KURSOR Journal
Research on Computing and its Aplications

Vol. 7, No. 1, January 2013 **ISSN 0216 – 0544**

SHARED SECRET-BASED KEY AND FINGERPRINT BINDING SCHEME

Tohari Ahmad
Department of Informatics, Sepuluh Nopember Institute of Technology
Gedung Teknik Informatika Lt. 2, Kampus ITS Sukolilo
Jln Raya ITS, Sukolilo, Surabaya, 60111
E-Mail:[a]tohari@if.its.ac.id

Abstrak

Salah satu permasalahan utama di dalam melakukan otentikasi pengguna dalam suatu sistem adalah keseimbangan antara akurasi dan kenyamanan. Akurasi berhubungan dengan kemampuan sistem untuk mengenali pengguna yang benar-benar diijinkan atau tidak diijinkan; sedangkan kenyamanan berhubungan dengan penerimaan pengguna terhadap keberadaan sistem. Kedua faktor tersebut merupakan faktor yang penting karena keduanya menentukan apakah sistem otentikasi tersebut dapat diaplikasikan atau tidak. Hal ini berdasarkan asumsi bahwa metode otentikasi yang sangat akurat tetapi sulit dilakukan, akan kurang dapat diterima oleh pengguna. Pada makalah ini, kami mengkombinasikan kemudahan penggunaan sidik jari dan akurasi kunci/kata sandi, sehingga pengguna dapat memanfaatkan kunci/kata sandi yang kuat tanpa harus menghafalnya, dengan menggunakan skema shared secret. Di dalam metode yang diusulkan ini, suatu kunci yang panjang dan random disebarkan ke titik-titik minutiae pada sidik jari, dan selanjutnya dibagi kepada masing-masing deskriptor, yang dalam hal ini dinyatakan dalam suatu vektor. Kunci tersebut hanya bisa direkonstruksi jika terdapat sejumlah deskriptor dan titik minutiae yang saling tumpang tindih. Percobaan yang dilakukan dengan menggunakan basis data publik (FVC2002DB2a) menunjukkan bahwa pendekatan yang diusulkan mempunyai performa yang bagus.

Kata kunci: Sidik Jari, Keamanan Data, Otentikasi, Biometrik, Kerahasiaan Data.

Abstract

One of critical issues in authenticating users to a system is thebalance between accuracy and convenience. The former relates to thecapability of the system to recognize authorized or unauthorizedusers; while the later relates to the user acceptance to the system.Both are important factors since they determine whether such authentication system is applicable or not. This is based on the assumption that an accurate but hard to use authentication method hasless user acceptance. In this research, we combine the convenience of the fingerprint and the accuracy of the password such that the users have a strong password without having to memorize it, by using theshared secret scheme. In this proposed scheme, a long and random password is shared among the fingerprint minutiae and further, shared among its descriptors (in this case, they are represented by vectors). The password can only be reconstructed if there is a substantial number of overlapping descriptors and minutiae points. The experiment which is conducted on the public database, FVC2002DB2a, shows that the proposed approach has a good performance.

Key words: fingerprint, data security, authentication, biometrics, data confidentiality.

INTRODUCTION

Password has been a popular tool for authenticating users. Its performance is relatively high, in terms of simplicity and accuracy. However, it can be the weakest point in the system [1]. This particularly happens if users create a ”bad” password which actually does not comply with the standard of the computer security policy. For example, a password/key must be random whose lenght is longer than 8 bytes and contains alphanumeric symbols.

Research in [2] finds that the word “password” has been a popular key used in the authentication system. This has made it easy for an adversary to compromise that key or even the whole system. There are some factors why users choose such simple key. For instance, users have a difficulty in memorizing a long and random key. So, there must be a mechanism which helps users to have a good key.

One of possible solution is to use biometrics which provides a relatively unique feature. The main advantage of biometrics is that the users do not have to memorize its pattern. This is because biometrics is part of their physical traits. Furthermore, biometrics also provides non-repudiation property because it is relatively difficult to be shared to others.

