Unclonable identification and authentication based on reference list decoding

Sviatoslav Voloshynovskiy*, Oleksiy Koval, Fokko Beekhof and Thierry Pun

University of Geneva, Department of Computer Science, 7 route de Drize, CH 1227,
Geneva, Switzerland
http://sip.unige.ch/

Abstract. In this paper we advocate a new approach to item identification based on physical unclonable features. Being unique characteristics of an item, these features represent a kind of unstructured random codebook that links the identification problem to digital communications via composite hypothesis testing. Besides this, complexity, memory storage and universality constraints should be taken into account for databases with several hundred millions entries. Therefore, we attempt to find a trade-off between performance, security, memory storage and universality constraints. A practical suboptimal method is considered based on our reference list decoding (RLD) framework. Similar setup is extended to the authentication problem.

1 Introduction

Historically protection of items (or physical objects) is based on the technologies, which use some features being difficult to duplicate, copy or clone. Authentication is considered as the process of verification of added specific overt, covert or forensic features to the item that allow to verify the item as genuine. Examples of these protection technologies include special magnetic taggants [9], invisible inks, crystals or powders with infrared (IR) or ultraviolet (UV) properties [11], optically variable materials, holograms and physical paper watermarks [2], *etc.* The verification of item authenticity is based either on the direct visual inspection of the added feature presence or special devices (detectors) that usually have proprietary character and are not assumed to be used in public domain by the fear of disclosing the proprietary technology secrets behind the used technology and physical phenomena.

The identification of items refers to the assignment of a special index to every item that can be used for its tracking and tracing. The assigned index is encoded and stored on the item in a printed form of a visible barcode [8] or invisible digital watermark [4] or overprinted sparse set of dots (usually yellow) [13] or specially designed storage device such as magnetic stripe, electronic smart cards or RFID [14]. Obviously, the information stored in such a way can be read and copied by any party even if the data are encrypted.

* The contact author is S. Voloshynovskiy (email: svolos@cui.unige.ch).

The security drawbacks of these approaches in authentication and identification applications are well known. Besides that the above techniques are mostly proprietary and kept secret by the security printing houses for years, they are still the most widely used. Contrarily, even the usage of the most advanced cryptographic techniques in the above identification protocol does not help a lot since the data are easily copied without the need to be decrypted. Additionally, new storage devices (electronic chips, RFIDs) are still quite expensive for large scale applications. Sometimes, there is also no possibility to embed such a device into the item structure or their presence is not acceptable due to various legal, commercial, marketing, ecological and technical reasons.

Moreover, another security drawback of both technologies is their *adhesiveness*. The protection mechanisms considered above are added to the item as independent objects or features, sometimes changing the properties, look, design and value of an item. This has very serious complications that were not considered in early protection systems. First, the added feature has nothing to do with the actual item and its unique features and physical properties. Secondly, all these protection features can be relatively easily reproduced by modern means.

It is well known for years that all objects and humans are unique due to the possession of special features that are difficult to clone or copy. These unclonable features are *random microstructures* for the physical object surfaces and *biometrics* (fingerprint, iris, *etc.*) for the humans. The unclonable features are formed by nature and naturally integrated into the items. Having a lot of advantages and being non-adhesive in the considered sense, nevertheless the unclonable features only recently become a subject of intensive theoretical investigation mostly thanks to the progress achieved in the design of cheap high resolution imaging devices. This mostly concerns biometrics while the usage of microstructures is still an open practical and challenging theoretical problem. In this paper, we will thus concentrate on the theoretical analysis and practical implementation of identification/authentication techniques based on random microstructures.

For the sake of generality, we will define *physical unclonable features* (a.k.a. fingerprinting in some contexts) as unique features carried out by the objects, products, or documents. The main properties of such features are: (a) they can be extracted and evaluated in a simple way, but (b) they are hard to characterize and (c) in practice cannot be copied (cloned). The unclonable features are based on the randomness created by nature that is present in practically all physical structured observed under the coherent or noncoherent excitation (light) in transparent or reflective modes. Sometimes, this randomness can be hand-made. The examples of unclonable features include the microstructures of paper, metal, plastic, fibers, speckle, *etc.* Therefore, the main application of unclonable features is anti-counterfeiting for identification purposes.

Although the robustness/invariance aspects of identification problem have received a lot of attention especially in computer vision, the issue of security still remains an open problem. New information-theoretic and detection-theoretic approaches to secure identification, as well as carefully designed attacks, should be proposed and investigated.

The design of efficient identification techniques is a challenging problem that should address the compromise among various conflicting requirements covering:

- *performance*, i.e., the ability of identification function to produce reliable results under the legitimate distortions applied to the data;
- *security*, i.e., the inability of attacker to reproduce the physical unclonable feature or to trick the identification using the leaked information about the protocol; this also includes the one-way identification property similar to hashing functionsand collision-free property;
- *complexity*, i.e., the ability to perform the identification with the lowest computational effort without the considerable loss in identification accuracy;
- *memory storage*, i.e., the memory needed to store the codebook or features used for the identification;
- *universality*, i.e., the aspects of optimal identification under the lack of statistics about input source distribution and channel distortions that are related to the machine learning framework and universal hypothesis testing.

