University of Wollongong

Research Online

Faculty of Informatics - Papers (Archive)

Faculty of Engineering and Information Sciences

20-8-2006

Cryptographic key generation from biometric data using lattice mapping

Gang Zheng
University of Wollongong, gavinzh@uow.edu.au

Wanqing Li
University of Wollongong, wanqing@uow.edu.au

Ce Zhan
University of Wollongong, czhan@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/infopapers

 Part of the Physical Sciences and Mathematics Commons

Recommended Citation

Zheng, Gang; Li, Wanqing; and Zhan, Ce: Cryptographic key generation from biometric data using lattice mapping 2006.
https://ro.uow.edu.au/infopapers/452

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Cryptographic key generation from biometric data using lattice mapping

Abstract

Crypto-biometric systems are recently emerging as an effective process of key management to address the security weakness of conventional key release systems using passcodes, tokens or pattern recognition based biometrics. This paper presents a lattice mapping based fuzzy commitment method for cryptographic key generation from biometric data. The proposed method not only outputs high entropy keys, but also conceals the original biometric data such that it is impossible to recover the biometric data even when the stored information in the system is open to an attacker. Simulated results have demonstrated that its authentication accuracy is comparable to the well-known k-nearest neighbour classification.

Disciplines

Physical Sciences and Mathematics

Publication Details

This article was originally published as: Zheng, G, Li, W & Zhan, C, Cryptographic key generation from biometric data using lattice mapping, 18th International Conference on Pattern Recognition, 2006 (ICPR 2006), 20-24 August 2006, 4, 513-516.

Copyright 2006 IEEE.

This conference paper is available at Research Online: https://ro.uow.edu.au/infopapers/452

Cryptographic Key Generation from Biometric Data Using Lattice Mapping

Gang Zheng, Wanqing Li and Ce Zhan
School of Information Technology and Computer Science
University of Wollongong, Australia
{gz207, wanqing, cz847}@uow.edu.au

Abstract

Crypto-biometric systems are recently emerging as an effective process of key management to address the security weakness of conventional key release systems using passcodes, tokens or pattern recognition based biometrics. This paper presents a lattice mapping based fuzzy commitment method for cryptographic key generation from biometric data. The proposed method not only outputs high entropy keys, but also conceals the original biometric data such that it is impossible to recover the biometric data even when the stored information in the system is open to an attacker. Simulated results have demonstrated that its authentication accuracy is comparable to the well-known k-nearest neighbour classification.

1. Introduction

Cryptographic systems have been widely used to secure information. Whether a symmetric cipher system or a public-key system, its security depends on the secrecy of the secret or private key. Traditional passcode or token based key release systems are not secure and convenient enough for ever increased security requirements of many applications since passcodes and tokens are easy to be forgotten, lost or stolen. Crypto-biometric system [11, 8] has recently been emerging as an effective means to address these issues faced by traditional passcode or token based systems. It intends to bind a cryptographic key with user's biometric information in a manner to meet the following requirements [4, 3, 8] of distortion tolerance, discrimination and security.

- Distortion tolerance is the ability to accommodate the variance of the biometrics. The system is expected to output the same key for the same user even if the biometrics is acquired at different time or under different conditions.
- Discrimination is the ability of the system to distinguish all users of the system and output different keys for different users.
- Security of the system means that neither the key, nor the user's original biometric information can be extracted or calculated when the stored information is compromised.

In general, there are two approaches to binding a cryptographic key with biometrics. One is biometric-based key release [11] where the key is hidden into a biometric template during enrollment and is released during authentication. A typical key release system [11, 1] employs a similar strategy to UNIX login password where a one-way hash of the biometric template is stored and the authentication is conducted in the hashing space.

The other is biometric-based key generation [11] in which the key is calculated directly from the biometric information. One of the main challenges in this approach is to maintain the entropy of the key and keep the security of the biometric information simultaneously. This paper is about a new cryptographic key generation method.

