Yale University
Department of Computer Science

P.O. Box 208205
New Haven, CT 06520–8285

Strong Theft-Proof Privacy-Preserving Biometric Authentication

Ewa Syta   Michael J. Fischer   Abraham Silberschatz
Gina Gallegos García   Bryan Ford

YALEU/DCS/TR-1455
May 25, 2012

Strong Theft-Proof Privacy-Preserving Biometric Authentication

Ewa Syta
Computer Science
Yale University
New Haven, CT 06520-8285
ewa.syta@yale.edu

Michael J. Fischer
Computer Science
Yale University
New Haven, CT 06520-8285
michael.fischer@yale.edu

Abraham Silberschatz
Computer Science
Yale University
New Haven, CT 06520-8285
abraham.silberschatz@yale.edu

Gina Gallegos García
National Polytechnic Institute
of Mexico
ggallegosg@ipn.mx

Bryan Ford
Computer Science
Yale University
New Haven, CT 06520-8285
bryan.ford@yale.edu

ABSTRACT

Biometric authentication offers many benefits ranging from strong security guarantees to user convenience, however, remote authentication poses unique challenges which are not fully addressed by biometrics alone. We propose a new remote authentication protocol that combines possession-based authentication and biometrics in a way that conquers the main weaknesses of both authentication methods. Our protocol offers strong protection to biometric data. It is theft-proof, guarding against attacks based on stolen or lost tokens. It is also privacy-preserving with respect to the users' biometric identities as well as actions performed using those identities.

In contrast to knowledge-based authentication, where passwords or PIN numbers may be updated freely, biometric data cannot be changed and therefore attacks on biometric templates are severe in consequences. To address this issue, our protocol handles biometric templates in a novel way - they are never directly stored, transmitted or made available to the verifying party. Identity verification is based on the difference between the biometric template provided in the enrollment phase and the one provided during verification. A user is authenticated only if the difference is sufficiently close to 0.

Authentication information is stored on a token, for instance a smart card, and is protected by biometric techniques to ensure that the token can only be used by its legitimate owner. User's identity is created with respect to a special blinding factor used to create a blinded biometric template, not the biometric data itself. Such approach offers two major benefits: biometric data protection and unlinkability of user's actions.

1. INTRODUCTION

The fast pace of technological advances bringing faster and cheaper computing devices and almost constant network access have changed the ways people utilize services available online. People use the Internet to perform more and more sensitive transactions like accessing their financial or medical records, making purchases, or communicating with others. Users' expectations of those transactions have changed too. They display a higher level of awareness of and need to protect themselves from threats, for example identity theft, arising from Internet activities [34, 36]. Furthermore, users not only expect their personal information to remain private but also the activities they engaged into while using protected resources requiring identity verification.

Almost every transaction require some form of identity verification to ensure that only legitimate users are granted access to protected resources. Authentication is a process of verifying user's identity based on some authentication factor presented by the user. From users' point of view authentication should be convenient, secure and privacy-preserving while a high level of assurance that users are indeed who they claim to be is especially important to the verifying parties.

The following properties define a successful remote authentication protocol and have proven to be challenging to achieve in practice.

1. *Security.* Authentication process needs to be robust and strong to resist various attacks, ranging from passive to active attacks.

2. *Privacy.* Users' personal information disclosed in order to authenticate themselves must be protected. Additionally, transactions performed with multiple service providers should remain *unlinkable* so users' actions cannot be tracked and later on linked together revealing personal information and possibly compromising their privacy.

3. *Assurance.* Authentication process must be based on reliable authentication factor(s) that are not easily compromised so that parties involved can trust that the authentication process will allow to successfully verify the identity of the proving party only if they are who they claim to be.

4. *Usability.* It's equally important for users and service providers that the authentication process can be completed in an efficient and uncomplicated manner. If the process is cumbersome or obtrusive, it will severely affect user acceptance.

There have been a number of authentication solutions proposed which can be categorized as knowledge-based, possession-based, or biometrics depending on the kind of factor they use to verify the claim of identity. Knowledge- based authentication ("something you know") is the most popular kind of authentication based on a shared secret (e.g., a password) that has to be provided in order to prove an identity claim. While passwords are easy to use, they pose serious security risks if used inappropriately. Possession-based authentication ("something you have") requires a user to demonstrate possession of a certain device (e.g., hardware token, smart card, or a mobile device). This method is convenient but imposes certain costs on users but more importantly quickly becomes obsolete if the device is lost. Biometric-based authentication ("something you are") uses biometric characteristics (e.g., a fingerprint, voiceprint, or hand geometry) for identity verification. Biometrics offer multiple benefits including non-repudiation and ease of use. Biometric data cannot be lost or forgotten and is constantly available. Certain biometric characteristics never change and can be measured quickly and unobtrusively. On the one hand, biometric data is unique to a person and therefore is an excellent way to define user's identity. On the other hand, the uniqueness of a biometric data poses a serious risk if it is ever compromised.

All three categories of authentication methods have their own advantages and disadvantages and have been extensively employed in different authentication systems over the years. Currently, passwords are the method of choice for user authentication and have been for many years [25]. Although password-based authentication does not provide the above properties, it is used to grant access to sensitive services, like online banking or medical records, because of a lack of better alternatives.

The major contribution of this paper is the new approach to remote authentication. The protocol we propose is strong, theft-proof and privacy-preserving. Biometrics coupled with possession-based authentication offer high security guarantees and usability. The new way to handle biometric data used for authentication gives higher privacy protection for users.

The rest of the paper is structured as follows. Section 2 gives an overview of biometrics and outlines challenges of remote biometric authentication. Section 3 introduces our protocol and details are given in Section 4. The security analysis of the protocol is included in Section 5. Section 6 outlines other solutions to remote biometric authentication and Section 7 concludes.

2. BIOMETRIC AUTHENTICATION

Biometrics have been long recognized as an excellent building block for authentication protocols. Biometric authentication offers a higher level of confidence that users are who they claim to be as well as convenience and usability as users are relieved from the need to remember multiple user names and passwords.