The above requirements are quite close to a robust perceptual hash function [6, 18, 23]. However, in the scope of this paper we will advocate a different approach that is based on the secure low-complexity multiple hypothesis testing in the secret domain defined by the key. This approach is also efficient in terms of memory storage requirements and is universal in terms of priors about the source distribution. That is why the goal of this paper is to introduce a decision-theoretic framework for the analysis and construction of unclonable identification that fits the above requirements.

This paper is organized as follows. In Section 2, we consider the existence of good identification codes in sense of achievable error exponent. The theoretical formulation of identification problem under certain ambiguity conditions is considered in Section 3.In Section 4, we consider a suboptimal algorithm, which trades performance for computational complexity. Finally Section 5 presents the results of experimental validation and Section 7 concludes the paper.

Notations We use capital letters to denote scalar random variables X, X^N to denote vector random variables, corresponding small letters x and x^N to denote the realizations of scalar and vector random variables, respectively. The superscript N is used to designate length-N vectors $x^N = [x[1], x[2], ..., x[N]]$ with k^{th} element $x[k]$. The i^{th} sequence is denoted as x_i^N. We use $X \sim p_X(x)$ or simply $X \sim p(x)$ to indicate that a random variable X is distributed according to $p_X(x)$. $p(x^N|H_m)$ denotes pdf/pmf of x^N under hypothesis H_m. Calligraphic fonts $\mathcal{X}$ denote sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ denotes the cardinality of set $\mathcal{X}$.

2 Unclonable identification as multiple hypothesis testing

The task of unclonable identification system is to infer the identity of the object based on its random physical features or fingerprint Y^N. The object fingerprint has to be matched against a possibly large number of fingerprints that are stored in the database (codebook) in order to assess 'the best match'. The index of the fingerprint m in the codebook is considered as an object ID or index.

The identification problem can be considered in the scope of pattern recognition [5] and digital communications [7]. Within the former theory typical approaches include the pattern classification based on the extraction of the unique features, their statistical characterization and classifier design with possible training. The performance analysis is generally based on experimental validation. The bounds on the system performance are deduced using Chernoff error analysis and large deviations.

Although, the problem resembles the evident similarity with the communication problem, at the same time it posses a number of significant differences. It concerns the encoding/decoding complexity as well as memory storage efficiency of modern capacity achieving coding strategies and error corrections codes based on structured codebooks. The considered identification problem is essentially based on random codewords, which is the key security element of such systems. However, simultaneously it complicates the design of efficient search/decoding and storage strategies. Taking into account the cardinality of codebook to be order of several millions, this can be a serious restriction for the on-line applications. Additionally, the distribution of the codewords in the codebook is fixed and determined by the nature and can not be chosen to achieve the capacity of the equivalent communication channel. Moreover, the communication problem assumes the presence of one of the codebook messages as the result of the decoding while the identification assumes the presence of a so-called zero-class, which is assigned if the observed codeword does not correspond to any of the messages.

Additionally, the identification is performed under geometrical desynchronization, which occurs during data acquisition. Finally, the aggressive behavior of the counterfeiter is not only reduced to the introduction of various distortions but it might target more severe attacks similar to the cryptographic ones.

An interesting extension of authentication protocol based on helper data was also presented in [**?**].

Thus, the identification problem can be considered as a stand-alone problem combining elements of pattern recognition, digital communications and cryptography. In this paper, we will consider the above problems in the scope of M-ary hypothesis testing known to be the theoretical basis for statistical classification, pattern recognition, machine learning and digital communications that also provides a nice link with the information theory [3, 7] and cryptography [17].

We will estimate the performance of hypothesis testing according to the average probability of error:

$$P_e = E_{p(y^N)} \left[1 - \max_{1 \le m \le |\mathcal{M}|} p(H_m|y^N) \right] = E_{p(y^N)} \left[\min_{1 \le m \le |\mathcal{M}|} p(H_m|y^N) \right], \quad (1)$$

where y^N is a vector of measured data and $p(H_m|y^N)$ denotes a posteriori probability of hypothesis H_m.

The test that minimizes the above error probability is the maximum a posteriori probability (MAP) decision rule:

$$\hat{m} = \psi_{MAP}(y^N) = \arg \max_{1 \leq m \leq |\mathcal{M}|} p(H_m|y^N) = \arg \max_{1 \leq m \leq |\mathcal{M}|} p(y^N|H_m)P(H_m), \tag{2}$$

where $p(y^N|H_m)$ is a likelihood term and $P(H_m)$ denotes a prior probability of each hypothesis.

In the case of known pmfs, the probability P_e is defined by the maximum pairwise probability of error $P_e^{i,j}$ between two hypothesis [15]:

$$P_e \leq \frac{|\mathcal{M}|(|\mathcal{M}|-1)}{2} \max_{i \neq j} P_e^{i,j}, \tag{3}$$

where $P_e^{i,j}$ depends on the minimum distance between two hypothesis:

$$\lim_{N \to \infty} \ln(\max_{i \neq j} P_e^{i,j}) = \min_{i \neq j} D_s(p(H_i|y^N), p(H_j|y^N)), \tag{4}$$

where $D_s(p_i, p_j)$ is the Chernoff distance defined as:

$$D_s(p(H_i|y^N), p(H_j|y^N)) = \max_{0 \leq s \leq 1} -\ln \int_{\mathcal{Y}} p(H_i|y^N) \left(\frac{p(H_j|y^N)}{p(H_i|y^N)} \right)^s dy^N. \tag{5}$$

3 Identification as composite multiple hypothesis testing

In the scope of this paper, the composite character of identification problem comes from a fact that the distribution of discrete memoryless source (DMS) $p_{X^N}(x^N)$ is not fully known. We will assume that the source generates the sequences x^N from the pmf $p_{X^N}(x^N|s_X^J)$, where s_X^J are the parameters of distribution given on a set $\mathcal{S}_X^J$ that is assumed to be discrete with the finite cardinality.

