Fundamental limits for privacy-preserving biometric identification systems that support authentication
Ignatenko, T.; Willems, F.M.J.

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/TIT.2015.2458961

Published: 01/01/2015

Document Version
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 21. Oct. 2018
Abstract—In this paper we analyze two types of biometric identification systems with protected templates that also support authentication. In the first system two terminals observe biometric enrollment and identification sequences of a number of individuals. It is the goal of these terminals to form a common secret for the sequences belonging to the same individual by interchanging public (helper) messages of all individuals such that the information leakage about the secrets from these helper messages is negligible. These secret keys are used for authentication purposes. Moreover, the second terminal should be able to establish the identity of an individual based on the presented biometric identification sequence and helper messages. It is important to realize that biometric data are unique for individuals and cannot be replaced if compromised. Therefore the helper messages should contain as little as possible information about the biometric data. In the second setting we consider the first terminal does not generate secret keys from biometric sequences of individuals but chooses them uniformly at random. These keys are conveyed to the second terminal by communicating the corresponding helper messages. In this paper we determine the fundamental trade-offs between secret-key, identification and privacy-leakage rates for both biometric settings.

Index Terms—Biometrics, identification, authentication, privacy leakage, secret sharing, multi-user information theory.

I. INTRODUCTION

Biometric identification systems were studied by O’Sullivan and Schmid [11] and Willems et al. [15]. They assumed storage of biometric enrollment sequences in the clear and determined the corresponding identification capacity. Later Tuncel [13] analyzed the trade-off between the capacity of a biometric identification system and the storage space (compression rate) required for the biometric templates. It should be noted that Tuncel’s method realizes a kind of privacy protection scheme.

A related concept for the study of biometric systems with privacy protection is the concept of secrecy capacity introduced by Ahlswede and Csiszár [1]. This notion can be regarded as the amount of common secret information that can be obtained in a biometric authentication system in which helper messages or data are (publicly) available. Helper messages facilitate generation of the secret information and are crucial in the biometric setting. We also call them protected templates. Interestingly this secrecy capacity, which is equal to the mutual information between enrollment and authentication biometric sequences in the biometric authentication setting, equals the identification capacity found by O’Sullivan and Schmid [11] and Willems et al. [15].

Important parameter of a biometric system is privacy leakage. Privacy leakage is the amount of information that is contained (leaked) about biometric enrollment sequences in the publicly available data, in this case - helper data. In [8] the fundamental trade-off between secret-key rate and privacy-leakage rate was studied for a biometric authentication system. In the present paper we concentrate on identification. Note that the same biometrics can be used for both authentication and identification purposes. Therefore we investigate biometric identification systems that also support authentication. These systems can play an important role in single sign-on type of systems, e.g. based on OpenID, to allow for a (single) biometric identity provider that performs both authentication and identification. Moreover, the setting with joint identification and authentication considered in this paper is similar to the situation when two separate biometric systems, i.e. for identification and authentication, where privacy leakage has to be minimized, are jointly analyzed. Works in this direction, where multiple authentication systems were analyzed, include [10] and [14]. In general, identification systems are intrinsically more complex than authentication systems, as an identification process involves one-to-many comparisons in contrast to one-to-one comparison performed during authentication. We will investigate the trade-offs between the amount of common secret information and privacy leakage that is achieved in an identification procedure with protected biometric templates. Unlike in biometric authentication systems, here we also take into account the identification rate.

In the first system that we investigate in this paper two terminals observe the enrollment and identification biometric sequences of different individuals. The first terminal forms a secret for each enrolled individual and stores the corresponding helper data in a public database. The secrets are used for authentication purposes. The helper data, on one hand, facilitate reliable reconstruction of the secret and, on the other hand, allow determination of the individual’s identity.
for the second terminal, based on the presented biometric identification sequence. All helper data in the database are assumed to be public. Since the biometric secrets produced by the first terminal are used for authentication, the helper data should provide no information on these secrets. On the other hand, since biometric data are unique for individuals and cannot be replaced if compromised, the helper data should also provide as little as possible information about biometric data. In our identification system the only reference data stored for identification is the helper data. Therefore these helper data are also called protected templates. In this paper we determine what identification, secret-key and privacy-leakage rates can be realized by such a biometric identification system that supports authentication.

We also consider identification systems that support authentication based on secret binding. This authentication setting corresponds to the one of the settings studied in [8]. Here the first terminal chooses a secret key uniformly at random for each enrollment sequence of individuals. The helper data in this setting is formed based on the chosen secret key and the enrollment sequence. Thus the secret key is bound to the enrollment sequence via the helper data. Just as before, helper data is used to reconstruct the secret key but also to determine individual’s identity upon observing the identification biometric sequence. Also in this setting, we determine the fundamental trade-offs between identification, secret-key and privacy-leakage rates.

The paper is organized as follows. In Section II we describe the model for the biometric sequences. Next, in Sections III and IV we present our models for biometric identification with secret generation and secret binding, respectively. Section V states our results; their proofs are provided in the appendices. In Section VI we discuss how the identification results found in this paper connect to some known, previously established results. Finally, Section VII concludes the paper.

II. BIOMETRICS

A biometric identification system is based on a biometric source \( \{ Q_s(x), x \in X \} \) and a biometric channel \( \{ Q_c(y|x), y \in Y, x \in X \} \). The system is designed to identify one out of |\( \mathcal{V} \)| individuals. For each individual \( v \in \{1, 2, \ldots, |\mathcal{V}| \} \) in the system, the biometric source produces a biometric enrollment sequence \( x^N(v) = (x_1(v), x_2(v), \ldots, x_N(v)) \) with \( N \) symbols from the finite alphabet \( X \). The enrollment sequence \( x^N \) occurs with probability

\[
\Pr\{X^N = x^N\} = \prod_{n=1}^{N} Q_s(x_n), \tag{1}
\]

thus the symbols \( \{x_n, n = 1, 2, \ldots, N\} \) are independent of each other and i.i.d. according to \( Q_s(\cdot) \). Note that the biometric sequences are independent of the individual’s identity.

During identification a biometric identification sequence \( y^N = (y_1, y_2, \ldots, y_N) \) with \( N \) symbols from the finite alphabet \( Y \) is observed. This sequence is the output of the biometric channel whose input was the enrollment sequence of this individual. If individual \( v \) was observed the sequence \( y^N \) occurs with probability

\[
\Pr\{Y^N = y^N|X^N = x^N(v)\} = \prod_{n=1}^{N} Q_c(y_n|x_n(v)), \tag{2}
\]

thus the biometric channel is memoryless.