Biometric authentication uses unique characteristics, physiological (e.g., a fingerprint or iris pattern) and behavioral (e.g., voice print or gait), of a human body to verify the identity of a person [12]. Biometric characteristics used in biometric systems are universal, unique, permanent, and collectable [7]. Such characteristics are exceptionally suitable for authentication purposes as they offer non-reputability, are constantly available and cannot be lost or forgotten.

2.1 Remote Biometric Authentication

A biometric system typically consists of five components: a sensor, feature extractor, template database, matcher, and a decision module. An attacker may choose to exploit individual system components or the communication channels between them [29, 14]. These attacks are much easier to accomplish in case of remote biometric authentication, which refers to the process of performing authentication over a network. However, such approach makes authentication more universal and flexible as the parties can be in different physical locations. While biometric authentication on a standalone system offers protection from many adversarial actions as normally users must appear in person and all system components are in the same, frequently attended location, it is unsuitable for many applications where authentication is needed.

Remote biometric authentication remains a very challenging task due to the fact that system components are typically distributed between the proving and verifying party, and biometric data is often transported over the network and made available to the verifying party.

2.2 Security of Biometric Templates

The perception and acceptance of biometric systems depends on the security of biometric data. Users expect high guarantees that their biometric data will remain secure [1]. A biometric authentication protocol consists of two phases, the enrollment phase and the verification phase. In the enrollment phase, a biometric template is created based on a biometric sample and stored for comparison purposes in the verification phase. During the verification phase, a fresh biometric sample is obtained, a new biometric template is created and then compared against the reference template.

The uniqueness of biometric data, a cherished feature of biometrics, is also the source of security and privacy concerns. Typically, biometric templates are made available to the verifying party for the purpose of comparison. This poses a risk of a serious attack in which biometric data is intercepted during transmission or stolen from the verifying party. Unlike passwords and other knowledge-based factors, biometric data cannot be reset or changed. Consequently, if compromised, it is potentially unusable for authentication purposes as it might be used to successfully impersonate an individual.

For this reason, biometric templates security becomes a crucial issue, especially in case of remote authentication, as those attacks are potentially most damaging to a biometric system. Attacks on templates can lead to the following vulnerabilities: a template can be replaced by an attacker's template to gain unauthorized access, biometric data can be retrieved from the template, or the stolen template can be replayed to the matcher [13].

There are two major factors that play an important role to template security: the way a biometric template is generated from a biometric sample and the location where the template is stored.

A biometric template is created based on raw biometric data and contains information about extracted features. Standard cryptographic solutions, like encryption or hashing, are unusable for protecting templates because even if two templates are generated using the same biometric data, they are never exactly be the same. For this reason, a common approach is to store transformed versions of templates.

There are two categories of schemes for protecting templates: features transformation and biometric cryptosystems [11, 13].

In feature transformation schemes, a transformation function f is applied to biometric data during the enrollment phase. In the verification phase, the same transformation function is applied and the transformed templates is compared against the transformed reference template. However, it has been shown that in some cases it is possible to recover biometric data from biometric templates [30, 33]. For this reason, template security cannot only be based on such approach.

In biometric cryptosystems templates are accompanied by helper data used to extract a cryptographic key from a template during verification. Matching is performed indirectly by verifying the correctness of the key extracted from the template submitted for verification. Helper data needs to be carefully designed as it based on, and therefore may leak, specific biometric features.

There are four main locations for storing biometric templates: portable tokens, central databases, sensors, and individual workstations [27], with the two former being the most popular options. A portable token, for example a smart card or mobile device, allows users to secure their biometric templates and gives them a sense of control over their personal data. However, issues arise when tokens are lost as their content is usually unsecured. A central database makes it possible for users to easily authenticate from multiple locations as templates are constantly available for verification. On the other hand, the database needs to be kept secure and may become a target of attacks because of its valuable content. Furthermore, central storage of templates causes privacy concerns because all authentication attempts go through a single point potentially revealing users' actions.

2.3 Privacy Concerns

It is impossible to avoid a disclosure of personal or personally identifiable information (PII) when it comes to authentication. Users are required to supply, often excessive, amounts of information in order to establish and then verify their identity. Such information may be easily misused and result in identity theft, information linkage across different providers, or secondary uses of supplied information [26].

These threats are especially evident in case of biometric authentication. A biometric template is based on characteristics which uniquely identify an individual and cannot be changed. Such template has user's identity "embedded" into it and therefore defenses in case of its compromise are frequently very limited [28].

Privacy issues in biometric systems are two fold. Firstly, user's biometric data might be compromised. Secondly, biometric data can be used to identify an individual and then successfully track his or her activities performed using the same biometric identity or identity based on the same features.

When it comes to remote biometric authentication, the fact that biometric templates are often sent over a network and made available to the verifying party is particularly troublesome and raises privacy concerns. For these reasons, privacy in biometric authentication has become a significant issue and must be addressed in any successful authentication protocol.

3. PROTOCOL OVERVIEW

We propose a novel approach to remote biometric authentication which builds upon two different kinds of authentication methods to take advantage of the benefits of *two factor* authentication. Two factor authentication adds an extra layer of security by requiring two different factors to be presented during verification. Our protocol combines possession-based authentication and biometrics. A possession-based factor, for example a smart card, stores authentication information. The token is protected from an unauthorized use by employing biometric techniques to ensure that only the legitimate owner can use it.

Two-factor authentication is hardly a new idea, and in fact, it has been widely used in practice. There have been a number of remote biometric authentication schemes using smart cards proposed as listed in Section 6. In contrast with most of those methods, our protocol does not require a password or a synchronized clock. Additionally, the computational and network requirements are minimal on both sides.

Our protocol provides a response to the remote authentication challenges identified in Section 2. It comes from the way we combine biometrics and possession-based authentication to take advantage of the benefits both methods provide and to conquer their weaknesses.

The main contribution and novelty of this protocol lies in the way biometric data is handled. Biometric templates are *never* directly stored, transmitted or made available to the verifying party. The templates are protected by employing blinding factors which are simultaneously created by the proving and verifying parties.

Identity verification is based on the difference between two biometric templates, the one provided in the enrollment phase and the one supplied in the verification phase, rather than the templates themselves. A blinding factor protects the stored template and binds the user to his token. Authentication information, including the blinded biometric template, is stored on a token, which is protected from an unauthorized use even if it is lost or stolen.