The channel is modeled as a cascade of a fixed discrete memoryless channel (DMC) given by the transition probability $p(v|x)$ and an invertible global mapping T_θ, which models a geometric transformation. We assume that the family $\{T_\theta, \theta \in \varTheta_N\}$ satisfies: (a) mapping invertibility $T_\theta : \mathcal{Y}^N \to \mathcal{V}^N$ for all N and for all $\theta \in \varTheta_N$; (b) restricted cardinality that at most is subexponential with N, i.e., $\limsup_{N \to \infty} \frac{1}{N} \ln |\varTheta_N| = 0$. This generic model was considered in details in [24] and here we will restrict our analysis to the case of DMC only (Figure 1).

We will estimate the performance of hypothesis testing according to the average probability of error (1) for a given set of source s_X^J parameters in $|\mathcal{M}|$-ary composite hypothesis testing and a chosen decision rule ψ:

$$P_e(s_X^J, \psi) = E_{p(y^N)} \left[\min_{1 \leq m \leq |\mathcal{M}|} p(H_m|y^N, s_X^J) \right]. \tag{6}$$

If the statistics of the source s_X^J are known, the test that minimizes the above error probability is the MAP decision rule (2) or the maximum likelihood (ML) decision rule, if all hypothesis are equiprobable, i.e., $p(H_m) = \frac{1}{|\mathcal{M}|}$:

$$\hat{m} = \psi_{ML}(y^N) = \arg \max_{1 \leq m \leq |\mathcal{M}|} p(y^N|s_X^J, H_m). \tag{7}$$

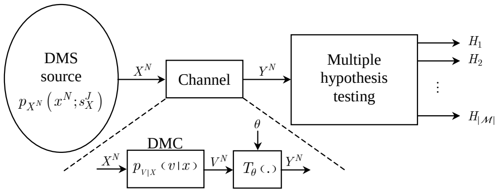

Fig. 1. Identification as multiple hypothesis testing.

It should be noticed that there exists generally no decision rule that achieves P_e, if the DMS parameters are not known. In practice, one attempts to construct *universal decision rules* that are independent of unknown parameters s_X^J. However, in general the performance will depend on them. Thus, in the scope of this paper we suppose that a universal test is *efficient*, if it achieves asymptotically close performance to the test with known statistics for all values of s_X^J:

$$\lim_{N\to\infty}\sup \max_{s_X^J \in \mathcal{S}_X^J} \Pr[\big|P_e(s_X^J, \psi) - P_e\big| \leq \xi] = 1, \tag{8}$$

where $\xi \to 0$ as $N \to \infty$. It should be pointed out, that there are other definitions of the efficient tests based on minimax rules.

In fact, considering the parameters of DMS s_X^J as random with some pmfs, one can apply Bayes approach using integration of $p(y^N|s_X^J, H_m)$ over the corresponding pmfs. However, this approach has some drawbacks related to: (a) the lack of knowledge of prior distributions; (b) once the realizations of parameters are drawn, they remain fixed through the experiment and (c) the integrals are difficult to compute in practice. Therefore, more often universal hypothesis testing based on generalized ML (GML) is used:

$$\psi_{GML}(y^N) = \arg \max_{1\leq m\leq |\mathcal{M}|} p(y^N|\hat{s}_X^J, H_m) \tag{9}$$

where $\hat{s}_X^J = \arg\max_{s_X^J \in \mathcal{S}_X^J} p(y^N|s_X^J, H_m)$ is the ML-estimate of s_X^J.

It is possible to show that the bound on difference in (8) is [22]:

$$P_e(s_X^J, \psi) - P_e \leq \frac{\sqrt{2\ln 2}}{|\mathcal{M}|} \sum_{m\in\mathcal{M}} \sqrt{D(p(y^N|H_m, s_X^J)||p(y^N|H_m, \hat{s}_X^J)}, \tag{10}$$

where $D(.||.)$ is the Kullback-Leibler distance (KLD). In the asymptotic case of large N, the ML estimate $\hat{s}_X^J$ yields the true parameters s_X^J thus approaching the exact match of $p(y^N|H_m, \hat{s}_X^J)$ and $p(y^N|H_m, s_X^J)$.

4 Practical identification

The considered $|\mathcal{M}|$-ary hypothesis testing is a problem that covers various aspects of the unclonable identification considered in Section 2. To fit the requirements of security and nondisclosure of the database, complexity of $|\mathcal{M}|$-ary hypothesis testing and memory storage requirements, we propose a practical identification technique shown in Figure 2. The main idea consists in the transforming the original codeword $x^N(m)$ into some secure domain of reduced dimensionality $\tilde{x}^L(m)$ using key-defined transform $\mathbf{\Phi}$. The obtained data are indexed in some hash table that can be public. At the identification stage, the observed data y^N is transformed into $\tilde{y}^N$ that is used for the matching with the data stored in the hash table. The result of the match is declared as the index $\hat{m}$.

