

A Security Model for Preserving the Privacy of Healthcare Cloud using a Fog Computing

Miss. Jambhulkar Chaitanya A¹ Miss. Jambhulkar Chaitanya A² Miss. Kature Nikita N³

^{1,2,3}Department of Information Technology

^{1,2,3}SND College of Engineering Yeola, Yeola, India

Abstract— Telemedicine is an emerging healthcare service where the healthcare professionals can diagnose, evaluate, and treat a patient using telecommunication technology. To diagnose and evaluate a patient, the healthcare professionals need to access the electronic medical record (EMR) of the patient, which might contain huge data including x-rays, ultrasounds, CT scans, MRI reports, etc. For efficient access and supporting mobility for both the healthcare professionals as well as the patients, the EMR needs to be kept in big data storage in the healthcare cloud. In spite of the popularity of the healthcare cloud, it faces different security issues for instance, data theft in this paper, the main focus has been given to secure healthcare private data in the cloud using a fog computing facility.

Keywords: EMR, HMAC, Fog Computing, DMBD Algorithm

I. INTRODUCTION

Big data in healthcare refers to sets of electronic medical health data that are large and complex. Due to their huge volume and complexity, it is difficult to manage those data sets using traditional software and Hardware. Telemedicine is one of the emerging fields for e-health research. In the telemedicine service, EMRs including MBD, images, and multimedia medical data are transmitted on the fly over insecure internet connections as they are required by the remote doctors. The healthcare cloud infrastructure would make it much easier to pull all different healthcare information together for a patient while the patient moves from one hospital to another; as a result, the patients' information can be managed and tracked easily. Similar to cloud computing, healthcare cloud computing has different issues related to its security, the most important of which are: legal and policy issues, data protection, privacy protection, lack of transparency, cyber security issues, absence of security standards, and software licensing. Each of these issues has different challenges that can be briefly discussed as follows. The challenges related to cloud computing's legal and policy issues are: liability, applicable law, compliance, copyright, data portability, and data protection. Speaking about protection, privacy protection means to protect the personally identifiable information (PII), by making it clear to the consumer how it is used and where it is stored. Usually, privacy issues are all about three things, which are trust, uncertainty, and compliance. Also, another issue related to the consumer is lack of transparency, which may appear through the consumer not knowing where his/her data are physically stored or what happens to it. On the other hand, another cloud security issue is cyber security.

II. BACKGROUND AND PRELIMINARIES

In this section we provide a few technical backgrounds that will help to ensure better understanding of our proposed

technique. 2169-3536 (c) 2017 IEEE. Translations and content mining are permitted for academic research only this article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2757844, IEEE Access IEEE ACCESS. VOL.XXX, NO.XX, 2017 3 A. Cloud Computing Cloud computing has different service models, which are divided into three categories: (1) IaaS, which allows users to take advantage of the infrastructure without mentioning the hardware running behind it; (2) PaaS, which builds on IaaS and provides clients with access to basic operating software and optional services to develop and use software applications without software installation; and (3) SaaS, which enables clients to use software applications without having to install them on their personal computer, by offering these as a service through the Internet. We can categorize cloud computing consistent with the deployment model into: (1) a public cloud, in which the resources are sold or rented to the public by the service provider, who at the same time is the owner; (2) a private cloud owned or rented by an organization; (3) community clouds, in which some closed communities share the same cloud resources; and (4) a hybrid cloud, which has the characteristics of two or more deployment models. Several features are available in cloud computing, for example: on demand broad network access, self-service, measured service, resource pooling, and rapid elasticity. Self-service means that the customers can manage and request their own resources. On the Internet or in private networks, the services offered are known as broad network access. In pooled resources, the customer draws from a pool of computing resources, usually in a remote data center. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information. This article has been accepted for publication in a future issue of this journal, but has not been fully edited.

A. Motivation and Our Contribution

MBD stored in the healthcare cloud typically reside in a shared environment collocated with data from other stakeholders. Data theft occurs when information is illegally copied or taken from a business or individual. Such theft usually involves user information, such as passwords, social security numbers, credit card information, other personal information, and confidential corporate information. A recent study done in the Saudi population sought to determine the most important contents in mobile phones. The results showed that 88% of the female respondents answered "personal photos," whereas 63% of the male participants cited "personal information" and "personal photos." The statistics clearly indicate that personal photos are the most important contents in mobile devices. This may also mean that types of

content which may be important to some people may not be as important for others.