Privacy protection comes from the ability to create multiple, unlinkable identities based on the same biometric template. User's identity is created with respect to the blinding factor which is known to the verifying party. This results in two benefits. First, the verifying party does not need to have access to the actual biometric data. Second, a user can create multiple identities with different verifying parties using the same biometric template.

A template is blinded with a blinding factor unique to the verifying party. During verification, the verifying party uses its unique factor to "unblind" a difference between the reference and verification templates, never the templates themselves. This approach allows to create multiple, fully independent *personas* based on the same biometric template. Each persona represents an identity as seen by the verifying party and can be used for transactions with that party. While each persona is based on the same biometric identity, none of the verifying parties can establish which one, even if they are colluding. This creates a separation and *unlinkability* of biometric identities and transaction performed using those identities. Users can utilize the same token and a biometric template to authenticate to different service providers without the fear of compromising their privacy as their transactions cannot be linked together between providers.

The protocol is generic enough to work with different kinds of biometrics. Depending on the desired level of security, cost and user convenience, the protocol can use biometric templates based on any distinctive and measurable characteristics that are suitable for biometric protocols.

The properties we wish the protocol to satisfy are defined as below.

- *It is strong.* The protocol handles biometric data in a secure way and offers strong protection of biometric templates. It guards against attack exploiting an unsecured communication channel, compromised proving party or compromised verifying party.

- *It is theft-proof.* A possession-based token stores authentication information and it is protected from an unauthorized use by biometrics. A user is bound to a particular token based on her biometric data. In order to authenticate, the proving party needs to be in possession of her token and the token can only be used if she is the legitimate user. If the token is lost or stolen, it is unusable. It cannot be used to falsely authenticate as the proving party or to retrieve biometric data.

- *It is privacy-preserving.* The protocol handles biometric data in a way that prevents it from being exposed as it is never directly stored or transmitted. Additionally, user's identity is created with respect to a special blinding factor, not the biometric data itself. Therefore, a user can create different identities based on the same biometric sample and safely use it for different online transactions.

Our protocol meets the properties defined in Section 1 in the following way.

1. *Security.* The protocol resists attacks based on information obtained after fully compromising either the proving or the verifying party. However, it does not offer protection if both parties are simultaneously compromised by the same attacker, which is a property extremely difficult, if not impossible, to achieve.

2. *Privacy.* The privacy protection is two-fold as explained above. Biometric data is protected from exposure because of the unique verification method. Moreover, users' actions are unlinkable across transactions with different service providers performed using the same biometric identities. Those identities are created with respect to a blinding factor, not the biometric data itself, so colluding providers are not able to uniquely identify and match users.

3. *Assurance.* Two authentication factors are used for identity verification. A user needs to be in possession of a token and present a biometric sample that binds her to the token. This approach gives a high level of confidence that the user is who she claims to be.

4. *Usability.* Even though users are required to keep tokens and provide biometric samples, it is not burdensome. Biometric characteristics are constantly available and can be measured quickly while tokens are small, easy to carry, and frequently already possessed by users (e.g., mobile phones or PDAs).

4. PROTOCOL DESCRIPTION

The authentication process is performed between a proving party (Peggy, the user) and a verifying party (Victor, the authentication server). The protocol consist of two phases: enrollment (Figure 1) and verification (Figure 2).

4.1 Enrollment Phase

Before Peggy can participate in the protocol, she must first be enrolled into the system by an *enrolling agent.* Depending on the application-specific security requirements, the enrolling agent can be an independent, trusted enrollment center or alternatively Victor can enroll users. In the following description we assume that Victor enrolls Peggy into the system.

During the enrollment process, Peggy and Victor securely establish a shared secret s, which will be used to seed a cryptographically secure pseudorandom number generator G. The generator consists of a pair (S, R), which defines the *next state* and *output* functions respectively. Details and requirements for the pseudorandom number generator as well as possible solutions for a secure seed exchange are given in Section 4.3.

Peggy obtains a token and "commits" to that token and authentication information stored on it by providing a biometric sample. The sample is used to generate a biometric template, which is in turn blinded and stored on the token. A biometric sample is obtained using an external sensor or a sensor built into the token depending on the kind of token used as described in Section 4.7.

1. Peggy and Victor agree on a random seed s to seed the output function R of a pseudorandom number generator G.

2. Victor sets $T_{s_0} = r_0$, where $r_0 = R(s)$ and keeps the next state of G denoted as $S(s_0)$.

3. Peggy obtains a biometric template P and calculates $T_{u_0} = P \oplus r_0$, where $\oplus$ is an exclusive-OR operation and $r_0 = R(s)$.

4. Peggy securely erases P and r_0, and keeps the blinded template T_{u_0} and the next state of G, $S(s_0)$.

Peggy's token stores T_{u_0} and $S(s_0)$ the next state of G. Victor keeps T_{s_0} and the next state of G, $S(s_0)$. If Victor is not the enrolling agent, he gets these information from the agent.

After the enrollment phase:

- Peggy has $T_{u_0} = P \oplus r_0$ and $S(s_0)$.

- Victor has $T_{s_0} = r_0$ and $S(s_0)$.

4.2 Verification Phase

To perform the verification phase, Peggy must be in possession of the token which was issued to her upon the enrollment into the system. In order to authenticate herself, she uses the token or the external reader to obtain a new biometric sample and generates a fresh template P'_i. Then, she calculates an authentication message $W_i = P'_i \oplus T_{u_{i-1}}$ for all authentication attempts $i = (1, \ldots, k)$ and sends it to Victor. After k authentications, Peggy and Victor reestablish the authentication information. The parameter k is based on application specific requirements.

