Geometrically robust perceptual fingerprinting: an asymmetric case

Oleksiy Koval, Sviatoslav Voloshynovskiy, Farzad Farhadzadeh, Taras Holotyak and Fokko Beekhof*

ABSTRACT

In this paper, the problem of multimedia object identification in channels with asymmetric desynchronizations is studied. First, we analyze the achievable rates attainable in such protocols within digital communication framework. Secondly, we investigate the impact of the fingerprint length on the error performance of these protocols relaxing the capacity achieving argument and formulating the identification problem as multi class classification.

1. INTRODUCTION

Recent advances in modern networking and multimedia technologies have open an access to an exponential and permanently increasing volume of multimedia data via various public services and social networks. In these circumstances, an urgent demand for efficient high multimedia volume managing and security systems arises. Therefore, scalability of existing design principles of such systems for large scale applications is raised by this demand.

In this paper we would like to analyze multimedia identification based on robust fingerprinting that can be considered as a unique solution to the identification problem when no modification of the content is admissible. Such a constraint is imposed while identifying art works, biometrics, medical data, making application of digital data hiding principles for this purpose unacceptable.

A digital fingerprint (a.k.a. robust perceptual hashing) provides a compact and robust representation of a content designed for its distinctive, computationally efficient and privacy protected management.

Recently, the domain of robust fingerprinting has performed a significant evolution. The main progress achieved at the side of practical algorithm development mainly concerns robust feature extraction techniques as well as elaboration of efficient matching strategies in large databases,[1].[2] Analysis of achievable rates of an identification system was firstly accomplished by Willems et. al.[3] in a classical information-theoretic formulation of infinitely long codeword transmission over the discrete memoryless channel (DMC). Several groups of authors analyzed the identification problem within the information-detection framework for various channel and communicated codeword length assumptions.[4–6]

In the most of the mentioned cases (besides a conjecture formulated in[4, 5]) it is explicitly assumed that the query is ideally synchronized with the database content. Such an assumptions while being valid for analysis of communications over the DMC could be too restrictive for the target application and might lead to inaccurate performance limit estimates. Moreover, omnipresence of desynchronizations at both multimedia database enrollment and query identification stages, makes an extension of existing analysis results for the DMC to the channels with desynchronization distortions an important research issue.

The first attempt to characterize the performance loss in terms of achievable rate introduced due to geometrical desynchronization over such a channel was recently performed in[7] where it is assumed that the database is composed of original/ideal multimedia data. Such a system design principle was firstly proposed by Willems et all.[8] for biometric identification and will be referred as symmetric in our paper. The presented analysis results are obtained assuming that geometrical desynchronization can be modeled as a parametric mapping defined over the set of finite cardinality.

*O. Koval, S. Voloshynovskiy, F. Farhadzadeh, T. Holotyak and F. Beekhof are with CUI-University of Geneva, Stochastic Information Processing Group, Battelle Batiment A, 7 route de Drize, 1227 Carouge, Switzerland. The contact author is O. Koval (email: Oleksiy.Koval@unige.ch). http://sip.unige.ch

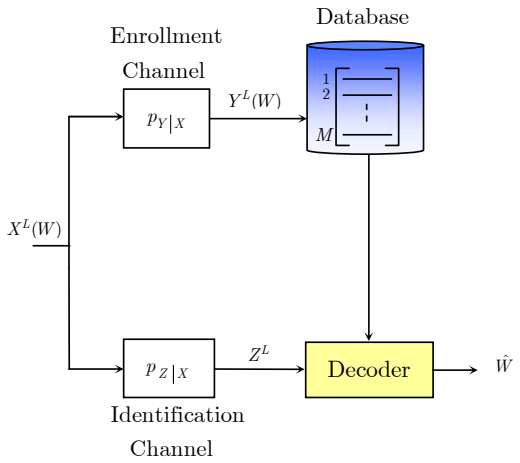

Figure 1. Multimedia object identification as a communications problem.

