PIXEL-BASED DISPERSAL SCHEME FOR MEDICAL IMAGE SURVIVABILITY AND CONFIDENTIALITY

NELMIAWATI

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Computer Science (Information Security)

Faculty of Computing
Universiti Teknologi Malaysia

JUNE 2014

This project report is dedicated to my wonderful husband Wahyudi Arifandi,
for his endless support and encouragement.

ACKNOWLEDGMENT

Alhamdulillaahirabbil'alamiin, deeply praise to Allah subhanahu wa ta'ala for help and ease me in all way to finally complete my study.

First and foremost, I would like to express my sincere gratitude to my supervisor Assoc. Prof. Dr. Mazleena Salleh, for the guidance, encouragement, and passion that help me a lot during my work on this project as well as during my study at UTM.

Also, my sincere gratitude to my mom and dad, my mother-in-law and father-in-law, sisters and brothers. Your prayer for me was what sustained me thus far. A special thanks to my wonderful husband who spent sleepless nights with and was always support me every time.

Not forgetting, my sincere gratitude to all my friends in UTM for your big support. Thank you so much.

ABSTRACT

Digital medical image survivability and confidentiality is an essential concern for healthcare institution in archiving medical image. Survivability issue in the current conventional archiving system happened when the server becomes temporarily inaccessible; meanwhile confidentiality happened when the data disclosed by unauthorized party. This project introduces a potential approach by integrating Rabin's Information Dispersal Algorithm (IDA) and Shamir's Secret Sharing Algorithm (SSA) for solving these issues. These approaches have been already known for providing data survivability and confidentiality in other fields. This project applies these secret sharing threshold schemes to a system called Medical Image Dispersal Picture Archiving Communication System (MIaDPACS) to provide digital medical image survivability and confidentiality. The pixels are extracted, encoded, and dispersed by using Rabin's IDA into distributed archiving system. The secret key used in Rabin's IDA is dispersed by using Shamir's SSA. Original image reconstruction requires only subset of the dispersed files with numbers equals to a defined threshold during dispersal. Experimental results showed that the reconstructed image was exactly identic with the original image both for no-salt and salt scheme. This is shown by Peak Signal to Noise Ratio (PSNR) value that is equal to infinity. In addition, random dispersed files does not reveal the original image. Thus, MIaDPACS is able to provide the survivability and confidentiality for digital medical image.

ABSTRAK

Ketahanan dan kerahsiaan digital imej perubatan merupakan usaha yang perlu bagi institusi kesihatan dalam pengarkiban imej perubatan. Isu ketahanan dalam sistem pengarkiban konvensional semasa berlaku apabila pelayan tak tercapai buat sementara waktu; sementara itu kerahsiaan berlaku apabila data didedahkan oleh pihak lain tanpa izin. Projek ini memperkenalkan pendekatan yang berpotensi dengan mengintegrasikan *Rabin Information Dispersal Algorithm* (IDA) dan *Shamir Secret Sharing Algorithm* (SSA) untuk menyelesaikan isu-isu ini. Pendekatan telah pun dikenali untuk menyediakan data ketahanan dan kerahsiaan dalam bidang lain. Projek ini menggunakan skim perkongsian ambang rahsia melalui sistem yang disebut *Medical Image Dispersal Picture Archiving Communication System* (MIaDPACS) untuk menyediakan ketahanan dan kerahsiaan digital imej perubatan. Piksel diekstrak, dikodkan, dan tersebar dengan menggunakan *Rabin's IDA* ke dalam pengarkiban sistem teragih. Kunci rahsia digunakan dalam *Rabin's IDA* disebar dengan menggunakan *Shamir's SSA*. Pembinaan semula imej asal hanya memerlukan subset fail berselerak dengan nombor sama dengan ambang yang ditakrifkan dalam penyebaran. Keputusan eksperimen menunjukkan imej yang dibina semula adalah tepat identik dengan imej asal untuk kedua-dua tanpa garam dan garam skim. Ini dinilai oleh *Peak Signal to Noise Ratio* (PSNR) yang mana hasilnya sama dengan infiniti. Tambahan, fail berselerak rawak tidak mendedahkan imej asal. Oleh itu, MIaDPACS mampu menyediakan ketahanan dan kerahsiaan bagi digital imej perubatan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACR	-	American College of Radiology
ASP	-	Application Service Provider
bits	-	binary digits
CCD	-	Charge-Coupled Device
CCIR	-	Consultative Committee on International Radio
CGA	-	Color Graphics Adapter
CMS	-	Clinical Management System
CMYK	-	Cyan Magenta Yellow black
CT	-	Computed Tomography
FDCT	-	Forward
DCT	-	Discrete Cosine Transform
DDR	-	Double Data Rate
DICOM	-	Digital Imaging and Communication in Medicine
DPI	-	Dot Per Inch
ECC	-	Error Correction Code
EGA	-	Enhanced Graphics Adapter
FDCT	-	Forward DCT
GF	-	Galois Field
GIF	-	Graphic Interchange Format
HIPAA	-	Health Insurance Portability and Accountability
HIS	-	Hospital Information System
HLS	-	Hue Lightness Saturation
HSV	-	Hue Saturation Value
IDA	-	Information Dispersal Algorithm
IDCT	-	Inverse DCT
IDE	-	Integrated Development Environment

