Using Multi-Cloud Approach to Enhance the Protection of Patient Privacy in Electronic Health Records

Naiem Yousefifard, Mitra Abbasi, Rasoul Jourmand

Abstract—Today, most hospitals and medical centers have a database to manage patients medical records. It is necessary to collect and access this information everywhere, to increase productivity and improve quality of health cares. Need to access electronic health information across the world and improve the quality of healthcare to patients will highlight importance of using cloud computing architecture in this area. However, despite the benefits of cloud computing applications for health care, the security challenges of cloud should be addressed.In this paper, we introduce and describe a new proposed model called EHR-MCDB (Electronic Health Record with Multi-Clouds Databases). This model will ensure the privacy of persons in a network and multi-cloud environment using cryptographic algorithm and distributed storage. This model is able to store any type of data such as number, string, image, etc. in a cloud environment. the EHR-MCDB prevent from illegal intrusion and access to data in cloud environment and provide services such as data integrity, confidentiality and permanent availability of data in a safe and secure way. The evaluation results of proposed EHR-MCDB show that despite increasing time of information storage and retrieval, System overhead is so that it can be ignored ,against the benefits that it provides, and it does not damage to Normal system activity. Also comparing the proposed model with data distribution method on different servers shows that it is improved almost "about 10% compared to the same multi-cloud model.

Index Terms— Cloud Computing, Electronic Health Records, Privacy, Multi-Clouds

1 Introduction

Cloud computing is a way of computing in a space related to the capabilities information technology. It is offered as a service to users and allows them to access the technology-based services in the Internet (Cloud) without specific information about this technology or control technology infrastructure that support them. For example, Google Apps provides online Public tools for business so that users can access to their software and data on the servers.

The public use the cloud as internet services such as Hotmail (since 1996), YouTube (since 2005), Facebook (since 2006) and Gmail (since 2007).

Hotmail was the first application on the cloud that allows people to store data, photos and files through remote access [1]. With the advancement of information technology most of the hospitals and health centers use software systems to register electronic patient's medical records [2].

There is a need for systematic and constant innovation for cost effectiveness and efficiency of health services and provides high quality services to patients. Many managers and experts predict that cloud computing can decrease cost of installing electronic health record systems [3].

Despite the benefits of cloud computing applications for health care health care, there are security challenges in the cloud systems especially public cloud systems that must be addressed which its infrastructure and computational resources are offered to Public access by a third party . In this regard, one of most important issues is privacy of patients that occurs in cloud architecture for data outsourcing. In this paper, we try to enhance the security and privacy of patients by using multi-cloud method and division of critical data on them.

2 Basic Definitions

2.1 EHR[1]

E-Health is a new field between medical informatics, public health and business that consider serving, distribution or clarification of Health and medical information via the Internet or related technologies. EHR includes systematic collection of Electronic records of individual health information at different periods. It contains illnesses reports, medicinal problems, vital signs, disease history, laboratory data and radiology reports. Quality of diagnosis and treatment of health care is highly dependent on the data about patient's condition. These records were usually stored locally within a certain collection and can be used. [5]. Today, most hospitals and medical centers have a database to manage their patients medical records But it is necessary to collect and access these information everywhere, to increase productivity and improve quality of health cares [2]. Required Information in HER:

- Information on Patients, including ID and personal details

- *Naiem Yousefifard M.Sc. Information Technology in E-Commerce from Shiraz University, Shiraz, Iran.*
  E-mail: n.yousefifard@mazums.ac.ir
- *Mitra Abbasi is currently pursuing master degree program in information technology from Mazandaran University of Science and Technology, Babol, Iran.*
  E-mail: mabbasi@mazums.ac.ir
- *Rasoul Jourmand is currently pursuing masters degree program in software engineering from Institute For Higher Education ACECR khouzestan, Ahwaz, Iran.*
  E-mail: rasoul.jourmand@gmail.com

[1] Electronic Health Record

- Information on doctors, including ID and personal details
- Patients health Information: including patient treatment ID and other clinical information (hospital or medical center name, treatment type, date and time, etc.)

2.2 Privacy

The same features that were considered as strengths of electronic registration of medical record create health information privacy concern [6]. The thread that information is kept out of the medical centers, as well as they are available everywhere, creates the problem of privacy and security of critical data. The three basic objectives of HER systems are seems essential: confidentiality, integrity and availability [5]. Health information privacy and security are always major concern for public, patients and health service providers [6].