Therefore, the main goal of this paper is to extend the results obtained in[7] to the asymmetric case,[8] where it is assumed that both a database entry and a query are DMC distorted versions of the ideal multimedia signal further desynchronized according to the above model.

The rest of the paper is organized as follows. The problem of multimedia object identification in channels with asymmetric distortions is formulated in Section 2. Performance analysis of multimedia object identification in channels with asymmetric distortions is accomplished in Section 3. The impact of the fingerprint length on the identification system design and its attainable performance limits are analyzed in Section 4. Finally, Section 5 contains conclusions and future research perspectives.

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 superscripts N and L are used to designate length-N or length-L vectors $x^N = [x[1], x[2], ..., x[N]]$, $x^L = [x[1], x[2], ..., x[L]]$, with i^{th} element $x[i]$. We use $X \sim p_X(x)$ or simply $X \sim p(x)$ to indicate that a random variable X is distributed according to $p_X(x)$. Calligraphic fonts $\mathcal{X}$ denote sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X}$.

2. PROBLEM FORMULATION

The setup for asymmetric multimedia object identification as a communication problem is presented in Fig. 1. Here we assume that we are given a set of M multimedia objects represented by corresponding indexes $w \in \{1, 2, ..., M\}$. Every object is associated to a corresponding raw fingerprint sequence $x^N = [x_1, x_2, ..., x_N]$ generated i.i.d. according to a certain distribution $p(x)$ from an alphabet $\mathcal{X}$, i.e.:

$$p(x^N) = \prod_{i=1}^{N} p(x[i]), \quad x^N \in \mathcal{X}^N. \tag{1}$$

Oppositely to the case considered in,[7] where $x^N(w), w \in \{1, 2, ..., M\}$, are assumed to be available without any distortions,[8] in the scope of this paper the distorted versions of these fingerprints are observed at the output of the memoryless channel $\{\mathcal{Y}, p(y|x), \mathcal{X}\}$, where $\mathcal{Y}$ is the enrollment channel output alphabet, in the **enrollment stage** according to a predefined feature extraction procedure, for example like in.[1,2] Therefore, one has:

$$p(y^N(w)|x^N(w)) = \prod_{i=1}^{N} p(y(w)[i]|x(w)[i]), \tag{2}$$

for all $y^N(w) \in \mathcal{Y}^N$. The enrolled vectors $y^N(w), w \in \{1, 2, ..., M\}$, are stored in the object database (Fig. 1).

In the **identification stage**, a raw fingerprint representing unknown multimedia object z^N is received at the output of the memoryless identification channel $\{\mathcal{Z}, p(z|x), \mathcal{X}\}$, where $\mathcal{Z}$ is the identification channel output alphabet:

$$p(z^N(w)|x^N(w)) = \prod_{i=1}^{N} p(z[i]|x(w)[i]), \tag{3}$$

An identity decoding is accomplished assuming the availability of all enrolled data using the following deterministic mapping:

$$g : \mathcal{Z}^N \to \{\delta, 1, 2, ..., M\}, \tag{4}$$

where δ denotes the decoder output for the case when the data irrelevant to the analyzed at enrollment one is observed by the system.

It is a known fact in information theory that the decoder (4) will produce a non-correct result with the following probability:

$$P_e^{max} = \max_{w \in \{\delta,1,2,...,M\}} \Pr[g(Z^N) \neq w | Y^N(W) = y^N(w)], \tag{5}$$

that is usually referred to as a maximum probability of error[9] defined in our case for a fixed rate of identification:

$$R_{id} = \frac{1}{N} \log_2 M. \tag{6}$$

Finally, capacity of the identification system C_{id} is defined as a supremum of the identification rates R_{id} such that $P_e \to 0$ for a sufficiently large L.[9]

Willems et al.[10] succeeded to determine the capacity of such an identification system. This result is stated in the following theorem.