Among known biometric modalities, fingerprint is common to use. As depicted in [2], fingerprint has a good performance, particularly in terms of permanence and distinctiveness. The former relates to the stability of the fingerprint pattern; while the later relates to the uniqueness of the fingerprint pattern. This means that fingerprint is a potential candidate to use in the authentication process, even though it has also disadvantages. Furthermore, fingerprint has a relatively high user acceptance level [2]. Users are convenient to use it because it is simple and easy to use. Therefore, in this paper we employ fingerprint as the representation of biometrics.

This paper proposes a scheme that combines convenience of the fingerprint with security of the strong key. Particularly, the key will be released if only the fingerprint query has a certain minimum number of minutiae points overlapping with those of the template. Once it has been released, the key can be used for any purpose, such as authenticating the user, decrypting the file, etc. This raises a challenge as it requires the fingerprint to produce exactly the same features while by the nature, the fingerprint feature itself is very likely to change from capture to capture due to the intra-class variability factor.

This paper will more focus on the binding the key to the fingerprint instead of the fingerprint templates protection. This paper is structured as follows. The next two sections explain the shared secret concept and the triangle-based feature extraction. These followed by the description of the proposed key binding design. The experiment and its results are provided in the next section, while the conclusion is drawn in the last section.

SHARE SECRET

In [3], Shamir proposes a method of how to share (divide) data between users such that it can only be reconstructed if there is cooperation between *some* of them, called a (*k*, *n*) *threshold scheme*, where *k* is the minimum number of users to reconstruct the data, and *n* is the total number of users. This has given an advantage as there is a balance between security and convenience, considering that $k = n$ results in secure but inconvenient case while relatively small *k* leads to insecure but convenient one. Let the data D_0 is shared among *n* users. This will generate sub-data D_1, D_2, ..., D_n. For any (*k*, *n*) threshold scheme, it needs to randomly generate $(k - 1)$ degree polynomial as specified in Equation (1), where $a_0,...,a_{k-1}$ are the coefficients.

$$\left.\begin{aligned} & f(x) = a_0 + a_1 x + ... + a_{k-1} x^{k-1} \\ & D_0 = a_o \\ & D_1 = f(1),...,D_y = f(y),...,D_n = f(n) \\ & 1 < y < n \mid y \in Z^+ \end{aligned}\right\} \quad (1)$$

It is also proposed in [3] to use modular arithmetic, by choosing a prime number *z*, where $z > D_0$ and $z > n$. The coefficients of the polynomial, a_1, ..., a_{k-1} which are integer values, are taken from (0, *z*). Note that a uniform distribution is used for choosing those coefficient values. In addition, the modulo *z* operation is also applied to D_y.

In the reconstruction stage, for any *k* pairs of (y,D_y) the coefficients of $f(x)$ can be found by

interpolation while D_0 is derived from $a_0 = f(0)$. The reconstruction of D_0 itself relies on the number of sub-data D_y being held, such that:

- If only $(k-1)$ or less of D_y are known, then reconstruction of D_0 will fail.
- If k or more of D_y are known, then reconstruction of D_0 will pass.

TRIANGLE-BASED FEATURE EXTRACTION

In [4], Germain, Califano and Colville propose a fingerprint indexing approach by using triplets of minutiae. Each triplet, as depicted in Figure 1, generates nine components, which consist of:

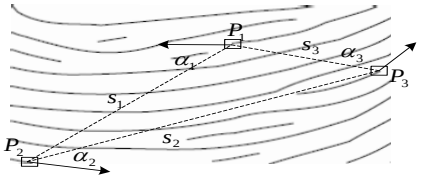

Figure 10. Minutiae Triplets of [4].

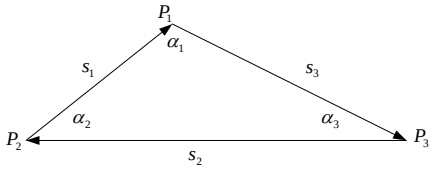

Figure 11. Minutiae Triplets of [5].

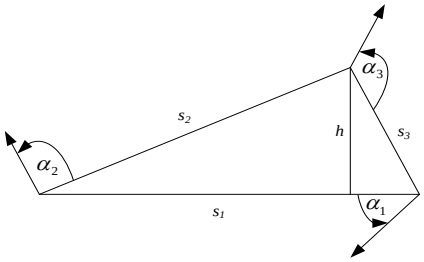

Figure 12. Minutiae Triplets of [6].

