UNIVERSITI PUTRA MALAYSIA

PRIVACY PRESERVING FOR ELECTRONIC HEALTH RECORD SYSTEMS

Zakariye Mohamed Yussuf

FSKTM 2019 17

PRIVACY PRESERVING FOR ELECTRONIC HEALTH RECORD SYSTEMS

BY

ZAKARIYE MOHAMED YUSSUF

Dissertation submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Information Security

June 2019

DEDICATIONS

This thesis is dedicated to my wonderful parents and beautiful family in the hope that it shows everything can be achieved given the right support.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the Degree of Master of Information Security

Privacy Preserving For Electronic Health Record Systems

By

ZAKARIYE MOHAMED YUSSUF

JUNE 2019

Supervisor: Dr. Mohd Taufik Abdullah
Faculty: Computer Science and Information Technology

Electronic Health Records (EHR) made it easier to manipulate and manage health records in health centers, it enables many medical institutions to exchange the EHR with ease. However, as the architecture of these services become complicated, it introduces new security threats, for instance privacy of patient's data and information when EHR is exchanged between institutions and users. In order to keep the patient's information private, many systems and methods have been proposed to implement access control to the health records of patients. However, most of the recent approaches don't state the importance of strong authentication, don't support fine-grained access control and also do not take into account the encryption of data in the server. Consequently, this research proposes an EHR system that works with attribute-based access control using PHP Laravel framework. The proposed system provides multi-factor authentication, access control and also encrypts EHRs.

ACKNOWLEDGMENTS

I take this opportunity to thank the almighty Allah for giving me the faith to believe in myself and carry on even when I almost gave up.

I would like to express my profound gratitude to Dr. Mohd Taufik Abdullah for his exemplary guidance, monitoring and constant encouragement towards completing this dissertation. He has been of great help and gave me guidance from time to time which I will be appreciating in my journey of life.

I thank and appreciate the valuable effort of my lecturers at Faculty of Computer Science and Information Technology (FSKTM), Universiti Putra Malaysia (UPM), Malaysia.

I am obliged to my family and friends for the valuable information and support provided by them through various stages of this dissertation. They were outstanding and stood for me for any help I needed. I am very much grateful for their cooperation and willingness to help me go through this stage without which this dissertation would not be possible.

This dissertation was submitted to the Information Security Department, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of **Master of Information Security**.

Dr. Mohd Taufik Abdullah
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia

Date: June 25, 2019

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

(Signature)

ZAKARIYE MOHAMED YUSSUF

(GS50400)

Date: June 25, 2019

Table of Contents

DEDICATIONS ..... **I**

ABSTRACT ..... **II**

ACKNOWLEDGMENTS ..... **III**

DECLARATION ..... **V**

TABLE OF FIGURES ..... **VIII**

TABLE OF TABLES ..... **IX**

CHAPTER ONE **1**

INTRODUCTION **1**

1.1 Background And Motivation 1
1.2 Problem Statement 3
1.3 Research Objective 3
1.4 Research Scope 3
1.5 Organization of Dissertation 4

CHAPTER TWO **5**

LITERATURE REVIEW **5**

2.1 Introduction 5
2.2 Definitions 5
2.2.1 Health Records 5
2.2.2 Electronic Health Record 5
2.2.3 Data privacy 6
2.2.4 Encryption 6
2.3 Authentication 6
2.4 Access Control 7
2.4.1 Types of Access Control 8
2.4.1.1 Mandatory access control (MAC) 8
2.4.1.2 Discretionary access control (DAC) 9
2.4.1.3 Role-based Access Control (RBAC) 9
2.4.1.4 Rule Based Access Control 9
2.4.2 Attribute-Based Access Control (ABAC) 10
2.5 History Of Electronic Health Records 10
2.6 Benefits Of Electronic Health Record Systems 11
2.7 Cons Of Electronic Health Record Systems 12
2.8 Attacks That Compromise the Privacy of Patient's Data 12
2.8.1 Password attack 13
2.8.2 SQL injection attack 13
2.8.3 Cross-Site Request Forgery 13

