An Hybrid Authentication System Using Biometrics and Text Compression

A.Sindhuja[*], B.Sri Nandhini[**], S.Bharathi[**], I.Kalaimathy[**]

[*] Research Scholar, Sri Manakula Vinayagar Engineering College, Pondicherry, Puducherry - 605 014, India
[**] Student, Department of IT, Sri Manakula Vinayagar Engineering College, Puducherry - 605 014, India

Abstract- Today, the Internet has melted into our daily lives with more and more services being moved on-line. While we enjoy the convenience, we are putting ourselves at risk. Most current commercial websites will ask their users to input their user identifications (IDs) and corresponding passwords for authentication. Once a user's ID and the corresponding password are stolen by an adversary, the adversary can do anything with the victim's account, leading to a disaster for the victim. As a consequence of increasing concerns over such risks, how to protect users' passwords on the web is becoming more and more critical. To solve this we propose an hybrid authentication system with fingerprint using Efficient Biohashing algorithm to maintain very low error rate than existing algorithms and text compression using AES algorithm.

Index Terms- Authentication system, Biometrics, Biohashing Algorithm, Biohash code, AES algorithm.

I. Introduction

Biometric have the benefit that it bases recognition on an intrinsic aspect of a human being who is to be authenticated to be physically present at the point of the authentication. Denial of service occurs because of high false rate rejection. Multimodal biometrics can reduce the probability of denial of access without sacrificing the false acceptation performance. In order to solve the problem of high false rejection, a novel two-factor authenticator based on iterated inner products between tokenized pseudo-random number generated by an Hash key and the user specific fingerprint features; in this way, a set of user specific compact codes can be produced which is named "BioHash code".

Direct mixing of pseudo-random number and biometric data is an extremely convenient mechanism with which to incorporate physical tokens. The main drawback of this method, is the low performance when an "impostor" B steals the Hash key or the pseudo-random numbers of A and tries to authenticate as A. When this problem occurs, the performance of BioHashing can be lower than that obtained using only the biometric data.

In this paper, we propose an Efficient BioHashing approach which results more robust than the base method also when an "impostor" steals the Hash key (worst case). Experimental analysis showed the weakness of the base approach in the length of the BioHash code which is bounded by the dimension of the feature space and we propose improvement in algorithm to overcome this problem.

II. RELATED WORK

BIOHASHING:

1.Generates a sequence of real numbers with the help of a given secret hash key to produce a set of pseudo random vectors $\mathbf{r}i \in \mathrm{R}n,\ i=1,...,m$. They are linearly independent , eventually discarding wrong ones since they are a basis of space . We adopt the Blum-Blum_Shub method eventhough there are a variety of pseudo-random bit algorithms.

2.Then applying the Gram-Schmidt ortho-normalisation procedure to transform the basis $\mathbf{r}i$ into an orthogonal set of vectors or i=1,....,m.

3. Compute the inner product between the biometric feature vector x and $\mathbf{or}i$ $(\langle \mathbf{x}|\mathbf{or}\ i \rangle)$, $i = 1,...,m$ and compute bi $(i = 1,...,m)$ as

$$bi = \{ \begin{array}{l} 0 \text{ if } \langle \mathbf{x}|\mathbf{or}\ \ i\rangle \leqslant \tau \\ 1 \text{ if } \langle \mathbf{x}|\mathbf{or}\ \ i \rangle > \tau \ \ \} \end{array}$$

Where τ is a preset threshold

The hash key is different among different users and different applications and is given to the user during the enrollment . The resulting bit vector b, named as "BioHashCode", is compared by the Hamming distance for the similarity matching.

III. Analysis

Our ideas to boost the performance of the BioHashing approach are based on these two considerations:
* Increasing the dimension of the hash code for improving the system security.
* Biohashing is an instable classifier.