IE	-	Information Entities
IED	-	International Electrotechnical Commission
IOD	-	Information Object Definition
IS	-	International Standard
ISO	-	International Standard Organization
ITU	-	International Telecommunication Commission
JFIF	-	Joint File Interchange Format
JPEG	-	Joint Photographic Expert Group
JRE	-	Java Runtime Environment
LPI	-	Line Per Inch
MRI	-	Magnetic Resonance Imaging
NEMA	-	National Electrical Manufacturers Association
PACS	-	Picture Archiving and Communication System
PAS	-	Patient Information System
PNG	-	Portable Network Graphics
PPI	-	Pixel Per Inch
RAM	-	Random Access Memory
RGB	-	Red Green Blue
RIS	-	Radiology Information System
RLE	-	Run Length Encoding
RS-ECC	-	Reed Solomon Error Correction Code
SC	-	Secondary Capture
SDK	-	Software Development Kit
SOP	-	Service Object Pair
SPI	-	Sample Per Inch
SSA	-	Secret Sharing Algorithm
SVGA	-	Super Video Graphics Array
TIFF	-	Tagged Image File Format
USG	-	Ultrasonography
VGA	-	Video Graphics Array
VR	-	Value Representation
XGA	-	Extended Graphics Array

CHAPTER 1

INTRODUCTION

1.1 Introduction

Digital data has become widely used in various aspects of human life. It offers cost-efficiency and flexibility on data manipulation, storage and transmission. It is classified into: numeric, text, audio, graphic, image, animation, and video. Radiology areas in healthcare institution have made significant revolution to adopt digital technology by using digital modalities such as X-Ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Ultrasonography (USG). Digital imaging in medicine provides radiologist, doctor, and patient to have instant access on it regardless their location and without afraid of degrading the image quality through time in a storage system as what happened on analog film imaging. Moreover, it also provides dynamic range of contrast, dynamic levels of gray, and several manipulation that could improve accuracy on image analysis during diagnose and treatment a disease (Vidar Systems Corporation, 2011).

1.2 Problem Background

Digital medical imaging technology has become an important part of the healthcare workflow. Today, it needs 50 to 100 TB for the average 150,000 to 200,000 imaging-related procedures done annually in average a 300-bed to 500-bed hospital (MgGee, 2009). It is common that the retention requirements to be decades, or even forever (Boyle, 2011). It is shown by several studies shown that a lot of number used digital medical image in healthcare institutions. A study concluded that there were significant rises in the use of CT and MR in one tertiary hospital in the United States from 1984 to 1993, while there were decrease in the use of conventional analog plain films and fluoroscopy (Khorasani, et al., 1998). Besides, about 26 million CT examinations were conducted in the United States in 1998, and then by 2008 there were more than 70 million CT examination were conducted (Watson & Odle, 2013). Moreover, the other study found that there has been also a growing number on usage digital medical imaging for developing countries (Mohd-Nor, 2011).

In addition, storing digital medical images only in local repositories as what is offered by Picture Archiving and Communication System (PACS) does not satisfied fast growing of digital medical image usages. As medical imaging technology improves and more clinicians demand access to patient's diagnostic quality images, healthcare institutions found that storage requirements are soaring (MgGee, 2009).

