Hybrid Cryptographic Access Control for Cloud-Based EHR Systems

Uthpala Premarathne, Alsharif Abuadbba, and Abdulatif Alabdulatif, RMIT University
Ibrahim Khalil, National ICT Australia
Zahir Tari, RMIT University
Albert Zomaya, University of Sydney
Rajkumar Buyya, University of Melbourne

A cryptographic role-based access control model for electronic health record (EHR) systems uses location- and biometrics-based user authentication and a steganography-based technique to embed EHR data in electrocardiography (ECG) host signals.
increase the gravity of security concerns and demands robust privacy-preserving measures.

Thus, the main research questions we address here are how to securely store and manage big EHR data, and how to ensure secure access to this data. Cloud-based utility services (such as storage) offer additional benefits to EHR systems—for example, they’re more cost effective, can be easier to manage (for example, access and retrieval), and support collaboration, with mobile technologies and devices to gather data.4,5 Electrocardiography (ECG) provides an appropriate host signal and a verifiably secure means to store big EHR data in the cloud. To ensure secure access to big EHR data, we propose a cryptographic role-based access control method that’s scalable to a large number of users.

Cryptographic Role-Based Access Control Model

Our comprehensive cloud-based architecture (Figure 1) facilitates secure access to EHR resources by enforcing cryptographic access control with context and location awareness. Based on existing large-scale ECG data-collection-based pilot studies,6 we assume that ECG of sufficient length are available to store EHR data.

To manage EHRs efficiently and securely, we propose a design based on steganography, which we use to hide confidential EHR data inside the ECG host data. Steganography offers more efficient and secure information concealment than traditional cryptography.7 Only authorized users can extract data based on their security parameters.8 Our steganography-based approach therefore improves the security of storage and retrieval of EHRs by hiding them inside ECG signals, and enhances performance through flexible feature adoption (such as dynamic policy changes).

In our approach, the health authority validates mobile users based on their identity and location attributes. A mobile user consults the domain server, which forwards the request to the health authority on the user’s behalf. Based on this validation, the domain server informs the cloud service provider (CSP) to allow the requested information to be transmitted.

In a mobile service environment, location awareness is vital in distributing privacy-sensitive data.9 Face biometrics can be used in an unobtrusive manner to verify a user in a domain.10 Thus, we couple these two features to identify and validate the user in a secure location known to the host network.

FIGURE 1. The proposed model. Electronic health records (EHRs) are hidden inside segments of electrocardiography (ECG) signals. Different users can retrieve different watermarked segments. (EHR-ResF: EHR resource facilitator)
The health authority is a trusted third party, such as a government institution. For each user, the health authority provides unique identity credentials (username and password). The identity authority, residing in the health authority, authenticates and validates the user. Both users and the system administrator share an offline process of generating users’ public/private keys using certificateless public key cryptography (CL-PKC) since the key-generating entity doesn’t need to know all the users’ private keys. These public/private keys provide secure communication between users and the health authority. The health authority is also responsible for extracting the required information from the encrypted data. Our role-based access control model maps access requests to generate the session keys. We use the Kerberos protocol to securely communicate the session keys to the CSP and the user.

The trusted CSP serves as the main datacenter for holding and organizing patient EHRs in a hierarchical manner to provide an efficient and secure mechanism for distributing session keys among users. The EHR resource facilitator performs two functions: Haar wavelet-based EHR data hiding in ECG host data using the patient key, and extraction of the encrypted data using the session key and transmits it.

### Authentication and Secure Session Establishment

The Kerberos protocol has two main parts: the authentication server and the ticket granting server (TGS), which uses role-based access control to manage users’ roles and distributes session keys to users (such as patients, physicians, nurses, and lab workers) to perform different tasks. We use role hierarchy to identify patients, healthcare professionals (such as physicians, surgeons, cardiologists, and gynecologists), and support staff (such as nurses and laboratory workers), who are assigned different permissions. The authentication server and TGS work in a complementary manner through secure communication to automate the authentication and authorization processes for users. The key distribution process checks users’ credentials during authorization and distributes appropriate session keys to users after receiving encrypted authorization information. The identity authority uses the user’s credentials as part of the authentication process to verify mobile users’ locations through communication with the domain server.

