A Survey of Distributed Biometric Authentication Systems

Neyire Deniz Sarier

Bonn-Aachen International Center for Information Technology
Computer Security Group
Dahlmann str. 2, 53113 Bonn Germany
denizsarier@yahoo.com

Abstract: In ACISP’07, Bringer et al proposed a new approach for remote biometric based verification, which consists of a hybrid protocol that distributes the server side functionality in order to detach the biometric data storage from the service provider. Besides, a new security model is defined using the notions of Identity and Transaction Privacy, which guarantees the privacy of the identity-biometrics relationship under the assumption of non-colluding servers. In this survey, we review the scheme of Bringer et al and the following biometric verification systems that improve upon it in terms of computation and communication complexity. In this context, we discuss about the recent result of Sarier, which describes a secure and efficient multi-factor authentication scheme with a different biometric storage method that results in reduced computation and database storage cost.

Keywords: Remote authentication, Biometric template security, Identity privacy, Distributed systems, Private Information Retrieval

1 Introduction

Biometric authentication systems are used in order to verify the claimed identity of a user based on his biometric characteristics. Although authentication information should be kept confidential, for biometrics this cannot be guaranteed since it is very easy to obtain biological information such as fingerprint, iris or face data through fingerprint marking or using a camcorder. In order to avoid the imitation attacks, biometric measurements should be performed in controlled environments, for instance under the supervision of an operator. Otherwise, spoof-resistant sensors and/or multi-factor authentication techniques should be employed that combine biometrics with token and/or password based authentication methods.

Biometric authentication could be categorized broadly as remote server or client end authentication, where in the first case, the remote server stores the reference biometric data and performs the matching. Although biometrics is assumed as public data, it should not be easy to obtain the biometric data by compromising the central server, where the biometrics of each user is often associated with his personal information. This also affects the social acceptance of the biometric systems especially when biometric data are stored in a central database which can be vulnerable to internal or external attackers.
The security and privacy protection of remote biometric-based verification systems is enhanced by implementing distributed biometric systems, where the goal is to detach the biometric data storage from the service provider and to guarantee the notions of identity and transaction privacy, which have been recently introduced as a new security model for biometric verification. In this model, the user $U$ registers its biometric template in cleartext or in encrypted form at the database $DB$. Besides, $U$ registers his personal information (i.e. identifier) and the index of the database storage location of his biometrics at the service provider $SP$. For biometric verification, $U$ encrypts his biometrics using a homomorphic encryption scheme and sends this to $SP$, which retrieves the index of $U$ to be used in a Private Information Retrieval (PIR) protocol between $SP$ and $DB$. Finally, a decision is made after decryption or in the encryption domain by exploiting the homomorphic properties of the underlying encryption scheme. Current systems implementing this approach provide provable security in this new model, however, the (public) biometric data are stored as encrypted using the relatively slow public key schemes to provide the privacy of the identity-biometrics relation resulting in high database storage costs due to ciphertext expansion. Besides, some systems require a detached verification unit $VU$ for the final decision, which increases the overall complexity of the system. Consequently, one has to design a secure and efficient remote biometric verification scheme for a distributed system with a detached biometric database, which minimizes the costs of storage, encryption and communication and thus, the scheme also becomes applicable to large scale systems. In this survey, we consider the schemes designed in the framework of Bringer et al.’s security model. The present contribution is largely based on the author’s paper presented at ICB’09 [13] with a special focus on the complexity of the PIR.