III. RELATED WORK

A decoy defense network can be deployed to bolster the security in different situations. In the authors discussed several scenarios in which a decoy can be used. One usage scenario involves using a decoy within a local computer, which means placing the decoy document within the same environment in which it was created. In another scenario, the decoy can be located on a network level. In both scenarios, the decoy is used to protect documents on different levels. However, a decoy can also be used to protect software, as by being made to look like a legitimate source code, decoy software can protect real software from unauthorized usage. Another decoy usage scenario applies a voicemail decoy to detect malicious activity; here the decoy is a legitimate voice message but contains false information. Lastly, a cloud-based decoy can be used to protect documents in the cloud against insider attacks. A few studies have focused on securing cloud data by using decoy documents. For instance, in the authors first carried out user behavior profiling to determine unauthorized access. When an attacker accesses the cloud, a decoy document is returned such that the real user's data are kept secure. Each decoy document header contains a hidden Hash-based Message Authentication Code (HMAC). Verification of whether or not the document is a decoy is done by calculating the HMAC based on the content of the document; if the two HMACs match, then the document is a decoy and an alert is issued. In this case, decoy documents are used for two purposes: first, to validate whether or not the data access is authorized when abnormal information access is detected, and second, to confuse the attacker by providing false documents. It should be noted that only decoy documents are used in this study, and these are selected manually and added into the file system by the user. In a similar technique carried out by Aruna, Prasad, and Reddy malicious insider attacks were prevented by using decoy information technology. When abnormal information access is detected, the decoy helps to validate whether or not the access is authorized

IV. PROPOSED SYSTEM

Now and in the subsequent sections, whenever we use the terms "Gallery/ Photo gallery" we mean "multimedia MBD gallery," and the gallery contains MBD. Using fog computing facilities and the decoy technique, a DMBD is created. This technique can be considered as an illusion technique, as it makes the attacker believe that he/she has accessed the user's MBD while in fact it is just a decoy gallery. In our proposed system, as shown in Figure 3, once the user accesses his/her 2169-3536 (c) 2017 IEEE. Translations and content mining are permitted for academic research only Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2757844, IEEE Access IEEE ACCESS. VOL.XXX, NO.XX, 2017 6 account, by default the DMBD is shown. Thus, both authorized and unauthorized users will be referred to the DMBD as the first step, while authorized legitimate users, as a second step, will be referred to the OMBD after being verified. We believe that by setting

the default value of the DMBD as shown and the OMBD as hidden, we keep the original MBD more secure. Also, we believe that verifying that the user is legitimate is much easier than detecting the attacker, which is why we tried to deal with the attacker in the first place by offering the DMBD as the first step. When the user accesses his/her account, whether he/she is a legitimate user or an attacker, his/her first step would be accessing the DMBD, which is located in the fog computing layer side by side with user profiling. User profiling is a familiar technique that can be applied to model in what way, at what time, and how considerable users access their information in the healthcare cloud. This method of behavior based security is commonly used in fraud detection application. The DMBD contains fake MBD, which are supposed to make an attacker believe that he/she has accessed the user's photos/medical image while in fact it is just a decoy gallery. The legitimate user already knows that the gallery he/she accessed is not his/her original one, so would move on to the next step. Moving to the next step, the legitimate user can access his/her OMBD after being verified by passing the security challenge. The security challenge might be a challenging security question or even a verification code. Thus, if he/she passes the security challenge, that means he/she is the legitimate user, so will be able to access the OMBD which is located on the cloud computing layer. In the event of the user accessing only the DMBD, an SMS or email will be sent to the legitimate user to inform him/her that his/her account has been accessed. The message will contain the attacker's information (e.g., access time and date and the IP address). Now, how do we ensure that the two galleries are similar to a large extent? Each time the legitimate user uploads a new photo/medical image on his/her account, a decoy photo from fog computing will be uploaded to his/her DMBD, as shown in Figure 4. When the user uploads the photo, he/she is supposed to recognize the photo category (ECG, X-ray, MRI, etc.), which will help fog computing to add the photo that belongs to the same category on the DMBD; this would make it closer to the original photo, so that the attacker would not differentiate between the real user's photo and the fake one