1.1. Related work

A handful of papers has been published so far in the area of key generation from biometric data. The proposed methods generally fall into two categories: error-correcting code [12] based and Shamir's key sharing scheme [7] based. In error-correcting code scheme, codewords and decoding functions are established from the biometric templates during the enrollment. The codeword or its hash value can be used either as a key or as a seed to a key generator. At the authentication stage, the biometric data is used to compute or retrieve the codeword. Typical error-correcting code based key generation methods are Juels' Fuzzy Commitment Scheme [4] and Fuzzy Vault Scheme [3], Dodis's Fuzzy Extractor [2] and Soutar's Bioscript for fingerprints [8].

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

IEEE COMPUTER SOCIETY

Shamir's key sharing scheme [7] provides another avenue to bind a key with biometric data. In this scheme, biometric data is first transformed into a binary sequence. During enrollment, the binary sequences of biometric templates are employed to generate the shares of a key. During authentication, binary sequences from the biometric data forms the knowledge (shares) of the key. Monrose et al. [6, 5] first employed this scheme to harden a password with keystroke dynamics [6] and then extended it to generate a key from voice [5]. Teoh et al. [10, 9] proposed a method to generate a key from fingerprints [10] and faces [9] by projecting features to a randomly selected orthogonal space.

Both schemes require the transformation from biometric data into binary sequences. This transform is critical not only to the accuracy of the authentication, but also the entropy of the generated keys.

1.2. Contributions of the paper

Following the principle of Juels' fuzzy commitment scheme and K-nearest neighbourhood (K-NN) classification, we propose in this paper a new method for key generation from biometric data by employing a set of error tolerant lattice functions to map biometric data from feature space into lattice spaces. The method does not require the original biometric data to be stored. Instead, it stores the lattice functions for the templates acquired during the enrollment. Original biometric features are not recoverable even the lattice functions are compromised. The method is generic and applicable to all types of biometric data with a comparable performance to K-nearest neighbour (K-NN) classification.

2. Lattice Mapping Based Fuzzy Commitment

2.1. Fuzzy Commitment Scheme [4]

Formally an n-bit commitment scheme consists of a function $F : \{0,1\}^n \times X \rightarrow Y$. To commit an n-bit codeword c, the sender chooses a *witness* $x \in X$, generally uniformly at random. The sender then computes $y = F(c, x)$, known as a *bolb*. It represents that the n-bit c is sealed in a "safe". To "open" or *decommit* the bolb y and produce the codeword c, it requires witness x. Notice that an n-bit witness x can be uniquely expressed in terms of the codeword (committed value) c along with an offset $\delta \in \{0,1\}^n$ such that $x = c + \delta$. The idea behind fuzzy commitment is to conceal c using a conventional hash function h, while leaving δ in the clear. The information δ provides tolerance in the witness required to open F. In particular, δ only provides partial information about x and the remaining information to specify x, namely the codeword c, is presented in a concealed form $h(c)$.

Let $h : \{0,1\}^n \rightarrow \{0,1\}^l$ be a hash (one-way) function, such as SHA-1. $F : (\{0,1\}^n, \{0,1\}^n) \rightarrow (\{0,1\}^l, \{0,1\}^n)$ can be defined as follows:

$$F(c, x) = (h(c), x - c). \tag{1}$$

To decommit $F(c, x) = (\alpha, \delta)$ using witness x', the receiver computes $c' = f(x' - \delta) = f(c + (x' - x))$. If $\alpha = h(c')$, then it has been successfully decomitted, with c' being the extracted commitment. Otherwise, x' is an incorrect witness. Here $f(\cdot)$ is a decoding function.

