Pattern Recognition Letters 80 (2016) 70–75

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

A new cryptographic primitive for noise tolerant template security☆

Koray Karabina[a,b,*], Onur Canpolat[a]

[a] ZebraPET LLC, Privacy Enhancing Technologies, Boca Raton 33433, FL, United States
[b] Florida Atlantic University, Boca Raton 33431, FL, United States

ARTICLE INFO

Article history:
Received 25 February 2016
Available online 7 June 2016

Keywords:
Template security
Biometrics
Physically Unclonable Functions
Noise tolerance

ABSTRACT

We introduce a new cryptographic primitive for noise tolerant template security. Our primitive consists of two algorithms. The first algorithm extracts a noise tolerant and secure template of fixed length binary data. The noise tolerance of our scheme follows from the second algorithm that determines whether two templates originate from a pair of data where the Hamming distance belongs to a priori-fixed error tolerance interval. The proposed scheme does not require any encryption infrastructure and its security does not rely on the secrecy of parameters. We discuss the security of our scheme with respect to the irreversibility and indistinguishability of templates and present concrete efficiency and security analysis for practical parameters under certain assumptions. Our techniques can potentially be used to increase security of systems and privacy of users in certain applications including biometric authentication systems and embedded security systems based on Physically Unclonable Functions.

1. Introduction

1.1. Motivation

In a generic biometric authentication system, biometric sample (e.g. images of a fingerprint) is captured from a user via a sensor during enrollment. Distinctive characteristics (e.g. fingerprint minutiae) of the sample are derived during feature extraction; and a digital representation of the feature, called *template*, is stored in a database. Templates are used during authentication as the basis for comparison. Therefore, it is critical to protect biometric templates to minimize the security and privacy risks. It should be information theoretically or computationally infeasible to construct the biometric data from its protected template (*irreversibility*), and to cross correlate two protected templates (*indistinguishability*). Conventional cryptosystems are of very limited use in securing biometric systems because a user's biometric samples are not likely to be identical during enrollment and authentication, unlike noise-free and repeatable measurements in password-based and token-based authentication schemes.

1.2. Review

Biometric cryptosystems (BC) and *cancelable biometrics* (CB) [13,19] are two main approaches to resolve security and privacy issues in biometric schemes. In BC, a cryptographic key is generated from (or bound to) a biometric input, and the biometric template in the system is replaced by its secure sketch. A secure sketch allows to reconstruct the original template (or to retrieve the correct key) only in the presence of the correct biometric input during a successful authentication. In CB, some distortion transform (typically non-invertible with a user specific secret) is applied to the biometric input, and the transformed template replaces the original one in the system. A successful authentication is achieved only if the correct biometric input and the transform parameters are presented during verification. In another direction of research in biometrics, secure multiparty computation, encryption, and private information retrieval techniques are deployed [1,3,4,7,22]. This class will be defined as *keyed biometrics* (KB). In addition to these two main techniques, there is *hybrid biometrics* (HB), that blend BC, CB, KB, and multifactor authentication (i.e., passwords/PINs and tokens/smart cards) [8,12].

A wide range of sophisticated attacks, such as cross correlation, statistical, false acceptance, decoding, hill climbing, and masquerade attacks have been successfully applied to both BC and CB. For an extensive list of attacks on BC and CB, see [10,16,19,21]. Biometric schemes built on error correcting code based secure sketches are inherently vulnerable to various attacks. For example, it is

☆ "This paper has been recommended for acceptance by Ajay Kumar."

* Corresponding author at: Florida Atlantic University, Boca Raton, FL, 33431, United States. Tel.: +1 561 297 0809; fax: +1 562 297 2436.
E-mail address: kkarabina@fau.edu (K. Karabina).

http://dx.doi.org/10.1016/j.patrec.2016.05.016
0167-8655/

shown in [21] that the schemes (in the class of BC) implemented in [6,23] are not secure with respect to the indistinguishability and irreversibility notions under plausible assumptions. In a recent paper [10], new masquerade attacks on CB are proposed. Biometric authentication schemes in the class of KB are not likely to suffer from attacks on BC and CB because the security of the scheme can naturally be reduced to the security of the underlying encryption mechanisms. More generally, the security of KB can be linked to the secrecy of the key in the system. However, KB is not as practical as BC and CB because of the computation and communication overheads.

1.3. Our contributions

As reviewed in Section 1.2, recent studies demonstrate that security and privacy of users can be severely compromised in biometric systems even though the underlying primitives are information theoretically or computationally secure in their defined settings. In this paper, we introduce a new technique to extract secure template of noisy data.

The key properties of our scheme that differ from previously known schemes are as follows. Our scheme does not require any kind of encryption infrastructure and its security does not rely on the secrecy of parameters. In particular, each system parameter and the database of secure templates can be made public at reasonable security levels. This may be an attractive feature for practical applications. We discuss the security of our scheme in detail in Section 3, and present concrete efficiency and security analysis and estimates under certain assumptions. Our scheme has a flexible setting for system parameters that offers various tradeoffs between security and efficiency. In Section 5, we explain how the security of our scheme in Section 2.3 can be enhanced at a cost of increasing the length of secure templates. In general, such flexibility may be hard to achieve especially in error correcting based schemes in BC. See Section 5 for more detail. Our scheme is highly non-linear compared to some previously proposed popular schemes in BC. Therefore, our scheme is potentially resistant against previously known cross correlation attacks that exploit the linear structure of the system; see Section 3. In Section 5, we show that the theoretical and practical security of our scheme can further be enhanced in several ways. In Section 2.3, we discuss several ways to increase the security of our scheme.