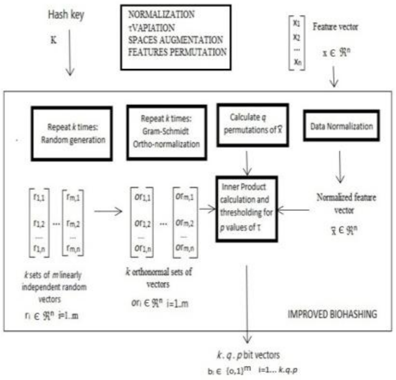
IV. EFFICIENT BIOHASHING ALGORITHM (MODIFIED BIOHASHING ALORITHM)

An improved version of the Biohashing method leads to Effective Biohashing algorithm as follows,

- **NORMALIZATION** : In this normalization , the biometric vectors normalized by their module before applying the BioHashing preocedure, such that the scalar product $<\mathbf{x}|\mathbf{or}_\mathbf{i}>$ is within the range [-1,1]
- **τVARIATION:** Instead of using a fixed value τ we use many values for τ and we combine with the "SUM rule"

the scores obtained varying τ between τ_{max} and τ_{min} with *p* steps of $\tau_{step} = (\tau_{max}-\tau_{min})/p$

- **SPACES AUGMENTATION**: Since the dimension of the projection space *m* cannot be increased, so we use more projection spaces to generate more BioHash codes per user. Let *k* be the selected number of projection spaces to be used, the BioHashing method is repeated k time on the same biometric vector in order to obtain *k* bit vectors b_i,i=1,…,*k*. Then the verification is carried out by combaining the classification scores obtained by each bit vector(Hash code).The random generation can be performed repeatedly , thus requiring a single Hash key **K** , in such a way that the random generator is not reinitialized bya new key unti the complete generation of the *k* bases in not performed.
- **FEATURES PERMUTATION**: Another way to generate more BioHash codes and without creating more projection spaces of the feature coefficients in **x** during the projection calculation: we use *q* bit vectors. As above the verification is carried out by combining the classification scores obtained by each bit vector.


Hash key
K
NORMALIZATION
τVAPIATION
SPACES AUGMENTATION
FEATURES PERMUTATION
x1
x2
...
xn
Feature vector
x ∈ ℜ^n
Repeat k times: Random generation
Repeat k times: Gram-Schmidt Ortho-normalization
Calculate q permutations of x̃
Data Normalization
r1,1 r1,2 ... r1,n
rm,1 rm,2 ... rm,n
or1,1 or1,2 ... or1,n
orm,1 orm,2 ... orm,n
Inner Product calculation and thresholding for p values of τ
Normalized feature vector
x̃ ∈ ℜ^n
k sets of m linearly independent random vectors
ri ∈ ℜ^n i=1..m
k orthonormal sets of vectors
ori ∈ ℜ^n i=1..m
IMPROVED BIOHASHING
k . q . p bit vectors
bi ∈ {o,1}^m i=1... k.q.p

The result of Efficient BioHashing procedure, if all the above solutions are exploited, is a set of *k.p.q* BioHash codes b_i , which are compared by the Hamming distance.

V. EXPERIMENT

Fingerprints:

The fingerprint databases *Db1*, *Db2*,*Db3* and *Db4* provided in FVC 2002 have been used for the experiments. We adopt the procedure named Finger-Code [21] for feature extraction algorithm, which requires to determine a reference point and region of interest for the fingerprint image, using a bank of Gabor filters and to compute the average absolute deviation from the mean of grey values in individual sectors in filtered images to define the feature vector. The result is a feature vector of a fixed size *n* = 640 which can be compared by the Euclidean distance.

For the verification task based on sole biometric data we adopt the method proposed in Ref. [22]. In our verification stage, the comparison of two fingerprints must be based on the same core point.

However, the comparison can only be done if both fingerprint images contain their respective core points, but 2 out of 8 impressions for each finger in FVC2002 [20] have an exaggerate displacement. In our experiments, as in Ref. [22], these two impressions were excluded, and hence, there are only 6 impressions per finger yielding 600 fingerprint images in total for each database.For the performance evaluation we adopt the equal error rate (EER) [16].

Table 1:
EER obtained of the fingerprint data sets using the following parameters: *m*=100, _max =0.1, _min =−0.1, *p* =5, *k* =5, *q* =5 (BEST hypothesis)

Method	Db1	Db2	Db3	Db4
BIO	5.5	5.2	18.3	8.4
BASE	1	0.8	5.4	1.4
τ VAR	0.9	0.6	4.6	1.4
SPAUG	0.5	0.4	3	0.8
FEATPERM	0.5	0.4	3	1
BIO + SPAUG	1	1	5	1.5

Table 2:
EER obtained of the fingerprint data sets using the following parameters: *m*=100, _max=0.1, _min=−0.1, *p*=5, *k*=5, *q*=5 (WORST hypothesis)