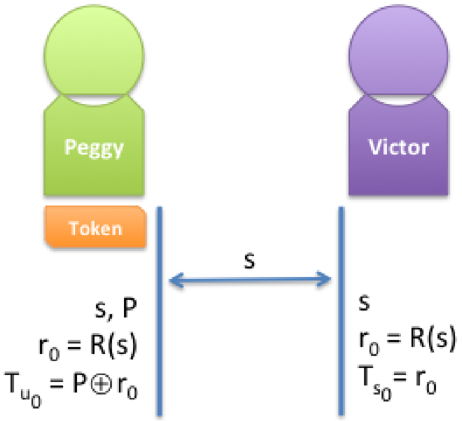

Figure 1: Enrollment Phase

Upon receiving W_i from Peggy, Victor computes $V_i = T_{s_{i-1}} \oplus W_i$ and accepts Peggy's claim of identity if $V_i \approx 0$. Assuming the generators are in sync,

$$\begin{aligned}
T_{s_{i-1}} &= r_0 \oplus \cdots \oplus r_{i-1}, \\
T_{u_{i-1}} &= r_0 \oplus \cdots \oplus r_{i-1} \oplus P, \\
V_i &= r_0 \oplus \cdots \oplus r_i \oplus P \oplus r_0 \oplus \cdots \oplus r_{i-1} \oplus P'_i = P \oplus P'_i, \\
V_i &= \Delta(P'_i, P),
\end{aligned}$$

the difference between the two biometric templates. Victor's goal is to establish whether the authentication message W_i came from Peggy. If two templates are created based on a biometric sample from the same user, they will be very similar. Therefore, if the difference between P and P'_i is sufficiently "small" according to a security threshold τ then authentication succeeds and Victor accepts Peggy's claim of identity. We write $\Delta(P, P'_i) \approx 0$ to denote that the difference between P and P'_i is sufficiently close to 0 if it is less than the security parameter τ. Section 4.5 discusses the issue of making the verification decision.

Protecting the blinded template

Peggy uses her blinded biometric template T_u for multiple authentication attempts. After each authentication attempt the template will be updated in order to protect it. After each successful attempt, both generators step to produce a new blinding factor, which is mixed into the blinded template. T_{u_i} is updated by generating the next value r_i in the sequence, and adding to the template as follows.

$$\begin{aligned}
T_{u_i} &= T_{u_{i-1}} \oplus r_i \\
T_{u_i} &= P \oplus r_0 \oplus r_1 \oplus \cdots \oplus r_{i-1} \oplus r_i
\end{aligned}$$

At this point, $T_{u_{i-1}}$ and r_i are both securely erased from memory and only T_{u_i} and the next state of G, $S(s_i)$ are retained. Similarly, Victor uses R to obtain the next value r_i to update T_{s_i}.

$$\begin{aligned}
T_{s_i} &= T_{s_{i-1}} \oplus r_i \\
T_{s_i} &= r_0 \oplus r_1 \oplus \cdots \oplus r_{i-1} \oplus r_i
\end{aligned}$$

As long as Peggy and Victor remain in sync, she will succeed to prove her identity. If they get out of sync, then resynchronization is needed. An approach to resynchronizing the generators is sketched in Section 4.6.

Steps performed by Peggy.

1. Obtain a fresh biometric template P'.
2. Calculate $W_i = T_{u_{i-1}} \oplus P'$ and send (ID, W_i) to Victor, where ID is Peggy's identifier.
3. If verification succeeded, update T_u: $T_{u_i} = T_{u_{i-1}} \oplus r_i$, where $r_i = R(s_{i-1})$.
4. Securely erase $P', W_i, T_{u_{i-1}}$ and r_i.

Steps performed by Victor.

1. Calculate $V_i = W_i \oplus T_{s_{i-1}}$.
2. Verify that $V_i \approx 0$ and if yes, accept Peggy's claim of identity.
3. If verification succeeded, update T_s: $T_{s_i} = T_{s_{i-1}} \oplus r_i$, where $r_i = R(s_{i-1})$.
4. Securely erase $W_i, T_{s_{i-1}}$ and r_i.

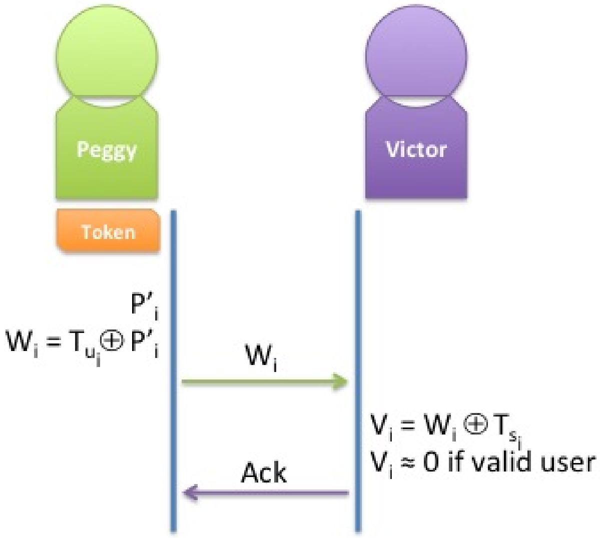
Peggy
Victor
Token
P'_i
$W_i = T_{u_i} \oplus P'_i$
W_i
Ack
$V_i = W_i \oplus T_{s_i}$
$V_i \approx 0$ if valid user

Figure 2: Verification Phase

4.3 Cryptographic Primitives

Secure seed exchange

The protocol relies on the parties' ability to establish a secret seed for pseudorandom number generation. The seed needs to be established in a secure manner in order to protect the sequences of blinding factors. The secret seed can be exchanged using one of the schemes to establish a shared secret, for example a key agreement protocol [10]. Alternatively, the seed can be sent through a secure channel.

Pseudorandom number generation

Blinding factors are randomly generated numbers. Both parties need to agree on a pseudorandom number generator. A pseudorandom number generator is a pair $G = (S, R)$, where S is the *next state* function and R is the *output* function. Let s be the secret seed. Then,

$$\begin{aligned} r_0 &= R(s) \\ s_0 &= S(s) \\ r_i &= R(s_{i-1}) \text{ for } i \geq 1 \\ s_i &= S(s_{i-1}) \text{ for } i \geq 1 \end{aligned}$$

We require G to be a cryptographically secure pseudorandom number generator that offers resistance to a strong kind of backtracking attack [16] that we call a *previous-outputs backtracking attack*. Specifically,

> *For every k, any algorithm that tries to predict r_k given s_k and $r_0, \ldots, r_{k-1}$ will succeed with only negligible advantage over random guessing.*

This property is needed to prevent an attacker who gains access to the card after stage k from recovering the blinding factor r_k that protects the biometric template P. The attacker gains all of the information stored on the card at the time of the attack, including the state s_k of G. We also have to assume that the attacker might have obtained $r_0, \ldots, r_{k-1}$ from observing the values W_i going over the channel. This information allows a previous-outputs backtracking attack to be carried out in an attempt to recover r_k.

