Oleksandr SIEVIERINOV[1], Svitlana KHALIMOVA[2]

Opiekun naukowy: Dmitro PROKOPOVICH-TKACHENKO[3]

SZABLONY ZABEZPIECZEŃ WYKORZYSTUJĄCE KOD SZUMOWY

Streszczenie: Artykuł rozważa metodę ochrony wzorców opartą na kodzie szumowym przy projektowaniu schematu rozmytych zobowiązań bez użycia algorytmów kryptograficznych. Pokazano, że zastosowanie krótkich kodów błędów korekcji dla szablonów biometrycznych szumów pozwala na tworzenie chronionych obrazów biometrycznych, odpornych na ujawnienie, z uwzględnieniem redundancji szablonu.

Słowa kluczowe: uwierzytelnianie biometryczne, zobowiązania rozmyte, kody błędów naprawczych

THE PROTECTION METHOD TEMPLATES ON THE BASIS OF THE CODE NOISING

Summary: The article considers the method of protection of patterns based on code noise in the design of a scheme of fuzzy obligations without the use of cryptographic algorithms. It is shown that the use of short codes for correction errors for noise biometric templates allows creating protected biometric images, resistant to the disclosure, taking into account the redundancy of the template.

Keywords: biometric authentication, fuzzy obligations, corrective error codes

1. Introduction

The success of using biometrics technology in cryptosystems determine by their security and error rate values in unauthorized access systems. Two basic requirements for protecting biometric information (ISO/IEC FCD 24745) are defined: irreversibility and inconsistency. Irreversibility determines the possibility of restoring

[1] Kharkiv National University of Radio Electronics, Candidate of Technical Sciences, Ukraine, Associate Professor, assistant professor of Department Security Information Technology, specialty: cybersecurity, oleksand.sievierinov@nure.ua

[2] Kharkiv National University of Radio Electronics, Ukraine, teacher, faculty of Computer engineering and management, specialty: protection of information, svitlana.khalimova@nure.ua

3 Candidate of Technical Sciences, University of Customs and Finance, Head of the Department of Cybersecurity and Information Technology, specialty: cryptography, cryptanalysis, information technology of leakage channel search, omega2417@gmail.com

the original biometric characters from the stored reference data, that is, the protected template. Recovery should be computationally complex, and creating a secure biometric template should be easy enough. The disconnected property determines that different versions of protected biometric templates can be created based on the same biometric data (renewable), and at that time, protected templates should not allow cross-linking of data (variety).

Because of the user-specific features in biometric features, providing security in biometric cryptosystems is a complex task. Since biometric characteristics are largely unchanged, attacking a biometric template can lead to the following vulnerabilities:

• the template can be replaced by a pretender template to gain unauthorized access;

• a physical copy can be created from the template to get unauthorized access to the system (as well as to other systems using the same biometric feature);

• a stolen template can be reproduced to gain unauthorized access;

• templates can be used for cross-reference in different databases to secretly track a person without his consent.

Standard encryption algorithms do not support the comparison of biometric templates in an encrypted domain due to user customization, and thus leave biometric templates unprotected during each authentication attempt.

The purpose of the article is to describe and analyze the method of protection of patterns based on code noise in the design of the scheme of fuzzy obligations without the use of cryptographic algorithms.

1.1. Organization of this paper

In section 2, we will present a review of the methods of biometric authentication based on a scheme of fuzzy obligations and a description of the related work. In Section 3, we will consider the method of protection of templates based on code noise and in the examples we perform an analysis of the basic characteristics of biometric authentication. In Section 4, we give an estimate of the cryptographic stability of the proposed method. Let's summarize the results and give suggestions for further research.

2. Biometric cryptosystem

Biometric template protection schemes are classified as biometric cryptosystems (schematics based on auxiliary data) and abolished biometrics (pattern conversion schemes).

