); and ii) Ticket granting server (TGS), which contains role-based access control that manages users' roles and distributes session keys

to users (e.g. patients, physicians, nurses, or lab workers),to perform different tasks.

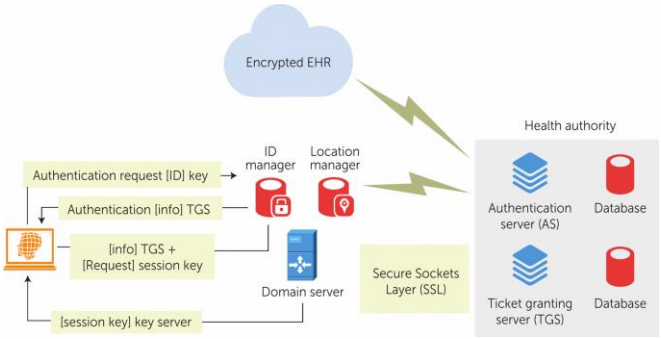
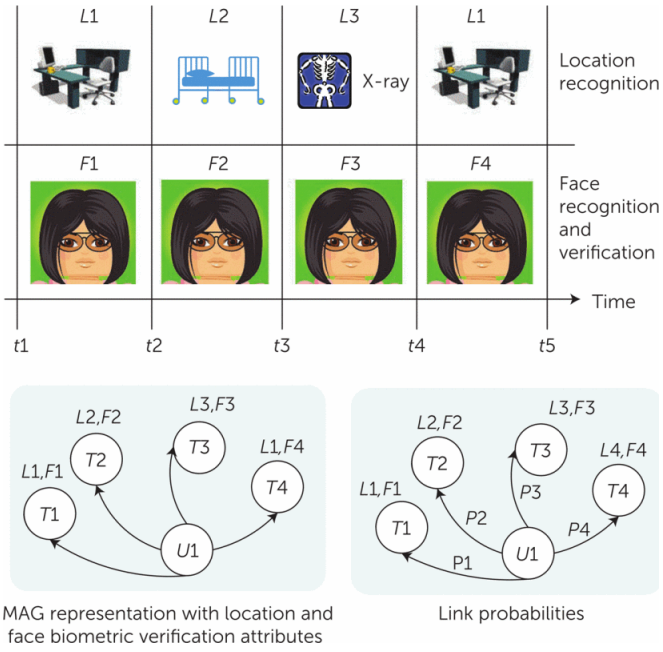


Figure 2. The Kerberos model infrastructure

The role hierarchy is used to identify patients, healthcare professionals (e.g. physicians, surgeons, cardiologists, gynecologists) and support staff (e.g. nurses, laboratory worker) assigned with different permissions. The AS and TGS work in a complementary manner through secure communication to automate, the authentication and authorization processes for users. The contribution of key distribution process is to check the users' credentials during the authorization process and distributes appropriate session keys for a user after receiving encrypted authorization information. The IU uses as a part of authentication process to verify mobile users locations though communication with domain server. In figure. 2, the first message pair is exchanged between users and the AS to prove the identities of users. This pair of messages are fully encrypted through the use of the AS credentials. The second pair of messages involves communication between the AS and the DS to verify users locations. In Section (A) we explain the process for location verification. Validation decision from DS to AS completes the authentication process. The third pair of messages exchanged between users and the TGS to complete the authorization process. The initial user request, validated credentials and the authentication details are encrypted and forwarded to TGS. TGS accepts and issues the session key based on the role-hierarchy. (A) Location Verification-The location validation is performed during the initial authentication of the user by the HA using a validation request to the DS and before delivering paragraphs must be indented. the extracted information content to the user. In the first instance location validation is vital to ensure the service request to be from a legitimate user belong to a secure location known to the respective DS. Before transmitting the extracted content to the user, HA contacts the DS to validate the current location of the user as a secure location. This validation is necessary as for the mobile user the location can change during the initial request was made and till the extraction process is completed. For example if the user is using an unmanaged WiFi network while travelling, then sensitive health information are susceptible to malicious eavesdropping and signal interceptions which can cause sever privacy breeches. Location-manager (see Figure 1) is able to verify the current location of the user and also can provide the trace based validity for the most recent 0n0 number of location changes. The affiliation of these n locations (i.e. the attributes) corresponding to two time instances (i.e. the nodes) can be computed using the multiplicative attribute graph (MAG) model [7]. The similarity between two sets of locations (known vs claimed locations) to declare the validity is interpreted as the affiliation between the attributes using the MAG model.When the Location-Manager gets the validation request from HA, it performs a proximity verification using the known location markers and the trace history information (Figure 3). Once this location markers are selected Location Manager further verifies with the Identity Manager by inferring the footage of the user by verifying the face biometric signature trace. Then the most likely location markers are selected using Multiplicative Attribute Graph (MAG) [7] (i.e. the link with the highest probability). If 0k out-of-n0 (where k=n > 60%) instances prove the users location and identity, then the user is validated as a legitimate user in a secure location known to the DS. Our approach is more robust as the verification is strongly coupled with the location information and face biometric trace within that time period.