Concerning protection of individual privacy in relation to electronic medical records, using aliases and Being unknown are the most common Protective measures. Confidentiality and integration between the patient and his medical records are the ways for Anonymity individual health information [2].

2.3 Cloud Computing

Cloud computing is a model to provide easy access to a set of changeable computer resources (e.g., networks, servers, storages, applications and services) based on user demand via a computer network. This type of access can be provided or released with minimum need to manage resources or direct involvement of service providers. Generally, cloud services consumers are not the owner of its physical infrastructure and to reduce the cost they do not need to spend a lot of money to buy expensive software or complex hardware. Because they can rent them third-party providers by using cloud services, including software services, infrastructure or platforms (SaaS[2], IaaS[3], PaaS[4]) and only pay the price for the resources that they use.

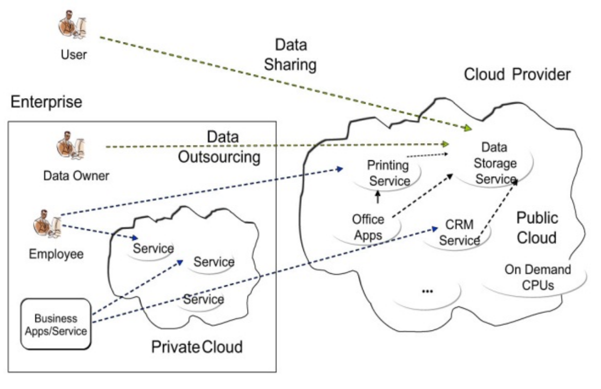

Fig. 1. Illustration of cloud computing process [7].

Some of the advantages of cloud computing:

- No need to prepare the equipment by the user
- No concern about loss of data on your local PC
- Data Sharing between various equipments
- Provide many services via the internet to users
- Fast and continuous software improvement

Today, Web 2.0 is the latest example of cloud computing. Using cloud-based platform enables the storage and exchange of medical records information between various hospitals and medical centers. In addition, using cloud-based platform, patients can get their medical information by an interface in an easy way, instead of searching in different hospitals and medical centers [2].

3 Challenges in Cloud Computing

The benefits of cloud computing are easy to identify. It will lead to more capacity and flexibility, and particularly reducing costs. All these benefits are needed to help growing a successful business. Clouds distinct advantages attract the attention of many organizations, but the aspect that makes many organizations to retreat against this technology is How to secure data in the cloud and ensure the safety of environment.

Outsourcing data and applications, sharing, multi-tenancy, virtualization, identification and authentication of users create challenges in cloud computing [8], [9], [10]. Data on the cloud consist of less critical public data, private and highly sensitive information such as social security numbers, medical records, or essential shipping information. Furthermore, identification of providers who are vulnerable can be a good target for hacker's attacks [11]. In addition, issues such as service availability and data transfer from one place to another (due to lack of resources caused by many customers' requests) can be considered as cloud computing challenges [1].

4 Cloud Computing Security

The Data generated by various people and institutions (such as email, medical records, photos, financial transactions, etc.) has been increasing rapidly and management of these data by the cloud will bring flexibility and economic saving. To protect information privacy and fight against unwanted access in the cloud, the owners of data before outsourcing sensitive data encrypt them [12]. To ensure confidentiality, integrity and availability of data, service providers of storage resources should contain at least the following capabilities:

- Stored data encryption schema
- Access control to prevent unauthorized access
- data backup plan
- Secure storage of data on backup media

National Institute of Standards and Technology (NIST) established cloud computing security group in order to create security policies on it. NIST recently published guideline to adoption and use of security protocols [11]. Considering that cloud services are built based on the internet, any matter that is related to Internet security will affect of cloud services. Cloud resources are accessible via the internet; as a result, even if the cloud providers focused on

[2] SaaS : Software as a Service
[3] IaaS : Infrastructure as a Service
[4] PaaS : Platform as a Service

cloud infrastructures security, the data is still transferred among the users through the internet that may be unsafe. In addition, using encryption techniques and secure protocols are not enough against data intrusion in the cloud.

We proposed a model based on multi-cloud databases, instead of common single-cloud model, to enhance the security of the stored data in the clouds. The main purpose of the proposed new model is to avoid malicious risk in the clouds and to prevent cloud services failure and security risks such as data integrity, data intrusion and service availability.

