Palm Vein Pseudonym for Anonymous Database Record

Nur Hafizah Mohd Nazari
Information Assurance & Security Research Group
University Technology Malaysia
Johor Bharu, Johor
nurhafizah.nazari@gmail.com

Assoc. Prof. Dr. Shukor Abdul Razak
Information Assurance & Security Research Group
University Technology Malaysia
Johor Bharu, Johor
shukorar@utm.my

***Abstract*—Every organizations nowadays need to collect and store tremendous amount of data for good purpose. Cloud service and wireless network service are widely used to manage those data. The services can reduce the time consuming to achieved production and ease people to manage their works. Typical services are using unique identifier to store the data in digital database. There are some weaknesses using unique identifier to store the data. The unique identifier is obviously linkable to the owner of the data such as name, Identity card number, address and so on. The adversary may use the unique identifier to steal the data. They may guess or eavesdrop to obtain the data. It can be leads to the data exposure and lack in preserving the data privacy. Therefore, data privacy become the important issue to take care of. This is because, during the transaction of the data and information using the current services, the data and information may be exposed and leaked to unwanted entity. Other than that, services attack also may be occurred such as forgery attack, spoofing attack, and many more during the data transaction. Thus, we proposed biometric authentication system using palm vein is used during the authentication process. To ensure the data in the database are preserved, pseudonym generation approach is used to make the database record is anonymous. Therefore, the adversary or unauthorized entity could not steal the data and information of authorized user.**

Keywords—Authentication, Palm Vein, ROI, Centroid Method, Pseudonym, Anonimous, Data Disintegration, Unlinkability, Data Preserving

INTRODUCTION

Nowadays, every organizations need to store tremendous number of data. The typical database system used must be well organized to ensure the data can be easily managed and protected. The system should assist the needs of the user such as owner of the data, the individual management and the national statistics, approving monitoring data record parameters and assisting administration and management. Other than that, the system need to upgrades the effectiveness database record services. In other words, the data is used to record the summary of personal information histories which can be shared and accessed by a wide range of user by using online method.

However, the current system used by most of organizations still have an issue with preserving the data privacy [1]. Most of the system using only one ID or called unique identifier to represent all attributes of the personal record. Those records are linkable from an organization with other organization. The records can be shared with any other organization for specific purpose. For example, a person one to buy new postpaid number with a telco company. Unfortunately, after checking his status using his ID number he could not buy the new number because his name already blacklisted with another telco company due to bills debt. This means that, all telco company can access the buyer status just by using the unique identifier. A personal records includes various of information. Some of those information are sensitive and confidential. From a privacy point of view, the typical services may lead to data exposure. Therefore, having such a unique identifier for every client is likewise a remarkable security danger: When information is lost or stolen, additionally any foe getting the information can utilize the unique identifier to interface all the distinctive datasets together. Likewise, collaborations of clients with various substances turn out to be effectively traceable [2].

Due to the privacy issue, numerous researches have been done to ensure the privacy of data record is preserved. Existing privacy preserving research can be divided into three categories which are privacy by policy, privacy by statistics, and privacy by cryptography [4]. Another research related to biometric cryptography also become popular in protecting personal data. However, these approaches did not compromise data anonymity [3]. Therefore, a robust biometric modalities and preserving data privacy method are needed to protect the confidentiality of the data.

This research is focused on pseudonym generation from palm vein image base on pseudonym system to ensure the data anonymity and to preserve the privacy of the data. Palm vein images are captured using palm vein scanner.

RESEARCH BACKGROUND

Preserving privacy of data is important to ensure the security and confidentiality of the information in the data. In the data records, there are a lot of valuable information. Researchers have been proposed numerous approaches in preserving data privacy. However, the size of the data are greatly increasing year by year and the data can be shared to

978-1-5386-6925-9/18/$31.00 ©2018 IEEE

the third parties for certain purpose. Furthermore, during the data sharing, the information of the data may disclose to irresponsible entities through adversary attacks. Therefore, most favorable approaches are needed to protect the data privacy.

Pseudonymous identification system is one of the approach to ensure the anonymity of the data and to preserve the privacy of the data. Pseudonymous identification system enhances the controllability of the information trades and furthermore abstains from forcing a unique identifier that makes the client traceable by default. This approach can prevent an adversary to recognize the real identity of an authorized user even though already succeeded to access the database. This is because, the adversary could not relate to whom the data is belongs to.

In a typical database record, one unique identifier such as user ID can assimilate the entire record. Therefore, it leads to disclosing the database information. Thus, a data disintegration mechanism able to avoid the integration of the patient's information. The determination of data disintegration is to preserve personal information from information leakage. In work [7], they mentioned that data integration mechanism is required to ensure the privacy preserving of healthcare information system.