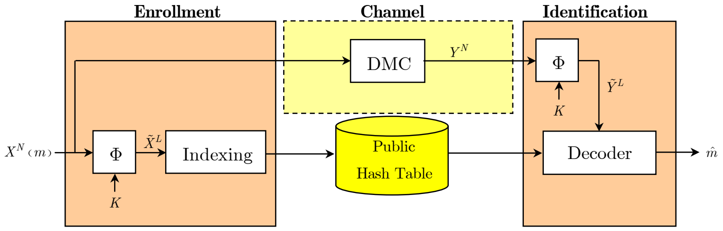

Fig. 2. Identification setup.

There are many possible selections of the key-defined transform $\mathbf{\Phi}$ that include but not limited by random projections and reference coordinate system. All these transforms can be considered as particular instances of generalized polynomial transform where random projections corresponds to the linear and reference coordinate system to the quadratic transforms (as one of possible distance measures). In the case of random projections, the basis vectors of $\mathbf{\Phi}$ are generated from some distribution preferably satisfying orthoprojection condition. However, general non-invertibility of the random projections transform unavoidably leads to the reduction of intercodewords distances and as the result of data processing inequality to the loss in performance.

For this reason we focus on the reference coordinate system transform. The main its idea is to perform the identification in the domain of pre-estimated candidates obtained by considering the observed codeword with respect to the reference coordinate system. This coordinate system is chosen according to the key-dependent selection of several reference codewords from the codebook considered to be the transform $\mathbf{\Phi}$. To cope with universality condition with respect to the knowledge of prior pdfs and channel state, we will apply the above considered GML framework.

We will first provide the practical consideration of the GML framework:

- the database of all samples to be identified or the codebook is known from the enrollment stage. Therefore, one can deduce the accurate estimation of source statistics $\hat{s}_X^J$ using the ML estimate (9);
- given the source statistics and assumed DMC, (9) can be rewritten as:

$$\psi_{GML}(y^N) = \arg \min_{1 \leq m \leq |\mathcal{M}|} \left[-\ln p(y^N | \hat{s}_X^J, H_m) \right] = \arg \min_{1 \leq m \leq |\mathcal{M}|} d^N(y^N, x_m^N), \tag{11}$$

assuming a family of the exponential distributions $p(y^N|s_X^J, H_m) = \frac{1}{Z_1(s_X^J, s_Z^J)} e^{-\frac{d^N(y^N, x_m^N)}{Z_2(s_X^J, s_Z^J)}}$, where the constants Z_1 and Z_2 include both the DMS s_X^J and DMC s_Z^J statistics.

- to compute $|\mathcal{M}| = 2^{NR}$ distances $d^N(y^N, x_m^N)$, R is the rate of equivalent communication system, one needs to accomplish $O(|\mathcal{M}|)$ vector computations that is prohibitively high for practical systems. Moreover, the codebook should be stored in the memory that might not fit to the memory storage conditions. Additionally, it needs to be stored on the server or shared that contradicts to the security requirements considered in Section 1.

The theoretical analysis of identification system performance (Appendix A) demonstrates the fundamental ability of the RLD to provide similar performance to direct complete search decoding. Here, we consider the practical aspects of RLD. According to the list decoding framework, one needs to identify the list of candidates and perform the decoding in this list. There are many possible algorithms capable of solving this problem. However, one needs to keep in mind the security issue according to which the entire codebook should not be disclosed at the identification stage. For this reason we have chosen a practical solution that is based on RLD. The main idea behind the RLD consists in the decoding based not on the original codewords but on the relative distances that can be considered as a sort of robust perceptual hash. In this case, it is not possible to recover the original codeword based on the relative distance to the key-defined reference codewords. This method consists of two main stages as discussed above, i.e., pre-processing of the codebook by computing the reference distances between all the codewords from the codebook and the reference codewords and identification itself that should find the best match between the distances computed for the observed codeword and those obtained for the list of candidates.

The **pre-processing stage** for a given codebook is summarized as follows:

1. Generate L codewords based on the secret key K (considered to be a seed for the random generator or random pointer to existing sequences) and denote them as $\{r_\ell^N\}$ for $1 \leq \ell \leq L$.
2. For all $m \in \mathcal{M}$ compute $d_{\ell,m} = d^N(r_\ell^N, x_m^N)$, where $d^N(.,.)$ denotes the distance between two sequences, and store these quantities in the *reference distance table*. This table is composed of a set of scalars and represents a sort of navigation map that can be efficiently stored and securely distributed.