In the multi-cloud model, data distribution depends on the number of providers (CSP) and the confidential data sharing is a major factor in this approach [13]. In this paper, we try to share and distribute important data in medical records in order to maintain and ensure the privacy of individuals.

5 Related Works

Various works are done on privacy and data security in cloud computing. Wang and et al [14] presented file distribution plan in different servers for increased data security in the cloud. Chuang and et al [15] have proposed a model for privacy by using of encryption algorithms as well as sharing data in various storage resources. AlZain and et al [13] have proposed multi-cloud model to increase security of stored data in the cloud. Zhou and et al [7] have provided a plan for protection of privacy in outsourcing data in cloud computing through production and management of encryption keys in the form of tree. Somani and et al [16] have used digital signature and RSA algorithm to increase data security in cloud computing. Feng and et al [17] have proposed a new encryption method to protect data in distributed storage environments. Haas and et al [5] have provided a model to maintain the privacy and protection of personal information in electronic health records. The proposed model consists of two subsystems for serving patients and storing data in the HER. Perera and et al [6] have done a research about the level of security risks acceptance and reception of making electronic medical records at a clinic in Ontario, Canada. The survey results state that 58% of patients and nearly 70% of doctors believe that medical records computerization benefits are greater than the risk of loss of confidentiality. King and Raja [18] with a focus on privacy laws in Europe and the United States, have offered framework to help protect the privacy and security of consumers sensitive data in the cloud. Li and et al [2] are trying to improve the anonymity of the patient’s in electronic medical record. Ahmed and Abdullah [19] by overview of rapid growth of cloud computing have examined telemedicine services. Yang and Oiao [20] have proposed privacy protection with a random break between data and maintaining their relationships. Jin and et al [21] suggested the use of fingerprints based on a string of bits to protect privacy.

It can be noted that there is no clear and specific evidence in individual privacy protection in electronic medical records on cloud computing. In this article, we will try to fix data security and privacy problem by using multi-cloud and data distribution method, and finally examine the efficacy.

6 Proposed Model

This section introduces and describes the new proposed model, called EHR-MCDB (HER with Multi-Cloud Databases). The new model will ensure privacy in a multi-cloud environment, using the encryption algorithm and distributed storage. This model is able to store any type of data such as number, string, image, etc. in a cloud environment. the EHR-MCDB prevent from illegal intrusion and access to data in cloud environment and provide services such as data integrity, confidentiality and permanent availability of data in a safe and secure way.

6.1 Overview of the EHR-MCDB Model

Overview of the EHR-MCDB model is shown in Figure 1. According to the figure, the private information of electronic health records have stored on several clouds instead of storing in a cloud. In addition, this storage is done as encrypted and distributed. In other words, the data has been encrypted by encryption algorithm and then in term of number of clouds are divided into several sections. Finally, each part is stored in a single cloud. On the other hand, data queries are sent from host computer to the server. The server by receiving and analyzing query from host application start to collect data from databases located on different clouds. After collecting data and decrypting them, the final information can be sent to host computer.

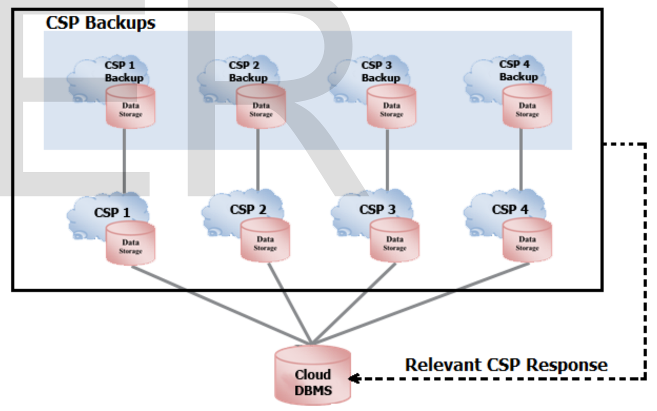

Fig. 2. Overview of the EHR-MCDB model.