2 Definitions and Preliminaries

2.1 Distributed Systems with Detached Biometric Storage

In recent years, the privacy protection and the secure storage of the biometric templates were addressed in a number of papers. As it is noted in [15], privacy protection not only means the attackers inability to compromise the biometric template but also the protection of the sensitive relationship between the identity and the biometric information of the user. To achieve this property, the storage of personal identity information should be separated from the storage of biometrics using the distributed structure of [4, 5, 6, 15, 13, 3], which is composed of the user $U_i$, the sensor client $SC$, the service provider $SP$ and the database $DB$. Some systems require the use of a smartcard for a multi-factor authentication [13] and/or a detached verification unit $VU$ (or a Matcher) [4, 3]. The entities of the system (i.e. $U_i$, $SC$, $SP$, $VU$ and $DB$) are independent (i.e. not colluding) of each other and they are all assumed to be malicious except for the sensor client. This way, $SP$ cannot obtain the biometrics of the user and can have business agreements with different parties that make the sensor client available to users at different locations. Also, $DB$ could function as a trusted storage for different $SP$’s. Since $SC$ captures the biometric data and performs the feature extraction, this component could be installed as a Trusted Biometric Reader or
biometric smartcard readers could be used as in [1].

2.2 Assumptions

- Liveliness Assumption: This is an indispensable assumption for any biometric system as it guarantees with high probability that the biometrics is coming from a live human user.
- Security link Assumption: To provide the confidentiality and integrity of sensitive information, the communication channel between $U_i, SC, SP, DB$ and $VU$ should be encrypted using standard protocols.
- Collusion Assumption: Due to the distributed system structure, we assume that $U_i, DB, VU$ and $SP$ are malicious but they do not collude. Additionally, the sensor client is always honest.

2.3 Security Requirements

2.3.1 Identity Privacy:
Informally, this notion guarantees the privacy of the sensitive relationship between the user identity and its biometrics against a malicious service provider or a malicious database even in case of multiple registrations of the same user with different personalized user-names. Briefly, it means that the service provider or the database (or an attacker that has compromised one of them) cannot recover the biometric template of the user [15].

2.3.2 Transaction Privacy:
Informally, transaction anonymity means that a malicious database cannot learn anything about the personal identity of the user for any authentication request made to the service provider [15].

The formal definition of the notions Identity and Transaction privacy could be found in [4, 5, 6, 15, 3].

2.4 Private Information Retrieval (PIR)

In order to provide Transaction Privacy, the systems in [4, 5, 6, 15, 13] employ a number-theory based PIR system, which allows the $SP$ to retrieve the $i$-th bit (more generally, the $i$-th item) from the $DB$ consisting of $n$ bits while keeping the value $i$ private. The PIR of [7] has an additional benefit of retrieving more than one bit, and in particular many
consecutive bits [10]. In this context, a Private Block Retrieval (PBR) protocol enables a user to retrieve a block from a block-database and the PIR/PBR setting of [5] consists of the DB containing a list of $N$ blocks $(R_1, ..., R_N)$ and the SP, which runs a PBR protocol to retrieve $R_i$ for any $i \in [1, N]$. The communication cost of the single database PIR system of [7] has currently the best bound for communication complexity of $O(\log(n) + b)$ for an $n$-bit DB, where $b$ is the bit-length of the block to be retrieved. However, the computational cost of number-theory based PIR’s is roughly a modular multiplication per bit of DB, which limits the usability of these schemes except for very small DB’s. In [8], the authors suggest to use batch codes to amortize the computational cost of PIR with a moderate increase in the communication cost, which is already very low. When the SP wants to retrieve $k$-bits (not necessarily consecutive) out of $n$-bit DB, batch code constructions can achieve $k^{1+o(1)}$ communication and $n^{1+o(1)}$ computation. Recently, [9] proposed a lattice-based PIR scheme, which is 100 times faster than number-theory based PIR’s and has reasonable communication.

2.5 Homomorphic Encryption

To construct a number-theory based PIR protocol and/or to make an authentication decision in the encryption domain based on a certain metric, we need a secure cryptosystem that is homomorphic over an abelian group. For a given cryptosystem with $(Keygen, Enc, Dec)$, the message space $M$ and the ciphertext space $C$ that are both groups, a homomorphic cryptosystem satisfies $Dec(Enc(a) * Enc(b)) = a * b$, where $a, b \in M$ and $*, \cdot$ represent the group operations of $M, C$ respectively.

2.6 Secure Sketches

Most of the schemes in the literature assume that the biometrics is represented as a fixed binary string, which is usually obtained by quantizing the original biometric template via a scaler quantizer and the resulting binary string is combined with a secure sketch or fuzzy extractor using binary error correcting codes. The main purpose of a secure sketch is to correct the noise in the biometric measurement by using some public information $PAR$, which is derived from the original biometric template $b$. A secure sketch scheme consists of two phases.