While it seems likely that many cryptographically strong pseudorandom number generators are resistant to a previous-outputs backtracking attack, we are not aware of any such generator that has been shown to enjoy this little-studied property.

4.4 Template Generation

A biometric template is a representation of the features from a biometric sample. A feature extractor is a component of a biometric system responsible for generating templates. During the feature extraction process, key features of the biometric sample are located, selected, measured, encoded and then stored in form of a template. The template quality directly impacts the performance of a biometric system. We require that a feature extractor produces templates of high quality. More specifically, we assume that two templates created based on a biometric sample from the same user are "sufficiently" similar to be suitable for authentication purposes. Similarly, we require that two templates created based on biometric samples from different users are "sufficiently" different. The goal is to achieve an acceptable false rejection rate (FRR) and more importantly a low false acceptance rate (FAR).

4.5 Verification Decision

In biometric systems, a matcher and decision module are the two components directly involved in making the verification decision.

A matcher takes two biometric templates, the reference template created in the enrollment phase and the freshly obtain template from the verification phase, as input. Then, it calculates a *match* score which shows how similar the two templates are [11]. In case of our protocol, the matcher functionality is embedded into the protocol. The verifying party calculates $V_i = T_{s_{i-1}} \oplus W_i$ which defines $\Delta(P, P'_i)$, the difference between two biometric templates. Therefore, the value of V_i defines the *difference score*. In other biometric authentication protocols, the authentication decision is based on the match score while in our it is based on the difference score. To express the difference score in terms of the match score we can say that the smaller the difference V_i, the higher the match score is.

The decision module takes a match score (in our case a difference score) as input and based on a predefined threshold parameter τ decides whether the two templates were created based on biometric samples from the same person. If the match score is greater than a predefined threshold τ, user's identity if verified. In our protocol, if the difference score if lower than τ, then the two templates are accepted as coming from the same user.

Choosing a proper value for τ is a challenging task. To have a high level of confidence that two templates were created based on samples from the same user, the difference should be very low. Hence, the value chosen for τ should reflect the desired level of security as well as the sensor and feature extractor's capabilities to create accurate templates. The goal is to balance the false rejection and false acceptance rates while ensuring a proper level of security.

4.6 Resynchronization

Biometric templates are protected by applying random blinding factors. Victor's ability to verify Peggy's identity depends on his ability to "unblind" the difference between two templates. This can only happen if the two pseudorandom number generators are in sync. While the values of T_u and T_s are updated only if an authentication succeeds, desynchronization issues are inevitable. The generators can go out of sync as a result of communication issues between Peggy and Victor during a legitimate authentication attempt or as a result of intentional attempts from an attacker.

If the generators are our of sync, Victor will not be able to correctly verify Peggy. The simple solution is to search forward in the sequence produced by G for some predefined distance n looking for a value of T_s that is the blinding factor. This approach will succeed if the authentication attempts fails due to Peggy's generator being ahead of Victor's. In less than n steps Victor will generate the proper blinding factor and correctly authenticate Peggy. After the correct authentication attempt, both generators will be in sync. However, if an attacker is trying to exploit the desynchronization of generators to gain unauthorized access, the resynchronization procedure will not give him an advantage greater than an advantage of random guessing the blinding factors, which is equal to an advantage of predicting an output of a cryptographically secure pseudo random number generator.

4.7 Tokens

The protocol makes use of possession-based authentication by using tokens to store and process authentication information. There are two different approaches to utilize tokens depending on the token's capabilities to obtain biometric samples.

A token without a sensor

In this case, the token is only used to store authentication information and to perform computations. It must be paired with a sensor to obtain a biometric sample and turn in into a template that is processed in the protocol. A token, depending on its kind, would be inserted or connected into a sensor. This implies certain level of trust that the sensor is not compromised and the communication channel between a sensor and token is protected. However, this approach makes it easy to utilize virtually any biometric characteristic or even use multiple characteristics to create a biometric template as different sensors could be paired with the same token.

Smart cards are the most obvious choice for such tokens. They have been extensively used for authentication applications. They are relatively cheap, small in size, and convenient to use [35]. Most smart cards offer enough computational power to easily perform the required operations, especially that the computational requirements of our protocol are minimal.

A token with a built-in sensor

Obtaining a biometric sample is a crucial part of the authentication process. If a token has a built-in sensor, it removes the assumption of a trusted communication channel and trusted sensor. Unfortunately, this approach precludes the usage of certain biometric characteristics which require specialized sensors that might not be combined with tokens because of their size, for example.

Mobile devices are an excellent option for such tokens. Most modern phones, PDAs, or tables, are equipped with a high resolution camera capable of capturing images suitable for authentication using several biometric characteristics like a fingerprint, facial geometry, or iris pattern [11]. Additionally, mobile devices make it possible to take advantage of less frequently utilized characteristics like voiceprint, keystroke or handwriting patterns, service utilization [6] or even gait [8].

5. SECURITY ANALYSIS

There is a plethora of attacks on authentication protocols. An attacker may try to compromise the proving party, verifying party or both. Our protocol offers protection from attacks on the verifying or proving party, but not on both simultaneously. We assume that all communication occurs over an unsecured channel. Below we consider our protocol in the context of attacks based on dishonest verifying party, compromised verifying party, compromised proving party, and attacks using authentication messages exchanged during successful authentication attempts.

Our main goal and concern is the security of biometric data as this is the most critical issue in case of biometric remote authentication.

5.1 Dishonest Verifying Party

Normally, verifying parties are trusted as they are the ones guarding protected resources to ensure that only legitimate users are allowed access. In case of knowledge-based authentication, the verifying party has no advantage in compromising prover's authentication factor because they are already in possession of it. Unlike biometric templates, passwords or PIN numbers do not carry any sensitive information. Therefore, in case of biometric authentication potential adversarial actions of the verifying party need be considered as well.