Study in [7] performed the partial pseudonym and pairing pseudonym for their mechanism in preserving the privacy of the medical data. They found that the partial pseudonym can improves the privacy preserving to the higher level. However, they stated that, it is necessary to continuously upgrades the data preservation. A mischievous user can be whoever even internal staff or family members.

In [5] studies, they have listed three existing approaches in preserving data privacy. The approaches are privacy preserving aggregation, de-identification and operations over encrypted data. However, they found that those approaches are infeasible and lack of anonymity of the data. But, they found that de-identification is the best approach in privacy preserving if an efficient and privacy-preserving algorithm can be developed.

Study in [6], has categorized two methods used in preserving the data privacy into cryptographic methods and non-cryptographic methods. However, the technology is arising greatly and an adversary always have its own advanced technique to collect the information illegally. Hence, they suggested that developing strong algorithm is needed for the technique used.

Based on previous researches, the approaches used are used to hide the identity of the patient. They want to ensure the owner of the data is unknown to malicious user. Therefore, they used the patient identity such as their name or their identity card number to transform it into unknown character or number such as pseudo number. However, identity of patient may be easy to obtain from adversary attack.

To overcome the weakness of previous research, study on biometric system acquiring more attention in protecting the data from any threat [8]. For verification and identification of an individuals, biometric system is used. Biometric traits of an individual are used in this system such as the physical features (palm print, fingerprint, voice, iris, palm vein, hand geometry, etc.) or behavioral features (signature, handwriting, keyboard typing, etc.). Since every single human have unique traits, it can be transform into a valuable data by capturing the traits using specific sensors and devices [9].

Based on previous study on biometric, researchers found that palm vein is the most secure and reliable for person identity verification and enable to protect the data privacy. This is because, palm vein cannot be easily damaged, changed or falsified [10]. Palm vein also has uniqueness and complexity of vein patterns for different each of a person [11].

The objective of this paper is to generate the pseudonym using palm vein for anonymous data.

PROPOSED SOLUTION

This paper focused on generating pseudonym in preserving the data privacy. Biometric trait, palm vein is used to generate the pseudonym. Figure 1 briefly shows the flow to generate pseudonym from palm vein image. Palm vein scanner is used to collect the palm vein sample. To find out the best region used to generate the pseudonym, Centroid Method is used. Every data of a patient may have more than one pseudonym. The pseudonym will be stored in the database. This method will make malicious user hard to predict the pseudonym belong to which patient and which data.

To ensure the proposed method will preserve the data privacy from adversary attack, two partial key will be generated from the palm vein ROI. After that, final pseudonym will be generate from the partial keys.

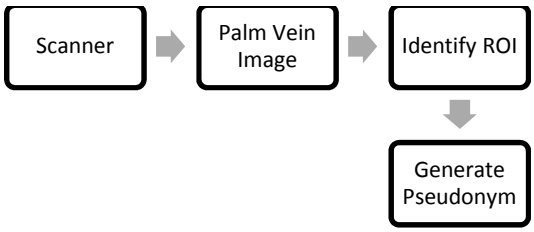

Fig.1. Flow to Generate Pseudonym

Collecting Data

The palm vein images are collected using palm vein scanner from Fujitsu. One of the scanner feature is contacless. Therefore, it is more hygiene an safe. Figure 2 shows an example of palm vain image captured using the scanner.

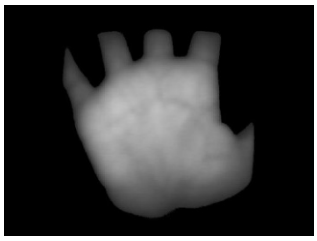

Fig.2. Palm Vein Image

Centroid Method for Palm Vein ROI

The robustness of ROI extraction based on centroid method for palm vain have been proved by [12]. The accuracy obtain is 99 %. This method will be used in this research to find out the best region can be used as an input to generate pseudonym. Figure 3 show the flow of ROI extraction of palm vein.

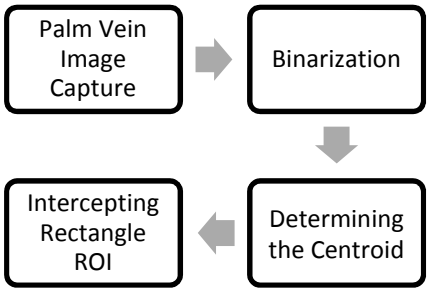

Fig.3. Flow of of ROI extraction of palm vein using centroid method [12]

The four specific step of the ROI extraction of the palm vein as follow:

1) Palm vein image is captured using palm vein scanner. The image captured already in greyscale. Figure 4 shows the image of palm vein.

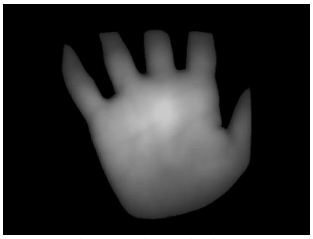

Fig.4. Palm vein image

2) During binarization processing, the ordinary gray image is used. The palm region and the background of the image are separated.