Theorem. The capacity C_{id} of a symmetric identification system operating over DMCs that model acquisition distortions between the original fingerprint and (a) the database entry and (b) the query, respectively, is given by $I(Y;Z)$, where $p(y,z) = \sum_{x \in \mathcal{X}} p(x)p(y|x)p(z|x)$ for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$ and $I(Y;Z)$ stays for a mutual information between two random variables Y and Z.

The proof of this theorem follows a random coding-based strategy and jointly typical decoding and is omitted in this paper for the sake of brevity and can be found in.[10]

2.1. Modeling desynchronization distortions in enrollment and identification channels

According to the asymmetric model for object identification in channels with geometrical desynchronization, we assume that distortions between the original fingerprints $X^L(w), w \in \{1, 2, ..., M\}$, and (a) database entries $Y^L(w), w \in \{1, 2, ..., M\}$, and (b) queries Z^L are modeled by channels that besides the DMC parts include desynchronization mechanisms. We model such mechanisms as a parametric mapper:

$$t_a : \mathcal{A}^J \times \mathcal{Y}^L \to \mathcal{Y}^L. \tag{7}$$

According to this definition, it is assumed that the mapper is parametrized by a set of J parameters, $a^J = (a[1], a[2], ..., a[J])$, taking their values in a finite set of cardinality $|\mathcal{A}|$. In the scope of this paper we assume that at enrollment and identification this mapper is governed by two parameter vectors a_e^J and a_i^J. Furthermore, we specifically suppose that such mappings impact only the coordinates of the elements of their input and does not modify this input cardinality. We suppose as well that the defined desynchronization preserves sample independence of its input.

Since, a malicious desynchronization falls outside of the scope of this paper, we assume that not all parameters from $\mathcal{A}^J$ can be used in our analysis due to technological or acquisition constraints to model desynchronizations introduced during enrollment and identification. Therefore, instead of considering all $|\mathcal{A}^J|$ possible parameters, we will constraint the entire set to a set of ϵ-typical desynchronization transformations defined over a corresponding set $\mathcal{A}_\epsilon^{(J)}$ with $|\mathcal{A}_\epsilon^{(J)}| \leq |\mathcal{A}^J|$.

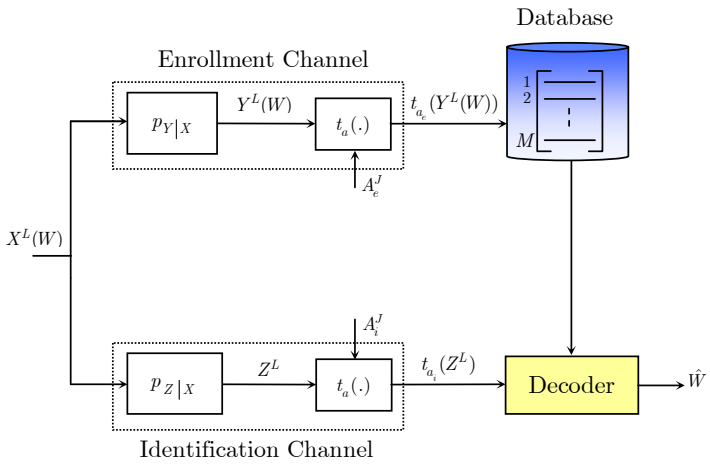

Figure 2. Multimedia object identification as a communications problem in channels with asymmetric desynchronizations.

If the parameters of $a^J = [a[1], a[2], ..., a[J]]$ are i.i.d. distributed according to $p(a)$, then, $|\mathcal{A}_\epsilon^{(J)}|$ could be upper bounded as[9]:

$$|\mathcal{A}_\epsilon^{(J)}| \leq 2^{J(H(A)+\epsilon)}, \tag{8}$$

where $H(A)$ is the entropy of A.

The next Section contains the impact analysis of the desynchronization mappings defined above on the performance limits attainable by the identification system.