The **identification stage** does not need the disclosure of the entire codebook but only assumes the availability of the reference codewords and the distances to them computed at the pre-processing stage. For the Gaussian set-up with the zero mean and variance σ_X^2, assuming $N \to \infty$, one can represent the identification procedure as the decoding on a sphere. Under such conditions, all codewords will be distributed on the surface of the sphere of radius $\sqrt{N\sigma_X^2}$ that is schematically shown in Figure 3. The identification procedure can be summarized as:

1. **Reference list estimation**:

   - For a given sequence y^N, compute the distances to the reference codewords $d_{\ell,y} = d^N(r_\ell^N, y^N) = ||r_\ell^N - y^N\|^2$ for $1 \leq \ell \leq L$ and $\|.\|^2$ is the ℓ_2-norm and denote it as a vector distance $\mathbf{d}_Y = [d_{1,y}, d_{2,y}, \cdots, d_{L,y}]$ that also corresponds to the vector $\tilde{y}^L$ in Figure 2.
   - Assuming the worst case channel state deduced for the allowable distortions (bounded variance σ_Z^2), we define $\delta = 4\sqrt{\sigma^2}$, where σ^2 is the variance of the difference $t_{\ell,m} = (d_{\ell,y} - d_{\ell,m})$ with $T_{\ell,m} \sim \mathcal{N}(N\sigma_Z^2,$ $2\sigma_Z^2(2\left\|x^N\right\|^2 + N\sigma_Z^2))$ for the correct m and for sufficient large N, to guarantee the convergence of χ^2-distribution to the Gaussian one. Find in the reference distance table those m satisfying $|d_{\ell,y} - d_{\ell,m}| \leq \delta$ and denote them as a set $\mathcal{C}$. It should be also noticed that this pre-selection can be efficiently organized based on the quantized distance table. For every candidate index $m \in \mathcal{C}$ define the corresponding distance vector as $\mathbf{d}_m = [d_{1,m}, d_{2,m}, \cdots, d_{L,m}]$ that corresponds to the vector $\tilde{x}^L$ in Figure 2. We will denote by $\mathbf{T}_m$ the corresponding vector $t_{\ell,m}$ for $1 \leq \ell \leq L$.

2. **Identification in the candidate list**: Find the codeword indexes that are in majority in the set $\mathcal{C}$ or simply find such an index that satisfies:

$$\hat{m}_{RLD} = \arg\min_{m \in \mathcal{C}} Var[\mathbf{T}_m], \tag{12}$$

   where $Var[\mathbf{T}_m] = E[\mathbf{T}_m^T \mathbf{T}_m] - E[\mathbf{T}_m]^2$.

The above decoding procedure is suboptimal in two aspects. To reduce the complexity of search we have assumed that the identification is based on the pre-estimated list of candidates that potentially reduces the performance with respect to (16). Relaxing complexity issue to the search in the entire codebook and letting $\Pr[E_1^m] = 0$, one needs to perform $|\mathcal{M}|$ checks over L-length distances versus N-length distance comparison in the classical communication setup. Thus, an additional loss in performance is caused by the drop in mutual information in (17) due to the data processing inequality [3]. The reduction in performance is a price for the security and storage advantages in the described sense versus the complete codebook need at the identification stage.

For the benchmarking purposes we will also consider the performance of *minimum reference distance* (MRD) decoder:

$$\hat{m}_{MRD} = \arg\min_{m \in \mathcal{M}} Var[\mathbf{T}_m], \tag{13}$$

for which $Pr[E^m] = 0$ and the the decoding is performed in the entire codebook of $|\mathcal{M}|$ reference distances $\{\mathbf{d}_m\}$.

5 Computer modeling

To confirm the theoretical findings we have performed the computer simulation on the Gaussian data $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \sigma_X^2 \mathbf{I}_N)$, with variance σ_X^2. The investigation was performed for codebook cardinalities of size 2^{10}, 2^{12} and 2^{13} codewords each of

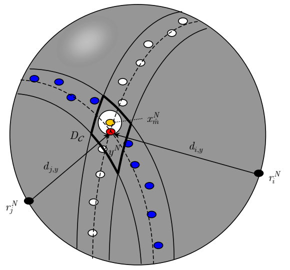

Fig. 3. RLD for the Gaussian case, $N \rightarrow \infty$: all codewords are located on the sphere surface while the codeword y^N is on the surface of noisy sphere with the radius $(N\sigma_Z^2)^{\frac{1}{2}}$ with the center at x_m^N that also represents the optimal decoding region for the MAP-rule. The candidate codewords are in the region $D_{\mathcal{C}}$ shown in bold boundaries.

length $N = 1024$. All codewords in three codebooks have been randomly indexed. Then the additive white Gaussian noise $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \sigma_Z^2 \mathbf{I}_N)$ was added to each codeword to obtain the resulting signal-to-noise ratio (SNR) $SNR = 10 \log_{10} \frac{\sigma_X^2}{\sigma_Z^2}$ equals -5dB, 5dB and 15dB to generate the system input vectors y^N.

The goal of the simulation was to evaluate the performance of different decoding strategies in terms of the average probability of error:

$$P_e = \frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} \Pr[\hat{m} \neq m | M = m], \tag{14}$$

and to compare the corresponding computational complexity measured in terms of the amount of decoder work defined as:

$$\mathcal{W} = \frac{|\mathcal{C}|}{|\mathcal{M}|} 100[\%]. \tag{15}$$

Since minimum distance decoding (MDD) produces zero probability for all considered SNRs and codebook sizes with $O(|\mathcal{M}|)$ checks, we do not plot it and use it as the bound for our performance evaluation. It should be also pointed out that the entire codebook of all codewords is assumed to be available for the identification that does not cope with the security and memory storage requirements.