We assume here that all individuals are equally likely to be observed for identification, hence

\[
\Pr\{V = v\} = 1/|\mathcal{V}|, \text{ for all } v \in \{1, 2, \ldots, |\mathcal{V}|\}. \tag{3}
\]

The enrollment and identification biometric sequences \( x^N(v) \) and \( y^N \) are observed by an encoder and decoder, respectively. During enrollment for each individual \( v \) the encoder produces an index \( m_v \in \{1, 2, \ldots, |\mathcal{M}|\} \), which is referred to as helper data or protected template. The helper data are stored at position \( v \) in a public database to make reliable identification possible and are used by the decoder. Biometric identification systems are supposed to identify individuals but also support authentication of the individuals. To realize authentication, the systems have to generate secret keys or bind uniformly chosen secret keys from/to the biometric enrollment sequence \( x^N(v) \) of each enrolled individual \( v \). Therefore we subdivide systems into those in which terminals identify individuals and generate secret keys for them and those in which terminals identify individuals and choose and bind chosen secret keys to them. The decoder’s estimate \( \hat{\mathcal{s}} \) of the individual’s identity takes on values from the set of individuals, i.e. \( \hat{\mathcal{s}} \in \{1, 2, \ldots, |\mathcal{S}|\} \). The generated or chosen secret \( s \) assumes values in \( \{1, 2, \ldots, |\mathcal{S}|\} \). The decoder’s estimate \( \hat{\mathcal{s}} \) of the secret \( s \) also assumes values from \( \{1, 2, \ldots, |\mathcal{S}|\} \). In identification systems with secret-key binding, the secret \( S \) is a uniformly distributed index, hence

\[
\Pr\{S = s\} = 1/|\mathcal{S}| \text{ for all } s \in \{1, 2, \ldots, |\mathcal{S}|\}. \tag{4}
\]

III. IDENTIFICATION WITH SECRET GENERATION

![Fig. 1. Model for protected biometric identification with secret generation.](http://dx.doi.org/10.1109/TIT.2015.2458961)
$X^N(v)$ of individual $v$ and encodes it into helper data $M_v$ and a secret $S_v$, hence

$$(M_v, S_v) = e(X^N(v)), \text{ for } v \in \{1, 2, \ldots, |V|\},$$

(5)

where $e(\cdot)$ is the encoder mapping. The helper data $M_v$ are then stored in a (public) database at position $v$. The secret $S_v$ is either a) stored in an encrypted way in this identification system, then the system can operate in both identification and authentication modes; or b) handed over to the individual who can use it for authentication purposes in another access control system, this system then can use the public helper data from the identification system to reconstruct the secret.

During identification, upon observing the biometric identification sequence $Y^N$, the decoder forms an estimate $\hat{V}$ of the identity label of the observed individual as well as an estimate of his secret $\hat{S}_v$, hence

$$(\hat{V}, \hat{S}_v) = d(Y^N, M_1, M_2, \cdots, M_{|V|}),$$

(6)

where $d(\cdot, \cdot, \cdots)$ is the decoder mapping. Note that authentication is a special case of identification and therefore in our analysis we can concentrate on identification during which both identity label and secret of an individual are estimated without the loss of generality.

Now we are interested to find out what identification, secret-key and privacy-leakage rates can be realized by such an identification system with negligible error probability, such that individuals’ secret keys are close to uniform in the entropy sense and that for each individual the helper data only provide negligible information on his secret. We use the following definition of achievability.

**Definition 1** A secret-key rate, identification rate, and privacy-leakage rate triple $(R_S, R_I, R_L)$ with $R_S \geq 0$ and $R_I \geq 0$ is achievable in a protected biometric identification setting with secret generation if for all $\delta > 0$ for all $N$ large enough there exist encoders and decoders such that

$$\Pr\{|\hat{V}, \hat{S}_v\} \neq \{V, S_v\} \leq \delta,$$

$$\frac{1}{N} \log |V| \geq R_I - \delta,$$

$$\frac{1}{N} H(S_v) + \delta \geq \frac{1}{N} \log |S| \geq R_S - \delta,$$

$$\frac{1}{N} I(S_v; M_v) \leq \delta,$$

$$\frac{1}{N} I(X^N(v); M_v) \leq R_L + \delta,$$

for all $v \in \{1, 2, \cdots, |V|\}$. (7)

Moreover, we define $\mathcal{R}_{bi,g}$ to be the region of all achievable secret-key, identification and privacy-leakage rate triples for a protected biometric identification system with secret generation.

**IV. IDENTIFICATION WITH SECRET BINDING**

In a biometric identification system with secret binding, see Fig. 2, the encoder observes the enrollment biometric sequence $X^N(v)$ of individual $v$. A secret key $S_v$ is chosen uniformly at random and independently of the biometric sequence, see (4). The encoder encodes these $X^N(v)$ and $S_v$ into helper data $M_v$, hence

$$M_v = e(S_v, X^N(v)), \text{ for } v \in \{1, 2, \cdots, |V|\},$$

(8)

where $e(\cdot, \cdot)$ is the encoder mapping. The helper data $M_v$ are then stored in a (public) database at position $v$. The secret key can be used for authentication purposes.

During identification, upon observing the biometric identification sequence $Y^N$, the decoder forms an estimate $\hat{V}$ of the identity of the observed individual as well as an estimate of his secret key $\hat{S}_v$, hence

$$(\hat{V}, \hat{S}_v) = d(Y^N, M_1, M_2, \cdots, M_{|V|}),$$

(9)

where $d(\cdot, \cdot, \cdots)$ is the decoder mapping. Again we can define achievable as follows.

**Definition 2** A secret-key rate, identification rate, and privacy-leakage rate triple $(R_S, R_I, R_L)$ with $R_S \geq 0$ and $R_I \geq 0$ is achievable in a protected biometric identification setting with secret binding if for all $\delta > 0$ for all $N$ large enough there exist encoders and decoders such that

$$\Pr\{|\hat{V}, \hat{S}_v\} \neq \{V, S_v\} \leq \delta,$$

$$\frac{1}{N} \log |V| \geq R_I - \delta,$$

$$\frac{1}{N} \log |S| \geq R_S - \delta,$$

$$\frac{1}{N} I(S_v; M_v) \leq \delta,$$

$$\frac{1}{N} I(X^N(v); M_v) \leq R_L + \delta,$$

for all $v \in \{1, 2, \cdots, |V|\}$. (10)

Moreover, we define $\mathcal{R}_{bi,c}$ to be the region of all achievable secret-key, identification and privacy-leakage rate triples for a protected biometric identification system with secret binding.

**V. STATEMENT OF RESULTS**

**A. Main Theorems**

In order to state our results we first define the region $\mathcal{R}_{bi}$ and then we present our theorems.
\( R_{bi} \triangleq \{(R_I, R_S, R_L) : 0 \leq R_I + R_S \leq I(U; Y), R_I \geq I(U; X) - I(U; Y) + R_T, R_I \geq 0, R_S \geq 0, \)

for some

\[ P(u, x, y) = Q_s(x)Q_c(y|x)P(u|x) \]

and \(|U| \leq |X| + 1\). \hfill (11)

**Theorem 1 (Biometric Identification, Secret Generation)**

\[ R_{bi,g} = R_{bi}. \hfill (12) \]

**Theorem 2 (Biometric Identification, Secret Binding)**

\[ R_{bi,c} = R_{bi}. \hfill (13) \]

The detailed proofs of these theorems are provided in Appendix A and Appendix B, respectively. Nevertheless, we sketch the main ideas of the proofs in the following section. Furthermore, we illustrate the properties of the determined regions by considering an example for a binary-symmetric double source.

From Thm. 1 and Thm. 2 we see that the information that can be reliably shared and/or reconstructed from biometric enrollment and identification sequences has to be distributed between identification rate and secret-key rate. Moreover, identification rate has an impact on privacy leakage, meaning that the higher identification rates are realized, the higher privacy leakage is. Note that the definition of region \( R_{bi} \) involves so-called test-channel \( \{P(u|x), x \in X, u \in U\} \) that specifies the auxiliary alphabet \( U \) and the mutual information \( I(U; X) \) and \( I(U; Y) \). This channel can be regarded as quantization channel. This quantization is essential component in achieving the optimal biometric systems with template protection. In general, we observe that for a given test-channel, the larger identification rate we need to achieve the smaller the secret-keys rates and the larger the privacy-leakage rates we can realize.

