Paper

Privacy Preserving and Secure Iris-Based Biometric Authentication for Computer Networks

Przemysław Strzelczyk

Research and Academic Computer Network (NASK), Warsaw, Poland

Abstract—The iris biometrics is considered one of the most accurate and robust methods of the identity verification. The unique iris features of an individual can be presented in a compact binary form which can be easily compared with the reference template to confirm identity. However in contrast to passwords and PINs biometric authentication factors cannot be revoked and changed as they are inherently connected to our characteristics. Once the biometric information is compromised or disclosed it became useless for the purpose of authentication. Therefore there is a need to perform iris features matching without revealing the features itself and the reference template. We propose an extension of the standard iris-based verification protocol which introduces a features and template locking mechanism, which guarantee that no sensitive information is exposed. Presented solutions can be easily integrated into authentication mechanisms used in modern computer networks.

Keywords—authentication, biometrics, iris recognition, privacy preserving.

1. Introduction

Among the biometric verification methods iris recognition is considered one of the most accurate and robust. Iris features can be easily extracted from eye images and they can be efficiently compared. However if the biometric reference template or set of biometric features are disclosed, the whole biometric system becomes useless for an individual, because the biometric information cannot be canceled or revoked as passwords. Therefore there is a need to perform iris features matching without revealing either the biometric data acquired during the verification process or the reference template from the database. In this paper we introduce a privacy preserving system for iris-based biometric identity verification. Firstly we propose a method in which iris biometric templates stored by the authenticator are locked with the keys known only to the data owner. In this scenario even when the database is compromised no private information is exposed. Secondly we introduce a privacy preserving verification protocol, which is based on the secure multi-party computation techniques.

2. Iris Biometrics

An iris is the colored ring around the pupil. Its structure is determined during the fetal development of the eye and remains unchanged. On contrary the color of the iris can change as a result of the variable pigmentation in tissues. The main role of the iris is to control the size of the pupil and adjust the amount of light which enters through the pupil into the eye interior. It is surrounded by the sclera, which is a white area of tissues and blood vessels, and it is covered by a transparent layer called cornea. The whole iris is visible only with eyes wide open, as eyelids and eyelashes usually occludes the lower and upper part of it. The possibility of using the iris to distinguish individuals is over 100 years old, but the first patent for the automated iris biometric system was obtained by Flom and Safir in 1987 [1]. However the most important work in the field of the iris recognition was due to Daugman [2], [3]. He introduced first practically verified methods for iris image segmentation, unique feature extraction and matching, which with slight modifications are used today world widely and which are the reference models for other algorithms.

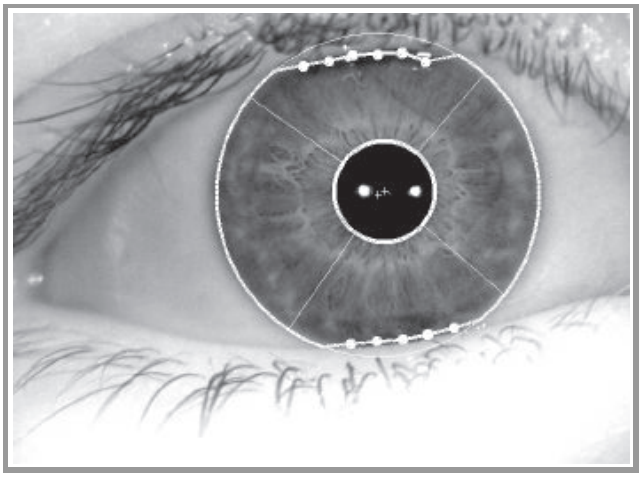

Fig. 1. An example eye image captured in infrared light with outer and inner boundaries approximsted by non-cocentric circles [www.biometriclabs.pl].

For the purpose of biometrics the eye image is captured in the near infrared light with the wavelengths between 700 – 900 nm. Usually special infrared illuminators and bandpass lens filters are used to acquire a image of good quality [4]. The infrared light reveals the detailed structure of the iris better than the visible light [5]. The iris is usually modeled as a ring with outer and inner border approximated by two circles which are nearly concentric.