The existence of cloud computing is one of current approach that is starting to be introduced to store digital medical images. It offers economic advantages to reduce cost storage, good quality of services and it could give economy advantage to healthcare institutions (Ribeiro, et al., 2011). However, this approach has several disadvantages related to the security of digital medical image in healthcare institutions:

i. Consumer lock-in and bankruptcy issue of cloud provider could lead to lost control of digital medical image store in a cloud services provided by external entities (Ribeiro, et al., 2011). The existence of digital medical image in the cloud service storage is questioned.
ii. Data stored in a cloud services belongs to external entities. Uploading data to a cloud makes the data no resistance for disclosure or unauthorized usage

(Ribeiro, et al., 2011). Hence, it introduces threat on its confidentiality and privacy.

There are methods used to address this archiving issue. Redundant Array of Independent Disk (RAID) stored the data into decentralized disks. However, the usage of RAID in a single local storage might cause vulnerability in terms of survivability. In addition, implementation of cloud storage for addressing the archiving issue still have some problems on the security as discussed above. Therefore, the archiving issue on digital medical image could be addressed by applying the concept of archive into decentralized servers through pixel-based dispersal scheme.

1.3 Problem Statement

The existence of digital medical image plays an important role in healthcare institution such as in diagnosing, study, and treatment of a patient, whereby losing or disclosure of these medical images caused a lot in healthcare business continuity. Therefore, digital medical image security is important to be concern from the threats such as storage malfunction, man-made attack, natural attack, and disclose data by unauthorized parties.

1.4 Research Questions

From the problem statement described in the Section 1.3, study of the project is referred to several research questions:

i. What is digital medical image structure and secret sharing threshold scheme?

ii. How to ensure survivability and confidentiality of digital medical images on the storage based on pixel-based dispersal scheme?

iii. How to implement pixel-based dispersal scheme based on digital medical image?

1.5 Project Goal

Digital medical image that is stored in a local storage nowadays are having survivability and confidentiality issues. Due to the storage malfunction or temporary down time storage caused by manmade attack or natural attack, these data are inaccessible. Besides, disclosure data by the third party caused unauthorized usage of these data. The data might be changed and not confidential anymore. Therefore, the goal of the project is to address digital medical image survivability and confidentiality issues in healthcare institution by using secret sharing threshold scheme through pixel-based dispersal scheme.

1.6 Project Objective

Based on the problem background described above, the objectives that have been identified to complete this project are:

i. To study structure of digital medical images and secret sharing threshold scheme.

ii. To design and analyze pixel-based dispersal scheme based on secret sharing threshold scheme.

iii. To implement and validate digital medical image survivability and confidentiality upon applying pixel-based dispersal scheme.

1.7 Project Scope

The scope that is included in this project:

i. The project is focused on Rabin's Information Dispersal Algorithm (IDA) (Rabin, 1989) (Wang, et al., 2011) and Shamir's Secret Sharing Algorithm (SSA) (Shamir, 1979).

ii. The implementation is supported on digital medical image data Digital Imaging and Communications in Medicine (DICOM) 2D data set (.dcm), produced by medical imaging modality such as MR, CT, X-Ray Angiogram (XA).

iii. The implementation is limited for single frame DICOM.

iv. The implementation does not support for DICOM with more than 1byte number of allocated byte for each band from its pixel.

v. There were four servers used as distributed storages and three threshold number used to reconstruct original medical image.

1.8 Significant of the Project

A pixel-based dispersal scheme using secret sharing threshold scheme that has been developed in this project gave the advantage to preserve the survivability and confidentiality issues in storing digital medical image. The digital medical image will survives in the storage every time whenever is needed by the doctor or patient even though there is a possibility of down time storage as well as another threats that effect the accessibility of patient's data. Digital medical image is dispersed into several distributed storage and could recover from a down time storage case from less amount of the distributed storages. Even though the third party can get the data from the dispersed data however, they could not reconstruct it and understand the content. Moreover, this scheme provides distributing bandwidth as well for fast accessed digital medical image into several devices and networks.

1.9 Contribution

Digital medical image is a sensitive data that is threaten due to loss and modification of data over server temporary down or even an insecure network. Therefore, survivability and confidentiality of digital medical image is an important issue to be addressed. Thus, this project has contributed on addressing survivability and confidentiality issues for digital medical image by using secret sharing threshold scheme.

1.10 Report Organization

The project report is managed into six chapters. Chapter 1 describes the main part of the project. It describes about the introduction of the project, including digital data especially digital imaging and get into digital medical imaging, identifies problem background and problem statement related to digital medical image archiving, and finally describe a propose alternative solution to overcome the current problem through project objective, scope and goal.