Figure 13. Minutiae Triplets of [7].

- The length of edges (S_1, S_2, S_3)
- The number of ridges crossed the edge between minutiae points (in Figure 1: $P_{1\_2} = 5$, $P_{2\_3} = 3$, $P_{3\_1} = 3$)
- The angle of the minutia point based on the fiducial side (α_1, α_2, α_3).

As only the relationship between minutiae points is extracted, this approach generates features which are invariant to rotation and translation. Yet, there exist some issues with this approach that leads to the performance degradation, which include [5-6]:

- The change of the length is significant to distortions.
- Stable angles and ridge counts much rely on the quality of the image.

Those factors require a relatively large bin for quantizing the invariance. On the other hands, this results to the increase of the intra-user variability level.

Still using the triplet-based approach, Bhanu and Tan [5] claim that theirs is more robust. This is based on the assumption that length of edges is invariant to translation and rotation, which is relative to scale, while angles are determined by the ratio of length. They argue that by using their transformation, angles have become invariant. Example features used in [5] are illustrated in Figure 2, which consist of:

- The maximum length of edges (S_2).
- The minimum and median angles based on the minutiae triplet (α_3, α_2).
- Triangle handedness, type and direction which are determined by the sign function and cross product of complex number differences according to the minutiae location; the type of minutiae which constructs the triangle; and the direction of each minutia within triangle.

It is found that the proposed model in [5] eliminates the minutiae orientation information. This is because their angels definition is based on the ones within the triangle only without involving the minutiae orientation. Also, according to [6], it much relies on the reference points selection, as the result of the use of relative translation and orientation features which also makes the quantization step more complex.

Farooq et al. [6] take seven invariants from a triangle, as illustrated in Figure 3 which is inspired by that of [4]. Three invariants are the angle between the edges and the respective

minutiae orientation, another three are the length of edges and the last feature is the 'height' which is the distance between a point to the largest edge. Suppose *s*, *a* and *h* represent the number of bits of edges, angles and the height, respectively. Each triangle has $3s + 3a + h$ bits. It is claimed that these invariants are more stable and easy to extract from the standard fingerprint representation.

Recently, Jin et al. [7] extract another feature set from the triangle by employing the number of minutiae within it. In particular, three vertexes are randomly generated on the fingerprint image. These are to be the location where the triangle is formed. Thus, it is not required for the vertexes to be same as the minutiae points or even the vertexes can be beyond the fingerprint boundary, as shown in Figure 4. In this case, the number minutiae points within the triangle is to be the invariant. Yet, this results in a relatively high error rate due to the reliability issue of the extracted feature.

KEY BINDING

Motivated by research in [3-9], we propose to bind a key to the fingerprint. For this purpose, the invariants are extracted from the fingerprint based on the minutiae triplets to have stable features. Each feature is associated with a sub-key which is used for reconstructing the key. The general process of this approach is shown in Figure 5.

Feature Extraction

Similar to [10-11], a set of minutiae points *BS* is chosen among those of extracted from the fingerprint, such that there are *p* points in it, where *p* is not necessary same between fingerprints. The invariants extraction is performed by selecting a point *i* in *BS* to be the reference point. Among $(p - 1)$ neighboring points, only *m* of them whose distance to the reference point is the smallest, are considered in the invariant extraction, where *m* is obtained from the experiment. Each of those *m* neighboring points are permuted to construct triangles, such that there are *q* possibilities of permutation *P* as depicted in Equation (2).

$$q = {}^{m}P_{3} = \frac{m!}{(m-3)!} \qquad (2)$$

In the proposed approach, the constructed triangle has to cover the reference point. The vertexes must be "around" the reference point to make it having a good performance, in particular, minimizing the inter-user similarity. Once this requirement is held, the invariants are extracted from the triangle *j*, as depicted in Figure 6, which consist of:

- Minutiae type (*t*, either bifurcation or ridge ending).
- Length of the edge (*s*).
- The difference of orientation between a minutiae pair (α).

This will generate a vector $v_{i\_j} = (t_1, t_2, t_3, s_1, s_2, s_3, \alpha_1, \alpha_2, \alpha_3)_{i\_j}$. The process is iteratively repeated such that all minutiae points in *BS* will have become the reference point.