Figure 1 shows an example image acquired in infrared light with the iris and pupil approximated by two circles. Because the iris area is hardly ever fully visible an additional mask is applied to the ring, which marks the areas where eyelids and eyelashes occlude the iris region. Optionally this mask also indicates image artifacts such as specular reflections from illuminators.

After the iris region is isolated, it is mapped into normalized pseudo-polar coordinate system [5]. Every point of the iris area is described by two coordinates: the angle α and the radial distance r. When the pupil and the iris centre overlap, r is the normalized distance from the iris pupil centre, and α is the angle of the line crossing the point and the pupil centre. Figure 2 shows an iris mapped into pseudo-polar coordinates. The resulting rectangular image has two important properties. Firstly the mapping models linear stretches of the iris, when the pupil contracts or dilates, which is regarded a good approximation of its nature [6]. Secondly the eye or head rotation is equivalent to the permutation of points in the α coordinate, so these effects can be easily compensated.

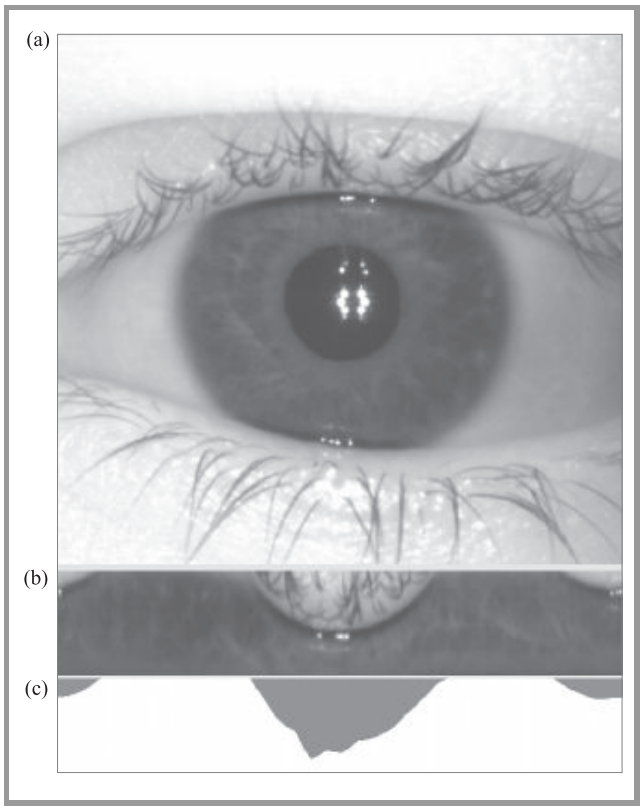

Fig. 2. An example iris image (a), iris image converted to pseudo-polar coordinates (b) and occlusions mask in pseudo-polar coordinates (c).

Most of the iris biometric methods convolve the transformed image with 2-dimensional filters designed to extract the unique iris texture patterns. Many different filters were tested but the best results were obtained for Gabor, Log-Gabor, Haar and Laplacian of Gaussian filters [3], [7]–[11]. The resulting values are re-sampled and quantized. Daugman proposed that only the sign of the filtered signal should be used as features. He introduced a binary iris representation which is summarized in 2048 bits vector, called an "iris code". Many other solutions follow this approach as it allows the iris codes to be compared efficiently using bitwise operations [2].

The iris code matching is based on the Hamming distance. The Hamming distance measures the fraction of corresponding bits of two binary vectors that disagree. Because not all of the bits of the iris code are valid due to occlusions and other disruptions, a modified Hamming distance must be introduced that take this into account. It can be implemented using three types of operations: bitwise XOR ($\oplus$), bitwise AND ($\wedge$) and bit counting as follows:

$$HD(x,w,y,m) = \frac{\sum_i (x_i \oplus y_i) \wedge w_i \wedge m_i}{\sum_i w_i \wedge m_i}, \qquad (1)$$

where x and y are two iris codes and w and m are the occlusions masks, whose bits are set if the corresponding bits of the iris codes are valid. To account for the eye rotation, the Hamming distance is computed for several different permutations of the bits corresponding to the different angles of rotations. At the end the minimal score is compared with the predefined threshold to check if the verification is successful.