6.2 The Scenario of EHR-MCDB Model

In the proposed model, we first determine that what kind of information is highly sensitive and they need protection because many data in electronic health records are worthless without availability of this sensitive information. For example, consider two different data tables that are used for the storing patient’s personal information and the drugs consumed by them, without personal information, consumed drugs information will not have much value for attackers and will not lead to disclosure of private information of patients. In the next step, sensitive data encrypt with DES asymmetric encryption algorithm. The key of this encryption is combination of social security number and patient’s ID in the system. After that, the encrypted data should be stored in the clouds. For this reason the length of encrypted data divided by the number of clouds. Each episode will be stored separately

TABLE 1
ENCRYPTED DATA STORAGE IN THE CLOUD SERVERS

Cloud Service Provider	*Stored data*
CSP 1	*F@*
CSP 2	*P&*
CSP 3	*%h*
CSP 4	*{9*

in single cloud. This will increase the amount of computation and somewhat will reduce the rate of Information storage and retrieval. Thus, we will make sure that if there was an illegal intrusion in one of the clouds, there will be no possibility to retrieve the entire data because each cloud contains part of original data. For instance, suppose we have four cloud service providers in our system, If the original data is "Cloud" and after encryption converted to "F@P&%h{9", this data will be store in the clouds according to Table 1.

Furthermore, for availability of the information any cloud service provider should have at least one backup server in order to avoid disruption in permanent data availability service. When the authorized user sends a query into the multi-cloud system to retrieve data is the next stage. In this case, information have been separately retrieved from cloud databases and integrated with together. After that, using the symmetric encryption key that is the combination of the user's social security number and ID in the system, the retrieved data can be decrypted. Finally, the decrypted data will be sent to the user.

6.3 The Benefits of EHR-MCDB Model

As mentioned in the previous section, in this model the integrity of data is preserved. If a hacker tries to access the users privacy information, unlike the single cloud method in which the intrusion causes to access to all of information, he or she requires to penetrate to all cloud servers. Even in this case only encrypted data will be found and decryption of it requires access to the encryption key. Therefore, in this model the information recovery for unauthorized person would be very complicated or even impossible.

Another advantage is that this model provides privacy by encrypting sensitive data. Therefore, without having these data, other information has no concept and value. On the other hand, Given that any cloud server in this model has a separate backup server, when the main server fails, the risk of failure in store and retrieval of data process will be disappeared. Consequently, this model always has the ability to service.

6.4 Implementation and Evaluations

In this section, we will discuss implementation and evaluation of EHR-MCDB model. The implementation of this model has been done with the C#.NET 3.5 and SQL Server 2008. A version of SQL Server has installed on each cloud server. The structure of database is the same in all cloud servers to store data in similar columns and rows. To evaluate the proposed model 60 thousand records containing 15 data field has been used. Data fields are types of string and integer. In addition, the size of used data was 15 MB. In Figure 3, the time required to store data on a variety of cloud servers is shown.

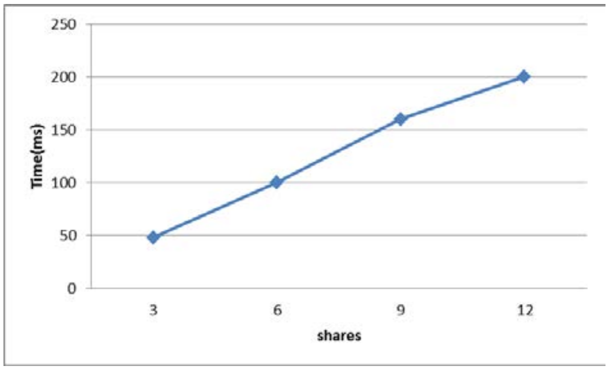

Fig. 3. Comparison of storing data on multi-cloud servers.

We use the same size of data and number of cloud servers in data retrieval evaluation. The results of this evaluation are shown in Figure 4.

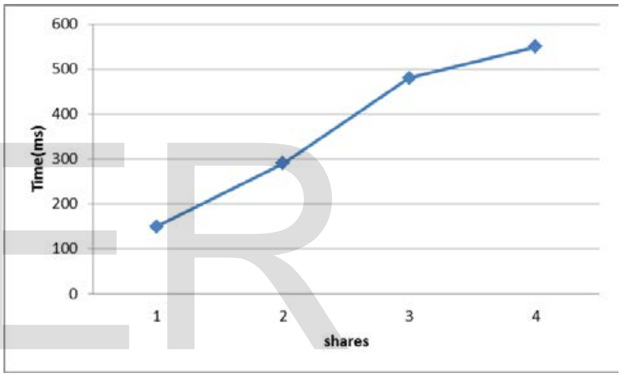

Fig. 4. Comparison of retrieval data on multi-cloud servers.