- The $Gen$ function takes the biometrics $b$ as input and returns the public parameter $PAR$;
- The $Rep$ function takes a biometric $b'$ and $PAR$ as input and computes $b$ if and only if $\text{dis}(b, b') \leq t$, where $\text{dis}()$ is the distance metric used to measure the variation in the biometric reading and $t$ is the error tolerance parameter.
An important requirement for such a scheme is that the value \( PAR \) should not reveal too much information about the biometric template \( b \). The first scheme of [5] and the schemes of [6, 15] implement a secure sketch protocol to test for equality using the homomorphic property of the encryption system.

### 3 Early Results

The first remote biometric verification scheme for distributed environments is described in [4], where the biometric template is assumed as a fixed binary string \( b = (b_1, ..., b_M) \) that is stored as a plaintext in \( DB \) during the registration phase. For authentication, a user \( U_i \) sends his fresh encrypted biometric template \( \epsilon(b') \) using Goldwasser-Micali scheme to \( SP \) resulting in a high transmission and computation cost due to individual encryption of each bit of \( b' \). Next, \( SP \) runs a PIR protocol using the index of the database location of \( U_i \) to obtain \( U_i \)'s encrypted biometric template \( \epsilon(b) \) computed by the \( DB \) during the PIR. Transaction privacy is guaranteed by employing this PIR scheme between the \( SP \) and the \( DB \) with the communication cost linear in the size \( N \) of the user’s in the \( DB \). Next, \( SP \) computes \( \nu_k = \epsilon(b'_k)\epsilon(b_k) \mod q = \epsilon(b'_k \oplus b_k) \) for \( k \in [1, M] \) due to the homomorphic property of Goldwasser-Micali scheme. Finally, a detached unit called \( Matcher \) with the secret key of the Goldwasser-Micali scheme decrypts the permuted \( \nu_k \)'s to compute the hamming weight and decides based on the threshold \( t \) to accept or reject the user \( U_i \).

#### 3.1 Analysis

The scheme of [4] is provably secure in the framework defined in section 2.3. However, a new attack with complexity exponential in \( N \) against this scheme is described in [3] that reveals the user’s biometric data to \( SP \). It is also noted that this attack can be avoided if the ciphertexts are re-randomized by the \( DB \). In [4, 3], an independent verification unit called \( Matcher \) is additionally required for the final decision, which increases the overall complexity of the system. As a result of the PIR system, the database performs \( O(N) \) exponentiations modulo \( q \), where \( q \) is an RSA modulus with \(|q|=2048 \) bits. Finally, the security of the system could be improved by storing the biometric data as encrypted as in the following schemes.

### 4 Improved Schemes

In [5], an extension to PIR system called as Extended Private Information Retrieval (EPIR) is presented, which is implemented for two different biometric verification schemes. In addition to the notion Identity Privacy (i.e. User Privacy), EPIR also satisfies the notion of Database Privacy, which means that the user (or the \( SP \)) does not learn anything about the other biometric entries. The main difference of this biometric authentication system is
the integration of a secure sketch scheme and the use of ElGamal encryption. This way, there is no need for a similarity metric for the final decision, instead the EPIR is used for equality testing. Particularly, the user $U_i$ registers by sending $R_i$, namely the ElGamal encryption of its biometric sketch to $DB$ and the parameter $PAR$ is publicly available for reconstruction used in the secure sketch scheme. For authentication, the $SC$ sends the encrypted biometric sketch $C$ using the $PAR$ and ElGamal encryption to $SP$, which is forwarded by $SP$ to $DB$. For each entry $i \in [1, N]$, the $DB$ selects a random $r_i$ and computes $T_i = (C/R_i)^{r_i}$, where $R_i$ is the ElGamal encryption of each user sketch stored in the system. Finally, $SP$ runs a PIR protocol to obtain the value $T_i$ corresponding to $U_i$ and decrypts it using his secret key. If the result is 1, $SP$ authenticates $U_i$, else rejects. In addition, [15] presents a slightly modified version of this scheme by simplifying the randomization step of the $DB$. Again, the same components, namely a PIR, secure sketch and ElGamal encryption scheme is considered. Apart from the computational cost of the PIR, the number of exponentiations computed by the $DB$ is reduced from $O(4N)$ as in [5] to $O(2N)$ due to the use of a single random number instead of two different random numbers for the randomization of the ciphertexts.