**B. Overview of the Proofs**

The proofs of both Thm. 1 and Thm. 2 consist of three parts, i.e. achievability part, converse part and the bound on cardinality of the auxiliary random variable. Our converse proofs are quite standard and are based on Markov property and Fano’s inequality. The bound on cardinality of \( U \) is proved using the Fenchel-Eggleston strengthening the Caratheodory lemma, see [19]. Finally, our achievability proofs are based on weak typicality, a concept which was introduced first by Shannon [12], and further developed by Forney [6] and then by Cover and Thomas [3].

It should be noted that the achievability proof for Thm. 1 is the most involved proof, while the achievability for Thm. 2 is its extension with an extra layer, see Fig. 3. In this layer the one-time pad is added to conceal a chosen secret key. Therefore in this section we outline the main idea of the achievability proof for Thm.1.

We start by fixing a conditional distribution \( \{P(u|x), x \in X, u \in U\} \) that determines the joint distribution \( P(u, x, y) = Q_s(x)Q_c(y|x)P(u|x) \), for all \( x \in X, y \in Y \), and \( u \in U \). Then we randomly generate roughly \( 2^{N(I(U; X))} \) auxiliary sequences \( u^N \). Each of these sequences gets a random \( s \)-label and a random \( m \)-label. These labels are uniformly chosen. The \( s \)-label can assume roughly \( 2^{N(I(U; Y) - R_T)} \) values, and the \( m \)-label roughly \( 2^{N(I(U; X) - I(U; Y) + R_T)} \) values.

During enrollment, the encoder observes \( |V| \) individuals. For each individual \( v \in \{1, 2, \cdots, |V|\} \) with the enrollment sequence \( x^N(v) \), the encoder finds a sequence \( u^N(v) \) that is jointly typical with \( x^N(v) \). It stores the helper-label \( m_v \) corresponding to this \( u^N(v) \) in a public database at the position \( v \). Moreover, the encoder issues the secret-label \( s_v \) corresponding to this \( u^N(v) \) to the individual \( v \).

During identification, the decoder observes an identification sequence \( y^N \). It checks all the records in the database to determine a unique individual with identity label \( \hat{v} \) such that the record \( \hat{v} \) of the database contains the helper-label \( m_v = m(u^N(\hat{v})) \) for which \( u^N(\hat{v}) \) and \( y^N \) are jointly typical. Upon finding such a record, the decoder issues the identity label \( \hat{v} \) and the secret estimate \( \hat{s}_v \). It can be shown that the decoder can reliably recover \( u^N(\hat{v}) \) and thus \( \hat{v} \) and \( \hat{s}_v \). Then, it is easy to check that the privacy leakage is not larger than \( I(U; X) - I(U; Y) + R_T \). Moreover, to prove the facts that secrecy leakage is negligible and that the secret is close to uniform we can use the property of the encoding procedure that \( u^N(\hat{v}) \) can be reliably reconstructed from \( s_v \) and \( m_v \).

**C. Example: Binary Symmetric Double Source**

Consider now an example with a binary symmetric double source with crossover probability \( 0 \leq q \leq 1/2 \). For this source we have that \( Q(x, y) = Q_s(x)Q_c(y|x) = (1-q)/2 \) for \( y = x \) and \( q/2 \) for \( y \neq x \). Now in order to illustrate the properties of the achievable region, we define privacy-leakage vs. secret-key and identification rate function

\[ R_{bi}(R_S, R_I) = \min\{R_L : (R_I, R_S, R_L) \in R_{bi}\}. \hfill (14) \]

Note that this function applies to both protected biometric identification system with secret generation and with secret binding.

Observe that for the binary symmetric double source we have

\[ I(U; Y) = 1 - H(Y|U), \]

\[ I(U; X) - I(U; Y) = H(Y|U) - H(X|U). \hfill (15) \]
From Mrs. Gerber’s Lemma [18] we know that if $H(X|U) = h(p)$ for some $0 \leq p \leq 1/2$, then $H(Y|U) \geq h(q \ast p)$, where $q \ast p = q(1 - p) + p(1 - q)$. Note that for binary symmetric $(U, X)$ with crossover probability $p$ the minimum $H(Y|U)$ is achieved, hence we obtain for identification rates $R_I \geq 0$ that

$$R_{bi}(R_S, R_I) = h(p \ast q) - h(p) + R_I,$$

for some $p$ satisfying $1 - h(p \ast q) - R_I = R_S$, and $R_I \leq 1 - h(p \ast q)$. (16)

Fig. 4. Rate function for the crossover probability $q = 0.1$.

In Fig. 4 we plot the resulting function for $q = 0.1$ and in Fig. 5-7 the corresponding projections to the identification rate and secret-key rate, identification rate and privacy-leakage rate, and secret-key rate and privacy-leakage rate planes, respectively. These figures demonstrate the trade-off between the three rates. Indeed, from Fig. 5 we observe that there exists a trade-off between identification and secret-key rates. In this way, the more individuals we would like to be able to reliably identify, the smaller secret keys we can assign to individuals for authentication purposes and thus the less secure the corresponding authentication system becomes. The latter means that it becomes easier to get access to the systems that deploy biometric authentication, since smaller secrets are easier to guess and also require less biometric information (smaller $U$) for their reconstruction. Fig. 5 shows that we have to sacrifice privacy and publish more data via protected templates, if we would like to achieve higher identification rate. Finally, from Fig. 7 we see that there is a similar trade-off between secret-key rate and privacy-leakage rate in the identification setting with secret keys, i.e. the higher secret rates one would like to achieve the more privacy has to be sacrificed.

VI. CONNECTION TO OTHER RESULTS

Observe that from the achievable region $R_{bi}$ we can derive a number of previously established results. In the following
we set $R_L = \infty$ in order to indicate that we exclude privacy leakage from our considerations.

**Corollary 1**  If we restrict ourselves to $R_I = 0$ and $R_L = \infty$

\[
R_{bi|R_I=0,R_L=\infty} = \{R_S : R_S \leq I(X;Y)\}. 
\]

(17)

This corollary gives us the Ahlswede and Csiszár [1] result for the amount of common secret information that can be generated by two terminals. Note that in the biometric setting the secrecy capacity can be achieved at privacy leakage rate of $H(Y|X)$.

**Corollary 2**  If we restrict ourselves to $R_I = 0$

\[
R_{bi|R_I=0} = \{(R_S,R_L) : 0 \leq R_S \leq I(U;Y), 
R_L \geq I(U;X) - I(U;Y), 
\text{for some } P(u,x,y) = Q_s(x)Q_u(y|x)P(u|x) 
\text{ and } |U| \leq |X| + 1\}. 
\]

The region in Cor. 2 corresponds to the region for biometric authentication with generated or chosen secret, derived in [8]. Clearly authentication is a special case of identification with one record in the identification database.

From the latter point of view it is interesting to see how identification procedure modifies the achievable proof for the authentication result of Cor. 2, see [8] for the proof details. Indeed the main difference comes from checking multiple biometric records. Thus in identification the decoder, upon observing identification sequence $y^N$, has to check all the records in the database to determine a unique individual with identity label $\hat{v}$ such that the record $\hat{v}$ of the database contains the helper-label $m_0 = m(\hat{v}U)$ for which $\hat{v}U$ and $y^N$ are jointly typical. This results in roughly $|V|$ extra error terms compared to authentication. To keep the coding error probability small, the helper-label rate has to be increased by identification rate, resulting therefore both in the reduced secret-key rate and in the increased privacy-leakage rate.