3. Preserving Iris Code Privacy

Vernam's cryptographic system, known also as one-time pad is a type of encryption method, which has been proven to be impossible to crack if used properly [12]. The main idea is that the plaintext is encrypted with a substitution cipher using a secret random key. The size of the key must be equal to the size of the plaintext. As long as the key is truly random, and as large as plaintext, the resulting cipher text is also truly random. For the binary data the substitution Vernamñs system can be implemented based on cryptographically secure pseudo-random number generator and XOR operation. The pseudo-random number generator is used to prepare a key, and the XOR operation is used to randomly invert the bit values based on the key bits. One-time pad is a symmetric encryption mechanism, as the XOR operation when used twice with the same key reveals the original plaintext. The main idea behind the privacy preserving iris based verification is that iris codes encrypted with one-time pad can be matched in the same way as unencrypted ones. The Hamming distance of two iris codes before and after encryption remain the same, as long as the keys used to encrypt them are identical. It is because the XOR operation used for Hamming distance computing nullifies the encryption as follows:

$$enc_{\oplus}(x,k) \oplus enc_{\oplus}(y,k) = x \oplus k \oplus y \oplus k = x \oplus y. \qquad (2)$$

The biometric features of the same individual differ each time the biometric measurements are done. For the iris biometrics this variability is mainly due to changing capture conditions and image processing. These processes introduce a noise to the iris code. Whether the noise is a user

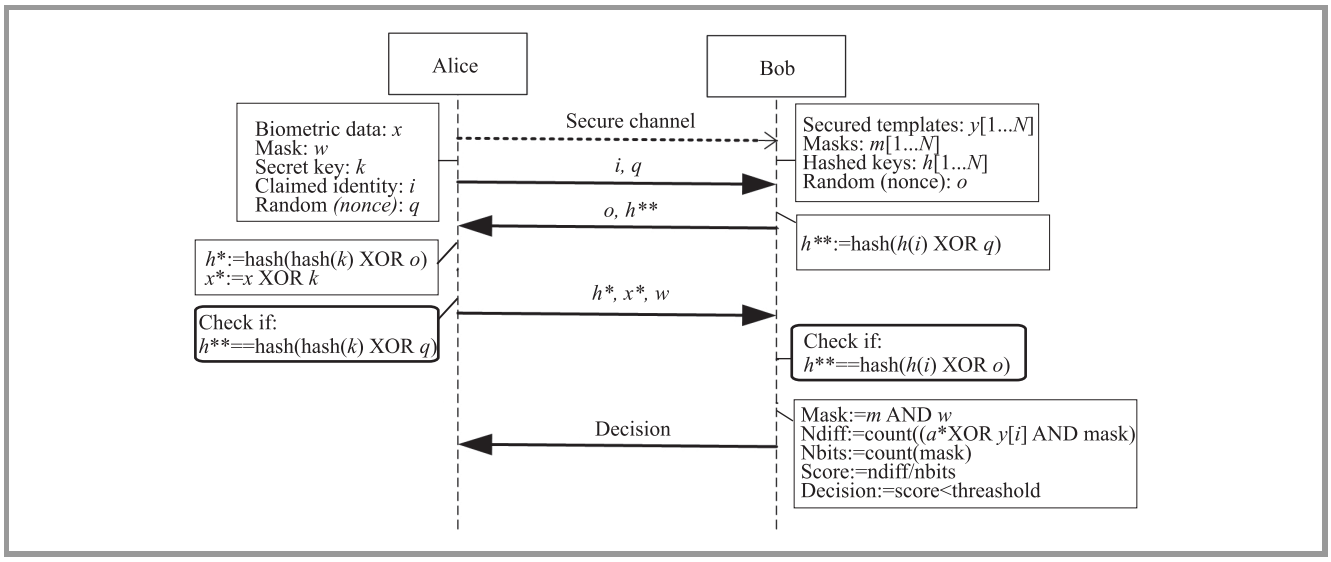

Fig. 3. A sequence diagram for the privacy preserving iris based biometric authentication protocol with two parties

characteristic or not is still an open question. Nevertheless we can assume that for each individual there exists an ideal iris code x and there is a measurement specific noise z. The real iris code can be expressed as a XOR operation between the ideal iris code x and the measurement noise z. In that situation the noise determines the Hamming distance between two iris codes of the same individual.