Evaluation results of proposed EHR-MCDB show that despite increasing time of information storage and retrieval, System overhead is so that it can be ignored ,against the benefits that it provides, and it does not damage to Normal system activity. In Figure 5, the comparison of storing data time of 15 MB data in a single cloud model is shown with EHR-MCDB model.

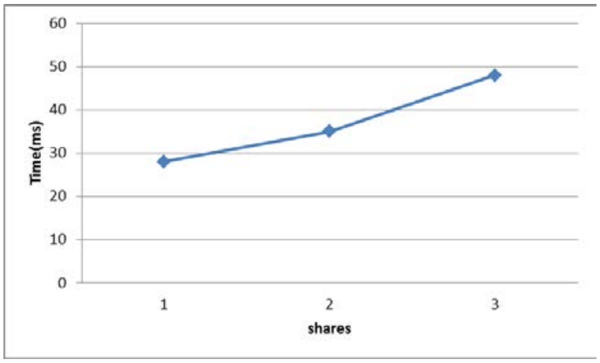

Fig. 5. Comparison of data storing time in a single cloud model with proposed multi- cloud for 2 and 3 servers.

In Figure 6, we compared proposed multi- cloud model with mentioned distribution methods and different distribution method. The results show that EHR-MCDB model has improved about 10%.

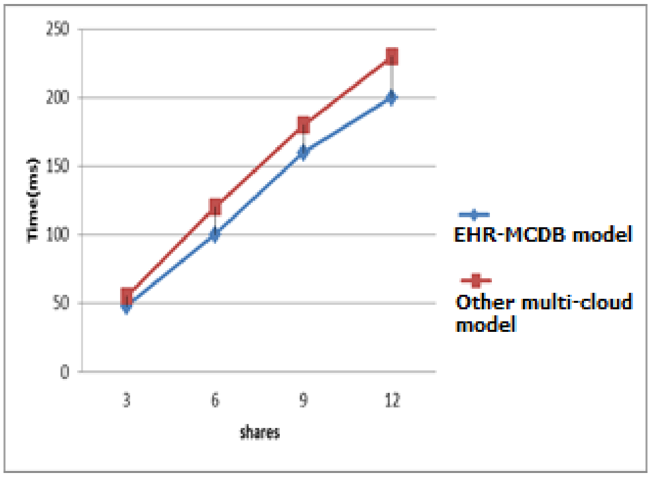

Fig. 6. Compression of storing data time in proposed multi model with similar multi-cloud models.

7 Conclusion

Need to access electronic health information across the world and improve the quality of healthcare to patients, Will highlight importance of using cloud computing architecture in this area. However, despite the benefits of cloud computing applications for health care, the security challenges of cloud should be addressed. In this paper, we introduce and describe a new proposed mode called EHR-MCDB (Electronic Health Record with Multi-Clouds Databases). This model will ensure the privacy of persons in a network and multi-cloud environment using cryptographic algorithm and distributed storage. In the proposed model, we first determine that what kind of information is highly sensitive and they need protection because many data in electronic health records are worthless without availability of this sensitive information. The evaluation results of proposed EHR-MCDB show that despite increasing time of information storage and retrieval, System overhead is so that it can be ignored ,against the benefits that it provides, and it does not damage to Normal system activity. Also comparing the proposed model with data distribution method on different servers shows that it is improved almost about 10% compared to the same multi-cloud model.

References

[1] Zaigham Mahmood,"Data Location and Security Issues in Cloud Computing", proceedings of the 2011 International Conference on Emerging Intelligent Data and Web Technologies, pp. 49-54, Tirana, Albania, September 2011.

[2] Zhuo-Rong Li, En-Chi Chang , Kuo-Hsuan Huang , Feipei Lai ,"A Secure Electronic Medical Record Sharing Mechanism in the Cloud Computing Platform", proceedings of the 2011 IEEE 15th International Symposium on Consumer Electronics, pp. 98-103, Singapore, June 2011.

[3] Alex Mu-Hsing Kuo, "Opportunities and Challenges of Cloud Computing to Improve Health Care Services", Journal of Medical Internet Research, vol. 13, no. 3, March 2011.

[4] Saleem-ullah Lar, Xiaofeng Liao, Syed Ali Abbas, "Cloud Computing Privacy & Security Global Issues, Challenges & Mechanisms", proceedings of the 2011 6th International ICST Conference on Communications and Networking, pp. 1240-1245, Harbin, China, August 2011.