In our protocol, at any given moment, Victor is in possession of the blinding factor T_{s_i}, S_{s_i}, the next state of G, and all messages W_i he received up to this point. Each message W contains only the difference between two biometric templates; the actual biometric data is never directly made available to him. For this reason, Victor cannot compromise Peggy's biometric data based on the information exchanged or given to him within the protocol.

5.2 Compromised Verifying Party

The are two ways in which Victor can be compromised. An attacker may get one time access to the server and obtain currently stored authentication information or he may get continuos access allowing him to obtain current authentication information and information coming from all future authentication attempts.

In the first case, if Victor is compromised, the attacker gets the current blinding factor T_{s_i} and $S(s_i)$, the next state of G. Given this information, he will be able to authenticate himself as Peggy knowing T_{s_i}, because he can prepare a fake message W' that is close to T_{s_i} so that verification will succeed on Victor's side: $V_{i+i} = T_{s_i} \oplus W' \approx 0$. However, no information about Peggy's biometric data is compromised because none of the information Victor has contains any information about P.

In the second case, the attacker gets the current blinding factor T_{s_i}, $S(s_i)$, and a sequence of messages $W_i, W_{i+1}, \dots, W_n$ sent by Peggy in attempt to authenticate herself since she will not know that Victor got compromised.

$$\begin{array}{lll}
W_i & = P' \oplus T_{u_{i-1}} & = P' \oplus P \oplus r_0 \oplus \cdots \oplus r_{i-1} \\
W_{i+1} & = P' \oplus T_{u_i} & = P' \oplus P \oplus r_0 \oplus \cdots \oplus r_{i-1} \oplus r_i \\
W_{i+2} & = P' \oplus T_{u_{i+1}} & = P' \oplus P \oplus r_0 \oplus \cdots \oplus r_{i-1} \oplus r_i \oplus r_{i+1} \\
\cdots & = & \\
W_n & = P' \oplus T_{u_{n-1}} & = P' \oplus P \oplus r_0 \oplus \cdots \oplus r_i \oplus \cdots \oplus r_{n-1}
\end{array}$$

Knowing the next state of G, the attacker can obtain the future values of r: $r_{i+1}, r_{i+2}, \dots, r_{n-1}$. Each message W contains the difference between two templates blinded with a sequence of blinding factors r. By knowing the future factors, starting at r_{i+1}, the attacker can partially unblind messages $W_{i+2}, \dots, W_n$. However, he will not be able to fully unblind messages $W_1, \dots, W_{i+1}$ as he cannot retrieve previous values of r from the current state of G. Even if the attacker was able to do so, all he would recover is a sequence of differences between two biometric templates, each very close to 0. Therefore, the attacker cannot compromise Peggy's biometric data even if he takes full control of Victor and Peggy keeps authenticating herself.

5.3 Compromised Proving Party

The token stores a blinded biometric sample T_{u_i} and the next state $S(s_i)$. If it is ever lost or stolen, and is in possession of an attacker, he may try the following attacks.

Recover P using previously sent messages W

Assume that the attacker compromised the token after eavesdropping on the channel Peggy and Victor used to exchange authentication messages. If that is the case, then he knows a number of W_i's Peggy sent to Victor. Assuming the worst case scenario, the attacker will know all messages $W_1, W_2, \dots, W_i$, where

$$W_1 = T_{u_0} \oplus P' = P \oplus P' \oplus r_0$$
$$W_2 = T_{u_1} \oplus P' = P \oplus P' \oplus r_0 \oplus r_1$$
$$\vdots$$
$$W_i = T_{u_{i-1}} \oplus P' = P \oplus P' \oplus r_0 \oplus r_1 \oplus r_2 \oplus \cdots \oplus r_{i-1}$$

and

$$T_{u_i} = P \oplus r_0 \oplus r_1 \oplus \cdots \oplus r_{i-1} \oplus r_i$$

Now, the attacker may try to recover P or P' using the above information: W_i and T_{u_i}.

$$\begin{aligned} & W_i \oplus T_{u_i} \\ &= (T_{u_{i-1}} \oplus P') \oplus T_{u_i} \\ &= P' \oplus (P \oplus r_0 \oplus \cdots \oplus r_{i-1}) \oplus (P \oplus r_0 \oplus \cdots \oplus r_i) \\ &= P' \oplus r_i \end{aligned}$$

However, r_i was securely erased when T_{u_i} was updated and it was never included in any of the messages sent. Additionally, r_i cannot by recovered using the state of G and $r_0, \ldots, r_{i-1}$ (assuming they are known). For this reason, the attacker will not be able to recover P' or P and hence will learn at most $P' \oplus P$.

Recover P using the stored state of G

To recover P from T_{u_i} an attacker would have to obtain a sequence of blinding factors $(r_0, r_1, \ldots, r_i)$, because $T_{u_i} = P \oplus r_0 \oplus r_1 \oplus \cdots \oplus r_i$. However, we assume that the pseudorandom number generator does not allow to recover past values from the current state of G. Hence, the attacker cannot recover P through $S(s_i)$.

Recover P from the stored biometric template

The blinded biometric template T_{u_i} consists of $P \oplus r_0 \oplus \cdots \oplus r_i$. To learn P, the attacker would need to recover a sequence of blinding factors r. As shown above, it is impossible to recover the sequence from the stored state of G, hence, T_{u_i} cannot be used to learn P.

Recover P after forcing $P' = 0$

Another potential attack is related to the way W_i is created: $W_i = T_{u_{i-1}} \oplus P'$. If the attacker is (somehow) able to force the feature extractor to produce biometric template $P' = 0$, then $W_i = T_{u_{i-1}} \oplus 0 = T_{u_{i-1}}$. This would allow the attacker to recover $T_{u_{i-1}}$, which is just a blinded biometric template and as shown above it does not reveal any information about P.