To satisfy these requirements and to obtain the lower empirical bound on the performance in the assumption of $O(|\mathcal{M}|)$ checks, i.e., the decoding in the secure domain defined by the reference codewords relaxing the constraint on the complexity, we implemented the MRD decoder (13) and performed the simulation for various number of reference codewords L. The same simulation was performed for the RLD method and the results are presented in Figure 4 for codebook of size 2^{13} for the sake of paper size. It is important to note that the results for the probability of error for the MRD and RLD practically coincide that confirms the above conclusion that the probability of miss of the correct codeword in the list of the candidates is negligibly small and two methods demonstrate the same performance. The probability of error for both RMD and RLD has increased in comparison to the decoding in the real domain (MDD). That is a price for the reduction of the length of the codewords and distances between them in the reference distance space. If the number of reference codewords L increases, the probability of error decreases exponentially. It is interesting to observe that for SNR of 5 dB and 15dB, already 51 and 11 reference codewords, respectively, are sufficient to obtain zero probability of error for all considered cardinalities. For low SNR, more than 200 reference codewords have been needed to $P_e = 0$. This also indicates that the RLD is capable to reach the performance of the MDD with especially smaller length of resulting vectors (in terms of distances) that is equal to L but without disclosure of the original codebook. Obviously, with the asymptotic increase of SNR, one would expect to receive a single codeword in the list of candidates and the RLD would be equivalent to the real value hashing.

The proposed RLD technique has also demonstrated the interesting performance in terms of the amount of work. As it was expected according to the theoretical consideration, the list of candidates is increasing with decrease of SNR due to the higher radius of ambiguity sphere. For example, for the negative SNR, all codewords are in the sphere of ambiguity that requires to consider all possible codewords in the list of candidates as it is demonstrated in Figure 4. For the positive SNR, smaller L is required to restrict the number of candidates. Finally, the amount of work deceases exponentially with L. The exponential decrease rate is also increasing with SNR. In particular, the considered zero-error probability for SNR=15dB requires 21 reference codewords that represent about 0.1-0.01% of the considered codebook size. The asymptotic increase of SNR will lead to a single candidate that corresponds to the classical hash. The increase of the number of candidates reflects the measure of robustness of the proposed identification. Finally, to compute the total amount of work, one needs to add the work for the calculation of L reference distances for the length-N vectors and the above work to find a right codeowrd in the list of candidates for the length-L vectors.

6 Authentication

In this section, we extend the considered framework to the authentication. The main challenge of this problem is to provide reliable authentication based on noisy observation that is different from those acquired at the enrollment stage and used for the extraction of authentication data. Obviously, the traditional cryptography-based authentication will produce a negative result even if a single bit is altered that is not suitable for this protocol. Additionally, the security leakages about the authentication protocol might cause an appearance of a number of attacks targeting to trick the authentication.

To resolve these robustness-security requirements, we propose to use a similar hypothesis testing framework for the evaluation of item authenticity considered for the identification. To enforce the security of the considered setup, we will also

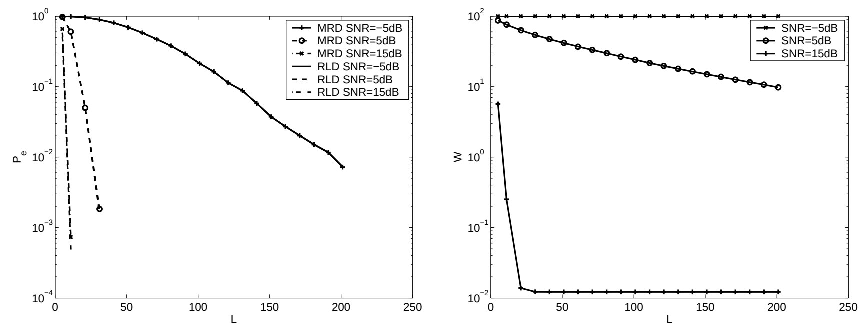

Fig. 4. Probability of error and amount of work for MRD and RLD strategies (left) and amount of work (right) for $|\mathcal{M}| = 2^{13}$.

use a projection of x^N into a secure key-defined domain. For this purpose, we will use a transform $\mathbf{\Phi}$ that besides the security insures the dimensionality reduction, complexity as well as memory storage. Additionally, the transform can be chosen in such a way to guarantee a certain robustness to the legitimate distortions. In our analysis, we will assume reference coordinate system transform.

The general block-diagram of the considered authentication is shown in Figure 5. In the scope of this framework, the helper data index m and the secret s are deduced at the enrollment stage based on the projected reference data $\hat{x}^L$. We assume here that the lossless distributed coding is used based on Slepian-Wolf framework [21]. The rate for index m communications is $\tilde{R}_X^{SW} \geq H(\tilde{X}|\tilde{Y})$. Similar in spirit approaches were firstly introduced by Maurer [16] and Ahlswede and Csiszar [1], where the index m was considered as a helper data for the common randomness extraction. [1] The index w is encrypted or hashed with the authentication key s that can be different even for the same x^N. The index m and the encrypted data $w \oplus s$ are communicated to the decoder possibly via public channel. The channel for x^N includes both the attacker, who can replace the sequence x^N by x'^N or use index m for x'^N, and the acquisition DMC channel $p_{Y|X}(y|x)$. At the authentication stage, one should make a decision about the item authenticity based on the observed vector y^N and the authentication data. For this purpose, the decoder retrieves the index $\hat{w}$ based on m and y^N. The estimate $\hat{w}$ is used for the decryption of s. The binary test produces the final decision by generating the hypothesis H_0, i.e., fake, or H_1, i.e., genuine.