The health authority performs location validation using a validation request to the domain server both during initial user authentication and before delivering the extracted information content to the user. In the first instance, location validation ensures that the service request is from a legitimate user at a secure location known to the respective domain server. Before transmitting the extracted content to the user, the health authority contacts the domain server to validate the user’s current location as secure. This validation is necessary since the mobile user’s location can change between when the initial request is made and when the extraction process occurs.

## FIGURE 2.

The Kerberos model infrastructure. Messages are exchanged between users and the authorization server and between the authorization server and domain server to verify users’ identities and locations. Session keys are issued by TGS based on the users’ roles.
The process is complete. For example, if the user is communicating over an unmanaged Wi-Fi network while traveling, sensitive health information is susceptible to malicious eavesdropping and signal interception, which can cause severe privacy breeches.

The location manager (see Figure 3) can verify a user’s current location and provide trace-based validity for the most recent \( n \) number of location changes. We can compute the affiliation of these \( n \) locations (that is, the attributes) corresponding to two time instances (the nodes) using the multiplicative attribute graph (MAG) model.\(^{13}\) We interpret the similarity between two sets of locations (known versus claimed) to declare the validity as the affiliation between attributes using the MAG model.

When the location manager receives a validation request from the health authority, it performs proximity verification using the known location markers and the trace history information (Figure 3). After selecting location markers, the location manager further verifies the location with the identity manager by interpreting the footage of the user by verifying the face biometric signature trace. It then chooses the most likely location markers using the MAG model (that is, the link with the highest probability). If \( k \) out of \( n \) (where \( k/n > 60\% \)) instances prove the user’s location and identity, the user is validated as a legitimate user in a secure location known to the domain server. Our approach is more robust since the verification is strongly coupled with the location information and face biometric trace within that time period.

**EHR Embedding and Retrieval**

In our model, we assume the health authority is fully secure and responsible for generating security parameters. Since we use lossy steganography,\(^{14}\) the data doesn’t increase in size because each bit of the original data is replaced by another bit of the hidden data. Simple bit replacement can significantly distort the original ECG. However, we minimize the distortion of the host ECG by applying wavelets.\(^{15}\) When we apply wavelet signal transformation, data is divided into many coefficients. We then randomly hide the sensitive data in the least significant coefficients to ensure minimum distortion.

Embedding EHR data involves organizing the EHR content as distinct sections into a tree structure and randomly allocating it to different portions of the ECG segments by specifying their indexes (\( I \)) and ends (\( E \)). After splitting an ECG segment, we apply a signal transformation technique using Haar wavelets. This signal transformation yields two sets of coefficients: coefficients approximation (\( CA \)) and coefficients detailed (\( CD \)). We use these coefficients to classify each segment as the most sensitive features of the original signal (\( CA \)) or the least significant features (\( CD \)), which can be freely used to hide EHR sections (see Figure 4).

Next, for each EHR section, we compute a hash value. To make the hiding process unique to the individual, we define a security key for each patient. We use this security key to encrypt each EHR section before hiding it, reshuffle coefficients (\( CD \)), and hide section bits in a certain set of coefficients. We then apply Haar wavelet recomposition on both \( CA \) and \( CD \). Consequently, a new watermarked segment is reconstructed. Next, we re-embed the watermarked segment into the full original ECG signal of the patient using its index and end. We repeat this process to hide all sections.

Finally, the health authority stores certain information, such as each segment’s index and end, the hidden section number and key, along with a unique patient ID, which is needed for retrieval. The health authority stores the watermarked ECG along with the generated number mapped to the patient ID on its cloud servers. Therefore, even if this information is intercepted, it won’t reveal anything.
For an authorized user, the health authority will extract the ECG segment from its cloud servers and perform extraction at its local servers. It uses the session key to encrypt the bits and send them to the user device.

**Security Analysis**

We analyze the security of the communication channel between users’ domain servers and the health authority and between the health authority and the CSP.

First, we provide a qualitative analysis of mimicry attack resilience on the communication channel between the health authority and the user.

Consider an instance where an intruder sends a request to the CSP via a secure domain using false identity and location information to pass as a legitimate user. Two security features will prevent the intruder from gaining access to the EHR system. First, the CSP checks with the particular domain server to validate the user identity and the location. The domain server validates the user trace based on the location and the associated face biometric over instances. Because the likelihood of the intruder forging the face biometric at all the traced locations is low, a mimicry attack can’t successfully gain access to the EHR system via the health authority. Moreover, the use of a public key infrastructure (PKI) to perform both authentication and authorization prevents man-in-the-middle attacks. All users must communicate with the health authority securely by applying both public and private keys to exchange messages. The health authority also uses time stamps on arriving messages. The PKI supports both confidentiality and integrity, and time stamps prevent reply attacks from intruders.