1.4. Applications

Our techniques can be adapted by biometric authentication schemes to increase security and privacy of users. This has been the main motive in Section 1.1. In Section 4, we present a simulated implementation to show that our algorithms can potentially be deployed in the particular case of a biometric authentication system that uses fixed length representation of fingerprints [23]. We provide concrete security and efficiency analysis in this particular case in Section 4 and Section 5. We should emphasize that, by construction (see our discussion on the correctness of Decomp), the integration of our scheme does not alter the performance of a system with respect to false accept and false reject rates. Our algorithms can also be adapted by embedded security systems based on Physical Unclonable Functions (PUFs) [17] to securely generate cryptographic keys from noisy data (see [5] and the references wherein for more detail on combining noise tolerant schemes and PUFs). The primitives we propose have strong potential to be applied in a wide range of scenarios including traditional authentication schemes (e.g. tolerating errors in human typed passwords), social media (e.g. friend matching applications), and location based services (e.g. finding nearby restaurants and friends).

2. NTT − Sec: The new construction

In this section, we propose a new method that provides **N**oise **T**olerant **T**emplate **Sec**urity of sensitive data x. We call our scheme NTT − Sec throughout the rest of this paper. In order to ease the presentation of our construction, we make the simplifying assumption that $x \in \{0,1\}^n$ is a binary string of length n for some positive integer n. We measure the noise between two data $x, y \in \{0,1\}^n$ by the usual Hamming distance function $d(\cdot,\cdot)$, where $d(x,y)$ counts the total number of indices at which the bits of x and y differ. This setting may be very restrictive for representing and comparing data in general. However, it is a valid and a common setting in practice as justified in several implementations of biometric systems that rely on fixed length representation of the biometric data [9,14,23].

2.1. Preliminaries

Let $\mathbb{F}_q$ be a finite field with q elements, where $q = p^m$ for some prime p and a positive integer m. For simplicity, we further assume that $p > 3$ and m is odd. Denote the order-$(q+1)$ cyclotomic subgroup of $\mathbb{F}_{q^2}^*$ by $\mathbb{G}$. Let $\mathbb{F}_{q^2} = \mathbb{F}_q[\sigma]/\langle f(\sigma)\rangle$, where $f(\sigma) = \sigma^2 - c$ such that $c \in \mathbb{F}_q$ is a quadratic non-residue. It is known that every non-identity element in $g = g_0 + g_1\sigma \in \mathbb{G}$ can be uniquely represented by an element $\alpha = (g_0+1)/g_1 \in \mathbb{F}_q$ such that $g = (\alpha+\sigma)/(\alpha-\sigma)$.

In particular,

$$\mathbb{G} \setminus \{1\} = \left\{ \frac{\alpha+\sigma}{\alpha-\sigma} : \ \alpha \in \mathbb{F}_q \right\}, \tag{2.1}$$

and given any $g_0 + g_1\sigma \in \mathbb{G} \setminus \{1\}$, the above representation can be obtained by setting $\alpha = (g_0+1)/g_1$; see [20].

Now, let $\mathcal{F} = \{\alpha_\sigma = (\alpha+\sigma)/(\alpha-\sigma) : \ \alpha \in \mathbb{F}_p\}$, and consider the k-product set

$$S_k = \left\{ \prod_{i=1}^{k} v_i : \ v_i \in \mathcal{F} \right\},$$

for some positive integer k. Clearly, $S_k \subset \mathbb{G}$ and so non-identity elements in S_k are of the form $x_\sigma = (x+\sigma)/(x-\sigma)$ for some $x \in \mathbb{F}_q$. Furthermore, each such element in S_k can symbolically be written as

$$x_\sigma = \prod_{i=1}^{k} \frac{\alpha_i+\sigma}{\alpha_i-\sigma} = \frac{f_0(e_1,\ldots,e_k) + f_1(e_1,\ldots,e_k)\sigma}{f_0(e_1,\ldots,e_k) - f_1(e_1,\ldots,e_k)\sigma} \tag{2.2}$$

$$= \frac{f_0/f_1 + \sigma}{f_0/f_1 - \sigma}, \tag{2.3}$$

where $f_0 = \sum_{j=0}^{\lfloor k/2 \rfloor} e_{k-2j} c^j$, $f_1 = \sum_{j=0}^{\lfloor (k-1)/2 \rfloor} e_{k-2j-1} c^j$, $e_0 = 1$, and $e_i = e_i(\alpha_1,\ldots,\alpha_k)$ is the ith elementary symmetric polynomial in $\alpha_1,\ldots,\alpha_k$. This identification verifies that given any $x_\sigma \in S_k$, one can efficiently recover $v_i \in \mathcal{F}$ with $x_\sigma = \prod_{i=1}^{k} v_i$ when $k \le m$ as follows:

1. Use Weil restriction to the equation $f_0 - f_1 x = 0$ and obtain m linear equations over $\mathbb{F}_p$ with k unknowns $e_1,\ldots,e_k$.
2. Find a solution $(e_1,\ldots,e_k)$ with $e_i \in \mathbb{F}_p$ to this linear system of equations[1].
3. Construct the polynomial
$$P(X) = X^k - e_1 X^{k-1} + e_2 X^{k-2} + \cdots (-1)^k e_k. \tag{2.4}$$
4. Determine the set of $\mathbb{F}_p$-roots (counted with multiplicities) of the polynomial P, and construct the ordered sequence

[1] The existence of a solution is guaranteed by the definition of S_k and the fact that $x_\sigma \in S_k$.