2.8.4 Unauthorized Access 13
2.8.5 Loss/theft and unencrypted records in the EHR 14
2.9 Related Work 16
2.10 Summary 23
CHAPTER THREE 24
METHODOLOGY 24
3.1 Introduction 24
3.2.1 Authentication 25
3.2.1.1 Username and Password 25
3.2.1.2 One-Time Password 25
3.2.2 Access Control 26
3.3 Base Work 27
3.4 System Components 28
3.5 Summary 28
CHAPTER FOUR 29
THE PROPOSED PRIVACY PRESERVING FOR ELECTRONIC HEALTH RECORD SYSTEMS 29
4.1 Medical Data (MIMIC III) 29
4.2 System Design 29
4.2.1 User Interface 31
4.2.2 Admin Interface 32
4.3 Tools Used 32
4.3.1 MySQL 32
4.3.2 Microsoft Project 32
4.3.3 Microsoft Visio 33
4.3.4 HTML 33
4.3.5 PHP 33
4.3.6 Laravel 33
4.3.7 JavaScript 33
4.3.8 Vue.js 34
4.3.9 Sublime Text 34
4.3.10 AES 34
4.3.11 CSS 34
4.4 System Features 35
4.4.1 Registration Page 35
4.4.2 Login Part1 35
4.4.3 Login Part2 36
4.4.4 Admin Page 37
4.4.5 Add Practitioner 37
4.4.6 Edit and Delete practitioner 38

4.4.7 Practitioner page 39
4.4.8 Add a Patient 40
4.4.9 Edit and Delete Practitioner 40
4.4.10 Medical Record Page 41
4.5 Summary 42
CHAPTER FIVE 43
RESULTS AND DISCUSSION 43
5.1 Introduction 43
5.2 Testing 45
5.2.1 Login Testing 46
5.2.2 Registration testing 47
5.2.3 Administrator 48
5.2.4 Practitioner Page 48
5.3 Summary 49
CHAPTER SIX 50
CONCLUSION AND FUTURE WORK 50
6.1 Introduction 50
6.2 Conclusion 50
6.3 Future Works 50
REFERENCES 51

Table of Figures

Figure 2. 1 Authentication categories 7
Figure 2. 2 Access control 8
Figure 2. 3 Number of Reported Data Breaches 12
Figure 2. 4 Unauthorized access/disclosure incidents 13
Figure 2. 5 exposed records in unauthorized access/disclosure 14
Figure 2. 6 Theft/loss incidents 14
Figure 2. 7 Records exposed in theft/loss 15
Figure 2. 8 PHR model (Qian H. et al, 2015) 16
Figure 2. 9 framework of using two different clouds. (Fatemeh et al, 2016) 18
Figure 2. 10 Framework of using ECG to conceal private EHR (Premarathne et al, 2016) 19

Figure 3. 1 Flowchart of the proposed system 24
Figure 3. 2 Multifactor authentication 25
Figure 3. 3 ABAC 26
Figure 3. 4 User sending request to the server 27

Figure 4. 1 Framework of the proposed model 30
Figure 4. 2 Use case of the model 30
Figure 4. 3 User interface 31
Figure 4. 4 Admin interface 32

Figure 4. 5 Registration form 35
Figure 4. 6 Login form 36
Figure 4. 7 One-time password 37
Figure 4. 8 Admin page 37
Figure 4. 9 Add practitioner 38
Figure 4. 10 Practitioner page 39
Figure 4. 11 Add patient 40
Figure 4. 12 Medical Record Page 41

Figure 5. 1 Code for encrypting data 45
Figure 5. 2 Encrypted data in the database 45

Table of Tables

Table 2. 1 healthcare data breaches (2009 – 2018) 15

Table 5. 1 Login testing 46
Table 5. 2 Registration testing 47
Table 5. 3 Edit and Delete testing 48
Table 5. 4 Add, edit and delete testing 48

CHAPTER ONE

INTRODUCTION

This chapter commences with an overview of the research topic and explains the motivation for this work. The problem statement is then identified followed by the research objectives and scope. At the end of this chapter the organization of the dissertation is provided.

1.1 Background And Motivation

With the rapid development of information technology, health organisations migrated from paper-based methods of keeping health records to digitised medical information system or electronic health records (EHR). With electronic health records, huge amounts of medical information can be managed easily in comparison to paper based methods. This migration comes with a handful of advantages, some of them are: improved quality of care, the promotion of evidence-based medicine and record keeping, a reduction in costs and mobility, Luis et al (2013). It's defined by Seol et al (2018) that EHRs are electronically stored digital forms consisting of all patient's health data which means that it's an electronic or digital version of a patient's record that holds elements regarding the patient's health like blood type, age, diseases, medications, diagnoses, laboratory and test results and so on. To improve the quality and effectiveness of health institutions as well as the accuracy of the diagnosis, patient's data are distributed across multiple-sites with different institutions and also are required to be shared and accessed by the medical practitioner that requires the history of the patient's medical record at that very moment.

The Electronic Health Record (EHR) has advanced to be in the focus of most European countries, and globally as stated by (Dipak et al, 2006). There are many standards that govern an EHR, such as, health level seven (HL7) which is a collection of standard formats that state the interfaces for electronic information exchange in healthcare environments between computer applications, health insurance portability and accountability act (HIPAA) which provides security means and protection measures to safeguard health information. HIPAA

defines some information as protected health information, these include social security number, address of home, credit card, mobile number, medical data and others.