Chapter 2 describes the literature review of the project. In the literature review, key components of the project has been studied and identified. It began with the explanation about the key concept of information security, especially on confidentiality and survivability. Security on digital medical image is concerned in this project. This chapter then continues describe about digital medical image and related component to digital medical image, such as DICOM and PACS. Moreover, secret sharing threshold scheme as a technique used to secure the data in this project also describe in this chapter. It involves Rabin's IDA and Shamir's SSA. Also, this chapter is presented along with current work on Rabin's IDA and Shamir's SSA.

Chapter 3 describes about research methodology used in the project. It identifies the progress of the project phase by phase through a research framework in accordance with the objectives of the project as described in the Chapter 1. Besides, the infrastructure of the work as in project environment including hardware, software, Application Programming Interface (API) setup and dataset are described.

Next, Chapter 4 describes design architecture and implementation of the project. It began by the explanation of environment deployment as in design architecture of the project development. Then, it discusses about pixel-based dispersal scheme implementation, especially main functions developed for the scheme implementation. The discussion also described pseudocode of the functions. It is then used in order to develop the project implementation. Finally, each menu of implemented application is discussed as well.

For the implementation part as in the previous discussion, Chapter 5 continues to the discussion of project validation including functional testing, testing result and result discussion. The section discuss about experiment setup of functional testing. Furthermore, this chapter discussed testing result and result discussion. Based on the result, the security of digital medical image survivability and confidentiality are identified and discussed.

Finally, Chapter 6 concludes the works. The achievements and milestones are listed. Besides, a suggested future work is described as well to improve the achievement of current result of this project.

REFERENCES

Bian, J. & Seker, R., 2013. The Jigsaw Secure Distributed File System. *Computers & Electrical Engineering,* XXXIX(4), p. 1142–1152.

Blitzer, H., Stein-Ferguson, K. & Huang, J., 2008. File Format and Compression. In: *Understanding Forensic Digital Imaging.* Indiana: Elsevier, pp. 241-256.

Boyle, M. A., 2011. Bursting at the Seams: Storage is Growing Problem for EHR Images. *Medical Economics*, 14 September.

Evans, D. & Davies, D., 2014. *dcm4che.org.* [Online] Available at: http://www.dcm4che.org/confluence/pages/viewpage.action?pageId=393 [Accessed 11 May 2014].

Forza Silicon Corporation, 2013. *History of Digital Imaging,* California: Forza Silicon Corporation.

Gutiérrez-Martínez, J., Núñez-Gaona, M. A., Aguirre-Meneses, H. & Delgado-Esquerra, R. E., 2012. A Software and Hardware Architecture for a High-Availability PACS. *Journal of Digital Imaging,* Volume 25, pp. 471-479.

Hood, M. N. & Scott, H., 2006. Introduction to Picture Archive and Communication Systems. *Journal of Radiology Nursing*, XXV(3), pp. 69-74.

Huang, H. K., 2011. Short History of PACS. Part I: USA. *European Journal of Radiology*, LXXVIII(2), pp. 163-176.

Intel, 2009. *Color Models,* US: Intel.

Jaganathan, S. & M., A. F., 2012. An Integrated Algorithm Supporting Confidentiality and Integrity for Secured Access and Storage of DICOM Images. *International Journal of Biometrics and Bioinformatics (IJBB)* , 6(2), pp. 47-57.

Kenney, A. R., Rieger, O. Y. & Entlich, R., 2003. *Moving Theory into Practice: Digital Image Tutorial,* New York: Cornell University Library Research Department.

Khorasani, R. et al., 1998. Trends in the Use of Radiology with Inpatients: What has Changed in a Decades?. *American Journal of Radiology,* April, 170(4), pp. 859-861.

Liu, B. J. et al., 2003. Trends in Image Storage and Archive. *Computerized Medical Imaging and Graphics,* XXVII(2-3), pp. 165-174.

Liu, Y.-L.et al., 2014. Performance Enhancement of a Web-Based Picture Archiving and Communication System Using Commercial Off-the-Shelf Server Clusters. *BioMed Research International,* Volume 2014, pp. 1-9.

McHugh, S., 2013. *Cambridge in Colour: Bit Depth Tutorial.* [Online] Available at: http://www.cambridgeincolour.com/tutorials/bit-depth.htm [Accessed 5 November 2013].

MgGee, M. K., 2009. Medical Imaging Advances Lead To Storage Headaches. *Clinical Information System*, 9 August.