Biometric cryptosystems are designed to safely bind a digital key to a biometric template, replacing the classic password creation scheme. Most biometric features provide the same level of security in a user group. Because of biometric dispersion, ordinary biometric systems perform "fuzzy comparisons" by applying decision thresholds that are established on the basis of the laws of the distribution of correspondence between genuine and non-genuine subjects.

Cancelable biometrics implements deliberate distortions of biometric patterns based on irreversible transformation violators that provide a comparison of biometric templates in a transformed domain [1]. Cancelable biometrics should ensure the irreversibility and discontinuity of biometric patterns [2].

A sufficiently complete analysis of the methods for constructing biometric

cryptosystems was performed in [3, 4]. The protection of patterns based on a scheme of fuzzy obligations was first proposed in [5,6]. So G.I. Davida, Y. Frankel, and B.J. Matt presented a schematic of storing a biometric template in an implicit, protected form, such that an error can be made when reading biometric characteristics. Protection is achieved by calculating control bits in a template using a linear error correction code and saving these control bits together with a pattern hash [5]. A subsequent analysis of their work can be found in [6]. A definite improvement of this method was made by A. Juels and M. Wattenberg in [7, 8]. The scheme of fuzzy obligations applied to biometric templates considers the template itself without any changes as a distorted codeword. This difference in perspective gives some advantages. Most importantly, this construction connects the number of codewords with the security parameter, while Davida et al binds it to a significantly larger message size (i.e., a pattern). As a result, in the design of A. Juels and M. Wattenberg, much smaller codes are used to correct errors than Davida et al. and provides significantly higher cryptographic stability. The design of the fuzzy obligation, thus, provides the idea of safe storage of biometric templates in practical applications.

The disadvantages of the proposed schemes are as follows. It is necessary to coordinate correction error codes with a long biometric template. Different codes will be optimal for different biometric characteristics. In the cryptographic authentication protocols, the block decoding procedure is implemented, which is particularly costly, since the number of checks is determined by the size of the user database. Practical implementation even for a small number of users is problematic. The template is transmitted over the communication channel and presented for inspection in an unsecured manner, unlike the fuzzy container that is stored on the authentication server. The transmission of an unprotected template is inadmissible, as the properties of irreversibility and inconsistency are violated.

For practical applications, templates that are transmitted in the communication channel and stored on the authentication server must be protected. If biometric characteristics of the biometric patterns are internal to the user, they should be provided not by cryptographic methods. The computational cost of authentication must be minimal and the required characteristics of the error values must be provided. To implement these requirements, a method for protecting templates based on code noise is proposed.

3. Method for protecting templates based on code noise

A classic template protection scheme based on error correction codes is the fuzzy obligation scheme proposed by A. Juels and M. Wattenberg [8]. The method includes the following steps.

Biometric characteristics are extracted from the biometric image and represented by a binary $S = (s_1 s_2 ... s_n)$ length n vector. A random binary secret key K is generated and encoded into a code word $C = (c_1 c_2 ... c_n)$.

The template $S = (s_1 s_2 ... s_n)$ and code word $C = (c_1 c_2 ... c_n)$ are added bit by bit $H = S \oplus C$ and this data is stored in the database. The hash of the code word $h(C)$ is also stored in the database.

During authentication by request, a biometric template $S' = (s'_1 s'_2 ... s'_n)$ is generated and compared with the code template $H = S \oplus C$ from the database

$Z = S' \oplus H$. Authentication is successful if S' close enough to the original pattern S. The degree of proximity is set when decoding z. When decoding z, the code word C' is restored, the hash $h(C')$ is calculated. Authentication is determined by matching hashes $h(C')$ and $h(C)$.

The assessment of the secrecy of the protection scheme is considered in [9]. The disadvantages of the template protection scheme are that the secrecy of the biometric template S is not ensured, the high computational complexity of decoding with recovery of code words, the need to coordinate parameters of error correction codes and templates.

The paper proposes a design for constructing secure templates based on noise by codes that detect errors. Let on the template S of l_s bits length to be superimposed on the code of the binary code of the corrective error (n, k, d). Thus, the number of such words equals $N = l_s/n$.