**Corollary 3**  If we restrict ourselves to $R_S = 0$ and $R_L = \infty$

\[
R_{bi|R_S=0,R_L=\infty} = \{R_I : R_I \leq I(X;Y)\}. 
\]

(19)

The special case given by the above corollary corresponds to the identification region for a biometric identification system without protected templates, derived in Willems et al. [15] and O’Sullivan and Schmid [11]. Indeed to achieve identification capacity we have to store all biometric information and thus cannot achieve any privacy protection as $R_L = H(X)$ then.

**Corollary 4**  If we restrict ourselves to $R_S = 0$

\[
R_{bi|R_S=0} = \{(R_I,R_L) : 0 \leq R_I \leq I(U;Y), 
R_L \geq I(U;X), 
\text{for some } P(u,x,y) = Q_s(x)Q_u(y|x)P(u|x) 
\text{ and } |U| \leq |X| + 1\}. 
\]

Also from the above corollary we can see that if we do not require a secret key and just concentrate on identification, then to achieve identification rate $I(U;Y)$ we have to store the template of rate $I(U;X)$ which results into the privacy-leakage rate $I(U;X)$. This is similar to the Tuncel result [13] if we assume that the underlying biometric source sequence corresponds to the enrollment biometric sequence. As such the Tuncel result [13] realized some kind of biometric template protection.

**Corollary 5**  If we restrict ourselves to $R_L = \infty$

\[
R_{bi|R_L=\infty} = \{(R_I,R_S) : 0 \leq R_I + R_S \leq I(X;Y)\}. 
\]

Finally, the last corollary corresponds to the identification setting with secret keys, where privacy leakage does not play a role, see also [9] and [16].

Now let’s take a look at the results discussed above from a different perspective. Observe that if we do not require privacy leakage to be as small as possible, then secret-key capacity, see Cor. 1, is the same as identification capacity, see Cor. 3. It is then not surprising to see that in the identification setting with secret keys, see Cor. 5, this capacity $I(X;Y)$ is distributed between secret-key rate and identification rate. However, if we take privacy leakage into account, then we see that although the maximum secret-key rate and identification rate given by Cor. 2 and Cor. 4 are the same, privacy leakage in the identification setting is larger then the one in the authentication setting, since helper-data rate need become larger and contain more information about biometric data in order to compensate for one-to-many comparisons and thus guarantee small identification error.

VII. CONCLUSIONS

In this paper we have considered biometric identification systems with protected templates. Biometric data used in such identification systems are also utilized in access control and authentication applications. These applications are typically based on biometric secrets. To create reliable identification systems, helper data of all enrolled individuals have to be accessed by the decoder. These data are assumed to be public. Thus public information of our biometric identification system should provide no information about biometric secrets, though facilitate reliable identification. Moreover, because biometric data cannot be replaced if compromised, the helper data should contain as little as possible information about biometrics.

In this paper we have analyzed what secret-key, identification and privacy-leakage rates can be realized by biometric identification systems with protected templates that support authentication. It appears that the larger identification rates we would like to achieve, the smaller secret keys we can generate and the more biometric information we have to leak. We also see that our results are strongly connected to the secret sharing concept of Ahlswede and Csiszár [1]; the biometric identification system without protected templates of Willems et al. [15] and O’Sullivan and Schmid [11]; the biometric identification system with restricted storage of Tuncel [13]; the biometric identification with secret keys of [9] and [7]; and
the biometric authentication system with privacy protection of [8]. These results can be derived as special cases in the biometric identification systems with protected templates considered here.

APPENDIX A
PROOF OF THM. 1

The proof of this theorem consists of three parts, i.e. the achievability, the converse and the bound on cardinality of $U$. We first present some concepts used in the first part of the proof.

A. Typical and Modified Typical Sets. Their Properties

Our achievability proof is based on weak typicality, a concept introduced by Forney [6], and further developed by Cover and Thomas [3]. Before moving to the actual proof, we define a modified typical set, that allows us to obtain a weak-typicality alternative for the so-called Markov Lemma that holds for the strong-typicality case, see Berger [2]. Strong typicality was first considered by Wolfowitz [17], but since then several alternative versions were proposed, see Berger [2], but also Csiszár and Korner [4] and Cover and Thomas [3]. The main advantage of weak typicality is that the results in principle also hold for non-discrete random variables. Therefore our proof generalizes e.g. to the Gaussian case.

Definition 3 Consider typicality with respect to distribution

$$\{P(u,x,y) = Q(x,y)P(u|x), u \in \mathcal{U}, x \in \mathcal{X}, y \in \mathcal{Y}\}.$$ Now the set $\mathcal{B}_c(N)(U|X)$ is defined as

$$\mathcal{B}_c(N)(U|X) = \{u^N : \Pr\{Y^N \in \mathcal{A}_c(N)(Y|u^N, x^N)\} \geq 1 - \epsilon\},$$

where $Y^N$ is the output of a “memoryless channel” $Q(y|x) = Q(x)/P(x)$ for $Q(x) = \sum_y Q(x,y)$, whose input is $x^N$. Moreover, $\mathcal{B}_c(N)(X|u) = \{u^N : (u^N, x^N) \in \mathcal{B}_c(N)(U|X)\}$ for all $x^N$.

Property 1 If $(u^N, x^N) \in \mathcal{B}_c(N)(U|X)$ then also $(u^N, x^N) \in \mathcal{A}_c(N)(U|X)$.

This follows from the fact that $(u^N, x^N) \in \mathcal{B}_c(N)(U|X)$ implies that there is at least one $y^N$ such that $(u^N, x^N, y^N) \in \mathcal{A}_c(N)(U|XY)$. This implies, by the definition of $\mathcal{A}_c(N)(U|XY)$, that also $(u^N, x^N) \in \mathcal{A}_c(N)(U|X)$.

Property 2 Let $U^N, X^N, Y^N$ be i.i.d. with respect to $P(u,x,y) = Q(x,y)P(u|x)$. Then for $\epsilon < 1$ and $N$ large enough

$$\sum\{(u^N, x^N) \in \mathcal{B}_c(N)(U|X)\} P(u^N, x^N) \geq 1 - \epsilon.$$

The statement follows from observing that

$$\Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_c(N)(U|XY)\} \leq \sum\{(u^N, x^N) \in \mathcal{A}_c(N)(U|XY)\} \Pr\{Y^N \in \mathcal{A}_c(N)(Y|u^N, x^N) \mid (U^N, X^N) = (u^N, x^N)\} \leq \sum\{(u^N, x^N) \in \mathcal{B}_c(N)(U|X)\} P(u^N, x^N) + \sum\{(u^N, x^N) \in \mathcal{B}_c(N)(U|X)\} P(u^N, x^N)(1 - \epsilon) = 1 - \epsilon + \epsilon \Pr\{(U^N, X^N) \in \mathcal{B}_c(N)(U|X)\},$$

or

$$\Pr\{(U^N, X^N) \in \mathcal{B}_c(N)(U|X)\} \geq 1 - \frac{1 - \Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_c(N)(U|XY)\}}{\epsilon}.$$ (24)

The weak law of large numbers implies that $\Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_c(N)(U|XY)\} \geq 1 - \epsilon^2$ for $N$ large enough. Then (23) follows from (24).

B. Achievability Part of Thm. 1

We start the achievability proof by fixing $0 < \epsilon < 1$ and a blocklength $N$. We also fix the auxiliary alphabet $\mathcal{U}$ and the conditional probabilities $\{P(u|x), u \in \mathcal{U}, x \in \mathcal{X}\}$.