Now if we use the same cryptographic key in the Vernam's cipher for different real iris codes of the same individual we may reveal the information about the noise but the ideal iris code remains encrypted. The iris code matching routine will not only nullify the encryption but also the ideal iris code. The adversary will be able to determine the character of the noise but will deduce nothing about the ideal iris code.

The presented encryption of iris code has some other interesting features. The experiments indicate that there is only about 100 to 200 bits of information in the iris code in the Shannon sense [13]. This suggests that the bits of iris codes are strongly dependent. The one-time pad encryption removes the bit dependencies in the ciphertexts if the key used has its bits independent. As a result the encrypted iris code will be indistinguishable from random data for an adversary.

4. Two-Party Protocol

The privacy preserving iris based biometric authentication protocol is based on the secure multi-party computation scheme and applies the one-time pad encryption of iris code presented in the previous section.

There are two parties in this process, which we will later call: Alice and Bob. Alice represents a user who is willing to undergo biometric verification. She does not want her biometric data (namely the iris code and the binary occlusions mask) to be disclosed to any other party. Bob represents an authentication server which verifies Alice identity based on her biometric data. Bob has access to the encrypted biometric templates database, which he does not want to disclose to anyone. The authentication protocol presented in this paper employs several proven cryptographic methods. The first one is the cryptographic one-way function called later hash function [14], [15]. It is a deterministic procedure which takes an arbitrary binary vector and transforms it into fixed size bit string, called a hash value. The transformation has some fundamental properties. It is easy to compute the hash, but it is infeasible to recover the binary vector from hash value. It is infeasible to find two binary vectors with the same hash or modify the binary vector without changing its hash. a sample hash function suitable for our solution could be SHA-2. The authentication scheme uses also a secure channel establishment techniques [14], [15]. There are many algorithms which can be used for that purpose. The protocol also requires cryptographically secure pseudo-random number generator (CSPRNG). It is a pseudo-random generator, which has a very long period, and which satisfies the "next-bit test" and withstands "state compromise extensions" [14], [15].

When Alice is enrolled to the biometric system, she is assigned a unique identifier i. Next she uses a trusted biometric device, which captures her iris images and generates her reference iris code $t(i)$ with optional occlusions mask $m(i)$. Then a cryptographically secure pseudo-random number generator creates a secret encryption key k. The key is used to encrypt the reference iris code $y(i) = enc(x(i),k)$. The hash value of the key is computed $h(i) = hash(k)$ and the key is released to Alice. The hash code is needed in the verification process to prove that Bob possesses the encrypted template of Alice. Bob does not have the access to the encryption key. Instead he receives encrypted reference iris code with associated mask and the hash value of the key $\{y(i), m(i), h(i)\}$ and he inserts them into his database, which is indexed with the identification numbers i. Afterwards the system is ready for biometric verification.

The verification process is shown in Fig. 3. When Alice wants to verify herself to Bob she initializes the secure communication channel. Then she generates a random number q and send it securely to the Bob together with her claimed identity i. Bob uses the random value q and the hash value of i-th key $h(i)$ to prove Alice that her template is in the database. Additionally he generates a random number o, which will be used by Alice to prove that she owns the proper key. Bob sends Alice the number o and the proof h^{**}. After the proof is checked by Alice, the biometric system captures her iris image and extract her iris code x with the associated occlusions mask w. The iris code is encrypted with the secret key k and the proof h^{*} based on o is prepared. The encrypted iris code x^{*} with the occlusion mask w and the proof h^{*} are sent to Bob. Bob verifies the proof and calculates the Hamming distance between the encrypted iris code from Alice and the corresponding encrypted template from database (using the bitwise XOR operation, bitwise AND operation and bit counting). If the resulting score is lower than a predefined dissimilarity threshold the verification is considered successful.