For patient's medical record to be exchanged to improve medical services, a common platform that is compatible with the different types of systems that hospitals and clinics use is required. Cloud computing is clearly the best platform simply because it can be accessed by anyone that pays for its services and it comes with a lot of benefits like speed, scalability, reduced cost and higher performance. As Microsoft defined, cloud computing is the delivery of computing services such as, servers, storage, databases, networking, software, and analytics over the Internet ("the cloud").

Electronic Health Records have come with a lot of advantages ranging from storing large amount of medical data to easily accessing patient's data from anywhere, but that comes with price, loads of attempt to safeguard the privacy of patient while not compromising usage of the information and corresponding healthcare services as discussed by Sharma et al (2018). Unauthorised exposure of sensitive health data violates HIPAA and also can have a big impact on the patient's social, health-related and economic life. To protect the patients' privacy of medical record is crucial in this era of technology where security issues arise regularly. Some examples of the attacks are, on January 29, 2015 anthem.inc which is a US health insurance giant experienced a massive data breach during which approximately 78.8 to 80 million Americans have had their personal information exposed to hackers as reported by Infosec Institute website. Sources stated that during the attack no medical information was stolen but personal information only, sources also stated that the insurance company didn't encrypt their files. Another attack reported by wall street journal that is an attack on a government computer that is used by insurance and brokers to directly enrol customers. The hacked computer interacts with healthcare.gov and an estimated of 75,000 files was compromised.

To protect all these cyber-attacks and to limit the exposure of sensitive medical records, this research proposes a EHR system that employs attribute based access control (ABAC) and can only be accessed by authorised users. The system also encrypts data in the database in case an intruder gains access, the data will not be plaintext and readable.

1.2 Problem Statement

Few methods to preserve patient's privacy of medical records were proposed in the last few researches. Seol et al (2018) proposed use of fine-grained access control called Attribute access control (ABAC), encryption and digital signature. In the mentioned research, partial encryption was deployed which reveals part of the data in the document therefore sensitive medical records can be exposed in case an intruder hacks the system and or the server is compromised.

Just like what happened to anthem.inc in 2015 in which their systems were hacked and approximately 80 million personal information stolen might be due to lack of privacy protection because their files and data were not encrypted.

1.3 Research Objective

- The main objective of this study is to design and implement EHR system that preserves the privacy of patient's medical records. To achieve this objective, the study will be guided by the following specific objectives:
    - To propose attribute based access control – access rights are granted to users through the use of policies which combine attributes together.
    - To propose mechanism to protect unauthorized user to access medical records by encrypting patient's medical records in the database.
    - To propose multifactor authentication to strengthen access to the system.

1.4 Research Scope

- The system will only be used with local data even though it can later be implemented in cloud infrastructure.
- This system will be used across the medical practitioners that require patients' health records to provide the necessary services.

1.5 Organization of Dissertation

This research comprises of five chapters. These chapters are as follows. Chapter 1 is the introduction chapter, covers the background of the study, identifies and discusses the problem statement, research objectives, explains the scope of research, ends with a brief description of the organization of the dissertation.

Chapter 2 introduces the literature review on privacy preserving on Electronic health record system, focuses on the threats that might compromise the security in terms of privacy, highlights and discusses different methods and techniques used to preserve privacy, identifies strength and weakness for the methods used. In addition to that, journals with studies related to the research topic are identified.

Chapter 3 explains the research methodology and presents the EHR framework in detail. What follows chapter 5 which explains the proposed research technique, and implementation by employing PHP Laravel version 5.8.

The results are presented and discussed in Chapter 5 while Chapter 6 offers conclusion of the study and makes recommendations for further related research

REFERENCES

Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. Á. O., & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. *Journal of biomedical informatics*, *46*(3), 541-562.

Seol, Kwangsoo, et al. "Privacy-Preserving Attribute-Based Access Control Model for XML-Based Electronic Health Record System." IEEE Access 6 (2018): 9114-9128.

Kalra, Dipak. "Electronic health record standards." Schattauer GMBH-Verlag, 2006. 136-144.

Sharma, Sagar, Keke Chen, and Amit Sheth. "Toward practical privacy-preserving analytics for iot and cloud-based healthcare systems." IEEE Internet Computing 22.2 (2018): 42-51.

https://resources.infosecinstitute.com/category/healthcare-information-security/healthcare-attack-statistics-and-case-studies/case-study-health-insurer-anthem/#gref

https://www.wsj.com/articles/hackers-breach-healthcare-gov-1539991262

Brakerski, Zvika, Craig Gentry, and Vinod Vaikuntanathan. "(Leveled) fully homomorphic encryption without bootstrapping." ACM Transactions on Computation Theory (TOCT) 6.3 (2014): 13.