Besides, the authors of [6] combine Goldwasser-Micali with Paillier encryption system in the Lipmaa’s PIR protocol, where the latter is used in this PIR system to encode the requested index of $U_i$. Each biometric template is stored as an encrypted sketch using Goldwasser-Micali scheme, which is the scheme used to encrypt the fresh biometric template during authentication. Next, $SP$ sends this data to the $DB$ and Lipmaa’s PIR protocol is applied by multiplying each of the $DB$’s elements with the encrypted fresh template and by exploiting the homomorphic properties of the two encryption systems. The detached verification unit decrypts the resulting ciphertexts using the keys associated to Paillier and Goldwasser-Micali schemes to obtain a codeword $c$ of $U_i$ and checks the hash of $c$ to the previously stored hash value for final decision. Similar to [5, 15], the scheme of [6] requires $O((M + 1)N)$ exponentiations modulo $q^s$ ($s = 2$ with Paillier) and stores for each user $|q|M$ bits as encrypted sketch, where $M$ is the bit-length of the sketch and $|q|$ is the size of an RSA modulus. Finally, another EPIR application for hamming weight is described in [5] using the BGN encryption system and a PIR, where the system does not employ a secure sketch.

5 Different Approaches

In [3], the authors describe a new distributed remote identification scheme by integrating a Support Vector Machine (SVM) to work as a multi-class authentication classifier. Particularly, the $|U|$-class SVM implemented in [3] is described as follows: For each user $U_i \in U$ with biometrics $b_i$, a mono classifier is trained using the remaining users $(U/U_i)$ as the rejected class after extracting the biometric feature vector $b_i$ of $U_i$. Next, a user profile $w_{ij}^*$ for each user $U_i$ is constructed. Each user profile $w_{ij}^*$ consists of support vectors $SV_{i,j}$ and their weights $\alpha_{i,j}$, where $i = 1...S, j = 1...|U|$. This will finish the registration phase of the system. For identification, each component of the feature vector $b_i$ is encrypted by $SC$ using Paillier encryption scheme and sent to the $SP$. $SP$ forwards the encrypted bio-
metric data to $DB$, which computes the SVM classification values $class$ in the encryption domain by using the homomorphic properties of Paillier encryption system. Specifically, $DB$ takes the profile data $w_{i|U}$ and computes for each class $j \in [1, |U|]$ the distance of $b_i$ to the $w_{j|U}$ in the encryption domain. Next, $DB$ re-randomizes the resulting ciphertexts and sends the final vector $class$ of size $|U|$ to $SP$, which permutes and re-randomizes this vector to $sclass$. Next, $VU$ decrypts each component of $sclass$ and finds the index $d$ of the maximum positive scaler contained in the decrypted vector. If there exists not such a positive index, $VU$ sends ⊥ to $SP$, else it sends $d$. Finally, $SP$ recovers the identity of $U_i$ using $d$ and the inverse of the permutation used in $sclass$. The communication cost of this scheme is $O(N)$ ($N = |U|$) and the computation cost is $O(N)$ exponentiations mod $q^2$.