In the context of key generation from biometric data, a user, U, presents biometric data x during enrollment. The system randomly selects a codeword c as the key or a seed to a key generator from a set $C \subseteq \{0,1\}^n$ of *codewords*, computes the fuzzy commitment $y_U = F(c, x)$ and stores it in a file for U. At authentication, a user purporting to be U presents biometric data x'. The system looks up y_U and check whether x' yields a successful decommitment. If so, the decommited c is the key or the seed to a key generator.

As seen from Equation 1, application of the fuzzy commitment scheme to a particular biometric data requires a transformation of the biometric data from its feature space into a binary sequence, a proper selection of the codewords and the decoding function $f(\cdot)$. Selection of the codewords decides the entropy of the key whereas the binary transformation and selection of decoding function will determine the authentication accuracy of the system. Though Juels [4] mathematically formulated the fuzzy commitment scheme with explanatory examples, no specification has been given for the selection of codewords and the decoding function. In the next section, we present a fuzzy commitment scheme based on a set of lattice functions.

2.2. Lattice Mapping

Let $\mathbf{x} = (x_1, x_2, \cdots, x_p)$ be a p-dimensional biometric feature vector and $x_i \in \Re, i = 1, \cdots, p$. We define the codeword c as a p-dimensional vector with each element being a random binary string. Let $c = (s_1, s_2, \cdots, s_p)$, $s_i \in_R \{0,1\}^q, i = 1, \cdots, p$, where $\in_R$ means uniform random selection from a set. We further treat the codeword c as the coordinate of $\mathbf{x}$ in a lattice space, L, mapped from its real feature space with δ being the half of the lattice grid size. This lattice space L can then be defined by its origin $O = (o_1, o_2, \cdots, o_p)$ and the grid size δ.

$$o_i = x_i - \delta - 2\delta\, s_i,\ i = 1, \cdots, p \tag{2}$$

With this arrangement, the decoding function $f(\cdot)$ becomes a simple mapping from $\mathbf{x}$ to c using the lattice system $L(O, \delta)$.

$$\begin{aligned} c &= f(\mathbf{x}) \\ s_i &= \left[\frac{x_i - o_i}{2\delta}\right],\ i = 1, \cdots, p \end{aligned} \tag{3}$$

IEEE COMPUTER SOCIETY

where, $[\cdot]$ is an operator taking the integer part of the input. It can be seen that any biometric data $\mathbf{x}'$ that lies within the hypersphere centered at $\mathbf{x}$ with radius δ will be mapped into the same grid (codeword c). In other words, δ serves as a parameter of distortion tolerance. In addition, it is impossible to calculate original biometric data $\mathbf{x}$ from the lattice system $L(O, \delta)$ and $\mathbf{x}$ is not required to decommit c from $\mathbf{x}'$. Therefore, the system only need to store $L(O, \delta)$ and the codeword c and x are secure even when $L(O, \delta)$ is open to an attacker.

3. Cryptographic key generation

3.1. Protocol

Let S denote the cryptographic key generation system based on the lattice mapping described above, U denote the user and K be a deterministic algorithm that takes input as seed and output a corresponding secret/public key pair (SK, PK). The protocol employing the proposed lattice mapping based fuzzy commitment is described as follows

- **Enrollment** The user U presents biometric data $\mathbf{x}$, the system S selects a codeword c, then S computes the origin of the lattice space O using the system parameters δ, and computes the secret key $sk = h(c)$ or the key pair $(SK, PK) = K(c)$. The system S stores the lattice system $L(O, \delta)$.

- **Authentication** A user purporting to be U presents a value of $\mathbf{x}'$ and decommit c' using Equation (3). If successful, she uses c' to compute the secret key $sk = h(c')$ or as a seed to K to derive (SK_U, PK_U).

The protocol describes general steps involved in the key generation and illustrates a special case where only one training sample of the biometric data is presented at enrollment. In practice, user U is often required to present their biometric data a few times in order to explore the possible variations of the biometric data.