Mohd-Nor, R., 2011. Medical Imaging Trends and Implementation: Issues and Challenges for Developing Country. *Journal of Health Informatics in Developing Countries,* 5(1), pp. 89-98.

Murphy, D. J. & Morgan, D. T. W., 2006. Availability, Reliability and Survivability: An Introduction and Some Contractual Implications. *The Journal of Defense Software Engineering,* pp. 26-29.

NEMA, 2011. *The DICOM Standard,* Virginia: NEMA.

Rabin, M. O., 1989. Efficient Dispersal of Information for Security, Load Balancing, and Fault Tolerance. *Journal of the ACM,* April, 36(2), pp. 335-348.

Rahman, M. M., Wang, T. & Desai, B. C., 2004. *Medical Image Retrieval and Registration: Towards Computer Assisted Diagnostic Apporach.* Beijing, s.n., pp. 78-89.

Ribeiro, L. S., Costa, C. & Oliveira, J. L., 2011. Current Trends in Archiving and Transmission of Medical Images. In: O. F. Erondu, ed. *Medical Imaging.* s.l.:InTech, pp. 89-106.

Rosset, A., 2013. *DICOM Sample Image Sets,* Switzerland: Antoine Rosset.

Seibert, J. A., 2006. *Archiving: Medical Image Data Characteristics,* Virginia: The Society for Imaging Informatics in Medicine.

Shamir, A., 1979. How to Share a Secret. *Magazine Communications of the ACM,* November, 22(11), pp. 612-613.

Slamanig, D. & Hanser, C., 2012. *On Cloud Storage and the Cloud of Clouds Approach.* London, IEEE, pp. 649-655.

Telagarapu, P., Naveen, V. J., Prasanthi, A. L. & Santhi, G. V., 2011. Image Compression Using DCT and Wavelet Transformation. *International Journal of Signal Processing, Image Processing and Platform Recognition,* IV(3), pp. 61-74.

Turner, J. E., 2002. *A Beginner's Guide to PACS,* Norwich: Crown.

Ulutas, M., Ulutas, G. & Nabiyev, V. V., 2011. Medical Image Security and EPR Hiding Using Shamir's Secret Sharing Scheme. *The Journal of Systems and Software,* pp. 341-353.

Vidar Systems Corporation, 2011. *The Transition to Digital Imaging in Medicine - The Integral Role of Film Digitizers,* Virginia: Vidar System Corporation.

Wang, S., Agrawal, D. & Abbadi, A. E., 2011. *A Compression Framework for Secure Query Processing on Relational Data in the Cloud.* Seattle, s.n., pp. 52-69.

Watson, L. & Odle, T. G., 2013. *Patient Safety and Quality in Medical Imaging: The Radiologic Technologist's Role,* Albuquerque: ASRT.

Wikipedia, n.d. *List of monochrome and RGB palettes.* [Online]
Available at:
http://en.wikipedia.org/wiki/List_of_monochrome_and_RGB_palettes#16-bit_RGB
[Accessed 27 October 2013].

Wylie, J. J. et al., 2000. Survivable Information Storage Systems. pp. 61-67.

Zain Alabdeen, A. A., 2013. *Enhancement Secret Sharing Scheme for Data Confidentiality and Integrity,* Johor Bahru: UTM.

APPENDIX A

ENVIRONMENT DEPLOYMENT INSTALLATION

1. Oracle JDK & JRE

   Steps by steps installation and configurations of Oracle JDK+JRE used for client and server side solution are given as following:

   i. Download jdk-7u55-windows-i586.exe from http://www.oracle.com/technetwork/java/javase/downloads/jdk7-downloads-1880260.html.

   ii. Double click the jdk-7u55-windows-i586.exe file to start the installation, and then follow the installation wizard.

   iii. Create JAVA_HOME environment variable with value the full path address of the JDK target installation, for example:

   JAVA_HOME = C:\Program Files\Java\jdk1.7.0_55

   iv. Append the value of `PATH` environment variable with %JAVA_HOME%\bin, for example:

   PATH = %JAVA_HOME%\bin;%PATH%

   v. Open new command prompt, and type following command to verify that the JDK and JRE were installed successfully.

   Verifying JDK:

```
prompt> javac –version
javac 1.7.0_55
```

   Verifying JRE:

```
prompt> java –version
java version "1.7.0_55"
Java(TM) SE Runtime Environment (build 1.7.0_55-b13)
Java HotSpot(TM) Client VM (build 24.55-b03, mixed mode,
sharing)
```