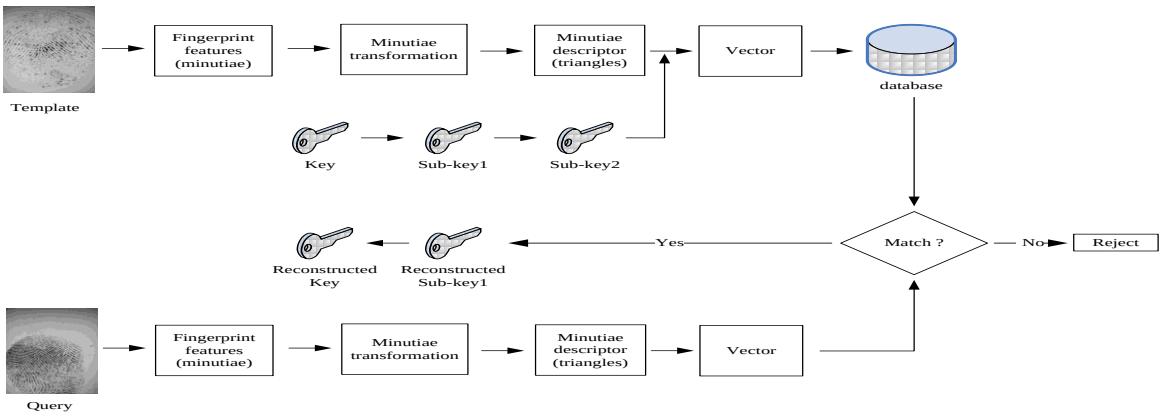

Figure 5. The General Architecture of The Proposed Approach

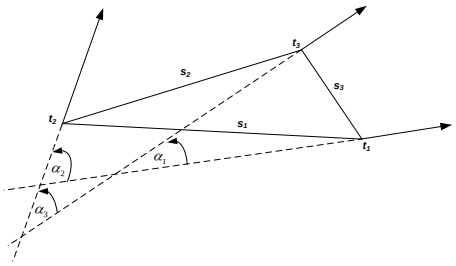

Figure 6. Feature Extraction of the Proposed Approach.

It is worth to note that not all of the reference points have the vector *v*, depending on its relative position to the triangles, as formulated in Equation (3). Also, the triangles which meet the requirement may have the same vertexes. In this case, such triangles are identical except the sequence of the vertexes in the vector, according to the permutation index. Thus, for all minutiae points in a fingerprint, the Equation (3) should be held.

$$\forall m_i \mid v_{i\_j} = \begin{cases} (t_1, t_2, t_3, s_1, s_2, s_3, \alpha_1, \alpha_2, \alpha_3,)_{i\_j}; i \in [1, p]; j \in [1, q]; i, j \in Z^+ \text{ if the triangle exists} \\ 0 \qquad \text{if no triangle exists} \end{cases} \quad (3)$$

Key Sharing

In the separate process, the system generates a random string which will be the key to associate with the finger. By using the method in [3], the key information is shared among minutiae points in the fingerprint. Here, only the minutiae with triangles description are given the sub-key. This is to be the sub-key level 1. In the next layer, the sub-key level 1 is shared among the triangles, by considering the permutation of the vertexes. It is to be the sub-key level 2.

Reconstructing sub-key level 1 requires a certain number of sub-keys level 2, and in turn, reconstructing the fingerprint key requires a certain number of sub-keys level 1. This sturcture can be viewed as the tree of keys, showing in Figure 7. At the verification (matching) stage, the fingerprint query is processed by similar procedure as that of the template to get a set of vectors v'_{a_b}. Matching is performed by comparing each v'_{a_b} of the query with v_{i_j} of the template. The fingerprint key can only be reconstructed if all of the matching requirements are met by the query.