Method	Db1	Db2	Db3	Db4
BIO	5.5	5.2	18.3	8.4
BASE	15	15	27	20
τ VAR	13	12	27	15.5
SPAUG	11	10	25	14
FEATPERM	11	10	25.1	14.4
BIO + SPAUG	7	6.8	22	9.1

The tests reported in tables 1 and 2 are aimed to compare the Efficient BioHashing method with the BioHashing(BASE) and the simple verification method based on sole biometric data (BIO). We test three configurations of the efficient method where the NORMALIZATION is always present: with “VAR” we denote the simple VARIATION solution , with “SPAUG” and with “FEATPERM” the other two solutions coupled with VARIATION. Since the last two solutions are quite similar each to other and do not produce sensible variations of performance we did not test them together. The last configuration we tested (BIO + SPAUG) is a combination by the SUM rule of the method-based sole on the biometric features and our best effecient BioHashing.We perform experiments both in the best hypothesis (BEST), when never an impostor steals the Hash key, and in the worst (WORST) and very unlikely hypothesis that always (in each match) an impostor steals the Hash key.

In conclusion, our experiments show that:

Our efficient BioHashing approaches (in all the proposed variants, and in particular in the best configuration SPAUG) dramatically improve the performance of the BioHashing method, mainly in the worst case that always an impostor steals the Hash key.

The fusion of the non-hashing verification approach and our SPAUG method (BIO + SPAUG) can be a good compromise: it reaches a nearly 0 EER in the best and most probable hypothesis of no-stealing of Hash keys and it improve the performance of a pure "BioHashing method" in the worst hypothesis of key stealing.

VI. CONCLUSION

Biometrics has the great advantage of basing recognition on an intrinsic aspect of a human being and thus requiring the person to be authenticated to be physically present. Unfortunately, biometrics also suffers from some inherent limitation: high false rejection of valid users, when the system works at a low false acceptation rate.We have proposed a modified BioHashing approach based on the several solutions for augmenting the length of the hash code, which gains performance improvement also in the worst case when always an impostor steals the Hash key. Moreover, we have shown that the fusion between our efficient BioHashing and a method trained using "only" the biometric data allows the performance to be further efficient.

REFERENCES

[1] 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.

[2] C. Roberts, "Biometric attack vectors and defences," *Computers and Security*, vol. 26, no. 1, pp. 14–25, 2007.

[3] M1.4 Ad Hoc Group on Biometric in E-Authentication, "Study report on biometrics in E-authentication," Tech. Rep. INCITS M1/07-0185rev, International Committee for Information Technology Standards (INCITS),Washington, DC, USA, August 2007.

[4] I. Buhan and P.Hartel, "The state of the art in abuse of biometrics," Tech. Rep. TR-CTIT-05-41, Centre for Telematics and Information Technology, University of Twente, Twente, The Netherlands, December 2005.

[5] A. K. Jain, A. Ross, and U. Uludag, "Biometric template security: challenges and solutions," in *Proceedings of the European Signal Processing Conference (EUSIPCO '05)*, Antalya, Turkey, September 2005.

[6] B. Cukic and N. Bartlow, "Biometric system threats and countermeasures: a risk based approach," in *Proceedings of the Biometric Consortium Conference (BCC '05)*, Crystal City, Va, USA, September 2005.

[7] K. Ishikawa, *Guide to Quality Control*, Nordica International, Tokyo, Japan.

[8] A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics*, Springer, Berlin, Germany, 2006.

[9] Biometric System Laboratory - University of Bologna, "FVC2006: the 4th international fingerprint verification competition,"2006, http://bias.csr.unibo.it/fvc2006/default.asp.

[10] C. Wilson, A. R. Hicklin, M. Bone, et al., "Fingerprint vendor technology evaluation 2003: summary of results and analysis report," Tech. Rep. NISTIR 7123, National Institute of Standards and Technology, Gaithersburg, Md, USA, June 2004.

AUTHORS

First Author – A.Sindhuja, Research Scholar, Sri Manakula Vinayagar Engineering College, Pondicherry, Puducherry - 605 014, India

Second Author – B.Sri Nandhini, Student, Department of IT, Sri Manakula Vinayagar Engineering College, Puducherry - 605 014, India

Third Author – S.Bharathi, Student, Department of IT, Sri Manakula Vinayagar Engineering College, Puducherry - 605 014, India

Fourth Author – I.Kalaimathy, Student, Department of IT, Sri Manakula Vinayagar Engineering College, Puducherry - 605 014, India