Random Code Construction, Encoding and Decoding: Random coding: For each index $j \in \{1, 2, \ldots, |J|\}$ generate an auxiliary sequence $w(j)$ at random according to $P(u) = \sum_{y \in \mathcal{Y}} Q(y|x)P(u|x)$, where there are $2^{N(I(U;X) + 4\epsilon)}$ such indices. Moreover, for each such an index $j$ (and the corresponding sequence $w(j)$) generate a secret-key label $s(j) \in \{1, 2, \ldots, |S|\}$ and a helper-data label $m(j) \in \{1, 2, \ldots, |M|\}$ uniformly at random. There are $2^{N(1 - 3\epsilon - 4\epsilon^2)}$ such secret-key-label and helper-data-labels, respectively, in a database with $|V|$ users. Note that the resulting code is public, and is available to both encoder and decoder.

Encoding (Enrollment): The encoder observes individual $v$ with biometric sequence $z(v)$. These observed biometric sequences $z(v)$ are only available to the encoder, and therefore private. Then the encoder finds the index $j$ such that $(w(j), z(v)) \in \mathcal{B}_c(N)(U|X)$. If such an index is found, the encoder produces a secret-key label $s(j)$ and helper-data label $m(j)$. Moreover, the encoder checks whether there is another index $j' \neq j$ such that $s(j') = s(j)$ and $m(j') = m(j)$. If not, the helper data $m(j)$ are stored at location $v$ in the database and the secret $s(j)$ is handed over to the individual. However, if index $j$ was not found or index pair $(s(j), m(j))$

2Observe that we use re-scaling of $\epsilon$ here.

3To get a more compact notation in this part of the proof we use $x^N$ instead of $x^{N_0}$, etc.

4Here we assume that $|S|, |M|, |J|$ and $|V|$ are integers. This might not be the case in general, but for the reasons of simplicity we ignore it here.
was not unique, the location stays empty, an index $j$ gets a random label from $\{1, 2, \ldots, |\mathcal{J}|\}$, and no secret is offered to the individual.

**Decoding (Identification):** The decoder observes an identification biometric sequence $y$, which is the output of the biometric channel. Then it checks all the records $v \in \{1, 2, \ldots, |\mathcal{V}|\}$ in the database and determines a unique individual $\hat{v}$ whose record contains $m_0 = m(j)$ for which there exists a unique index $\hat{j}$ such that $(\hat{w}, \hat{y}) \in A^{(N)}(UY)$. If such a unique individual $\hat{v}$ and a unique index $\hat{j}$ can be found, the decoder outputs the estimate of individual’s identity label $\hat{v}$ and the estimate of his secret $\hat{s}_v = \hat{s}(\hat{j})$. If not, an error is declared.

**Events, Error Probability:**

**Events:** Let $V$ be the identity label of the actual individual, $X$ be the corresponding biometric enrollment sequence and $Y$ be the resulting observed biometric identification sequences. Moreover, let $J$ be the index determined by the encoder, $S(j)$ and $M(j)$ the random labels assigned to $j \in \{1, 2, \ldots, |\mathcal{J}|\}$, $M_v$ the helper data stored at location $v \in \{1, 2, \ldots, |\mathcal{V}|\}$, and $S$ and $M$ the actual labels corresponding to individual $V$. Now we define the events:

- $A_j \triangleq \{(U(j), X) \in B^{(N)}(UX)\}$,
- $B_j \triangleq \{S(j) = S \land M(j) = M\}$,
- $C_j \triangleq \{(U(j), Y) \in A^{(N)}(UY)\}$,
- $D_v,j \triangleq \{M_v = M_j\}$,
- $E_j \triangleq \{(U(j), X, Y) \in A^{(N)}(UXY)\}$.

**Error probability:** We obtain the following upper bound for the resulting error probability $P_E$ averaged over the ensemble of codes. We assume that $v$ runs over $\{1, 2, \ldots, |\mathcal{V}|\}$ and $j$ runs over $\{1, 2, \ldots, |\mathcal{J}|\}$.

\[
\Pr\left\{ \bigcap_j A^c_j \cup \bigcup_{j \neq j} B_j \cup C^c_j \cup \bigcup_{v \neq V, j \neq J} (C_j \cap D_v,j) \right\}
\leq \Pr\left\{ \bigcap_j A^c_j \right\} + \Pr\left\{ \bigcup_{j \neq j} B_j \right\} + \Pr\left\{ \bigcup_{j \neq j} (A_j \cap C^c_j) \right\} + \Pr\left\{ \bigcup_{v \neq V, j \neq J} (C_j \cap D_v,j) \right\}
\leq \Pr\left\{ \bigcap_j A^c_j \right\} + \sum_{j \neq j} \Pr\{B_j\} + \Pr\left\{ \bigcup_j A_j \right\} \cap E_j \right\} + \sum_{v \neq V, j \neq J} \Pr\{C_j \cap D_v,j\},
\]

where we used union bound in the first step, and $E_j \Rightarrow C_j$ in the last step.

**First term:** As in Gallager [5], p. 454, we write

\[
\Pr\left\{ \bigcap_j A^c_j \right\} = \sum_{u \in X^N} \sum_{j} Q_s(u) \prod_j \left(1 - \sum_{w \in B^{(N)}(U)} P(u|w)\right) \leq \sum_{u \in X^N} \sum_{j} Q_s(u) \left(1 - \sum_{w \in B^{(N)}(U)} P(u|w)\right)^{|\mathcal{J}|} \\
\leq \sum_{u \in X^N} \sum_{j} Q_s(u) \left(1 - \sum_{w \in B^{(N)}(U)} P(u|w)\right) + \exp(-|\mathcal{J}| \cdot 2^{-N(I(U;X)+\varepsilon)}) \\
\leq \sum_{u \in X^N} P(u, x) + \sum_{u \in X^N} Q_s(x) \exp(-2^N \varepsilon) \\
\leq 2\varepsilon,
\]

for $N$ large enough. Here (a) follows from the fact that for $(u, x) \in B^{(N)}(UX)$, using Property 1, we get

\[
P(u, x) \leq P(u, x) \frac{Q(x)P(x)}{P(x)} \leq P(u, x) 2^{-N(I(U;X)+\varepsilon)} \\
\geq P(u, x) 2^{-N(I(U;X)+3\varepsilon)},
\]

(b) from the inequality $(1 - \alpha \beta)^K \leq 1 - \alpha + \exp(-K \beta)$, which holds for $0 \leq \alpha, \beta \leq 1$ and $K > 0$; and (c) from Property 2 and since we have $|\mathcal{J}| = 2^{N(I(U;X)+4\varepsilon)}$.