### 5.1 An Efficient System

At ICB’09, Sarier proposed a new approach for a multi-factor biometric verification designed for distributed systems, which stores a random pool of features instead of the biometric templates of each user. Specifically, biometrics of a user is considered as a set of features and set overlap is used as the distance metric, where the threshold $t$ represents the error tolerance in terms of minimal set overlap. Furthermore, the features of each user are randomly located as a separate entry in the central database instead of storing the biometric template (in cleartext or in encrypted form) of a user, which is a different technique from all the existing schemes, since each feature is stored only once by detecting the common features that are already stored in the database. Specifically, each of the features of arbitrary length are hashed using some collision-resistant hash function or mapped to an element of $\mathbb{Z}_p^*$ as in [2, 12] and stored in $DB$. Before this mapping, a secure sketch similar to the design of [14] could be implemented to improve the accuracy. The security of each feature is provided due to one-way hash function and the security of the communication channel is also provided via encryption. For this purpose, an Identity Based Encryption
(IBE) scheme such as Boneh-Franklin IBE to encrypt a random session key for AES and an efficient PIR protocol [7] is used, which allows SP to retrieve an item from the DB without revealing which item SP is retrieving. Based on this different approach for the database storage, the author presents a new remote biometric-based verification system achieving reduced storage and computational cost compared to the existing schemes.

**Registration Phase:** The registration phase consists of the following initialization of the components.

1. The four components of the system, namely, $U_i$ with a smartcard, $SC$, $SP$ and $DB$ are initialized by the Private Key Generator (PKG) of the IBE system with the private keys $d_i, d_{SC}, d_{SP}, d_{DB}$, respectively. The secret key $d_i$ of $U_i$ is stored in the smart card of the user.

2. The user $U_i$ presents its biometrics to the sensor client which extracts the feature set $B_i = (\mu_1, ..., \mu_k)$, where $\mu_i \in \mathbb{Z}_p^*$ of the user.

3. The user picks some random indexes $i_m \in \mathbb{Z}$ where $1 \leq m \leq k$ and registers his features at these locations of the database.

   If some of the locations are already occupied by other features, then the user selects other random indices. Also, if some of the features of the user are already stored in $DB$, then $DB$ returns the indices of the common features. Thus, common features are not stored more than once, which decreases the total storage cost of $DB$.

4. The user $U_i$ registers its personalized username at the service provider and stores the index list $\text{Index}_i = (i_1, ..., i_k)$ as encrypted with the public key of the $SP$ in his smart card.

**Verification Phase:** The following figure shows the workflow of this phase.

In this phase, $U_i$ inserts his smart card into the terminal of $SC$ and presents its biometrics. The transmission of the biometric data between the reader $SC$ and $U_i$'s smartcard is secured using IBE for session key generation and AES for encryption similar to the system in [11]. Next, $U_i$ sends a re-encryption of the stored $\text{Index}_i$ data to $SP$, which decrypts it to obtain the index list of $U_i$ to be used in the PIR protocol between $SP$ and $DB$. In Figure 2, the abbreviations denote the following: $B'_i = (\mu'_1, ..., \mu'_k)$ is the fresh template and $E_{ik}$ is the re-encryption of the encrypted index list $i_k \in \text{Index}_i$ of $U_i$. Using his biometric features $\mu_l$, the user is able to compute the encryption of $H(r_l)$ as $R_l$ for $l \in \text{Index}_i$, which are sent as encrypted to $SP$ for final decision based on the threshold $t$.

Here, $E^1_l = r_l \oplus \mu_l$ and $E^2_t = H(r_t \oplus \mu_l, H(r_l))$ for $t \in [1, N]$. Finally, $M_l = r_l \oplus \mu_l$ for $l \in \text{Index}_i$.

**5.2 Analysis of the Protocol**

- **Identity-biometric template relation:** At the registration phase, a user selects a random number for each feature of his biometrics and each feature is stored as a sep-
arate entry using the randomly selected index. Hence, even if the database is compromised, the attacker would not be able to find an index that points to a biometric template stored as cleartext or encrypted. This also provides security against the database since it only stores a randomly ordered pool of features from different users, where each feature is hashed using a specific cryptographic hash function before it is stored in the database. Besides, when the same user registers at the service provider using different personalized (pseudorandom) usernames, than the service provider is not even aware of this situation since it does not store any index number corresponding to the database storage location.

- No single point of failure: In order to impersonate a user, the attacker needs to obtain both the biometrics and the smart card that stores the private key and the index list of the user. Besides, the user has to store only a private key for IBE and some index numbers in the smart card instead of his biometrics. When the user’s smart card is lost or stolen, the user can obtain a new secret key from PKG and the index list by re-registering to the database.