Let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m\}$ be the m copies of the biometric data presented by the user U during enrollment. The system S computes m lattice systems $LS = (L_i(O_1, \delta), \cdots, L_m(O_m, \delta))$ such that all $\mathbf{x}_i, i = 1, \cdots, m$ will be mapped to the same codeword c. LS is then stored in the system with respect to U.

During authentication, when a user claiming to be U presents a copy of their biometric data $\mathbf{x}'$, the system S decommit m copies of c'. Let $c'_1, \cdots, c'_m$ represent the m copies of c' and $c'_i, \forall i$ is the decommited value from the corresponding lattice system $L(O_i, \delta)$. Depending on how far away between $\mathbf{x}'$ and $\mathbf{x}_i$, c'_i can be either same as c or different from c.

$$c'_i = \begin{cases} c & \text{if } \|\mathbf{x}' - \mathbf{x}_i\| \leq \delta \\ \text{not } c & \text{otherwise} \end{cases} \quad (4)$$

A proper selection of the tolerance parameter δ will allow more than one c'_i to be identical to c. A majority vote will decide which c'_i should be output by the system S.

4. Simulated results

In this section, we present simulated results to demonstrate the performance of the proposed method. Instead of using artificially generated data, we used the well-known Iris plant data. The data has 3 classes, 4 real numeric attributes and 150 samples, which simulated 3 users.

Figure 1 shows the total authentication error rate of the proposed method and K-NN at k=7 when different number of training samples were used. The training samples were randomly selected from the data set from each class (user) and the rest of the samples was used as test data. When $\delta = 0.7$, the proposed method performed comparable to K-NN classification.

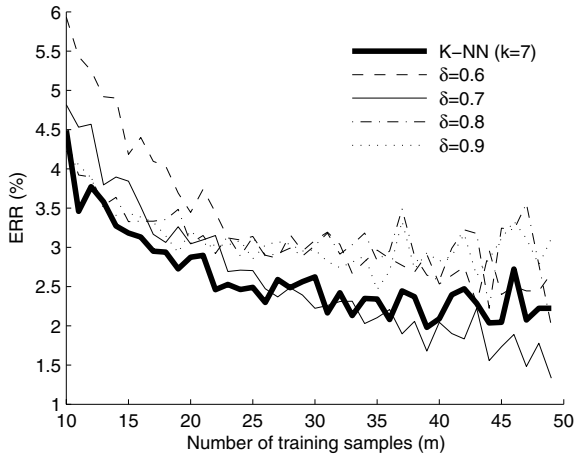

Figure 1. Error rate of K-NN classification and authentication by the proposed method at various values of δ when different number of training samples are used

Figure 2 shows the ROC of the system as the tolerance δ varies. The system achieved the best results for the Iris data set when $\delta = 0.7$

Like all biometric system, selection of the parameter δ is critical to the performance of the system, since the value of δ is also a reflection of the distribution of the training samples used in the enrollment. For a real system, the best performance may be achieved by a proper selection of the

IEEE COMPUTER SOCIETY

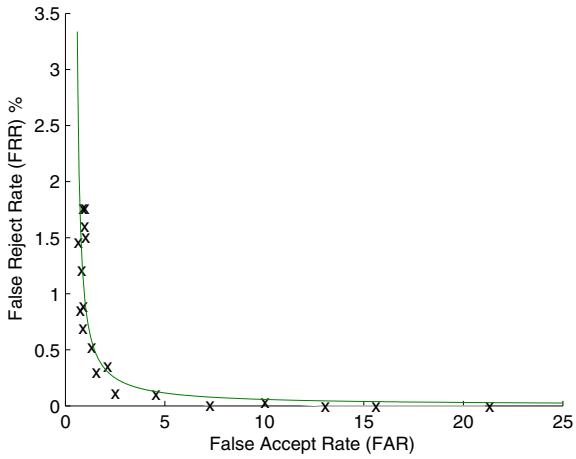

Figure 2. ROC of the proposed system

training samples and selection of δ through statistical analysis of sample distribution.