3. PERFORMANCE ANALYSIS OF MULTIMEDIA OBJECT IDENTIFICATION UNDER ASYMMETRIC DESYNCHRONIZATIONS

The identification system under our analysis is presented in Fig 2. The difference with the setup discussed in Section 2 consists in the desynchronization operator $t_a(.)$ parametrized by a_e^J and a_i^J that acts according to (7). Therefore, the main goal of this Section consists in evaluating the impact this operator might have on the performance limits of the identification system with respect to the Theorem presented in the previous Section. In the achievability part of our analysis we will try to limit the average probability of the decoder failure (5) that can be redefined according to the desynchronization parts of the channels in the following way:

$$\begin{aligned} P_e^G &= \frac{1}{M|\mathcal{A}_\epsilon^{(J)}|} \\ &\times \sum_{w=1}^{M} \sum_{a_i \in \mathcal{A}_\epsilon^{(J)}} \Pr[g(t_{a_i}(Z^L)) \neq w | t_{A_e}(Y^L(w)) = t_{a_e}(y^L(w))], \qquad a_e^J \in \mathcal{A}^L, a_i^J \in \mathcal{A}^L, \end{aligned} \tag{9}$$

taking into account the cardinality of $\mathcal{A}_\epsilon^{(J)}$. Here we use the fact that the desynchronization should be estimated and compensated with respect to fingerprints stored in the database $t_{a_e}(Y^L(w))$. Therefore, one can bound this probability using the properties of jointly typical sequences[9] by:

$$\begin{aligned} P_e^G &\leq |\mathcal{A}_\epsilon^{(J)}| 2^{LR_{id}} 2^{-L(I(Y;Z)-\epsilon)} \\ &\leq 2^{L\frac{1}{L}\log_2|\mathcal{A}_\epsilon^{(J)}|} 2^{LR_{id}} 2^{-L(I(Y;Z)-\epsilon)} \\ &\leq 2^{L(\frac{1}{L}\log_2|\mathcal{A}_\epsilon^{(J)}|+R_{id}-(I(Y;Z)-\epsilon))}, \end{aligned} \tag{10}$$

where it is assumed that decoding was performed for all elements of $\mathcal{A}_\epsilon^{(J)}$. Therefore, if R_{id} satisfies

$$R_{id} \leq I(Y;Z) - \epsilon - \frac{1}{L}\log_2|\mathcal{A}_\epsilon^{(J)}|, \tag{11}$$

$P_e^G \to 0$ as $L \to \infty$ and $\epsilon \to 0$. Moreover, since only typical desynchronization parameter sets are considered, $\frac{1}{L}\log_2|\mathcal{A}_\epsilon^{(J)}| \leq \frac{J(H(A)+\epsilon)}{L}$ that vanishes for $L \to \infty$, one can conclude that

$$R_{id} \leq I(Y;Z) - \epsilon', \tag{12}$$

where ϵ' is a positive constant that can be made arbitrarily smaNl. This result coincides with the achieveability part of the Theorem in Section 2.

The proof of the converse part of the multimedia object identification theorem over channels with desynchronizations present at enrollment and identification, can be summarized as follows:

$$\begin{aligned} LR_{id} &= H(W) = H(W|t_{a_i}(Z^L)) + I(W;t_{a_i}(Z^L)) \leq H(W|t_{a_i}(Z^L)) + I(t_{a_e}(Y^L(W));t_{a_i}(Z^L)) \\ &\leq H(W|t_{a_i}(Z^L)) + I(Y^L(W);Z^L) \leq 1 + P_e^G LR_{id} + LC_{id}, \end{aligned} \tag{13}$$

where in the first inequality we used the fact that $W, t_{a_e}(Y^N(W))$ and $t_{a_i}(Z^N)$ constitute a Markov chain that reflects the design of the database and assumes information transmission over the equivalent communication channel between $t_{a_e}(Y^N(W))$ and $t_{a_i}(Z^N)$; the following Markov chain $t_{a_e}(Y^N) \to Y^N \to X^N \to Z^N \to t_{a_i}(Z^N)$ is used in the second one and Fano inequality is exploited in the last one.[9] Therefore, normalizing by N and assuming $N \to \infty$, one has:

$$R_{id} \leq \frac{1}{N} + P_e^G R_{id} + C_{id} \tag{14}$$

that approaches C_{id} due to the vanishing character of the first two terms for $N \to \infty$ and $P_e^G R \to 0$.