B. EHR Embedding And Retrieval

In this model, we assume that HA is fully secure and responsible for generating the security parameters.Since we use lossy steganography, it does not increase the size as each bit of the original data is replaced by another bit of the hidden data. Simple a user U1 the face biometric features



MAG representation with location and face biometric verification attributes

Link probabilities

Figure 3. Identity verification using the location trace of for the time period (t1 t5) is represented using MAG model. bit replacement can significantly distort the original ECG. However, minimize the distortion of the host ECG by applying wavelet. When apply wavelet signal transformation, data gets divided into lots of coefficients. Then, we randomly hide the sensitive data into the least significant coefficients to ensure minimum distortion to the data. EHR embedding include, i) organizing the EHR content as distinct sections on to a tree structure and ii) randomly allocating them to different portions of the ECG segments by specifying their Indexes I and Ends E. After splitting ECG Segment, signal transformation technique using Haar wavelet is applied. The result of signal transformation yeild two sets of coefficients: CA and CD. The coefficients are used to classify each Segment a either: 1) the most sensitive features of the original signal (i.e. coefficients approximation CA), and 2) the least significant features (i.e. coefficients detailed CD) which can be freely used in hiding EHR Sections (See Figure 4). Next, for each EHR Section, a hash value H is computed. In order to make the hiding process unique to each patient, a security key (K) will be defined for every patient which will be used to 1) encrypt each Section of EHR before hiding it, 2) reshuffle coefficients (CD), and 3) hide Section bits in a certain set of coefficients. Haar Wavelet recomposition will then be applied on both CA and CD. Consequently, a new watermarked Segment is reconstructed. Next, re-embed the watermarked Segment into the full original ECG signal of that patient using its I and E. The same process will be repeated to hide all Sections. Finally, HA stores certain information such as I and E of each Segment, the hidden Section number and K along with unique Patient ID which is necessary for retrieving process. HA then stores the watermarked ECG along with generated number mapped to the patient ID on its cloud servers. Therefore, even these information are intercepted, they will not reveal anything. For an auhorized user, HA will extract ECG Segment from its cloud servers and accomplish the extraction process at its local servers. Then, it uses the session key to encrypt the bits and send them to the user device.

IV. SECURITY ANALYSIS

In this section, we present a security analysis for the communication channel i) between users' domain and HA, ii)between HA and cloud provider. Based on the analyses we demonstrate the resilience of the proposed system against man-in-the-middle (MiM) and reply attacks.

A. Communications Between User Domain And HA

It provides a qualitative analysis to demonstrate robustness of the communication channel between HA and the user. Consider an instance where an intruder sends a request to CCSP via a secure domain pretending to be a legitimate user with false identities and location information. The intruder will not succeed in gaining access to the EHR system due to two security features. Firstly, CCSP validates with the particular DS to re-validate the user identity and the location. DS validates the trace of the user based on the location and the associated face biometric over n instances. So the likelihood of the intruder to forge the face biometric at all the traced locations is highly unlikely. Therefore, it is clear that a mimicry attack cannot be successful to gaining access to the EHR system via the HA. Moreover, MiM attack is prevented through utilize public key infrastructure (PKI) to perform both authentication and authorization processes. All users have to communication with the HA securely through applying both public/private keys to exchange messages combined with time stamps. The PKI support both confidentially and integrity and the time stamps are prevent any reply attacks form intruders.

B. Communications Between HA To Cloud Based EHR Database

Assumption that HA is fully secure for generating and securely storing the key K and PID for all patients. Consider a scenario, where an intruder has access to the watermarked ECG at the cloud servers or during the transmission between cloud DB and HA.The number of all possible combinations of ECG segment hosting a particular EHR section can be defined as follows (Equation 1) ,

$$P=\prod_{n=1}^{n} S * \sum_{r=1}^{r} R * \sum_{c=1}^{c} C * N^{L} \quad \text{.....(1)}$$

where P is the total number of possible combinations, n is the number of samples in the ECG, r and c are the rows and columns numbers in the reshuffled coefficients CD, L is the key length and N is its possibilities. Assume n = 1000(i.e. an ECG of 10 second length), r =128 and c = 32(i.e. the size of reshuffled CD) and the key character set is 256 and its symbols length is 256 (See Eq2),

$$p=1000!*\sum_{r=1}^{128} R! * \sum_{c=1}^{32} C! * 256^{256} \rightarrow p=\infty \quad \text{.....(2)}$$

Thus, it proves that it is highly improbable to find the intended EHR section in reasonable time as shown in Equation 2. In conclusion, the channel between the HA and the cloud based EHR host is resilient to attacks such as man-in-the-middle attack.