We assume that the health authority is fully secure for generating and securely storing the security key and patient ID 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 storage and the health authority. The number of possible combinations of ECG segments hosting a particular EHR section can be defined as
\[ P = \prod_{n=1}^{\infty} S \times \sum_{r=1}^{R} R \times \sum_{c=1}^{C} C \times N^k, \]

where \( P \) is the total number of possible combinations, \( n \) is the number of samples in the ECG, \( r \) and \( c \) are the row and column numbers in the reshuffled coefficients \( CD \), \( L \) is the key length, and \( N \) is its possibilities.

Assume \( n = 1,000 \) (that is, a 10-second-long ECG), \( r = 128 \), and \( c = 32 \) (that is, the size of the reshuffled \( CD \)), the key character set is 256, and its symbol length is 256:

\[ p = 1000! \times \sum_{r=1}^{128} R! \times \sum_{c=1}^{32} C! \times 256^{256} \Rightarrow p = \infty \]

Thus, it’s highly improbable that an attacker would find the intended EHR section in a reasonable time. Thus, the channel between the health authority and the cloud-based EHR host is resilient to man-in-the-middle and similar attacks.

Our future work will focus on a robust key exchange management between various parties involved. We’ll also consider key revocations and risk mitigation strategies.

References

UTHPALA SUBODHANI PREMARATHNE is a PhD student and a National ICT Australia scholarship holder at RMIT University, Melbourne, Australia. Her research interests include access control, privacy, and trust negotiations in distributed systems. Premarathne has an MSc in computer science from the University of Moratuwa. Contact her at uthpala.s.p@gmail.com.

ALSHARIF ABUADBBA is a PhD student at RMIT University, Melbourne, Australia. His research interests include big data security, steganography, signal processing, high sensor streams, and the Internet of Things. Abuadbba has a master’s degree in computer science from RMIT University, Australia. Contact him at alsharif.abuadbba@rmit.edu.au.

ABDULATIF ALABDULATIF is a third year PhD student at RMIT University, Melbourne, Australia. His research interests include encryption and decryption techniques, distributed systems and networks, data
mining, and remote healthcare. Alabdulatif has a master’s degree in computer science from RMIT University. Contact him at abdulatif.alabdulatif@rmit.edu.au.

IBRAHIM KHALIL is a staff member at National ICT Australia (NICTA), Victoria Research Laboratory, Melbourne. His research interests include large network provisioning, management software, access control, network security, scalable efficient computing in distributed systems, and wireless body sensor networks. Khalil has a PhD in computer science from the University of Berne, Switzerland. Contact him at ibrahim.khalil@rmit.edu.au.

ZAHIR TARI is a professor in the School of Computer Science and Information Technology at RMIT University, Melbourne. His research interests include a special focus on the performance of Web servers, SOAP-based systems, SCADA system security, and Web services (specifically, protocol verification and service matching). He’s a senior member of IEEE. Contact him at zahir.tari@rmit.edu.au.

ALBERT ZOMAYA is the chair professor of high-performance computing and networking, an Australian Research Council Professorial Fellow in the School of Information Technologies, and director of the University of Sydney’s Center for Distributed and High Performance Computing, all at the University of Sydney. His research interests include parallel and distributed computing and complex systems. Contact him at albert.zomaya@sydney.edu.au.

RAJKUMAR BUYYA is a professor of computer science and software engineering, a Future Fellow of the Australian Research Council, and director of the Cloud Computing and Distributed Systems (CLOUDS) Laboratory at the University of Melbourne, Australia. He also serves as the founding chief executive officer of Manjrasoft. His research interests include cloud, grid, distributed, and parallel computing. Buyya has a PhD in computer science from Monash University. He’s a Fellow of IEEE. Contact him at rbuyya@unimelb.edu.au.

What’s Trending?

The information you need and only the information you need. Industry intelligence delivered on your terms, when and how you want it.

• myCS—delivers your publications your way
• myComputer—customizable mobile app delivering targeted information specific to your specialty
• Computing Now—this award winning website features industry news and developments.

Learn something new.
Check out these resources today!

Stay relevant with the IEEE Computer Society
More at www.computer.org