The respective vertexes (minutiae points) of template and query triangles must have the same type (e.g., ridge ending or bifurcation). There are two resulting possibilities of this comparison. First, the vectors do not match. In this case, the comparison is skipped and proceed to the next vector. Second, the vectors match. If it is, then *s* and α differences between template and query must not exceed the threshold γ_s and γ_α as represented in Equation (4).

$$\left.\begin{array}{l} (t_{i\_j})_r = (t'_{a\_b})_r \\ |(s_{i\_j})_r - (s'_{a\_b})_r| \leq \gamma_s \\ |(\alpha_{i\_j})_r - (\alpha'_{a\_b})_r| \leq \gamma_\alpha \\ i, a \in [1, p]\, j, b \in [1, q]; r \in [1,3]; i, j, a, b, r \in Z^+ \\ i \text{ and } j \text{ are not necessary to be same as } a \text{ and } b, \text{ respectively} \end{array}\right\} \quad (4)$$

Due to the many-to-many vector comparison, this may result to duplicate matched vectors. On the other hand, the relationship between template and query vectors must be injective but is not necessary to be bijective. In case there are duplicate matched vectors, the elimination step of such vectors is conducted by considering the difference factor σ. Suppose wgh_s and wgh_α are weight factors of *s* and α differences, respectively, only that with the least difference factor σ will be selected, as defined in Equation (5).

$$\begin{aligned} \sigma = wgh_s * (|(s_{i\_j})_1 - (s'_{a\_b})_1| + |(s_{i\_j})_2 - (s'_{a\_b})_2| + |(s_{i\_j})_3 - (s'_{a\_b})_3|) + \\ wgh_\alpha * (|(\alpha_{i\_j})_1 - (\alpha'_{a\_b})_1| + |(\alpha_{i\_j})_2 - (\alpha'_{a\_b})_2| + |(\alpha_{i\_j})_3 - (\alpha'_{a\_b})_3|) \end{aligned} \quad (5)$$

In case the number of matched vectors is not less than the specified threshold τ_2, the sub-key 1 can be reconstructed. In turn, if there are at least τ_1 sub-keys 1 have been successfully reconstructed, then the full key will be recovered. Note that in case there are duplicate matched center (reference) points, as the result of the many-to-many comparison, the number of matched vectors between (v_{i_j}) and (v'_{a_b}), and average of difference values (σ) are considered, equivalent to that in [10].

RESULT AND DISCUSSION

The experiment is conducted in a public database FVC2002DB2a [12]. Among the available fingerprint impression sets, we use the first as a template and the second as a query, similar to that of [8, 10, 13-14]. The fingerprint minutiae is obtained by using the Verifinger software [15] and the shared-secret implementation is based on [16].

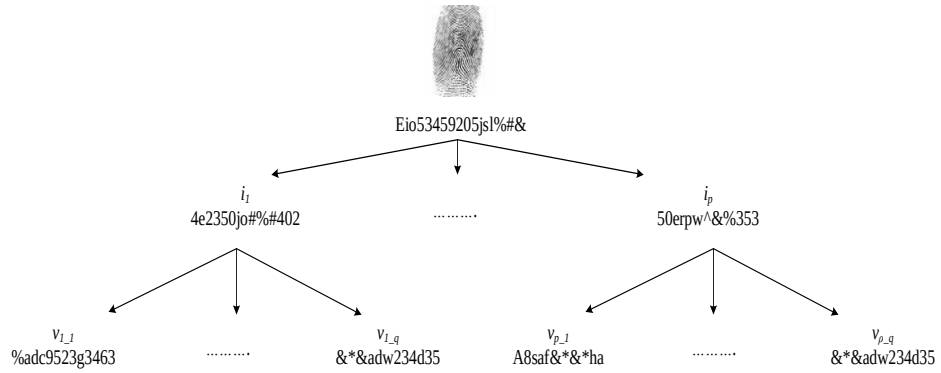

Figure 7. Key Sharing Among Minutiae Points and Vectors.

The performance of this proposed method is evaluated by measuring the number of successful key reconstruction in the genuine and imposter testings which are quantified into GAR (Genuine Acceptance Rate), FRR (False Rejection Rate), GRR (Genuine Rejection Rate) and FAR (False Acceptance Rate) values. The legitimate user testing itself is done by matching every fingerprint in the query set with its relating fingerprint in the template set; the illegitimate user testing is carried out by matching every fingerprint in the query set with those in the template set other than its relating fingerprint.