3) Centroid algorithm is used to extract the image centroid. A intercepted rectangle will be mark as center of the origin, thus the centroid image is captured as shown in figure 5.

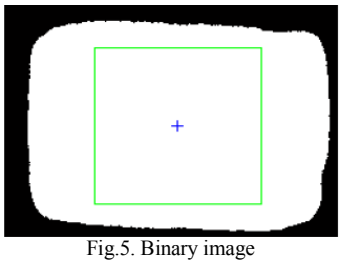

Fig.5. Binary image

4) To obtained the ROI image of the palm vein, the 256x256 sub image is cut out in gray image as shown in figure 6.

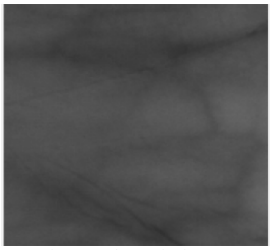

Fig.6. Palm vein image

Pseudonym Generation

Research by [15] already proposed the pseudonym system. Digital pseudonym is used where a public key is used to authenticate the signature of an anonymous sender who has signed the message with his or her private key. However, the method is not good enough to evading attacker from eavesdrop and link messages sent with the similar pseudonym. Moreover, the unlinkable feature is not good enough in preserving the privacy of the user data. Therefore, improvement in using the pseudonyms is made.

The idea of generating the pseudonym is adapted from [7]. Figure 3 shows the process of collecting the pseudonym.

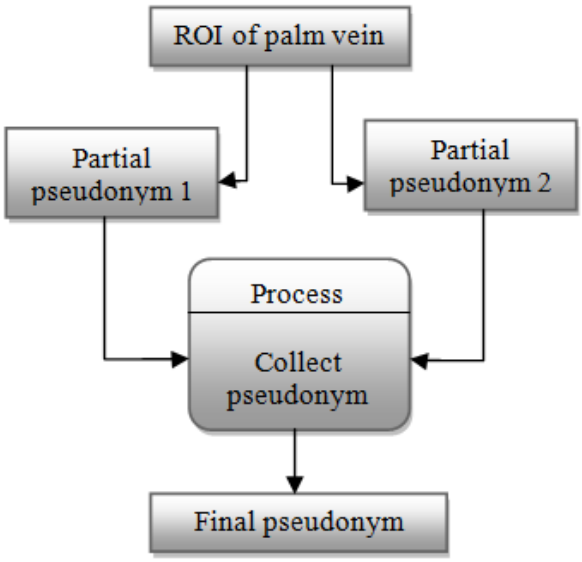

Fig.3. Pseudonym process

To collect the pseudonym, two partial pseudonym will be generated from the ROI of palm vein. After that, both of the partial pseudonym will be combined to produced final pseudonym. Collecting of partial pseudonym will be increase the level of security [7].

To ensure the pseudonym is unlinkable, during the process of generating partial pseudonyms and collection of final pseudonym must be randomized [14]. During the phase of generating the partial pseudonyms, random secret key will be generated.

A pseudonym is generating by implement the keyed-hash message authentication code (HMAC) to secure the user's information that already registered in the system.

Follows are the step to generate pseudonym:

1. Let begins with a real secrete id (ID). This secret provided by the server after user register in application using palm vein.
2. HMAC is generates using the ID. If this secret is disclosed, then the security of real name is compromised.
3. A cryptographic HMAC is applied N times to ID, thereby producing a hash chain of pseudonym. The pseudonym (PID) are the result of the application of the cryptographic hash function.

Requirement for the Pseudonym System

To ensure the pseudonym system can be generated successfully, there are a few requirements need to be fulfilled. The following are the requirements needed.

1. **Each authenticated pseudonym corresponds to unique user:**

   An authorized Turing machine is called identity extractor is needed. For any usable nym, the Turing machine can have reversible access thus permitting the user to be authenticated as the owner of the nym with non-negligible prospect.

2. **Security of the user's master secret key:**

   The system wants to ensure that user's master secret key is not exposed by his public key or by the user's interaction with the pseudonym system. Hence, anything that can be computed about the user's secret key as a result of the user's interaction with the system, can be computed from his public key alone is needed.

3. **Credential sharing implies master secret sharing:**

   A user with valid credentials might help another user to inappropriately obtain any rights the credential brings. The user may do so by revealing the master key to another user so that the another user can personate in all regards. This type of attack cannot be prevent but a scheme is require where whenever a user disclose some information that allows another user to use the credential or nyms, the user is efficiently disclosing the master secret key to another user. its means that, there exists an extractor such that if a user succeed to the user's pseudonym, then the secret key of another user who does have a valid credential can be extracted by having rewindable access to the user.