Thus, the result of the converse part of the multimedia object identification theorem over channels with desynchronization present at both enrollment and identification coincides with the one obtained in Section 2 under the assumed conditions.

We conclude this Section by a discussion on the types of implementations of theoretical decision making strategies that are capable of achieving the identification capacity in channels with desynchronizations. Our two major candidates can be formulated as follows. In the first scenario, one assumes that the database (Fig. 2) is extended at the enrollment stage by generating all possible desynchronized versions of $2^{N(I(Y;Z))}$ original codewords off-line. Therefore, there are $|\mathcal{A}_\epsilon^J|2^{N(I(Y;Z))}$ finally stored entries in the extended database. At the identification stage, the direct match of the observed query is performed versus the generated database.

A possible alternative consists in keeping the database unmodified at enrollment and generating a set of all $|\mathcal{A}_\epsilon^J|$ possible desynchronized versions of the query at identification. According to this scenario, the match of two data sets is performed at decoder.

Table 1 contains a comparison of computational complexities and storage memory requirements of these two strategies. Since, in the big $\mathcal{O}(.)$ notation the complexities of both approaches coincide and the storage memory of the latter one is significantly smaller, one can conclude the advantages of this approach over the former one.

Table 1. Comparison of computational complexities and storage memory of two strategies achieving identification capacity.

	Storage	$\mathcal{O}(.)$
Extended database	$2^{N(I(Y;Z)+\epsilon)}2^{J(H(A)+\epsilon)}$	$N2^{N(I(Y;Z)+\epsilon)}2^{J(H(A)+\epsilon)}$
Extended query	$2^{N(I(Y;Z)+\epsilon)}$	$(1+2^{N(I(Y;Z)+\epsilon)})2^{J(H(A)+\epsilon)}$

4. ANALYSIS OF THE DATABASE ENTRY LENGTH IMPACT ON IDENTIFICATION SYSTEM DESIGN AND PERFORMANCE

The analysis of the achievable rates of object identification in channels with asymmetric desynchronizations was accomplished using a capacity achieving argument that explicitly assumes infinite length of database entries. This makes justification of this system design parameter impact on its performance problematic. Therefore, for the sake of such a justification we formulate this problem as a multiple hypothesis testing problem with a fixed