Suppose $\delta(c,c')$ is the closeness value between a template fingerprint c and a query fingerprint c', similar to that in [10-11]. A fingerprint pair is authentic if only their similarity is at least same as the specified threshold Φ, such that $\delta(c,c') \geq \Phi$.

In the experiment, the parameter m is empirically set to seven. Therefore, each minutiae point is described by at most 210 triangles (note, the order of the vertexes is considered) and the minimum number of matched vectors (τ_2) is set to two.

Figure 8 shows the ROC curve when γ_s is fixed to 12. The highest GAR can be achieved by γ_α=15 is 94% when its FAR is 3.29%. This maximum GAR is lower than that of others, which is 95%. At this GAR level, γ_α =16 has the lowest FAR, which is about 4.25%. Thus, γ_s = 16 can be the reference.

The performance of the proposed approach for various γ_s when γ_α=16 is depicted in Figure 9. Similar trend to Figure 8, small γ_s delivers better performance. Nevertheless, GAR of γ_s=10 is only up to 93%. γ_s = 11 can achieve 94% while γ_s=12, 13 and 14 can have 95%. At this GAR level, γ_s=12 shows the lowest FAR which is 4.25%. This FAR is inconsiderably higher than that of γ_s=11 or even 10.

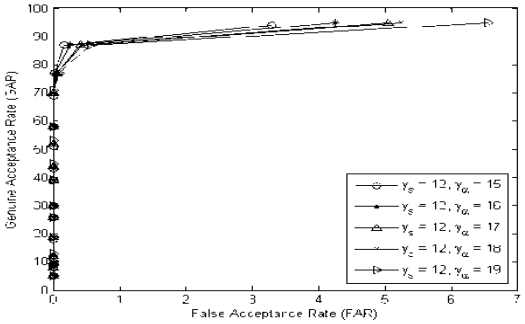

Figure 8. ROC Curve when γ_s is Fixed and $1 \leq \tau_1 \leq 15$.

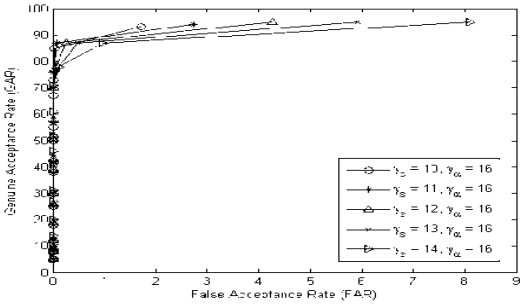

Figure 9. ROC Curve when γ_α is fixed and $1 \leq \tau_1 \leq 15$.

In terms of the minimum number of minutiae needed to reconstruct the key, smaller has better performance as represented in Figure 10. It describes that τ_1=1 has better performance than τ_1=2 or 3, where GAR=95% can be reached by FAR=4.25%. When τ_1 = 2, the best performance is reached at GAR=90% and FAR=3.14% while τ_1=3 is at GAR=83% and FAR=2.27%. Likewise, Figure 11 shows

an equivalent pattern. τ_1=1 can achieve 95% of GAR when FAR is about 4.2%. The least FAR can be held is about 3.2% when its GAR is 94%. On the other hands, τ_1=2 has stable GAR at 87% while its FAR is very low, which is less than 1%. τ_1=3 has also very low FAR, however, its GAR is lower than that of τ_2, which is 77%. Therefore, τ_1=2 is more appropriate to use in case the security is the main concern, moreover, its performance is still relatively high.

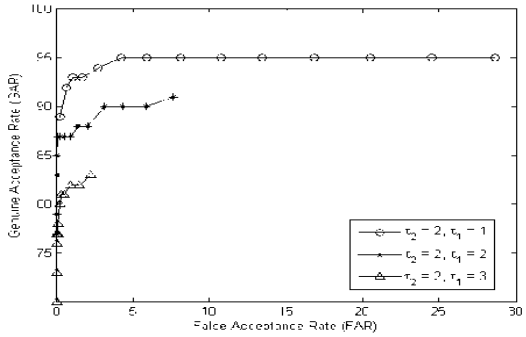

Figure 10. ROC Curve when γ_α is Fixed and 7 $\leq \gamma_s \leq$15

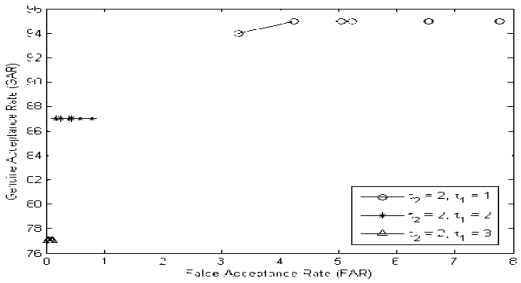

Figure 14. ROC curve when γ_s is fixed and 15 $\leq \gamma_\alpha \leq$20.