[5] Sebastian Haas, Sven Wohlgemuth, Isao Echizen, "Aspects of privacy for electronic health records", International Journal of medical informatics, vol. 80(2), pp. 26-31, Elsevier, October 2010.

[6] Gihan Perera, Anne Holbrook, Lehana Thabane, Gary Foster, Donald J. Willison," Views on health information sharing and privacy from primary care practices using electronic medical records", international journal of medical informatics, vol. 80, no. 2, pp. 94-101, February 2011.

[7] Miao Zhou, YiMu, WillySusilo, JunYan, LijuDong, "Privacy enhanced data out sourcing in the cloud", Journal of Network and Computer Applications , vol. 35, no. 4, pp. 1367–1373, July 2012.

[8] Hassan Takabi, James B.D. Joshi, Gail-Joon Ahn," Security and Privacy Challenges in Cloud Computing Environments", Copublished by the IEEE Computer and Reliabiliy Societies, vol. 8, no. 6, pp. 24–31, Nov.-Dec. 2010.

[9] Jianfeng Yang, Zhibin Chen," Cloud Computing Research and Security Issues", proceedings of the 2010 International Conference on Computational Intelligence and Software Engineering (CiSE), pp. 1-3, Wuhan, December 2010.

[10] Fu Wen, Li xiang," The Study on Data Security in Cloud Computing based on Virtualization", proceedings of the 2011 International Symposium on IT in Medicine and Education (ITME), pp. 257 - 261, Cuangzhou, January 2012.

[11] Lori M.Kaufman, "Data Security in the World Of Cloud Computing", Copublished by the IEEE computer and Reliabiliy Societies, Vol 7, No 4, pp. 61–64, July-Aug. 2009.

[12] Cong Wang , Kui Ren, Wenjing Lou, Jin Li, "Toward Publicly Auditable Secure Cloud Data Storage Services", Network, IEEE press, Vol 24, No 4, pp. 19–24 , July-August 2010.

[13] Mohammed A. AlZain, Ben Soh, Eric Pardede," MCDB: Using Multi-Clouds to Ensure Security in Cloud Computing", proceedings of the 2011 Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing, pp.784-791, Sydney, Australia, Dec 2011.

[14] Cong Wang, Qian Wang, Kui Ren , Wenjing Lou, "Ensuring Data Storage Security in Cloud Computing", proc. Quality of Service. 17th International Workshop on 2009, Publications ieee, pp.1-9, July 2009.

[15] Hsun Chuang , Syuan-Hao Li , Kuan-Chieh Huang, Yau-Hwang Kuo," An Effective Privacy Protection Scheme for Cloud Computing", proceedings of the 2011 13th International Conference on Advanced Communication Technology (ICACT), pp. 260- 265, Seoul , Feb 2011.

[16] Uma Somani, Kanika Lakhani, Manish Mundra," Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing", proceedings of the 2010 1st International Conference on Parallel, Distributed and Grid Computing (PDGC - 2010), pp. 211-216, Solan, India, Oct 2010.

[17] Jun Feng, Yu Chen, Wei-Shinn Ku, Zhou Su," D-DOG: Securing Sensitive Data in Distributed Storage Space by Data Division and Out-of-order keystream Generation", proceedings of the 2010 IEEE International Conference on Communications (ICC), pp.1-6, September 2010.

[18] Nancy J. King, V.T. Raja," Protecting the privacy and security of sensitive customer data in the cloud", Computer law & security views , pp. 308–319, March 2012.

[19] Shaftab Ahmed, Azween Abdullah," Telemedicine in a cloud – A Review", proceedings of the 2011 IEEE Symposium on Computers & Informatics , pp. 776 - 781, Kuala Lumpur, Malaysia, March 2011.

[20] Weijia Yang, Sanzheng Qiao," A novel anonymization algorithm: Privacy protection and knowledge preservation", Journal of Expert Systems with Applications, Vol 37, No 1, pp. 756–766, January 2010.

[21] Zhe Jin, Andrew Beng Jin Teoh, Thian Song Ong, Connie Tee," Fingerprint template protection with minutiae-based bit-string for security and privacy preserving", Journal of Expert Systems with Applications, Vol 39, No 6, pp. 6157–6167, May 2012.