Chen, Feng, et al. "Perfectly Secure and Efficient Two-Party Electronic-Health-Record Linkage." IEEE internet computing22.2 (2018): 32-41.

Yang, Ji-Jiang, Jian-Qiang Li, and Yu Niu. "A hybrid solution for privacy preserving medical data sharing in the cloud environment." *Future Generation Computer Systems* 43 (2015): 74-86.

Khan, Shahidul Islam, and Abu Sayed Latiful Hoque. "Privacy and security problems of national health data warehouse: a convenient solution for developing countries." *Networking Systems and Security (NSysS), 2016 International Conference on*. IEEE, 2016.

Lu, Yang, et al. "Privacy-Preserving Access Control in Electronic Health Record Linkage." *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*. IEEE, 2018.

Dagher, Gaby G., et al. "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology." *Sustainable*

Cities and Society 39 (2018): 283-297

Rezaeibagha, Fatemeh, and Yi Mu. "Distributed clinical data sharing via dynamic access-control policy transformation." *International journal of medical informatics* 89 (2016): 25-31.

Premarathne, Uthpala, et al. "Hybrid cryptographic access control for cloud-based EHR systems." *IEEE Cloud Computing* 4 (2016): 58-64.

Yang, Kan, et al. "Time-domain attribute-based access control for cloud-based video content sharing: A cryptographic approach." *IEEE Transactions on Multimedia* 18.5 (2016): 940-950.

Sandhu, R. S., & Samarati, P. (1994). Access control: principle and practice. *IEEE communications magazine*, *32*(9), 40-48.

Evans, R. S. (2016). Electronic health records: then, now, and in the future. *Yearbook of medical informatics*, *25*(S 01), S48-S61.

Qian, H., Li, J., Zhang, Y., & Han, J. (2015). Privacy-preserving personal health record using multi-authority attribute-based encryption with revocation. *International Journal of Information Security*, *14*(6), 487-497.

Qian, H., Li, J., Zhang, Y., & Han, J. (2015). Privacy-preserving personal health record using multi-authority attribute-based encryption with revocation. *International Journal of Information Security*, *14*(6), 487-497.

Narayan, S., Gagné, M., & Safavi-Naini, R. (2010, October). Privacy preserving EHR system using attribute-based infrastructure. In *Proceedings of the 2010 ACM workshop on Cloud computing security workshop* (pp. 47-52). ACM.

Danwei, C., Xiuli, H., & Xunyi, R. (2009, December). Access control of cloud service based on ucon. In *IEEE International Conference on Cloud Computing* (pp. 559-564). Springer, Berlin, Heidelberg.

Yuan, E., & Tong, J. (2005, July). Attributed based access control (ABAC) for web services. In *IEEE International Conference on Web Services (ICWS'05)*. IEEE.

Bell, D. E. and La Padula, L. J. 1976, Secure computer system: Unified exposition and multics interpretation, Tech. rep., MITRE CORP BEDFORD MA.

Qian, X. and Lunt, T. F. 1996. A MAC policy framework for multilevel relational databases. IEEE Transactions on Knowledge and Data Engineering 8 (1): 3{15.

Kuhn, D. R., Coyne, E. J., & Weil, T. R. (2010). Adding attributes to role-based access control. *Computer*, *43*(6), 79-81.

Kofler, M. (2001). What Is MySQL? (pp. 3-19). Apress

https://en.wikipedia.org/wiki/Microsoft_Project

https://www.groovypost.com/howto/enable-gmail-input-tools/

Bakken, S. S., Suraski, Z., & Schmid, E. (2000). *PHP Manual: Volume 2*. iUniverse, Incorporated

https://laravel.com/docs/4.2/introduction

Raggett, D., Le Hors, A., & Jacobs, I. (1999). HTML 4.01 Specification. *W3C recommendation*, *24*.

https://developer.mozilla.org/en-US/docs/Web/JavaScript

https://vuejs.org/v2/guide/

Johnson, A. E., Pollard, T. J., Shen, L., Li-wei, H. L., Feng, M., Ghassemi, M., ... & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific data*, *3*, 160035.

Abbas, A., & Khan, S. U. (2014). A review on the state-of-the-art privacy-preserving approaches in the e-health clouds. *IEEE Journal of Biomedical and Health Informatics*, *18*(4), 1431-1441.

Sandhu, R., Ferraiolo, D., & Kuhn, R. (2000, July). The NIST model for role-based access control: towards a unified standard. In *ACM workshop on Role-based access control*(Vol. 10, No. 344287.344301).