7 Conclusion

In this paper, we proposed a suboptimal identification technique for the unclonable item features that trade-offs security-performance-memory storage-universality requirements. The proposed technique is based on the identification in some

[1] A lossy coding can be used as well. In this case, m can be considered as a hash obtained with the corresponding randomized codebook generation.

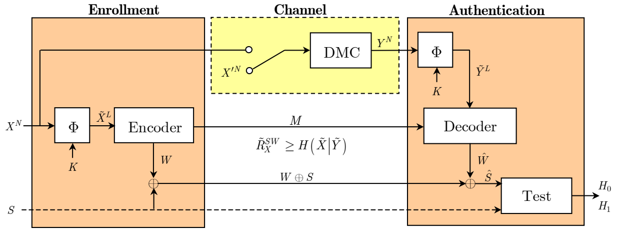

Fig. 5. Authentication setup.

secure domain defined by the list of key-dependent reference codewords. The sub-optimality of the scheme concerns the achievable performance in terms of average error probability that is higher in comparison to the direct decoding. Thus, in future we will investigate the optimal selection of the number of codewords as the function of the SNR and optimize the construction of reference system of codewords to minimize the error probability and number of reference codewords. We will also investigate the security of the system as a function of number of reference codewords and the corresponding attacks.

Acknowledgments

This paper was partially supported by SNF Professeur Boursier grants PP002–68653, 114613 and 200021–111643, by the European Commission through the IST Programme under contract IST-2002-507932-ECRYPT, FP6-507609-SIMILAR and Swiss IM2 projects.

The information in this document reflects only the authors' views, is provided as is and no guarantee or warranty is given that the it is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

Appendix A: Theoretical analysis of performance

First, we will perform the theoretical analysis of identification assuming the possibility to perform $O(|\mathcal{M}|)$ vector computations and disregarding the constraint on the memory storage and security due to the codebook disclosure. In this case, the identification system is equivalent to the communications one, where the decision about m should be deduced based on y^N applying any type of decoding that fits the considered MAP-rule including the ML- or jointly typical decoding [3]. By the symmetry of the codebook construction, the average probability of error does not depend on the particular index m. Thus we can assume $M = 1$ and denote the average probability of error as $P_e(M = 1)$. We will denote the event of observing jointly typical pair x_i^N, y^N to be $E_i = \{(X_i^N, Y^N) \in A_\epsilon^N\}$, for $i \in \{1, 2, \cdots, 2^{NR}\}$, with A_ϵ^N to the a set of jointly typical pairs with $\epsilon \to 0$ as $N \to \infty$ [3]. Then, the average probability of error is equal to the case during the decoding when the true codeword and the observed codeword are not jointly typical or $E_2 \bigcup \cdots \bigcup E_{2^{NR}}$ occurs, i.e., when a wrong codeword from the codebook is jointly typical with the observed sequence.

$$P_e(M=1) = \Pr\left[E_1^c \bigcup E_2 \bigcup \cdots \bigcup E_{2^{NR}}\right] \leq \Pr\left[E_1^c\right] + \sum_{i=2}^{2^{NR}} \Pr\left[E_i\right], \tag{16}$$

where the inequality follows from the union bound for probabilities. Using asymptotic equipartition property (AEP), one obtains [3]:

$$P_e(M=1) \leq \epsilon + (2^{NR} - 1)2^{-NI(X;Y)-3\epsilon}, \tag{17}$$

that is asymptotically small for sufficiently large N and $R < I(X;Y) - 3\epsilon$ or $|\mathcal{M}| < 2^{NI(X;Y)-3\epsilon}$.

At the second stage, we will try to trade-off the above requirements of complexity-memory storage-security by making the problem asymmetric in sense of re-allocation of computation at the pre-processing stage and identification. It should be pointed out that the pre-precessing stage should be done once for a given codebook. In the case of a new entry to the codebook, the pre-computation is introduced for new codewords and the results are updated accordingly without the need to perform the re-computation for the entire codebook.

The main idea consists in defining a computationally feasible set $\mathcal{C}_m$ of possible candidates (reference list) for the index m and then to perform the decoding only in this set. Obviously, this procedure is suboptimal in sense of achievable provability of error but might lead to the solution to the above trade-off. First, we will consider the performance of this scheme in terms of achievable average probability of error assuming the availability of the codebook and relaxing security constraint and then will advocate a practical algorithms in the scope of this framework where the entire codebook is not needed for the identification.