The pattern is broken into N words $S = (S_1, S_2 ... S_N)$. Binary random secret keys $K_1, K_2, ..., K_N$ are generated and encoded into code words $C_1, C_2, ..., C_N$.

The conversion is determined by the operation of the bitwise word combination of the template S_i and the code C_i

$$S_i \oplus C_i = SC_i, , \ i = \overline{1, N}.$$

The database for each template S contains $SC_1, SC_2, ..., SC_N$ words.

During authentication by request, a biometric template is generated with $S' = (S'_1, S'_2 ... S'_N)$ words, a random set of code words $C'_1, C'_2, ..., C'_N$ is generated, and, are calculated $S'_i \oplus C'_i = SC'_i, \ i = \overline{1, N}$.

The decision-making scheme for identification compares the noisy images that are stored in the database on the server SC_i with those accepted SC'_i for authentication. The comparison based on the operation of bitwise addition for pattern words

$$SC_i \oplus SC'_i = SC_i^{rez}, \ i = \overline{1, N}.$$

When comparing we get the result

$$SC_i^{rez} = \left(S'_i \oplus C'_i\right) \oplus \left(S_i \oplus C_i\right) = \left(S'_i \oplus S_i\right) \oplus \left(C'_i \oplus C_i\right), \ i = \overline{1, N}.$$

The sums of code words in the patterns are code words of the linear block code (n,k,d). We get

$$SC_i^{rez} = \left(S'_i \oplus S_i\right) \oplus C_i^{rez}, \ i = \overline{1, N},$$

where are the code words $C_i^{rez} = C'_i \oplus C_i$ of the linear block code (n, k, d) and $S'_i \oplus S = S_i^{rez}$ are noise vectors for the code words.

The number of code words C_i^{rez}, $i = \overline{1, N}$ in which errors are detected determines the degree of conformity of the template presented for authentication and the template that is stored in the database.

An advantage of the proposed template protection scheme is the low cost of computing code words for authentication and securing biometric templates when stored in the database and during authentication.

Consider the estimates for the biometric identification system - errors of the first kind, when the probability of a false access denial to a client with the right of access

FRR (False Rejection Rate) and errors of the second kind, as the probability of erroneous access, when the system mistakenly identifies someone else as their FAR (False Acceptance) is determined Rate).

The code (n,k,d) has the length of the code words n, the length of the information words k, and the code distance d. Thus, the number of code words equals $N_c = 2^k$.

The code distance determines the ability to correct errors by multiplicity $t = (d-1)/2$ and determine the errors with the multiplicity of $d-1$. Code words are generated by random values of information words k_i, $i=1,N$. The key sequence that generates code words should be random and have lengths equal to $K = kN = kl_s / n$ bits. Value k/n determines the speed of the code.

There are two situations.

1. Comparison of templates of different users.

The distribution of the appearance of 1 and 0 of the resulting template $S'_i \oplus S = S_i^{rez}$, $i = \overline{1,N}$, for independently distributed patterns will be equiprobable $p_0 = p_1 = 0.5$. The code (n,k,d) will not find only errors that will match the code words. Such code words equal to $N_c = 2^k$. Total number of bit combinations on the code length equals to $N_b = 2^n$. The probability that errors will occur during decoding (discrepancy S'_i and S_i) determines as $p_e = 1 - N_c / N_b$.

During authentication, the number of code words in which errors are defined should not exceed the threshold value T. In this case, we obtain an estimate of the probability of false access FAR which equals to

$$FAR = \sum_{j=0}^{T} C_N^j p_e^j (1-p_e)^{N-j}.$$

2. Comparison of the user templates and their template stored on the server.

The distribution of the appearance of 1 and 0 in the resulting template $S'_i \oplus S = S_i^{rez}$, $i = \overline{1,N}$, will depend on estimates of errors in reading the biometric image. For one person, biometric images will be different. On average, this can be determined by the probability p_{dist}. This probability will determine the probability of occurrence of 1 in the resulting template $S'_i \oplus S = S_i^{rez}$. If the images coincide completely S_i^{rez}, the result is a null sequence.