**Second term:** Since $|S| = 2^{N(I(U;X)+3\varepsilon)} - \log |\mathcal{V}|$ and $|M| = 2^{N(I(U;X)-I(U;X)+3\varepsilon)} + \log |\mathcal{V}|$, then for all large enough $N$ we get

\[
\sum_{j \neq j} \Pr\{B_j\} \leq \frac{|\mathcal{J}|}{|S| \cdot |M|} \leq 2^{-N \varepsilon} \leq \varepsilon.
\]

**Third term:** Now for this term we get

\[
\Pr\left\{ \bigcup_j A_j \right\} \leq \max_{(u, x) \notin B^{(N)}(U)} \frac{\Pr\{Y \notin A^{(N)}(Y|u, x)|(U, X) = (u, x)\}}{\varepsilon},
\]

where the last step follows directly from the definition of $B^{(N)}(UX)$. The final version of record is available at http://dx.doi.org/10.1109/TIT.2015.2458961
Fourth term: For a fixed $y$

$$\Pr\{U \in A^c_x(U|y)\} = \sum_{u \in A^c_x(U|y)} P(u)$$

$$\leq \sum_{u \in A^c_x(U|y)} P(u)P(Y|u)$$

$$\leq 2^{-N(I(U;X)-3\varepsilon)} \sum_{u \in A^c_x(U|y)} P(u)$$

$$\leq 2^{-N(I(U;Y)-3\varepsilon)}.$$

Now, for $N$ large enough we have

$$\sum_{v \neq V} \sum_{\tilde{J} \neq J} \Pr\{C \cap D_{v,j}\} \leq \sum_{v \neq V} \sum_{\tilde{J} \neq J} \frac{1}{|M|} \max_{u \neq V} \Pr\{U \in A^c_x(U|y)\}$$

$$\leq \frac{|V| \cdot |J|}{|M|} 2^{-N(I(U;Y)-3\varepsilon)}$$

$$\leq 2^{-N\varepsilon} \leq \varepsilon. \quad (29)$$

Wrap up:

Identification rate, secret-key rate and error probability:

For all $N$ large enough there exist codes in the ensemble of codes ($u$ sequences and $s(\cdot)$ and $m(\cdot)$ labels) having error probability $P_E \leq P$. Here $P$ denotes the error probability in the sense of (25). For such a code

$$\Pr\{U, \hat{S}_V \neq (V, S_V)\} \leq P \leq \epsilon,$$  \quad (30)

$$\log |V| + \log |S| = N(I(U;Y) - 3\varepsilon), \quad (31)$$

for our fixed $0 < \varepsilon < 1$. This follows from (25) and our code construction.

Secrecy leakage: First, note that if no error occurs then $(u(j), \hat{u}(v)) \in A^c_x(U|X)$, and $J$ uniquely defines $\hat{U} = \hat{U}(J)$. Moreover, if an error occurs, then $J$ is some index from $\{1, 2, \cdots, |J|\}$ and consequently $\hat{U}$ is some sequence from $U^N$. Note also that $|A^{(N)}_x(X|u(j))| \leq 2^{N(H(X|U)+2\varepsilon)}$, then for all $v \in \{1, 2, \cdots, |V|\}$

$$H(X(v)) \leq H(U(j), \hat{U}(j))$$

$$= H(U(j)) + H(X(v)|U(j))$$

$$\leq H(U(j)) + P E \log |X|^N$$

$$(1 - P) \log 2^{2N(H(X|U)+2\varepsilon)}$$

$$\leq H(U(j)) + 5N \log |X| + N H(X|U) + 2N \epsilon,$$  \quad (32)

and therefore, since $H(X(v)) = NH(X)$,

$$H(U(j)) \geq N(I(U;X) - 5\varepsilon \log |X| - 2\varepsilon). \quad (33)$$

Next consider

$$H(S_v, M_v) = H(U(j), S_v, M_v)$$

$$- H(U(j)|S_v, M_v)$$

$$\geq H(U(j)|S_v, M_v, \hat{U})$$

$$\geq H(U(j)|S_v, M_v, \hat{U}) - H(U(j)|\hat{U})$$

$$\geq H(U(j)|S_v, M_v) - P \log |J| - 1$$

$$\geq N(I(U;X) - 5N \varepsilon \log |X| - 2N \varepsilon$$

$$- 5N(I(U;X) + 4\varepsilon) - 1$$

$$\geq 7N \varepsilon + 10N \varepsilon \log |X| + 20N \varepsilon^2 + 1.$$  \quad (34)

where in step (a) we used the fact that for a unique label pair $(S_v, M_v)$ there is a unique index $J$, which defines $\hat{U}$, in (b) Fano’s inequality, in (c) we used (33), and in (d) the fact that $I(U;X) \leq H(X) \leq \log |X|$.

Finally, using (34), we obtain for the secrecy leakage.

$$I(S_v; M_v)$$

$$= H(S_v) + H(M_v) - H(S_v, M_v)$$

$$\leq N(I(U;Y) - 3N \varepsilon \log |V| + N(I(U;X)$$

$$- N(I(U;Y) + N \varepsilon \log |V| - N(I(U;X)$$

$$+ 10N \varepsilon \log |X| + 2N \varepsilon + 20N \varepsilon^2 + 1$$

$$\leq 7N \varepsilon + 10N \varepsilon \log |X| + 20N \varepsilon^2 + 1.$$  \quad (35)

Uniformity: The uniformity of the secret key $S_v$ follows from the following inequality where we use (34).

$$H(S_v)$$

$$= H(S_v) - H(M_v)$$

$$\geq H(S_v, M_v) - H(M_v)$$

$$\geq N(I(U;X) - 10N \varepsilon \log |X| - 2N \varepsilon - 20N \varepsilon^2 - 1 -$$

$$N(I(U;Y) + N \varepsilon \log |X|$$

$$\geq N(I(U;Y) - \log |V| + 10N \varepsilon \log |X|$$

$$- 10N \varepsilon - 20N \varepsilon^2 - 1$$

$$= \log |S| - 10N \varepsilon \log |X| - 7N \varepsilon - 20N \varepsilon^2 - 1.$$  \quad (36)

Privacy leakage: Note that from the code construction it immediately follows that

$$I(X^N(v); M_v)$$

$$\leq H(M_v)$$

$$\leq N(I(U;Y) - I(U;X) + 8\varepsilon) + \log |V|.$$  \quad (37)

Conclusion: We now conclude the proof by letting $\varepsilon \downarrow 0$ and $N \to \infty$ and observing that the achievability follows from (30), (31), (35), (36), and (37).

C. Converse Part of Thm. 1

We start by considering the joint entropy $H(V, S_v)$ of the individual’s identity label and his secret. We use that $(V, S_V) = d(Y^N, M_1, M_2, \cdots, M_{|V|})$ and Fano’s inequality
then if we take a time-sharing variable $T(c)$, we define $U_n$ such that

$$H(V, S_V | \tilde{V}, \tilde{S}_V) \leq F,$$

where $F \triangleq 1 + \operatorname{Pr}\{\tilde{V}, \tilde{S}_V \neq (V, S_V)\} \log(|V| \cdot |S|)$. Then

$$H(V, S_V) := I(V, S_V; M_1, M_2, \ldots, M_{|V|}, Y^N) + H(V, S_V; M_1, M_2, \ldots, M_{|V|}, Y^N, \tilde{V}, \tilde{S}_V) \leq I(V, S_V; M_1, M_2, \ldots, M_{|V|}, Y^N) + H(V, S_V | \tilde{V}, \tilde{S}) \leq I(V, S_V; M_1, M_2, \ldots, M_{|V|}, Y^N) + I(V, S_V; Y^N | M_1, M_2, \ldots, M_{|V|}) + F$$

$$\text{(a)} \quad I(S_V; M_1, M_2, \ldots, M_{|V|}) + \sum_{n=1}^{N} I(Y_n; V, S_V, M_V, Y^{n-1}) + F \leq I(S_V; M_1, M_2, \ldots, M_{|V|}) + \sum_{n=1}^{N} I(Y_n; V, S_V, M_V, Y^{n-1}, X^{n-1}(V)) + F$$

$$\text{(b)} \quad I(S_V; M_1, M_2, \ldots, M_{|V|}) + \sum_{n=1}^{N} I(Y_n; V, S_V, M_V, X^{n-1}(V)) + F \leq I(S_V; M_1, M_2, \ldots, M_{|V|}) + \sum_{n=1}^{N} I(Y_n; V, S_V, M_V, Y^{n-1}, X^{n-1}(V)) + F$$

$$\text{(c)} \quad \frac{1}{|V|} \sum_{v=1}^{|V|} I(S_v; M_v) + NI(U; Y) + F,$$  

(38)

where step (a) follows from the fact that the biometric sequence $Y^N$ is independent of all the helper data other than the helper data corresponding to the actual individual’s identity; (b) holds since $Y^{n-1} - (V, S_V, M_V, X^{n-1}(V)) = Y_n$. Finally, to obtain (c), we define $U_n \triangleq (V, S_V, M_V, X^{n-1}(V))$ for $n = 1, 2, \ldots, N$, then if we take a time-sharing variable $T$ uniform over $\{1, 2, \ldots, N\}$ and independent of all other variables and set $U \triangleq (U_n, n), X \triangleq X_n$, and $Y \triangleq Y_n$ for $T = n$, we get

$$\sum_{n=1}^{N} I(V, S_V, M_V, X^{n-1}(V); Y_n) = \sum_{n=1}^{N} I(U_n; Y_n) = NI(U_T; Y_T|T) = NI(U_T; Y_T) = NI(U; Y).$$  

(39)

Finally, note that $U_n - X_n = Y_n$ and, consequently, $U - X = Y$.

Now if we assume that a triple $(R_S, R_T, R_L)$ is achievable, then $F \leq 1 + \delta \log(|Y| \cdot |S|)$ and we obtain that

$$\log(|Y| \cdot |S|) \leq \log |Y| + \min_{v=1;2;\ldots;|V|} H(S_v) + N \delta \leq H(V) + H(S_V | Y) + N \delta \leq 2N \delta + NI(U; Y) + 1 + \delta \log(|Y| \cdot |S|),$$

(40)

and finally that

$$R_I + R_S - 2\delta \leq \frac{\log(|Y| \cdot |S|)}{N},$$

(41)

for some $P(u, x, y) = Q_s(x)Q_e(y|x)P(u|x)$. Now we continue with the privacy leakage.

$$I(X^N(V); M_1, M_2, \ldots, M_{|V|})$$

$$= I(X^N(V), V; M_1, M_2, \ldots, M_{|V|}) - H(X^N(V), V, S_V) + H(V) + H(N^{X^N(V)}(V), S_V)$$

$$= H(X^N(V), V, S_V) - H(V) - H(N^{X^N(V)}(V), S_V) + H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, Y^N)$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, X^{n-1}(V))$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, Y^{n-1})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, X^{n-1})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, Y^{n-1})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, X^{n-1})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, Y^{n-1})$$

$$\geq H(V) - H(V, S_V | M_1, M_2, \ldots, M_{|V|}, X^{n-1})$$

$$\geq \log |Y| + NI(U; X) - NI(U; Y) - F,$$  

(42)

for the joint distribution $P(u, x, y) = Q_s(x)Q_e(y|x)P(u|x)$, mentioned before. Here step (a) holds, since $S_V$ is a function of $X^N(V)$.

Now for achievable triples $(R_S, R_T, R_L)$ we get

$$R_L + \delta \geq \frac{1}{N} \max_{v \in \{1;2;\ldots;|V|\}} I(X^N(v); M_v)$$

$$\geq \frac{1}{|V|} \sum_{v=1}^{|V|} I(X^N(v); M_v)$$

$$\geq \frac{1}{|V|} I(X^N(V); M_1, M_2, \ldots, M_{|V|})$$

$$\geq \frac{1}{|V|} (\log |V| + NI(U; X) - NI(U; Y) - F)$$

$$\geq R_I - \delta + I(U; X) - I(U; Y).$$
\[ -\frac{1}{N} (\delta \log(|V| \cdot |S|) + 1) \]
\[ \geq I(U; X) - \frac{1}{1-\delta} I(U; Y) + R_I \]
\[ - \frac{1}{|1-\delta|N} \frac{\delta(1+\delta)}{1-\delta}, \]  
(43)

here we used Fano’s inequality and (41).

Finally, if we let \( \delta \downarrow 0 \) and \( N \to \infty \), then we obtain the converse from both (41) and (43).

D. Bound on the Cardinality of \( U \)

To find a bound on the cardinality of the auxiliary variable \( U \) let \( D \) be the set of probability distributions on \( \mathcal{X} \) and consider the \(|\mathcal{X}|+1\) continuous functions of \( P \in \mathcal{D} \) defined as

\[ \phi_x(P) = P(x) \] for all but one \( x \),
\[ \phi_X(P) = H_P(X), \]
\[ \phi_y(P) = H_P(Y), \]

where in the last equation we use \( \Pr\{Y = y\} = \sum_x P(x) Q_{\mathcal{G}}(y|x) \). By the Fenchel-Eggleson strengthening of the Caratheodory lemma, see Wyner and Ziv [19], there are \(|\mathcal{X}|+1\) elements \( P_u \in \mathcal{D} \) and \( \alpha_u \) that sum to one, such that

\[ Q(x) = \sum_{u=1}^{1+|\mathcal{X}|} \alpha_u \phi_x(P_u) \] for all but one \( x \),
\[ H(X|U) = \sum_{u=1}^{1+|\mathcal{X}|} \alpha_u \phi_X(P_u), \]
\[ H(Y|U) = \sum_{u=1}^{1+|\mathcal{X}|} \alpha_u \phi_Y(P_u). \]

(45)

The entire probability distribution \( \{Q(x,y), x \in \mathcal{X}, y \in \mathcal{Y}\} \) and consequently the entropies \( H(X) \) and \( H(Y) \) are now specified and therefore also both \( I(U; X) \) and \( I(U; Y) \) are.

This implies that cardinality \(|U| = |\mathcal{X}|+1\) suffices for our region \( \mathcal{R}_bi \).

**APPENDIX B**

**Proof of Thm. 2**

The achievability proof for this theorem is based on the achievability proof for Thm. 1. The converse part is also an adapted version of the converse for Thm. 1. The proof for the bound on the cardinality is the same as for Thm. 1 and is therefore omitted.

A. Achievability Part of Thm. 2

The achievability proof corresponding to this theorem is based on the achievability proof of Thm. 1. Here however we make use of the masking layer, also applied in the achievability proofs for authentication systems. In this layer the generated secret \( S^0_v \) of each individual \( v, v \in \{1,2,\ldots,|V|\} \) is used to conceal an independently chosen secret \( S^0 \) in a one-time pad system. Now we denote by \( \oplus \) addition modulo \( |S| \) and by \( \ominus \) subtraction modulo \( |S| \). Then each individual \( v \) is given an additional helper data

\[ M^0_v = S^0_v \oplus S^0. \]  
(46)

The secret key can be reconstructed as

\[ \hat{S}^0_v = M^0_v \ominus \hat{S}^0 = S^0_v \ominus (S^0_v \ominus \hat{S}^0). \]  
(47)

Note that \( M^0_v \) need not be stored in the identification public database as it only required for authentication purposes. However, these helper data might be stored in a public database corresponding to an authentication system. Therefore, we will consider an aggregated helper data \( (M^0_1, M^0_2) \) in our proof when we analyze secrecy and privacy leakage.

Now taking into the account that \( \hat{S}^0 \) is uniform on \( \{1,2,\cdots,|S|\} \) and independent of \( X^N(v) \), the generated secret \( S^0_v \), and corresponding helper data \( M^0_v \), we obtain

\[ I(S^0_v; M^0_1, M^0_2) \]
\[ = I(S^0_v; M^0_1, S^0_v \oplus S^0_2) \]
\[ = I(S^0_v; M^0_1) + I(S^0_v; S^0_v \oplus S^0_2|M^0_2) \]
\[ \leq H(S^0_v \oplus S^0_2) - H(S^0_v \oplus S^0_2|M^0_2) \]
\[ = \log |S| - H(S^0_v | M^0_2, S^0_2) \]
\[ \leq \log |S| - H(S^0_v) + I(S^0_v; M^0_2) \]  
(48)

\[ I(X^N(v); M^0_1, M^0_2) \]
\[ = I(X^N(v); M^0_1, S^0_v \oplus S^0_2) \]
\[ = I(X^N(v); M^0_1) + I(X^N(v); S^0_v \oplus S^0_2|M^0_2) \]
\[ \leq I(X^N(v); M^0_1) + H(S^0_v \oplus S^0_2) \]
\[ - H(S^0_v \oplus S^0_2|M^0_2, X^N(v), S^0_2) \]
\[ \leq I(X^N(v); M^0_1) + \log |S| \]
\[ - H(S^0_v | M^0_1, X^N(v), S^0_2) \]
\[ = I(X^N(v); M^0_1) + \log |S| - \log |S| \]
\[ = I(X^N(v); M^0_1). \]  
(49)

Thm. 1 states that for all \( \delta > 0 \) and \( N \) large enough there exist encoders and decoders for which \( \Pr\{(\hat{V}, \hat{S}^0_v) \neq (V, S^0_v)\} \leq \delta \), and

\[ \frac{1}{N} \log |V| \geq R_I - \delta, \]
\[ \frac{1}{N} H(S^0_v) + \delta \geq \frac{1}{N} \log |S| \geq R_S - \delta, \]
\[ \frac{1}{N} I(S^0_v; M^0_2) \leq \delta, \]
\[ \frac{1}{N} I(X^N(v); M^0_1) \leq R_L + \delta, \]

for all \( v \in \{1,2,\cdots,|V|\}. \)  
(50)

Note that using the masking layer implies that \( \hat{S}^0_v \) is independent \( S^0_v \) only if \( \hat{S}^0_v = S^0_v \), and thus \( \Pr\{(\hat{V}, \hat{S}^0_v) \neq (V, S^0_v)\} \leq \delta \), and also

\[ \frac{1}{N} \log |V| \geq R_I - \delta, \]
\[ \frac{1}{N} H(S^0_v) = \frac{1}{N} \log |S| \geq R_S - \delta, \]
\[ \frac{1}{N} I(S^0_v; M^0_2) \leq 2\delta, \]
\[ \frac{1}{N} I(X^N(v); M^0_1) \leq R_L + \delta, \]

for all \( v \in \{1,2,\cdots,|V|\}. \)  
(51)
Consequently secret-key, identification and privacy-leakage rate triples \((R_S, R_I, R_L)\) that are achievable for protected biometric identification systems with secret generation are also achievable for systems with secret binding.

**B. Converse Part of Thm. 2**

Note that for systems with secret binding, the secret key and the individual’s identity label are uniform and independent of each other. Now just as in the converse for Thm. 1 we obtain

\[
\log(|V| \cdot |S|) = H(V, S) + I(U; Y) + F,
\]

where we used the Markov property \(Y^{n-1} = (V, S, U, X^n, Y^n)\) that also applies here. Just as before we define \(U_n \triangleq (V, S, U, X^n)\) for \(n = 1, 2, \ldots, N\) and take a time-sharing variable \(T\) uniform over \(\{1, 2, \ldots, N\}\) and independent of all other variables and set \(U \triangleq (U_n, n), X \triangleq X_n,\) and \(Y \triangleq Y_n\) for \(T = n\). Then \(U_n = X_n - Y_n\) and, consequently, \(U = X - Y\) hold. Then for achievable triples \((R_S, R_I, R_L)\) we get that

\[
R_I + R_S - 2\delta \leq \frac{\log(|V| \cdot |S|)}{N} \leq \frac{1}{1 - \delta} (I(U; Y) + \delta + \frac{1}{N}),
\]

for some \(P(u, x, y) = Q_u(x)Q_s(y|x)P(u|x)\).

For the privacy leakage we obtain as before

\[
R_L + \delta \geq I(U; X) - \frac{1}{1 - \delta} I(U; Y) + R_t - \frac{1}{(1 - \delta)N} - \delta \left(1 - \frac{\delta}{1 - \delta}\right),
\]

for the joint probability \(P(u, x, y)\) mentioned before.

Now letting \(\delta \downarrow 0\) and \(N \to \infty\), we obtain the converse from both (53) and (54).

**REFERENCES**


---

Tanya Ignatenko (S’06-M’08) was born in Minsk, Belarus, in 1978. She received the M.Sc. degree in applied mathematics from Belarussian State University, Minsk, in 2001. She received the F.D.Eng. and Ph.D. degrees from Eindhoven University of Technology, Eindhoven, The Netherlands, in 2004 and 2009, respectively. From 2008 till 2010, she was a postdoctoral researcher at Eindhoven University of Technology; from 2011 till 2013, she was a research scientist at Philips Research Laboratories. Currently she is an assistant professor at the Electrical Engineering Department, Eindhoven University of Technology. Her research interests include biometric systems privacy and security, privacy-preserving information processing, and multi-user information theory.

Dr. Ignatenko is a co-recipient of the 2011 IEEE Signal Processing Society Best Paper Award for the paper “Biometric Systems: Privacy and Security Aspects.” She has published more than 30 journal, conference papers and book chapters, and she holds a number of patents.
Frans M.J. Willems (S’80-M’82-SM’05-F’05) was born in Stein, The Netherlands, in 1954. He received the M.Sc. degree in Electrical Engineering from the Technische Universiteit Eindhoven, Eindhoven, The Netherlands, and the Ph.D. degree from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1979 and 1982 respectively.

From 1979 to 1982 he was a research assistant at the Katholieke Universiteit Leuven. In 1982 he joined the Electrical Engineering Department of the Technische Universiteit Eindhoven, where he is an associate professor now. His research contributions are in the areas of multi-user information theory, noiseless source coding, data-embedding, and biometrics. Since 1999 he is an advisor for Philips Research Laboratories for subjects related to information theory.

Dr. Willems received the Marconi Young Scientist Award in 1982. From 1988 to 1990, he served as Associate Editor for Shannon Theory for the IEEE Transactions on Information Theory. From 2002 to 2006 he was Associate Editor for Information Theory for the European Transactions on Telecommunications. He is co-recipient of the 1996 IEEE Information Theory Society Paper Award for a paper in which the Context-Tree Weighting Algorithm was proposed. From 1998 to 2000 he was a member of the Board of Governors of the IEEE Information Theory Society. Since 2005 Dr. Willems is an IEEE Fellow. He is Counselor of the IEEE Student Branch Eindhoven and Chairman of the IEEE Benelux Chapter on Information Theory. Recently Dr. Willems received a 2011 Best Paper Award from the IEEE Signal Processing Society for the paper “Biometric Systems: Privacy and Security Aspects.”

Dr. Willems has contributed more than two hundred journal and conference papers. He holds several patents.