CONCLUSION

This paper has proposed a method which combines the key with the fingerprint by using the shared secret scheme whose features are extracted from triplets (triangles). It uses two layers of the key reconstruction, namely: triangle (vector) and minutiae point layers. So, there are two steps should be met in order to reconstruct the key. This proposed method is tested in a public database whose result, in general, depicts that it has a relatively low error value. The actual number of the minimum points to reconstruct the key depends on the environment, whether security, convenience or the between is preferred.

REFERENCES

[1] U. Uludag, S. Pankanti, and A. K. Jain, "Fuzzy Vault for Fingerprints," in *Proceedings of Audio- and Video-based Biometric Person Authentication*, pp. 310-319, Rye Brook, NY, July 2005.

[2] A.K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security," *IEEE Transactions on Information Forensics and Security*, vol. 1 no. 2, pp. 125 – 143, 2006.

[3] A. Shamir "How to share a secret". *Communication of the ACM*, vol. 22 no. 11, pp. 612 – 613, 1979.

[4] R.S. Germain, A. Califano, and S. Colville, "Fingerprint matching using transformation parameter clustering," *IEEE Computational Science & Engineering*, vol. 4 no. 4, pp. 42–49, 1997.

[5] B. Bhanu and X. Tan, "Fingerprint indexing based on novel features of minutiae triplets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 616 – 622, 2003.

[6] F. Farooq, R.M. Bolle, J. Tsai-Yang, and N. Ratha, "Anonymous and revocable fingerprint recognition", In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, Minneapolis, 2007.

[7] Z. Jin, A.B.J. Teoh, T.S. Ong, and C.Tee, "Secure minutiae-based fingerprint templates using random triangle hashing," *LNCS* 5857, pp 521–531, 2009.

[8] K. Xi and J. Hu, "Biometric mobile template protection: A composite feature based fingerprint fuzzy vault". In

Proceedings of IEEE International Conference on Communications (*ICC*) 2009, pp. 1–5, Dresden, 2009.

[9] K. Xi, A. Tohari, J. Hu and F.Han, "Fingerprint based bio-cryptographic security protocol designed for client/server authentication in mobile computing environment," *Journal of Security and Communication Networks*, vol 4 no 5, pp. 487-499, John Wiley & Sons, Ltd, 2011.

[10] A. Tohari, J. Hu, and S. Wang, "Pair-polar coordinate based cancelable fingerprint templates," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2555-2564, Elsevier, 2011.

[11] A. Tohari and F.Han, "Cartesian and polar transformation-based cancelable fingerprint template," *The 37th Annual Conference of the IEEE Industrial Electronics Society* (*IECON* 2011), Melbourne, Australia, pp 373-378, 2011.

[12] FVC2002, Fingerprint verification competition, 2002.

[13] A. Tohari and J.Hu "Generating cancelable biometric templates using a projection line," in *Proceedings of The 11th IEEE International Conference on Control Automation Robotics & Vision* (*ICARCV* 2010), pp. 7-12, Singapore, 2010.

[14]` A. Tohari, J. Hu and S.Wang, "String-based cancelable fingerprint templates, " in *Proceedings of The 6th IEEE Conference on Industrial Electronics and Applications* (*ICIEA* 2011), pp. 1028-1033, Beijing, China, 2011.

[15] Neurotechnology, Verifinger, version 5.0.

[16] Y. Wu (2011, January) "*Shamir's secret sharing code*". Matlab Central. [online]. Available: url:http://www.mathworks.com/matlabcentral/fileexchange/29989-shamirs-secret-sharing.. [Accessed: April 2011].