Decoding of code words of the code (n,k,d) allows guaranteed detection of errors with the multiplicity d-1. Thus, the code detects errors with probability

$$p_{iden} = \sum_{j=1}^{d-1} C_n^j {p_{dist}}^j (1-p_{dist})^{n-j}.$$

During the authentication process, the number of code words in which errors are defined should not exceed the threshold value T. Denial of access to a user entitled to access occurs if the number of detected code words with errors exceeds a threshold. We obtain an estimate of the error of the first genus of FRR.

$$FRR = \sum_{j=T+1}^{N} C_N^j {p_{iden}}^j (1-p_{iden})^{N-j} = 1 - \sum_{j=0}^{T} C_N^j {p_{iden}}^j (1-p_{iden})^{N-j}.$$

Example 1.

Let's perform the calculations using the Hamming code (8,4,4). Let the length of the biometric template be ls = 256 bits. Accordingly, the number of code words for 256 bit noise will be $N = l_s / n = 256/8 = 32$. Thus, 32 code words of the Hamming code are superimposed on the biometric template of 256 bits. Code words of 8 bits are generated independently by random 4-bit words.

We calculate the probabilities of errors of the first and second kind. We substitute in the formulas for FAR and FRR the values of the code parameters (8,4,4) $n=8$, $k=4$, $d=4$, number of noisy code words $N=32$, mismatch frequency in the bit sequence $S_i^{rez} = S_i^{chan} \oplus S_i^{serv}$ $p_{dist} = 0.05$. We obtain the following expressions.

$$p_e = 1 - N_c / N_b = 1 - 2^k / 2^n = 1 - 2^4 / 2^8 = 15/16$$

$$FAR = \sum_{j=0}^{T} C_N^j p_e^j (1-p_e)^{N-j}$$

$$FAR = \sum_{j=0}^{T} C_{32}^j 0.938^j 0.063^{32-j},$$

when T - decision threshold of stranger/registered user.

$$p_{iden} = \sum_{j=1}^{3} C_8^j {p_{dist}}^j (1-p_{dist})^{8-j}.$$

$$FRR = \sum_{j=T+1}^{32} C_{32}^j {p_{iden}}^j (1-p_{iden})^{32-j} = 1 - \sum_{j=0}^{T} C_{32}^j {p_{iden}}^j (1-p_{iden})^{32-j}.$$

Example 2.

We perform calculations of the probabilities of errors of the first and second genus for the container with the repeated code (2,1,2), with the lengths of the templates 256, 512, 1024, 2048 bits. The corresponding code word words will equal $N = 128, 256, 512, 1024$.

The probability of determining the erroneous code words of the code (2,1,2) when comparing the patterns of different biometric images in the case that the bits of the sequences coincide / do not coincide with the probability of 0.5 is determined by the relation

$$p_e = 1 - N_c / N_b = 1 - 2^1 / 2^2 = 0.5.$$

$$FAR = \sum_{j=0}^{T} C_N^j p_e^j (1-p_e)^{N-j}$$

$$FAR = \sum_{j=0}^{T} C_N^j 0.5^j 0.5^{32-j} = 0.5^{32} \sum_{j=0}^{T} C_N^j,$$

where T - decision threshold of stranger/registered user.

$$p_{iden} = \sum_{j=1}^{3} C_8^j \, p_{dist}{}^j (1 - p_{dist})^{8-j} .$$

$$FRR = \sum_{j=T+1}^{32} C_{32}^j p_{iden}{}^j (1 - p_{iden})^{32-j} = 1 - \sum_{j=0}^{T} C_{32}^j p_{iden}{}^j (1 - p_{iden})^{32-j} .$$

4. Estimations of the secrecy of a biometric template with code noise

The evaluation of the cryptographic stability of a biometric template can be determined based on the general theory of cipher stability, which was determined by Shannon through the uniqueness distance of the cipher. The uniqueness distance is determined by the relation

$$l_0 = \frac{H(K)}{(1 - H_{real} / H_{\max})} / \log_2 L ,$$

where $H(K)$ - entropy of the key source, L - message source alphabet, H_{real} - message source entropy, $H_{\max}$ - the entropy of the source of messages if the symbols of the alphabet had a uniform distribution and were independent in the text.