- No need for PKI: Our scheme uses an efficient and anonymous IBE scheme such as Boneh/Franklin IBE for the generation of session keys for AES, hence, an eavesdropper (or a malicious database) on the communication channel cannot discover the identity of the user $U_i$ since the ciphertext does not reveal anything about the identity of the recipient (and the sender for authenticated Boneh/Franklin IBE scheme) of the ciphertext since Boneh-Franklin IBE is an anonymous IBE scheme. Also, our design does not require a Public Key Infrastructure (PKI).

- Efficient memory storage: Since each feature is stored as a separate entry in the

---

**Figure 2: Verification phase of the Protocol [13]**

<table>
<thead>
<tr>
<th>$SC$ (with Reader)</th>
<th>$E(B'_j)$</th>
<th>$U_i$ (with Smartcard)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extract feature set $B'_j$ &amp; compute $E(B'<em>j) = Enc</em>{sk}(B'_j, \text{Sign}(H(B'_j)))$</td>
<td>$H(B'_j)$</td>
<td>Decrypt $E(B'_j)$ and verify the signature</td>
</tr>
<tr>
<td>$U_i$ Encrypt the stored index $i$</td>
<td>$(B_i), T_i$</td>
<td>Compute Index from $B_i$'s</td>
</tr>
<tr>
<td>$DB$ PIR</td>
<td></td>
<td>$SP$ Get $E$'s of Index, via PIR</td>
</tr>
<tr>
<td>For $j \leq M_k$, pick a random $r_j$ and compute $E_j = (B_j, B_j^f)$</td>
<td>$(B_i), T_i$</td>
<td>$SP$ for each $i_k$, Index, compute $E_i = Enc_{sk}(M_i, \text{Sign}(H(M_i)))$</td>
</tr>
<tr>
<td>$U_i$ Decrypt $(E_j)$, $e$, compute $R = \text{Encrypt}(H(M_i \oplus \mu_i))$</td>
<td>$(R), T_i$</td>
<td>$VI$, decrypt $(R)$ and compute $H(M, H(M(\oplus \mu_i)))$</td>
</tr>
<tr>
<td></td>
<td>CK/NCK</td>
<td>Apply threshold &amp; decide</td>
</tr>
</tbody>
</table>

---

137
database, there could be common features belonging to different users. Thus, during registration phase, the database could check for this situation and could return the indices of the previously stored features. This way, the size of the registered feature set and the total storage in the database could be smaller. Besides, since no biometric template is stored as an entry, there is no need to apply a public key encryption scheme such as ElGamal to store the biometric data as encrypted, where the ciphertext size is twice the plaintext size as in [15, 5]. Finally, the choice of the system parameters of [6, 4] result in a constraint on the size of the database, whereas our design is also suitable for a large scale central database that stores biometric data.

- Lower computational cost: In [6, 4], the database performs $O(N)$ exponentiations modulo $q^2$ [6] and modulo $q$ [4], where $q$ is an RSA modulus with $|q|=2048$ bits. Similarly, the schemes of [15, 5] require $O(N)$ exponentiations in group $G$, on which the ElGamal public key scheme is defined. The computational cost of our scheme is dominated by the $O(N)$ random number selections and $O(N)$ hash computations in order to encrypt each feature stored in the database using one time pad. Except for the session key generations, we use symmetric key encryption and lightweight cryptographic primitives, hence, our scheme is suitable for user’s with smart cards. In the following table, we summarize various remote biometric-based authentication schemes that satisfy the security model described in section 2.