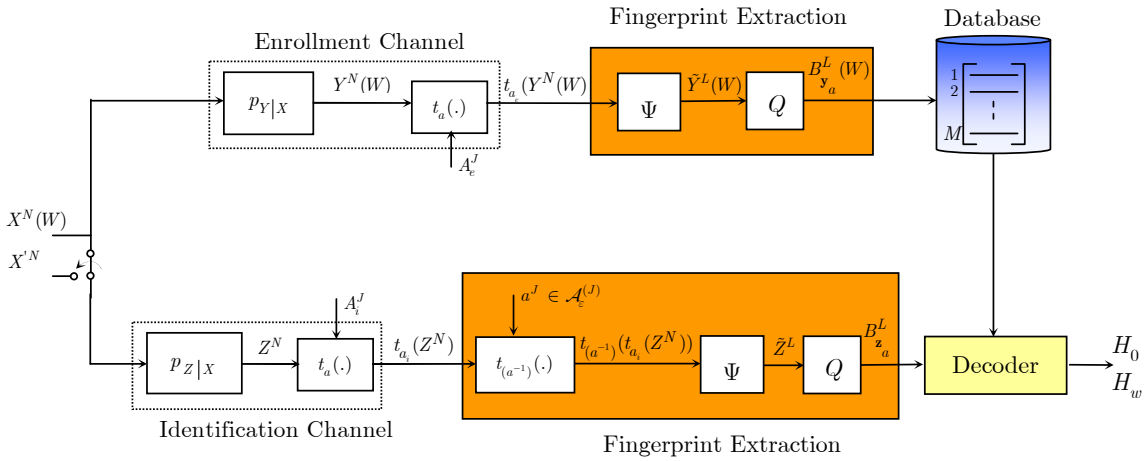

Figure 3. Multimedia object identification as a communications problem in channels with asymmetric desynchronizations.

number of alternatives and analyze it using large deviations. In such a definition, we are interested in the analysis of a probability of error of multimedia object identification over channels with asymmetric desynchronizations rather than in maximization of the achievable rate considered in the previous Section. Targeting quantitativeness of the analysis, we deviate from the identification setup considered in earlier Sections of this paper to a specific formulation.

A block diagram of multimedia object identification system of interest is presented in Fig. 3, where it is assumed oppositely to[11] that both outputs of enrollment $Y^N(w), w \in \{1, 2, ...M\}$, and identification Z^N are distorted and desynchronized versions of original raw fingerprints $X^N(w), w \in \{1, 2, ...M\}$.

According to this system design it is assumed that binary fingerprints are generated using the following two-stage procedure. First, in order to coop with complexity, storage and privacy constraints of multimedia object identification, the object relevant data $t_{a_e}(y^N(w))$ of reduced dimensionality $\tilde{y}^L(w)$ are generated using random projections as follows:

$$\tilde{y}^L(w) = \mathbf{\Psi} t_{a_e}(y^N(w)) \tag{15}$$

at the enrollment stage. We assume that $\mathbf{\Psi} \in \mathbb{R}^{L \times N}$, where $L \leq N$ and $\mathbf{\Psi} = (\psi_1, \psi_2, \cdots, \psi_L)^T$, denotes a dimensionality reducing operator. It is supposed that the elements of $\mathbf{\Psi}$, $\psi_{i,j}$, are independently and identically generated from a Gaussian distribution, i.e., $\Psi_{i,j} \sim \mathcal{N}(0, \frac{1}{N})$. The parameters of the generating distribution are adjusted to guarantee that $\mathbf{W}$ is an approximate orthoprojector, i.e., $\mathbf{\Psi}\mathbf{\Psi}^T \approx \mathbf{I}_L$. The selection of the Gaussian distribution for the generation of $\mathbf{\Psi}$ is justified by a desired property of such a projection output to have the Gaussian statistics too. It should be admitted that such a property will be attributed to other statistical designs of $\mathbf{\Psi}$ for a sufficiently long input vectors due to the Central Limit Theorem.

In order to improve the mentioned storage memory, computational complexity as well as amplify the overall system privacy, a random projection output is converted to a length L binary fingerprint according to:

$$b_{\mathbf{y}_a}[i] = sign(\psi_i^T t_{a_e}(y^N(w))), i \in \{1, 2, ..., L\}, \tag{16}$$

where ψ_i stays for the i^{th} row of $\mathbf{\Psi}$. Generated in such a way binary fingerprints are stored in the database.

The strategy of the design of the identification part of the system follows the analysis of the identification capacity achieving strategies presented in the previous Section. In order to be able to properly address the issue of desynchronizations present in the identification and enrollment channels, our approach here is dedicated to the brute force estimate of the desynchronization parameters that introduce geometrical misalignment between the inputs to the enrollment and identification stages. Such an approach corresponds to the decision making strategy with distorted query (Table 1).