Falsely authenticate as Peggy using the stored state of G

Knowing $S(s_i)$, the attacker can get r_{i+1}, the next r in the sequence. Then, the attacker may try to use this value to falsely authenticate as Peggy. Recall, that both pseudorandom number generators need to be in sync. One of the solution in case of desynchronization is to check a couple of the next values of r to see it they work. In this situation, an attacker can "force" the number generators to go out of sync and then send a value r_{i+1}. Then, Victor would calculate $V_i = r_{i+1} \oplus T_{s_{i-1}}$. However, given the way we update T_u and T_s, we have $T_{s_i} = r_0 \oplus r_1 \oplus \cdots \oplus r_i$, which is the difference between all blinding factors used so far. Thus, knowing the next value r in sequence does not help to falsely authenticate as Peggy because of the previous blinding factors used.

Falsely authenticate as Peggy using the stored blinded template

The token stores the blinded biometric template T_{u_i} and the attacker may try to use it to falsely authenticate as Peggy. The attacker sets $W' = T_{u_i}$ and sends it to Victor, who calculates $V_{i+1} = W' \oplus T_{s_i} = T_{u_i} \oplus T_{s_i} = P \oplus r_0 \oplus r_1 \oplus \cdots \oplus r_i \oplus r_0 \oplus r_1 \oplus \cdots \oplus r_i = P$.

If Victor is not collaborating with the attacker, he will not know that $V_{i+1} = P$ because he has only seen $\Delta(P, P')$, never P or P'. Therefore, he will disregard V_{i+1} and assume an unsuccessful authentication attempt. However, if Victor and the attacker are collaborating, then Peggy's biometric data is compromised. The same is always be true if both parties are simultaneously compromised by the same attacker or collaborating attackers.

5.4 Both Parties Compromised

If both parties are compromised, Peggy's biometric template will be compromised. The attacker will get T_{s_i} from the verifying party and T_{u_i} from the proving party. Then, $T_{s_i} \oplus T_{u_i} = P$.

5.5 Compromised Communication Channel

A passive attacker can always listen to messages being exchanged as we assume communication over an unsecured channel. He will see a series of $W_1, \ldots, W_n$, which are blinded differences between two biometric templates. Even if the attacker manages to "unblind" some of W's, he will only learn the difference between two templates, which is close to 0. The same attacker may try to reuse some of the W's to falsely authenticate as Peggy. However, each W is created with a fresh value r and therefore, using an already used authentication message W will not allow for successfull authentication.

6. RELATED WORK

Combining biometric and possession-based authentication is a very popular approach to remote biometric authentication, which was proposed in response to unsuccessful attempts to create a secure smart card and password-based remote authentication scheme. [19] was the first scheme which combined biometrics with a smart card and a password. However, the scheme was shown to succumb to masquerade [24] and conspiring [5] attacks. Later, the improved scheme of [24] was shown by [18] to be vulnerable to server spoofing attacks. The scheme was further improved by [20]. Another, more efficient scheme proposed by [21] enabled users to change their passwords and removed the requirement of a synchronized clock between the proving and verifying parties. However, the scheme was shown by [23] not to provide proper authentication and to be susceptible to the man-in-the-middle attacks. The resulting scheme was broken and then improved upon by [15].

Chaos-based cryptosystem proposed by [17] provides template privacy and security by applying chaotic encryption scheme using unique per transaction keys that are randomly and dynamically generated. The scheme was shortly shown by [39] to be vulnerable to a privileged insider's attacks and impersonation attacks by using compromised devices.

[2, 3, 9, 22, 37] provide two primitives, a fuzzy extractor and secure sketch, for turning biometric information into keys usable for any cryptographic application, and reliably and securely authenticating biometric data. Secure sketches paired with fuzzy extractors are widely used for biometric systems based on different biometric characteristics.

ZeroBio [31] is a zero-knowledge proof based approach to biometric remote authentication which allows the verifying party to authenticate the proving party while concealing prover's biometric data. The scheme improved upon the original ZeroBio protocol by lowering computational complexity and network traffic at the cost of a small decline of security level.

[4] showed how to achieve a strong privacy-preserving biometric-based authentication schemes by employing extended private information retrieval and proposed a different framework for remote biometric authentication, which separates biometric template storage from its processing during authentication. [38] formalized the concepts of identity privacy and transaction anonymity and proposed an approach to addressing privacy concerns in biometric remote authentication schemes by employing private information retrieval and homomorphic encryption. [32] surveyed different solutions that fit into the distributed biometric authentication model. While the scheme provide good privacy and security guarantees under the assumption of non-colluding servers, they result in rather high communication and computational costs. [32] also proposed a distributed biometric scheme that achieved reduced computational and database storage costs.

7. CONCLUSIONS

Remote biometric authentication faces significant challenges related to the uniqueness of biometric data. While biometrics are exceptionally suitable for authentication purposes, biometric templates carry sensitive information as user's identity is embedded into them. Consequently, biometric data protection is of utmost importance.

The protocol we proposed defines a new approach to remote biometric authentication. It combines biometric- and possession-based authentication in a way that secures tokens in case of a loss or theft, protects biometric data from exposure, and allows the reuse of biometric templates for authentication with multiple parties without the fear of privacy compromise.

8. REFERENCES

[1] O. Bernecker. Biometrics: security: An end user perspective. *Information Security Technical Report*, 11(3):111 – 118, 2006.

[2] X. Boyen. Reusable cryptographic fuzzy extractors. In *ACM CCS 2004, ACM*, pages 82–91. ACM Press, 2004.

[3] X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky, and A. Smith. Secure remote authentication using biometric data. In *In EUROCRYPT*, pages 147–163. Springer, 2005.

[4] J. Bringer, H. Chabanne, D. Pointcheval, and Q. Tang. Extended private information retrieval and its application in biometrics authentications. In *Proceedings of the 6th international conference on Cryptology and network security*, CANS'07, pages 175–193, Berlin, Heidelberg, 2007. Springer-Verlag.

[5] C.-C. Chang and I.-C. Lin. Remarks on fingerprint-based remote user authentication scheme using smart cards. *SIGOPS Oper. Syst. Rev.*, 38(4):91–96, Oct. 2004.

[6] N. Clarke and S. Furnell. Authentication of users on mobile telephones - a survey of attitudes and practices. *Computers and Security*, 24(7):519 – 527, 2005.