V. CONCLUSION

Secure and efficient storage, access control and retrieval are crucial for big data EHR systems. In the article a hybrid cryptography based access control model for cloud based EHR systems. We use strong user authentication based on biometrics and location with Kerberos protocol to issue encrypted session tickets. The approach ensures a fail-proof two level user authentication for more robust secure interactions with the EHR system. Wavelet based steganographic technique to store the EHR data in a ECG signal which acts as a fail-proof host data to ensure the data to be accessed only to legitimate users. Furthermore, the security analysis proves the robustness of solution in terms of its resilience to some common yet notorious attacks. Thus, our solution offers a significant contribution to developing a large scale distributed EHR system with strong security features to preserve the privacy of the EHR data.

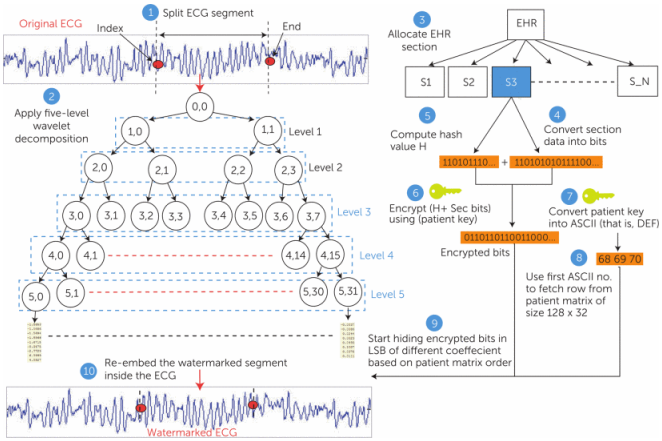


Figure 4. Block diagram shows the steps of hiding a section of EHR inside a segment of ECG host signal before re-embedding the segment in the watermarked ECG.

VI. REFERENCES

[1]. D. OLeary, Articial intelligence and big data, Intelligent Systems, IEEE, vol. 28, no. 2, pp. 96–99, 2013.

[2]. E. E. Schadt, M. D. Linderman, J. Sorenson, L. Lee, and G. P. Nolan, Computational solutions to large-scale data management and analysis, Nature Reviews Genetics, vol. 11, no. 9, pp. 647–657, 2010.

[3]. Australian Burea of Statistics, http://www.abs.gov.au/ausstats/abs%40.nsf/94713ad445ff 425ca25682000192af2/1647509ef7e25faaca2568a900154b63? OpenDocument.

[4]. S. Pandey, W. Voorsluys, S. Niu, A. Khandoker, R. Buyya, An autonomic cloud environment for hosting ECG data analysis services, Future Generation Computer Systems, vol. 28, no. 1, pp. 147–154, 2012.

[5]. U.S. Premarathne, I. Khalil, Z. Tari, and A. Zomaya. ”Cloud-based Utility Service Framework for Trust Negotiations using Federated Identity Management.”, IEEE Transactions on Cloud Computing, vol.99, pp.1-14, 2015.

[6]. J. Marek, V. Bufalino, J. Davis, K. Marek, A. Gami, W. Stephan, and F. Zimmerman. ”Feasibility and findings of large-scale electrocardiographic screening in young adults:data from 32,561 subjects”, Heart Rhythm, vol.8, no.10, pp. 1555-1559, 2011.

[7]. M. Kim and J. Leskovec, Multiplicative attribute graph model of real-world networks, Internet Mathematics, vol. 8, no. 1-2,pp. 113–160, 2012.
[8]. N. Kumar, A. Berg, P. N. Belhumeur, and S. Nayar, Describable visual attributes for face verification and image search, IEEE Transactions on Pattern Analysis and Machine Intelligence,vol. 33, no. 10, pp. 1962–1977, 2011.
[9]. M. S. Kirkpatrick, G. Ghinita, and E. Bertino, Privacypreserving enforcement of spatially aware rbac, IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 5, pp. 627–640, 2012.
[10]. F. Magrabi, M.-S. Ong, W. Runciman, and E. Coiera, Using fda reports to inform a classification for health information technology safety problems, Journal of the American Medical Informatics Association, vol. 19, no. 1, pp. 45–53, 2012.
[11]. F.A.P. Petitcolas, R.J. Anderson, and M.G. Kuhn. ”Information hiding-a survey.” Proceedings of the IEEE, special issue on Protection of Multimedia Content, vol. 87, no.7,pp. 1062-1078, 1999.
[12]. S. P. Miller, B. C. Neuman, J. I. Schiller, and J. H. Saltzer, Kerberos authentication and authorization system, in In Project Athena Technical Plan. Citeseer, 1987