The probability of error of list decoding algorithm consists of two parts, i.e., the probability of error of missing the right index $(M = 1)$ in the preselected set of indexes $\mathcal{C}_1$ denoted as $\Pr[E_1^m]$ and probability of decoding error among the list of candidates in the set $\mathcal{C}_1$ of cardinality $|\mathcal{C}_1|$. Using AEP, the probability of error of list decoding algorithm can be bounded as:

$$P_e(M=1) \leq \Pr\left[E_1^m\right] + (1 - \Pr\left[E_1^m\right])\left(\epsilon + (|\mathcal{C}_1| - 1)2^{NI(X;Y)-3\epsilon}\right), \tag{18}$$

where the second term corresponds to the decoding error in the set $\mathcal{C}_1$. The fact of missing the right codeword at the reference list has a crucial impact on the performance of the whole decoding procedure. If $\Pr\left[E_1^m\right] \rightarrow 1$, then $P_e(M=1) \leq 1$ and no reliable decoding is possible. Contrarily, to guarantee that the right codeword will not be missed, i.e., $\Pr\left[E_1^m\right] = 0$, one should assume that $|\mathcal{C}_1| \leq |\mathcal{M}|$ that requires to check all codewords in $\mathcal{C}_1$ and converges to the exhaustive search decoding considered above. Finally, for any finite small $\Pr\left[E_1^m\right] \leq \xi$, one obtains the intermediate situation of bounded decoding error probability and reduced complexity decoding. It should be pointed out that under the assumption of reliable decoding in the case of (16), one can assume that $|\mathcal{C}_1| < |\mathcal{M}| \leq 2^{NI(X;Y)-3\epsilon}$ that leads to the bound $P_e(M=1) \geq \Pr\left[E_1^m\right]$ as $N \rightarrow \infty$.

References

1. R. Ahlswede and I. Csiszar. Common randomness in information theory and cryptography - part i: secret sharing. *IEEE Trans. Inform. Theory*, 39(4):1121–1132, 1993.
2. G. Colgate. *Document Protection by Holograms*. Optical Document Security, R. van Renesse, Ed., Artech House, Norwood, MA, 1993.
3. T. Cover and J. Thomas. *Elements of Information Theory*. Wiley and Sons, New York, 1991.
4. I. J. Cox, M. L. Miller, and J. A. Bloom. *Digital Watermarking*. Morgan Kaufmann Publishers, Inc., San Francisco, 2001.
5. R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley, 2000.
6. J. Fridrich and M. Goljan. Protection of digital images using self embedding. In *Symposium on Content Security and Data Hiding in Digital Media*, New Jersey Institute of Technology, USA, May 1999.
7. R.G. Gallager. *Information theory and reliable Communication*. J. Wiley and Sons, 1968.
8. C. K. Harmon. *Lines of Communication: Bar Code and Data Collection Technologies for the 90S*. Helmers Pub, 1994.
9. T. D. Hayosh. Apparatus and method for enhancing check security. US Patent 6600823, Issued on July 29, 2003, filed July 1997.
10. H. Hel-Or, Y. Yitzhaki, and Y. Hel-Or. Geometric hashing techniques for watermarking. In *ICIP 2001*, page Watermarking i, 2001.
11. W. A. Houle. Light sensitive invisible ink compositions and methods for using the same. US Patent 6513921, Issued on February 4, 2003, filed February 2000.
12. S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall Signal Processing Series, 1993.
13. Z. F. Keith, T. Knox, and E. J. Schneider. Method and apparatus for detecting photocopier tracking signatures. US Patent 6515764, Issued on February 4, 2003, filed February 1998.
14. J. Koniecek and K. Little. *Security, ID Systems and Locks: The Book on Electronic Access Control*. Elsevier Computers/Computer Security, 1997.
15. C. Leang and D. Johnson. On the asymptotics of m-hypothesis bayesian detection. *IEEE Trans. on Information Theory*, 43(1):280–282, January 1997.
16. U. Maurer. Secret key agreement by public discussion from common information. *IEEE Trans. Inform. Theory*, 39:733–742, 1993.
17. U. Maurer. A unified and generalized treatment of authentication theory. In *Proc. 13th Symp. on Theoretical Aspects of Computer Science (STACS'96)*, volume 1046 of *Lecture Notes in Computer Science*, pages 387–398. Springer-Verlag, February 1996.
18. V. Monga and B. L. Evans. Robust perceptual image hashing using feature points. In *ICIP 2004*, pages 677–680, 2004.
19. J. G. Proakis. *Digital Communications*. McGraw-Hill, 1995.
20. M. Schneider and S. Chang. A robust content based digital signature for image authentication. In *Proceedings of the IEEE International Conference on Image Processing*, pages 227–230, Lausanne, Switzerland, September 1996.
21. D. Slepian and J.K. Wolf. Noiseless encoding of correlated information sourcea. *IEEE Trans. Information Theory*, 19:471–480, July 1973.
22. N. Vasconcelos. Minimum probability of error image retrieval. *IEEE Transactions on Signal Processing*, 52(8):2322–2336, 2004.

23. R. Venkatesanan, S. Koon, M. Jacubowski, and P. Moulin. Robust image hashing. In *ICIP 2000*, Vancouver, BC, Canada, September 2000.

24. S. Voloshynovskiy, O. Koval, F. Beekhof, and T. Pun. Robust perceptual hashing as classification problem: decision-theoretic and practical considerations. In *IEEE 2007 International Workshop on Multimedia Signal Processing*, Chania, Crete, October 1-3 2007.