### Table 1: Comparison of distributed remote authentication systems

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Computation Cost</th>
<th>Storage Cost at DB index</th>
<th>Storage Cost per user</th>
</tr>
</thead>
<tbody>
<tr>
<td>System 1 [4]</td>
<td>$M$ exponentiations + $(MN)/2$ multiplications</td>
<td>$M$ bits</td>
<td>$M$ bits</td>
</tr>
<tr>
<td>System 2 [6]</td>
<td>$O(N)$ exponentations</td>
<td>$</td>
<td>q</td>
</tr>
<tr>
<td>System 3 [15]</td>
<td>$O(N)$ exponentations</td>
<td>$2M$ bits</td>
<td>$2M$ bits</td>
</tr>
<tr>
<td>System 4 [5]</td>
<td>$O(N)$ exponentations</td>
<td>$2M$ bits</td>
<td>$2M$ bits</td>
</tr>
<tr>
<td>System 5 [3]</td>
<td>$O(N)$ exponentations</td>
<td>$</td>
<td>q</td>
</tr>
<tr>
<td>Our System</td>
<td>$O(N)$ random number + hash computations</td>
<td>$</td>
<td>\mu</td>
</tr>
</tbody>
</table>

Abbreviations: $N=$total number of entries in the database; $k=$dimension of the feature vector of a user; $M=$ bit-length of the biometric template; $|\mu|=$ bit-length of a stored feature; $c =$ number of common features of a user; $|q|=$size of an RSA modulus

### 5.3 Complexity of the PIR

The communication cost of the systems evaluated in Table 1 is dominated by the PIR, which is usually instantiated using the number-theory based PIR systems such as [7], which has currently the best bound for communication complexity of $O(\log(n) + b)$,
where $b$ is the bit-length of the block to be retrieved from an $n$-bit $DB$. We assume that $M \approx k \cdot |\mu|$, where $M$ is the size of the secure sketch.

Since the system of [13] has to retrieve $k$ non-consecutive blocks of size $|\mu|$, a naive solution is to just run the PIR solution of [7] with complexity $PIR$ independently $k$ times, which results in the complexity of $k \cdot PIR$. However, in [10], the solution to the problem of retrieving $k$ items that are not necessarily consecutive is presented using hashing. This way, the complexity is much smaller than the na"ive solution, namely $s \cdot PIR$, where $s = \sigma \log(k\mu)$ for $\mu \in \mathbb{Z}_p^*$. Furthermore, better performance is derived via explicit batch codes instead of hashing, since small values of $k$ do not work with hashing. The reader is referred to [10] for a more detailed discussion of application of batch codes for amortizing the time complexity of PIR. Recently, [9] introduced an efficient noise-based PIR scheme, which is 100 times faster than all of the number-theory based PIR systems. The communication cost of [9] is not optimal as of [7], however, communication cost is not the main performance measurement of PIR as shown in the following table due to the enormous computational cost at the $DB$-end for number-theory based PIR schemes [9].

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Query size</th>
<th>Download size</th>
<th>Bandwidth usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipmaa’s PIR</td>
<td>162 Kb</td>
<td>0.16s</td>
<td>33h</td>
</tr>
<tr>
<td>Gentry and Ramzan’s PIR [7]</td>
<td>3 Kb</td>
<td>\approx 0s</td>
<td>17h</td>
</tr>
<tr>
<td>Noise-based PIR [9]</td>
<td>19Mb</td>
<td>19s</td>
<td>10min</td>
</tr>
</tbody>
</table>

6 Conclusion and Future Directions

In this paper, we evaluated new designs for remote biometric based authentication protocols that follow the state-of-the-art security model for biometric authentication. In addition to the systems that store encrypted biometric sketches, we review the schemes with different database storage mechanisms that involve a SVM or a random pool of features, where the latter results in reduced storage cost even in small databases due to the single storage of the common features. Besides, this system could be applied to a variety of biometrics that could be represented by a feature vector. Also, the size of the stored biometric data is much smaller than existing systems that store biometrics as encrypted with public key encryption. We note that the compromise of the database (namely, a random pool of features) would not help any attacker in the recovery of a user’s template, which could otherwise only be guaranteed by storing the biometric templates as encrypted. An interesting future work could be to improve the schemes that require a PIR using efficient storage methods and encryption systems.

Acknowledgement

The author is grateful to her supervisor Prof. Dr. Joachim von zur Gathen for his valuable support, encouragement and guidance.
References