For a scheme with code noisy, we have

$$H(K) = kN = kl_s / n ;$$

$$L = 2 ;$$

$1 - H_{real} / H_{\max}$ - message redundancy factor.

The cipher is definitely persistent if the uniqueness distance is greater than the length of the cryptogram. In this case, there is no method by which you can uniquely find the key sequence of the cipher.

Example 3.
Let the Hamming code (8,4,4) of a template with a length of 256 bits be used for noise. We get

$$H(K) = kN = \frac{k}{n} l_s = \frac{4}{8} 256 = 128 ; \qquad L = 2 ;$$

If $1 - H_{real} / H_{\max}$ redundancy of the message text, for example, is equal to 0.5, we obtain the uniqueness distance of the cipher $l_0 = \frac{H(K)}{(1 - H_{real} / H_{\max})} / \log_2 L = 256$ bits. Considering that the entropy of the key $H(K) = kN = \frac{k}{n} l_s = \frac{4}{8} 256 = 128$, it can be argued that a crypto algorithm can be found to crack a noisy biometric template. Given that the entropy of the key is 128 bits, a brute-force attack based on a full search will have a complexity rating 2^{128}.

5. Conclusions

The presented design of fuzzy containers based on noise with codes that detect errors makes it possible to create biometric authentication systems that are resistant to inside user variations of biometric characteristics with properties of irreversibility and incoherence. The scheme of the fuzzy obligation, based on the noise of codes, ensures the security of the transmission and storage of biometric templates. Protection of biometric templates is realized not by cryptographic methods and implementation is computationally simple. For practical applications, the use of short error detection codes to clutter biometric patterns provides small values of errors of the first and second genus. An error estimate of the first genus of false denial of access to a client having the access right is the upper boundary, because it was calculated for the condition of a statistically uniform distribution of error errors for sequences of biometric templates.

REFERENCES

1. RATHA N., CONNELL J., BOLLE R.: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst J 2001, 40:614-634.
2. CAVOUKIAN A, STOIANOV A: Biometric encryption: the new breed of untraceable biometrics. Biometrics: Fundamentals, Theory, and Systems Wiley, London, 2009.
3. RATHGEB C, UHL A. A survey on biometric cryptosystems and cancelable biometrics/ EURASIP Journal on Information Security 2011, 2011:3
4. JAIN A, NANDAKUMAR K, NAGAR A: Biometric template security. EURASIP J Adv Signal Process 2008, pp. 1-17.
5. DAVIDA G., FRANKEL Y., and MATT B. On enabling secure applications through off-line biometric identification. Proc of IEEE, Symp on Security and Privacy 1998, 148-157.
6. DAVIDA G.I., FRANKEL Y., MATT B.J. On the relation of error correction and cryptography to an off-line biometric based identification scheme. Proc of WCC99, Workshop on Coding and Cryptography 1999, pp. 129-138.
7. JUELS A, SUDAN M.: A fuzzy vault scheme. Proc 2002 IEEE Int Symp on Information Theory 2002, p.408.
   8. Juels A., Wattenberg M. A Fuzzy Commitment Scheme. September 25, 2013 p.21
8. LAFKIH M., MIKRAM M., GHOUZALI S., EL HAZITI M., ABOUTAJDINE D.: Biometric Cryptosystems based Fuzzy Commitment Scheme: A Security Evaluation. The International Arab Journal of Information Technology, Vol. 13, No. 4, July 2016 pp.443-449