4. **Unlinkability of pseudonym:**

   The nyms of a user must not be linkable at any time better than by random guessing.

5. **Unforgeability of credential:**

   A credential may not be distributed to a user without the organization's collaboration.

6. **Pseudonym as a public key for signature and encryption:**

   The pseudonym system is able to sign with one's nym, as well as encrypt and decrypt messages.

DISCUSSION

The propose method is robust to protect the privacy of the data from any system attacks such as, forgery attacks, offline password guessing, insider attacks, stolen-verifier attacks and spoofing attacks. This propose method can overcome the problem in preserving the data privacy such as information leakage. Using palm vein as biometric traits is more reliable than other traits because of its unique characteristics. Palm vein pattern is hard to damaged, changed or falsified. Therefore, it is secure and reliable for person identity verification. Table 1 shows the different biometric modality with the accuracy level of authentication.

ACCURACY LEVEL FOR DIFFERENT BIOMETRIC MODALITY

Biometric Modality	Table Column Head
Fingerprint	99%
Palmprint	>95%
Hand Geometry	>95%
Vein Pattern	99%
Face	95%
Ear	>95%

The propose ROI extraction method which is based on Centroid method is capable to obtained high accuracy of authentication which is 99%. From the ROI extraction result, pseudonym will be generated. The propose pseudonym generation may increase the privacy of the data from any threat. The unlinkable feature is important to ensure another user hard to predict to who the data belongs to. Therefore, partial pseudonyms are generated before go to the final pseudonym. Those pseudonym will be different for every data a user have.

CONCLUSION

This paper focus on determining the reliable region of palm vein can be used in the authentication system. To ensure the privacy of the data is preserved, new pseudonym generation technique is proposed. These propose solution also may increase the security of the data from various system attacks.

ACKNOWLEDGEMENT

This research is sponsored by Ministry of Education Malaysia under TRGS grant number (4L844).

REFERENCES

[1] Shrestha, N. M., Alsadoon, A., Prasad, P. W. C., Hourany, L., & Elchouemi, A. (2016, April). Enhanced e-health framework for security and privacy in healthcare system. In Digital Information Processing and Communications (ICDIPC), 2016 Sixth International Conference on (pp. 75-79). IEEE.

[2] Camenisch, J., & Lehmann, A. (2015, October). (Un) linkable Pseudonyms for Governmental Databases. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (pp. 1467-1479). ACM.

[3] Gkoulalas-Divanis, A., Loukides, G., & Sun, J. (2014). Publishing data from electronic health records while preserving privacy: A survey of algorithms. Journal of biomedical informatics, 50, 4-19.

[4] Yang, J. J., Li, J. Q., & Niu, Y. (2015). A hybrid solution for privacy preserving medical data sharing in the cloud environment. Future Generation Computer Systems, 43, 74-86.

[5] Lu, R., Zhu, H., Liu, X., Liu, J. K., & Shao, J. (2014). Toward efficient and privacy-preserving computing in big data era. IEEE Network, 28(4), 46-50.

[6] 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.

[7] Samsi, M. S., & Razak, S. A. (2014, August). A mechanism for privacy preserving in healthcare organizations. In Biometrics and Security Technologies (ISBAST), 2014 International Symposium on (pp. 208-213). IEEE.

[8] Balakumar, P., & Venkatesan, R. (2012). A Survey on Biometrics based Cryptographic Key Generation Schemes. International Journal of Computer Science and Information Technology & Security, 2(1), 80-85.

[9] Verma, I., & Jain, S. K. (2015, March). Biometrics security system: A review of multimodal biometrics based techniques for generating crypto-key. In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on (pp. 1189-1192). IEEE.

[10] Wu, K. S., Lee, J. C., Lo, T. M., Chang, K. C., & Chang, C. P. (2013). A secure palm vein recognition system. Journal of Systems and Software, 86(11), 2870-2876.

[11] Raut, S. D., & Humbe, V. T. (2014). Review of biometrics: palm vein recognition system. IBMRD's Journal of Management & Research, 3(1), 217-223.

[12] [11] Lin, S., Xu, T., & Yin, X. (2016, October). Region of interest extraction for palmprint and palm vein recognition. In Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), International Congress on (pp. 538-542). IEEE.

[13] Lysyanskaya, A., Rivest, R. L., Sahai, A., & Wolf, S. (1999, August). Pseudonym systems. In Selected Areas in Cryptography (Vol. 1758, pp. 184-199).

[14] Camenisch, J., & Lehmann, A. (2017, April). Privacy-Preserving User-Auditable Pseudonym Systems. In Security and Privacy (EuroS&P), 2017 IEEE European Symposium on (pp. 269-284). IEEE.

[15] Chaum, D. (1985). Security without identification: Transaction systems to make big brother obsolete. Communications of the ACM, 28(10), 1030-1044.