5. Conclusions

The presented methods allow to authenticate an individual based on the iris biometrics without revealing the information about the biometric features. These methods use the specific proprieties of the iris code and its matching routines, and employ proven and verified cryptographic techniques. The presented protocol can be incorporated into existing authentication schemes used in the computer networks. The author of this paper is working on the specification and the reference implementation of authentication server and client, which will use the Extensible Authentication Protocol (EAP) family adapted to the presented biometric requirements. Our future work will focus on investigating the iris code noise properties. It is essential for the presented protocol to discover how much individual-specific information is still present in the noise, and what can be done to make the noise more random.

Acknowledgements

This work was co-financed by the Ministry of Science and Higher Education grant OR00 0026 07 “Platform for secure implementation of the biometric systems for verification and identification”.

References

[1] L. Flom and A. Safir, “Iris Recognition System”, February 3, 1987. US Patent 4,641,349.

[2] J. G. Daugman, “High confidence visual recognition of persons by a test of statistical independence”, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 11, pp. 1148–1161, 1993.

[3] J. G. Daugman, “Biometric Personal Identification System Based on Iris Analysis”, US Patent 5,291,560, March 1, 1994.

[4] A. Czajka and A. Pacut, “Iris recognition with adaptive coding”, *Rough Sets and Knowledge Technology*, *Lecture Notes in Artifical Intelligence*, vol. 4481, Springer, 2007, pp. 195–202.

[5] J. Daugman, “How iris recognition works”, *IEEE Trans. Circ. Sys. Video Technol.*, vol. 14, no. 1, pp. 21–30, 2004.

[6] H. J. Wyatt, “A minimum-wear-and-tear’meshwork for the iris”, *Vision Res.*, vol. 40, no. 16, pp. 2167–2176, 2000.

[7] W. W Boles and B. Boashash, “A human identification technique using images of the iris and wavelet transform”, *IEEE Trans. Signal Process*, vol. 46, no. 4, pp. 1185–1188, 1998.

[8] L. Chenhong and L. Zhaoyang, Efficient iris recognition by computing discriminable textons”, in *Proc. Int. Conf. Neural Netw. Brain ICNN&B’05*, Beijing, China, 2005, vol. 2, pp. 1164–1167

[9] C. T. Chou, S. W. Shih, W. S. Chen, and V. W. Cheng, “Iris recognition with multi-scale edge-type matching”, *Pattern Recogn.*, vol. 4, pp. 545–548, 2006.

[10] J. Thornton, M. Savvides, and B. V. K. Kumar, “An evaluation of iris pattern representations”, in *Proc. First IEEE Int. Conf. Biometrics: Theory, Appl. Systems BTAS 2007*, Washington, DC, USA, 2007, pp. 1–6.

[11] P. Yao, J. Li, X. Ye, Z. Zhuang, and B. Li, “Iris recognition algorithm using modified log-gabor filters”, *Pattern Recogn.*, vol. 4, pp. 461–464, 2006.

[12] C. E. Shannon, “Communication theory of secrecy systems”, *AT&T Bell Labs Tech. J.*, vol. 28, pp. 656–715, 1949.

[13] J. Daugman, “Results from 200 billion iris cross-comparisons”, Tech. Rep. UCAM-CL-TR-635, University of Cambridge, 2005.

[14] B. Schneier, *Applied cryptography*. Wiley, 1996.

[15] P. C. Van Oorschot, A. J. Menezes, and S. A. Vanstone, *Handbook of applied cryptography*. Boca Raton, Fl, USA: Crc Press, 1996.

Przemysław Strzelczyk received his M.Sc. degree in Information Technology in 2005 from the Warsaw University of Technology, and he is currently a Ph.D. candidate in the same field. Since 2005 he is with Reseach and Academic Computer Network (NASK) working for Biometric Laboratories, and within 2007 and 2009 he was research assistant at the Warsaw University of Technology. His main interests include: biometrics, security and methods of artificial intelligence.
e-mail: przemek.strzelczyk@nask.pl
Research and Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland