Diagnostic Category Leakage in Helper Data Schemes for Biometric Authentication

Joep de Groot[1], Boris Škorić[2], Niels de Vreede[2] and Jean-Paul Linnartz[1]
[1]*Signal Processing Systems, Eindhoven University of Technology, Eindhoven, The Netherlands*
[2]*Security and Embedded Networked Systems, Eindhoven University of Technology, Eindhoven, The Netherlands*

Keywords: Biometrics, Secrecy Leakage, Privacy Leakage, Helper Data Scheme, Template Protection.

Abstract: A helper data scheme (HDS) is a cryptographic primitive that extracts a high-entropy noise-free secret string from noisy data, such as biometrics. A well-known problem is to ensure that the storage of a user-specific helper data string in a database does not reveal any information about the secret. Although Zero Leakage Systems (ZSL) have been proposed, an attacker with a priori knowledge about the enrolled user can still exploit the helper data. In this paper we introduce diagnostic category leakage (DCL), which quantifies what an attacker can infer from helper data about, for instance, a particular medical indication of the enrolled user, her gender, etc. The DCL often is non-zero. Though small per dimension, it can be problematic in high-dimensional biometric authentication systems. Furthermore, partial a priori knowledge on of medical diagnosis of the prover can leak about the secret.

1 INTRODUCTION

Nowadays authentication and identification applications rely more and more on biometrics, since it is considered a convenient solution and difficult to forge. Contrary to passwords and tokens biometrics cannot be forgotten or lost, but are inherently bound to the user. They truly identify who someone is, rather than what somebody knows or possesses.

It will be difficult to keep biometrics strictly secret. For example, a face or iris can be captured as a photographic image unnoticed. Whereas a user's fingerprints can be found on many daily objects. However, in-the-clear storage of templates extracted from the biometric is out of the question, since that will make it very easy for an adversary to misuse them.

Roughly the problems introduced by storing biometric features can be split into security and privacy risks (Jain et al., 2005). The former include the reproduction of fake biometrics from the stored features, e.g. rubber fingers (Matsumoto et al., 2002). These fake biometrics can cause serve security issues, e.g. to obtain unauthorized access to information or services or to leave fake evidence at crime scenes. Such actions are commonly known as identity theft.

On the other hand there are privacy risks bound to the application of biometrics (Labati et al., 2012). The most sensitive are: (i) some biometrics might reveal diseases and disorders of the user and (ii) unprotected storage allows for cross-matching between databases.

Helper data schemes (HDS) (Juels and Wattenberg, 1999; Linnartz and Tuyls, 2003; Dodis et al., 2004; Juels and Sudan, 2006; Chen et al., 2007) have been proposed to ensure that hashes of biometrics can be stored, such that even during verification no in-the-clear biometric templates can be retrieved from a database. These schemes exploit a prover-specific variable, called the helper data to ensure reliable exact digital reproducibility of a biometric value.

Zero Secrecy Leakage (ZSL) helper data schemes have been proposed (Verbitskiy et al., 2010; de Groot and Linnartz, 2011; de Groot and Linnartz, 2012), to ensure that the mutual information between the helper data and the secret key is zero. However, it has been recognized that this property does not fully ensure total protection of the prover's privacy.

Ignatenko and Willems (Ignatenko and Willems, 2009) introduced the notion of privacy leakage, defined as the mutual information between helper data and the biometric value it self as opposed to the helper data and the secret. Yet we are not aware of any paper that confirms the severity of the theoretical privacy leakage in terms of how much valuable information the attacker actually gets about the prover. If for instance the biometric is the length of a person, many helper data schemes, such as (de Groot and Linnartz,

de Groot J., Skoric B., de Vreede N. and Linnartz J..
Diagnostic Category Leakage in Helper Data Schemes for Biometric Authentication.
DOI: 10.5220/0004524205060511
In *Proceedings of the 10th International Conference on Security and Cryptography* (SECRYPT-2013), pages 506-511
ISBN: 978-989-8565-73-0

2011), leak that the last decimals of the value, for instance are 593, but give no clue about whether it is an 1.63593 meter small person or 1.93593 meter tall person. In this paper we address the question whether such leakage is serious. For instance if we know from the helper data of a cyclist that his heart rate is equal to an unknown integer plus some known fraction, how much does that tell us about the likelihood of an enlarged EPO concentration in his blood. In this paper we analyze such questions.

Another form of key or privacy leakage (de Groot and Linnartz, 2011) can occur when the attacker has a priori knowledge about the prover, or about any person in the data base. For instance that the cyclist is a 28 year old female.

Our current paper has been motivated by an implementation project that records data from epileptic patients from body sensor networks, with biometric configuration of the radio links. Here we encountered the question of how severe such issues are for practical biometrics.

We perform a security analysis for three important scenarios. (i) The case of a mismatch between the true distribution of the features x and the distribution used for creating helper data w. The attacker is assumed to know the true distribution. (ii) An attacker who has partial information about enrolled users, e.g. a medical indication or gender, and tries to learn something about the stored secret. (iii) An attacker who tries to learn something about the enrolled user's characteristics by exploiting the public helper data and some a priori partial information about the user.

These scenarios lead to a mismatch between the distribution as seen by the attacker and the distribution used to make w. The question is how much the ZSL helper data w leaks under these circumstances, in addition to the already existing leakage. We prove an upper bound on this additional leakage.

2 ZERO SECRECY LEAKAGE SCHEME

We consider a commonly accepted verification scheme which consists of an enrollment and verification phase. In the enrollment phase the prover provides his biometric data $\underline{x} = (x_0, \ldots, x_{M-1})$. From this data, the system extracts a secret $\underline{s} = Q(\underline{x})$, which the system stores safely in the hashed form $(h(\underline{s}\|z), \underline{w})$, where $\underline{w}$ is the helper data, which is generated as $\underline{w} = g(\underline{x})$ and z is the salt. The salt is a system and/or user specific random string to prevent cross-matching between different databases. In the verification phase the prover provides his correlated biometric data $\underline{y} = (y_0, \ldots, y_{M-1})$ to prove his identity. All variables, except for the salt z, are length M vectors extracted by some means of preprocessing, to ensure that the components are (nearly) independent, but not necessarily identically distributed. Independence can be obtained by for example applying a principle component analysis (PCA) to the raw data.

Analysis will be carried out per dimension since we have assumed the features to be independent. In this case the total leakage in a verification scheme will be a summation of the leakage per dimension. For clarity notation of the biometric feature x, secret s and helper data w will be without subscript i.

Initially, leakage elimination has been studied (Verbitskiy et al., 2010) for secret values that are equiprobable (Fuzzy Extractor). Each interval belonging to a secret is then subdivided in equiprobable intervals to define the helper data. The helper data intervals are repeated for each interval of the secrets. This construction yields helper data whose probability is independent of the enrolled secret.

Meanwhile, it has been argued that verification performance highly depends of effective quantization of the analog (continuous valued) biometrics and continuous-valued helper data within the quantization intervals (Linnartz and Tuyls, 2003; Chen et al., 2007). Also in this domain, leakage is a concern (de Groot and Linnartz, 2011; de Groot and Linnartz, 2012). Instead of demanding equiprobable *discrete* values as helper data, helper data w is defined as a continuous variable that indicates the relative position of the enrollment feature x within a quantization interval belonging to a secret s. To achieve ZSL the scheme has to take into account the probability density of the features. ZSL is achieved in this case by

$$s = Q(x) = \lfloor N \cdot F_X(x) \rfloor, \tag{1}$$

$$w = g(x) = N \cdot F_X(x) - s \tag{2}$$

in which N is the number of quantization intervals and F_X is the cumulative distribution function (CDF) of feature x. The number of quantization intervals N does not necessarily have to be a power of 2.

The above construction yields a continuous helper data w that reveals no information about the enrolled secret s. In fact one can only reconstruct N possible x values, each in a different quantization interval. This reconstruction is given by

$$x_s(w) = F_X^{-1}\left(\frac{s+w}{N}\right) \tag{3}$$

In this work we will limit ourselves to a leakage analysis on the continuous scheme only, since the discrete scheme can be considered a special case of the continuous version.

3 LEAKAGE ANALYSIS

For the leakage analysis presented in this section we will make a distinction between a priori leakage and additional leakage due to the public helper data. The former is solely due to the assumed improved understanding of the biometric features by the attacker and assumes the attacker does not yet exploit the information in the helper data. Whereas the latter is this possible "bonus" due to exploiting the combination of a priori knowledge and public helper data.

3.1 Mismatch Between the Real and Assumed Distribution

The distribution $f_{\text{sys}}(x)$ used by the authentication system is not exactly equal to the real distribution f_X of X. When the system is set up, the statistical knowledge about X is based on a finite number of observations, from which f_{sys} is derived. Due to finite size effects a (small) mismatch between f_X and f_{sys} arises. It is prudent to assume that attackers have full knowledge of f_X, e.g. due to scientific progress after the system has been fixed. Given this mismatch, the probabilities for S and W are derived as follows. First of all we can derive the joint density of the helper data and secret as

$$\chi(s,w) = \frac{1}{N}\frac{f_X(x_s(w))}{f_{\text{sys}}(x_s(w))} \tag{4}$$

which follows from $f_X(x)\mathrm{d}x = \chi(s,w)\mathrm{d}w$ and $\mathrm{d}x = \mathrm{d}w/[N f_{\text{sys}}(x)]$ evaluated at $x = x_s(w)$. The probability of the secrets follows from integrating f_X between the boundary points that correspond to $S = s$, hence

$$\chi(s) = \mathbb{P}(S = s) = F_X(x_s(1)) - F_X(x_s(0)). \tag{5}$$

Finally, the marginal

$$\chi(w) = \frac{1}{N}\sum_{s=0}^{N-1}\frac{f_X(x_s(w))}{f_{\text{sys}}(x_s(w))} \tag{6}$$

follows from (4) by summing over s. These probability functions can subsequently be used to determine the leakage

$$I(S;W) = \sum_s \int_0^1 \chi(s,w)\log_2\frac{\chi(s,w)}{\chi(s)\chi(w)}\,\mathrm{d}w, \tag{7}$$

in which I stands for mutual information.

An example for such leakage is given in Figure 1. This particular example assumes both distributions, real and assumed, to be Gaussian and the number of quantization intervals $N = 4$.

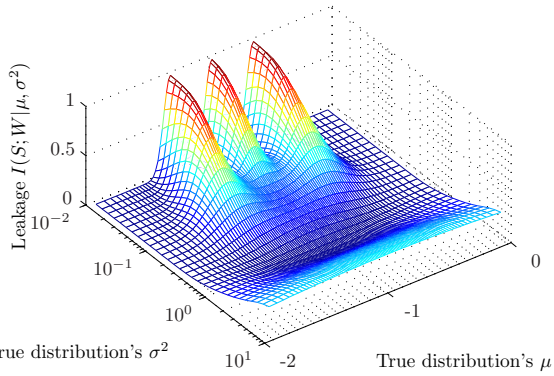


Figure 1: Additional leakage due to mismatch between real distribution and assumed distribution for $N = 4$. Only for $(\mu,\sigma^2) = (0,1)$ the leakage is zero.

3.2 Related Property Known by Attacker

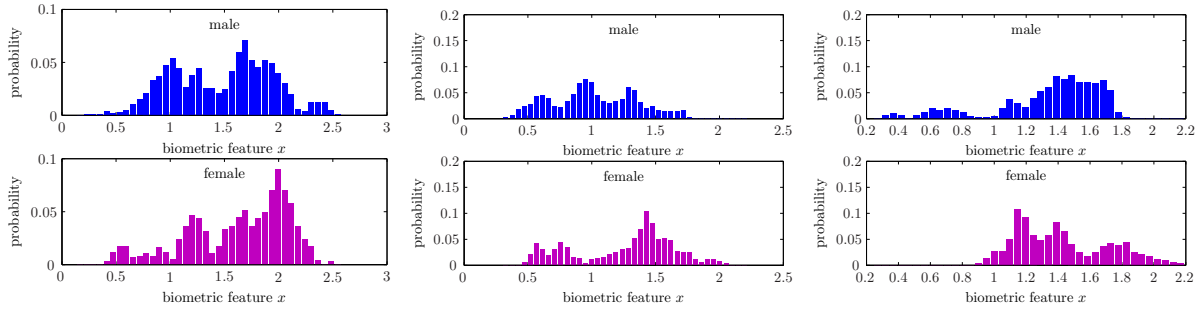
There is another source of leakage. It may happen that the statistics of the measured quantity x depend on, e.g. the gender of the enrolled users, skin color, medical diagnosis, or some other (discrete) property.

This idea has been motivated by results from a biometric verification experiment with ECG signals. The extracted features showed a clear divergence when sorted by gender. A few striking examples are depicted in Figure 2. These features were obtained by calculating autocorrelation (AC) on 1 minute epochs and subsequently applying a discrete cosine transform (DCT) (Agrafioti and Hatzinakos, 2008).

We will consider a general discrete category $C \in \mathcal{C}$. We ask ourselves the question whether an attacker can gain an advantage from some observation $\tilde{C} \in \mathcal{C}$ which yields (partial) knowledge about the category C. One can think of $\tilde{C}$ as an estimate derived from an observation for a specific person, e.g. gait or height, or an observation of the whole enrolled population, e.g. the percentage of men vs. women. For example the observation could be a 1.9m tall person, which might give rise to the assumption it is a man, since men are usually taller than women. However, we might be dealing with an exceptionally tall woman.

We will investigate two attack scenarios:

1. *Secret Estimation*
   The attacker wants to leverage the side information to derive a better guess for an enrolled person's secret S. In this scenario the mutual information $I(S;W,\tilde{C})$ is the quantity of interest.
2. *Category Estimation*
   Based on the side information, the attacker wants to diagnose an enrolled person's category C (medical indication). In fact we generalize this to any



(a) DCT coefficient 2. (b) DCT coefficient 4. (c) DCT coefficient 5.

Figure 2: Examples of deviating distributions per gender in a ECG based verification experiment.

privacy sensitive category including gender, race, etc. Here the quantity of interest is $I(C;W,\tilde{C})$.

For given c, we have to consider the $f_X(x)$ in (4) to $f_{X|C}(x|c)$. The $f_X(x)$ remains unchanged, since the enrollment is done without regard to categories. Note that S and W have no additional dependence on C. By using the joint probability $Q_{c\tilde{c}}$ of c and $\tilde{c}$ and the chain rule $\chi(s,w,c,\tilde{c}) = Q_{c\tilde{c}}\,\chi(s,w|C=c)$ we can write

$$\chi(s,w,c,\tilde{c}) = Q_{c\tilde{c}}\,\frac{1}{N}\,\frac{f_{X|C}(x_s(w)|c)}{f_X(x_s(w))}. \tag{8}$$

From (8) we can derive all the marginal distributions that are necessary for computing $I(S;W,\tilde{C})$ and $I(C;W,\tilde{C})$. Some examples assuming Gaussian distributions are depicted in Figure 3(a) and Figure 4(a).

3.2.1 Bound on the Secrecy Leakage

We will show that the total amount of information that can be obtained is very limited. The expression of the mutual information between enrolled secret s and public data, i.e. helper data w and category estimate $\tilde{c}$, can be split in two terms

$$I(S;W,\tilde{C}) = I(S;W) + I(S;\tilde{C}|W). \tag{9}$$

Since the scheme is a zero leakage key extraction scheme, i.e. $I(S;W) = 0$, it follows that

$$I(S;W,\tilde{C}) = I(S;\tilde{C}|W) \leq H(\tilde{C}|W) \leq H(\tilde{C}) \tag{10}$$

where H stands for Shannon entropy and I for mutual information. Therefore we can conclude that the secrecy leakage satisfies

$$I(S;W,\tilde{C}) \leq H(\tilde{C}). \tag{11}$$

This bound, which limits the amount of information about the secret that can be obtained in a ZSL scheme, is limited by the entropy in the category estimate and is independent of the public helper data and the type of ZSL scheme. If an attacker for example knows the gender of an enrolled user, he can never learn more than 1 bit even if the secret is more than 1 bit.

4 TOY EXAMPLE: GAUSSIAN DISTRIBUTIONS

4.1 Secret Estimation

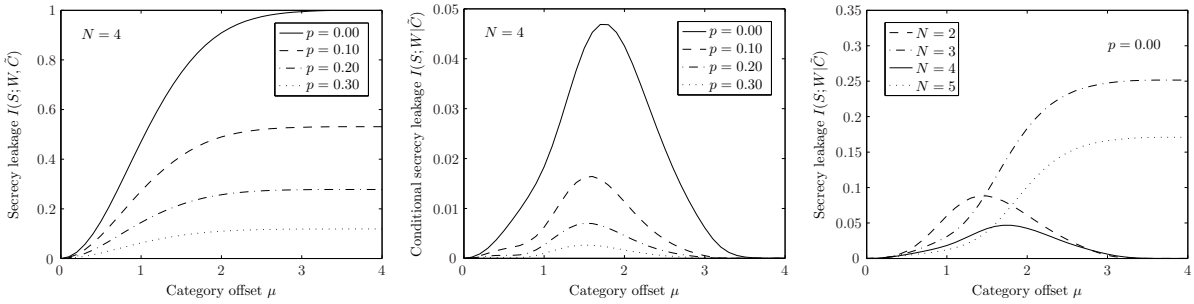
In this section we study the situation that emerges when an attacker knows $\tilde{C}$, i.e. an estimate of the category C of the enrolled user. For this particular example we construct the category as a single bit. For example "0" is male and "1" is female. The estimate is derived from the actual category with some error p, which is modeled as a Binary Symmetric Channel (BSC) with cross-over probability p. We assume that a priori both categories are equiprobable, thus

$$Q_{c\tilde{c}} = \begin{cases} \frac{1}{2}(1-p) & c = \tilde{c} \\ \frac{1}{2}p & c \neq \tilde{c} \end{cases}. \tag{12}$$

For the feature distribution we assume a Gaussian Mixture Model (GMM) with two distributions, which represent the two categories. The parameters for this model are set to $\mu_0 = -\mu$, $\mu_1 = \mu$ and $\sigma_0^2 = \sigma_1^2 = 1$. This mean value parameter $\mu \geq 0$ will be varied together with error probability p to study the emerging leakage in the system.

To calculate $x_s(w)$ in Eq. (8) we need to calculate the inverse CDF of the Gaussian mixture as given by Eq. (3). This has been solved by applying Newton's method to the given PDF and CDF of the Gaussian mixture. For arguments smaller than $1/2$, μ_0 was used as initial guess and for arguments larger than $1/2$, μ_1, which ensured a rapid convergence and accurate results.

The inverse CDF allows us to calculate the joint probability density function $\chi(s,w,\tilde{c})$ as a function of w. This marginal is derived from Eq. (8). Subsequently we can calculate the secrecy leakage in terms of mutual information as



(a) Leakage by category estimate $\tilde{c}$ and helper data w.

(b) Component of the leakage due to helper data w given $\tilde{c}$.

(c) Component of the leakage due to helper data w given $\tilde{c}$ for different number of quantization intervals N.

Figure 3: Leakage of secret S in secret estimation scenario.

$$I(S;W,\tilde{C}) = \sum_{s,\tilde{c}} \int_0^1 \chi(s,w,\tilde{c}) \log_2 \frac{\chi(s,w,\tilde{c})}{\chi(s)\chi(w,\tilde{c})} \mathrm{d}w. \tag{13}$$

At increasing value of μ we observe a clear saturation for the total leakage $I(S;W,\tilde{C})$. Moreover, the better the estimate ($p \to 0$), the more information an attacker obtains. However, even for $\mu \gg 0$ and $p = 0$, i.e. a perfect category estimate, there is a maximum leakage of 1 bit, which agrees with the bound found in Section 3.2.1. The results of this calculation for different values of p can be found in Figure 3(a).

A distinction can be made between leakage by a priori knowledge of the category $I(S;\tilde{C})$ irrespective of the helper data and "bonus" leakage $I(S;W|\tilde{C})$ caused by the category estimate $\tilde{c}$ combined with knowledge of the helper data w. So

$$I(S;W,\tilde{C}) = I(S;\tilde{C}) + I(S;W|\tilde{C}) \tag{14}$$

and by doing the numerics for

$$I(S;W|\tilde{C}) = \sum_{s,\tilde{c}} \int_0^1 \chi(s,w,\tilde{c}) \log_2 \frac{\chi(\tilde{c})\chi(s,w,\tilde{c})}{\chi(s,\tilde{c})\chi(w,\tilde{c})} \mathrm{d}w \tag{15}$$

we can assess the amount of leakage actually caused by the helper data scheme.

In the special case of symmetric distributions and an even number of quantization intervals (as assumed in Figure 3(b)), for $\mu \gg 0$ it holds that $I(S;W|\tilde{C}) \to 0$. This effect is caused by the fact that the two category distributions become favorably located over the quantization intervals. However, for more unfortunate choices, e.g. odd N, this favorable effect is not present, as can be seen in Figure 3(c).

We conclude that leakage can only be severe for a pre-informed attacker who has specific a priori knowledge. However, such a situation closely resembles a situation in which an attacker possesses the biometric feature x itself and not a single ZSL scheme can protect against such well informed attackers, as in the limiting case the attacker knows as much as the verifier.

4.2 Category Estimation

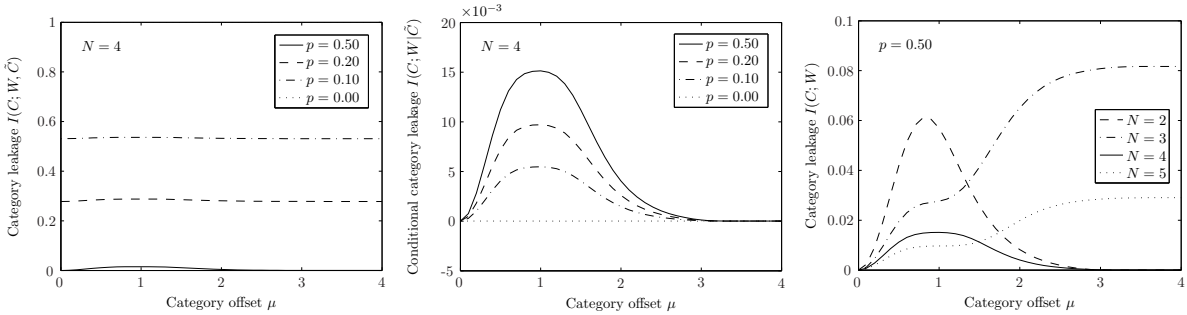
For the scenario that an attacker tries to extract privacy–sensitive information about a category (e.g gender, race, epileptic indications, use of certain mediation or drugs) to which the prover belongs, we can obtain similar results. The total information about C can again be split in a part from the estimate $\tilde{C}$ and a part caused by the helper data W as follows

$$I(C;W,\tilde{C}) = I(C;\tilde{C}) + I(C;W|\tilde{C}). \tag{16}$$

Most information about the category is obtained from the category estimate $\tilde{C}$. Since we modeled this estimate as a BSC this equals $1 - h(p)$. In this equation $h(p)$ is the binary entropy function. This effect can also be seen in Figure 3(a). The contribution of the helper data is only partial as confirmed by Figure 4(b). Also the convergence to zero for $\mu \gg 0$ only applies for even N as can be seen in Figure 4(c). In this example we have set $p = .5$, which effectively removes the a priori knowledge on $\tilde{C}$.

However, the leakage as show in Figure 4(b) and Figure 4(c) might seem small, but this is a leakage per dimension. An authentication scheme will in general use more the one dimension and it is not unlikely that the category under consideration will have influence on more than a single dimension, as is also confirmed in Figure 2. In case one wishes to determine a binary quantity, e.g. gender, with high probability this could be possible by combining the information from all available dimensions.

In fact, biometric secret extraction of 64 bits or more may typically require several tens of dimensions. Although such a system can be secure in terms of key entropy, it may inadvertently reveal privacy-sensitive information about the subject and even give



(a) Leakage by category estimate $\tilde{c}$ and helper data w.

(b) Component of the leakage due to helper data w given $\tilde{c}$.

(c) Component of the leakage due to helper data w given $\tilde{c}$ for different number of quantization intervals N.

Figure 4: Leakage of category C in category estimation scenario.

the attacker almost certainty about certain (binary) medical diagnoses. Using more dimensions from improved biometric feature extraction thus creates a privacy issue.

5 CONCLUSIONS

We have studied and quantified two kinds of leakage. The first due to a mismatch that can emerge due to improved understanding of feature distributions after the system has been set up, and the second if the attacker knows an enrolled user belongs to a specific category with a specific feature distribution. We for the latter we distinguished between the leakage about the enrolled secret and about the (medical diagnostic, racial, etc.) category.

From the results we can conclude that most of the leakage is caused by a priori information and only little information is revealed by the helper data. Only situations in which very specific information is known to the attacker can cause more serious key leakage. We believe that the Diagnostic Category Leakage (DCL), which has been introduced in this paper, can serve as a practical measure for privacy-sensitive leakage of biometric systems.

REFERENCES

Agrafioti, F. and Hatzinakos, D. (2008). ECG based recognition using second order statistics. In *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*, pages 82 –87.

Chen, C., Veldhuis, R., Kevenaar, T., and Akkermans, A. (2007). Multi-bits biometric string generation based on the likelihood ratio. In *Proc. IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems*.

de Groot, J. and Linnartz, J.-P. (2011). Zero leakage quantization scheme for biometric verification. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*

de Groot, J. and Linnartz, J.-P. (2012). Optimized helper data scheme for biometric verification under zero leakage constraint. In *Proc of the 33st Symp on Inf Theory in the Benelux*.

Dodis, Y., Reyzin, L., and Smith, A. (2004). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *LNCS*. Springer.

Ignatenko, T. and Willems, F. (2009). Biometric systems: Privacy and secrecy aspects. *Information Forensics and Security, IEEE Transactions on*, 4(4):956 –973.

Jain, A. K., Ross, A., and Uludag, U. (2005). Biometric template security: Challenges and solutions. In *Proceedings of European Signal Processing Conference*, pages 1–4.

Juels, A. and Sudan, M. (2006). A fuzzy vault scheme. *Des. Codes Cryptogr.*, 38:237–257.

Juels, A. and Wattenberg, M. (1999). A fuzzy commitment scheme. In *CCS '99: Proceedings of the 6th ACM conf on Comp and comm security*.

Labati, R. D., Piuri, V., and Scotti, F. (2012). Biometric privacy protection: Guidelines and technologies. In *E-Business and Telecommunications*, pages 3–19. Springer.

Linnartz, J.-P. and Tuyls, P. (2003). New shielding functions to enhance privacy and prevent misuse of biometric templates. In *Audio- and Video-Based Biometric Person Authentication*. Springer.

Matsumoto, T., Matsumoto, H., Yamada, K., and Hoshino, S. (2002). Impact of artificial "gummy" fingers on fingerprint systems. *Optical Security and Counterfeit Deterrence Techniques*, 4677:275–289.

Verbitskiy, E. A., Tuyls, P., Obi, C., Schoenmakers, B., and Škorić, B. (2010). Key extraction from general nondiscrete signals. *Information Forensics and Security, IEEE Transactions on*, 5(2):269 –279.