A. DMBD Algorithm

DMBD is used as a trap gallery that makes it not of direct relevance to the legitimate user but it is used to secure his/her OMBD by distracting the attacker. As shown in Figure 5, the DMBD is placed in the fog computing as a honeypot to secure the original one, which is located in the cloud. As noted, a number of anomaly-detection systems are provided by fog computing such as user profiling and a decoy file system. Therefore, for each newly uploaded MBD in the OMBD, a decoy one will be placed on the DMBD.

B. User Profiling Algorithm

User profiling can help to determine whether a user is legitimate or not based on certain parameters, such as the user search behavior, amount of downloaded data, nature of operations, division of tasks, and IP address. Knowing how a legitimate user deals with his/her cloud data based on these parameters will help determine whether or not the user is malicious. There are three different types of user profiling, each with different advantages and disadvantages based on

the techniques used. The type that we will use in our system is the hybrid user profile, which is a combination of explicit and implicit user profiles. The explicit user profile usually contains high-quality information because it is gathered from the user him/herself, but it requires a lot of effort from the user to update his/her profile information. On the other hand, the implicit user profile is automatically updated with minimal user effort; however, a large amount of interaction between the user and the content is required before an accurate user profile can be created. Thus, combining the two types into a hybrid user profile should reduce the weak points and enhance the strong points of each technique used to monitor the cloud data access and detect any unusual data access pattern.

V. CONCLUSION

As a part of securing the cloud data mission, this paper focuses on securing user's multimedia data within the cloud by using cryptography, fog computing and steganography. To this end, two photo galleries are generated. The OMBD is kept secretly in the cloud and the DMBD is used as a honeypot and is kept in the fog. Therefore, instead of retrieving the DMBD only when any unauthorized access is discovered, the user, by default, accesses the DMBD.

ACKNOWLEDGMENT

We felt great pleasure in submitting this partial project report on 'a security model for preserving the privacy of healthcare cloud using a fog computing'. At the outset, we wish to express our deep sense of gratitude towards our guide Prof. K. A. Bhoir, for guidance and encouragement, without which it would have been not possible. She has always been constantly helping in preparing this project by clearing our doubts. We would like to thank prof. P. P. Rokade, (H.O.D. IT department) for his encouragement, guidance and allowing me to use the all facilities in the department. We are also grateful to Principal Prof. H. N. Kudal, who has been a constant source of our inspiration. We express our immense pleasure and thankfulness to all the teachers and staff of the department of the information technology for their cooperation and support. Last but not the least we thank all others, and especially thankful to all our friends who has helped us in the successful completion of this partial project report. They are one who is always ready to stand behind me at any cost. We owe our all success to them.

REFERENCES

- [1] M. Chen, J. Yang, Y. Hao, S. Mao, K. Hwang, "A 5g Cognitive System For Healthcare ", *Big Data And Cognitive Computing*, Vol. 1, No. 1, Doi:10.3390/Bdcc1010002, 2017
- [2] Frost Sullivan: Drowning In Big Data? Reducing Information Technology Complexities and Costs for Healthcare Organizations.
- [3] M. Chen, S. Mao, Y. Liu, "Big Data: A Survey", *Mobile Networks and Applications*, Vol. 19, No. 2, Pp. 171-209, April 2014.

- [4] M. S. Hossain, And G. Muhammad, "Healthcare Big Data Voice Pathology Assessment Framework," *IEEE Access*, Vol. 4, No. 1, Pp. 7806-7815, December 2016.
- [5] M. Chen, Y. Had, K. Hwang, L. Wang, L. Wang, "Disease Prediction By Machine Learning Over Big Healthcare Data", *IEEE Access*, Vol. 5, No. 1, Pp. 8869-8879, 2017.
- [6] M. Chen, P. Zhou, G. Fortino, "Emotion Communication System", *IEEE Access*, Vol. 5 Pp. 326-337, 2017.