5. Security analysis

The proposed method is a specific implementation of Juels' fuzzy commitment and fuzzy vault scheme [4, 3] using lattice mapping functions with an extension from binary string $\mathbf{x}$ in [4, 3] to $\mathbf{x} \in \Re$. Theorems and Lemmas on security proved in [4, 3] are still valid to the proposed method.

Since the codeword c is randomly generated, the security of the derived key depends on the hashing function $h(c)$ or the key generator $K(c)$. Many existing hashing and key generation algorithms can be employed. In addition, the lattice space depends on the number of bits used for s_i and dimension, p, of the biometric features, which can be virtually large, considering that p could be as large as over 200 in voice and thousands in fingerprints and faces. Assume s_i is a l-bit string, then $c \in \{0,1\}^{lp}$.

Besides the security of the key, the biometric data is also secure as discussed in Section 2.2. Our system stores only the lattice mapping parameters $L(O, \delta)$, not the original biometric features and it is impossible to calculate the original biometric feature $\mathbf{x}$ from $L(O, \delta)$.

The release of δ reveals partial information about $\mathbf{x}$, this may reduce the entropy of the produced key, however, this reduction will be small [2].

Finally, users are able to use the same biometric data to generate different keys for different applications. This means if one system is compromised, other systems accessed by the same user are still safe.

6. Conclusion

This paper presents a lattice mapping based fuzzy commitment method to compute a cryptographic key from biometric data without revealing the biometric data. The method not only meet the security requirements, but also produce comparable accuracy to well-known K-NN classification method. Experiments on real biometric data, particular fingerprints and voice, are being conducted and will be reported in the near future.

7. Acknowledgment

The work was funded by University of Wollongong Research Council Small Grant 2005.

References

[1] G. L. Davida, Y. Frankel, and B. J. Matt. On enabling secure applications through off-line biometric. In *Proc. 1998 IEEE Symposium on Privacy and Security*, pages 148–157, 1998.

[2] Y. Dodis, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In C. Cachin and J. Camenisch, editors, *Proc. Eurocrypt 2004*, volume 3027 of *LNCS*, pages 532–540, 2004.

[3] A. Jules and M. Sudan. A fuzzy vault scheme. In *Proc. IEEE International Symposium on INformation Theory*, pages 408–421, 2002.

[4] A. Jules and M. Wattenberg. A fuzzy commitment scheme. In *Proc. ACM Conf. Computer and Communication Security*, pages 28–36, 1999.

[5] F. Monrose, M. K. Reiter, Q. Li, and S. Wetzel. Cryptographic key generation from voice. In *Proc. IEEE Symposium Security & Privacy*, pages 202–213, 2001.

[6] F. Monrose, M. K. Reiter, and S. Wetzel. Password hardening based on keystroke dynamics. In *Proceedings of* $6'th$ *ACM Conference on Computer and Communications Security*, pages 73–82, 1999.

[7] A. Shamir. How to share a secrete. *Communications of the ACM*, 22(11):612–613, 1979.

[8] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. V. K. V. Kumar. Biometric encryption. In R. K. Nichols, editor, *ICSA Guide to Cryptography*, pages 649–675. McGraw-Hill, New York, 1999.

[9] A. B. J. Teoh, D. C. L. Ngo, and A. Goh. Personalised cryptographic key generation based on FaceHashing. *Computer & Security*, 23:606–614, 2004.

[10] A. T. B. Teoh, D. N. C. Ling, and A. Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition*, 37:2245–2255, 2004.

[11] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain. Biometric cryptosystems: Issues and challenges. *Proceedings of IEEE*, 92(6), June 2004.

[12] S. A. Vanstone and P. C. van Oorshot. *An introduction to error correcting codes with applications*. Kluwer Academic Publishers, 1989.

IEEE COMPUTER SOCIETY