[7] R. Clarke. Human identification in information systems: Management challenges and public policy issues. *Information Technology & People*, 7(4):6–37, 1994.

[8] M. Derawi, C. Nickel, P. Bours, and C. Busch. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on*, pages 306 –311, oct. 2010.

[9] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM J. Comput.*, 38(1):97–139, Mar. 2008.

[10] S. Goldwasser and M. Bellare. Lecture notes in cryptography, 2001.

[11] A. Jain, A. Ross, and K. Nandakumar. *Introduction to Biometrics*. Springer, 2011.

[12] A. K. Jain and D. Maltoni. *Handbook of Fingerprint Recognition*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.

[13] A. K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP J. Adv. Signal Process*, 2008:113:1–113:17, Jan. 2008.

[14] A. K. Jain, A. Ross, and S. Pankanti. Biometrics: A tool for information security. *IEEE Transactions On Information Forensics and Security*, 1(2):125–143, 2006.

[15] S. Q. Jian-Zhu Lu, Shaoyuan Zhang. Enhanced biometrics-based remote user authentication scheme using smart cards. Cryptology ePrint Archive, Report 2011/676, 2011.

[16] J. Kelsey, B. Schneier, D. Wagner, and C. Hall. Cryptanalytic attacks on pseudorandom number generators. In S. Vaudenay, editor, *Fast Software Encryption*, volume 1372 of *Lecture Notes in Computer Science*, pages 168–188. Springer Berlin Heidelberg, 1998.

[17] M. Khan and J. Zhang. Implementing templates security in remote biometric authentication systems. In *Computational Intelligence and Security, 2006 International Conference on*, volume 2, pages 1396 –1400, nov. 2006.

[18] M. K. Khan and J. Zhang. Improving the security of 'a flexible biometrics remote user authentication scheme'. *Comput. Stand. Interfaces*, 29(1):82–85, Jan. 2007.

[19] J. K. Lee, S. R. Ryu, and K. Y. Yoo. Fingerprint-based remote user authentication scheme using smart cards. *Electronics Letters*, 38(12):554, 2002.

[20] Y. Lee and T. Kwon. An improved fingerprint-based remote user authentication scheme using smart cards. In *Proceedings of the 2006 international conference on Computational Science and Its Applications - Volume Part II*, ICCSA'06, pages 915–922, Berlin, Heidelberg,

2006. Springer-Verlag.

[21] C.-T. Li and M.-S. Hwang. An efficient biometrics-based remote user authentication scheme using smart cards. *J. Netw. Comput. Appl.*, 33(1):1–5, Jan. 2010.

[22] Q. Li, Y. Sutcu, and N. Memon. Secure sketch for biometric templates. In *In Asiacrypt*, pages 99–113. Springer-Verlag, 2006.

[23] X. Li, J.-W. Niu, J. Ma, W.-D. Wang, and C.-L. Liu. Cryptanalysis and improvement of a biometrics-based remote user authentication scheme using smart cards. *J. Netw. Comput. Appl.*, 34(1):73–79, Jan. 2011.

[24] C.-H. Lin and Y.-Y. Lai. A flexible biometrics remote user authentication scheme. *Computer Standards and Interfaces*, 27(1):19 – 23, 2004.

[25] J.-C. Liou and S. Bhashyam. On improving feasibility and security measures of online authentication. *Int. Journal of Information and Communication Technology*, 2(4):6–16, 2010.

[26] L. Millett and S. Holden. Authentication and its privacy effects. *Internet Computing, IEEE*, 7(6):54 – 58, nov.-dec. 2003.

[27] A. Patric. Usability and acceptability of biometric security systems. In *Proceedings of the Financial Cryptography Conference (FC04)*, volume 3110 of *Lecture Notes in Computer Science*, 2004.

[28] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition: Security and privacy concerns. *IEEE Security and Privacy*, 1:33–42, 2003.

[29] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.*, 40(3):614–634, Mar. 2001.

[30] A. Ross, J. Shah, and A. Jain. From template to image: Reconstructing fingerprints from minutiae points. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):544 –560, april 2007.

[31] T. Sakashita, Y. Shibata, T. Yamamoto, K. Takahashi, W. Ogata, H. Kikuchi, and M. Nishigaki. A proposal of efficient remote biometric authentication protocol. In *IWSEC'09*, pages 212–227, 2009.

[32] N. D. Sarier. A survey of distributed biometric authentication systems. In A. BrŽmme, C. Busch, and D. H§hnlein, editors, *BIOSIG*, volume 155 of *LNI*, pages 129–140. GI, 2009.

[33] W. J. Scheirer and T. E. Boult. Cracking fuzzy vaults and biometric encryption. In *in Proceedings of Biometrics Symposium*, pages 1–6, 2007.

[34] H.-P. Shih. An empirical study on predicting user acceptance of e-shopping on the web. *Information & Management*, 41(3):351 – 368, 2004.

[35] Smart cards and biometrics. A Smart Card Alliance Physical Access Council White Paper, March 2011. Publication Number: PAC-11002.

[36] B. Suh and I. Han. The impact of customer trust and perception of security control on the acceptance of electronic commerce. *Int. J. Electron. Commerce*, 7(3):135–161, Apr. 2003.

[37] Y. Sutcu, Q. Li, and N. Memon. How to protect biometric templates. In *Proceedings of the SPIE Conference on Security, Steganography and Watermarking of Multimedia Contents IX*, volume 6505, January 2007.

[38] Q. Tang, J. Bringer, H. Chabanne, and D. Pointcheval. A formal study of the privacy concerns in biometric-based remote authentication schemes. In *Proceedings of the 4th international conference on Information security practice and experience*, ISPEC'08, pages 56–70, Berlin, Heidelberg, 2008. Springer-Verlag.

[39] E.-J. Yoon and K.-Y. Yoo. A secure chaotic hash-based biometric remote user authentication scheme using mobile devices. In K. Chang, W. Wang, L. Chen, C. Ellis, C.-H. Hsu, A. Tsoi, and H. Wang, editors, *Advances in Web and Network Technologies, and Information Management*, volume 4537 of *Lecture Notes in Computer Science*, pages 612–623. Springer Berlin / Heidelberg, 2007.