In particular, the identification stage of our system is organized as follows (Fig. 3). First, the output of the identification channel, that is a DMC $p(z|x)$ distorted and desynchronized ($t_{a_i}(Z^N)$) version either of one of the enrolled objects or of an input that has no relevance to the database content, is further transformed according to $t_{(a^{-1})}(t_{a_i}(Z^N)$ for all $a^J \in \mathcal{A}_\epsilon^{(J)}$. Then, in order to preserve symmetry with the enrollment stage, one applies:

$$\tilde{Z}^L = \mathbf{\Psi} t_{(a^{-1})}(t_{a_i}(Z^N)), a^L \in \mathcal{A}_\epsilon^{(J)} \tag{17}$$

$$b_{\mathbf{z}_a}[i] = sign(\tilde{Z}^L), a^L \in \mathcal{A}_\epsilon^{(J)}, i \in \{1, 2, ..., L\}. \tag{18}$$

Therefore, the main purpose of $t_{a^{(-1)}}(.)$ application in fingerprint extraction module (17) over the entire set $\mathcal{A}_\epsilon^{(J)}$ is to find such an a^L that compensates the desynchronization part of the equivalent channel between $t_{a_e}(Y(W)^N)$ and $t_{a_i}(Z^N)$ and converts it to a simple DMC.

In the case the elements of X^N are i.i.d. zero-mean distributed, one can demonstrate that binary fingerprints generated at enrollment and identification will have a Binomial distribution with 0.5 probabilities of both binary events due to particularities of the fingerprint generation.

We include in the setup analysis of this Section the cases of system irrelevant inputs X'^N since such situations are not rare in identification setups. These inputs were not analyzed in details in the previous Sections due to the capability of the jointly typical decoder to reliably eliminate them from consideration with high probability for the infinite codeword length case.[9]

Therefore, one can formulate the multimedia object identification problem in channels with asymmetric desynchronizations as a multiple hypothesis testing problem with corresponding prior probabilities[7]:

$$\begin{cases} H_0: & B_{\mathbf{z}_a}^L \sim p(b_{\mathbf{z}_a}^L | b_{\mathbf{y}'}^L), \\ H_w: & B_{\mathbf{z}_a}^L \sim p(b_{\mathbf{z}_a}^L | b_{\mathbf{y}_a}^L(w)), w = 1, \cdots, M, \end{cases} \tag{19}$$

where $b_{\mathbf{y}'}^L$ denotes a fingerprint generated based on x'^N. Since the probabilistic models on alternative hypotheses in the above case can be defined by a Binary Symmetric Channel (BSC) with a crossover probability P_{b_e}, (19) can be modified accordingly:

$$\begin{cases} H_0: & B_{\mathbf{y}}^L \sim \frac{1}{2^L}, \\ H_w: & B_{\mathbf{y}}^L \sim P_{b_e}^{d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w))} (1 - P_{b_e})^{L - d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w))}, \end{cases} \tag{20}$$

where $d_H(.,.)$ designates the Hamming distance.

Having access to the database content and $b_{\mathbf{z}_a}^L$, the decoder should decide which one out of $M+1$ alternatives is present at the input of the identification system. We assume that it operates according to the Bounded Distance Decoding Rule (BDD):

$$d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w)) \leq L\gamma, \tag{21}$$

defined for a certain threshold γ. The optimal selection of γ for the case of the BSC distortion model was considered in.[11] Using probability of error analysis, it was demonstrated that the optimal value of the threshold is a function of the channel and:

$$\gamma_{\text{opt}} = \frac{1 - R_{id} + \log_2(1 - P_{b_e}) - 1/L}{\log_2\left(\frac{1 - P_{b_e}}{P_{b_e}}\right)}. \tag{22}$$

Therefore, according to,[11] the threshold value that determines the performance of the identification system in terms of probabilities of error coincides with the classical BDD for large L.

In the remaining part of this paper, we will try to demonstrate how desynchronization impacts these asymptotics.

Similarly to,[11] we will analyze probabilities of error of two kinds assuming that decoding is performed at every point of an ϵ-typical desynchronization transformations.

For the probability of false acceptance P_f one has:

$$\begin{aligned} P_f &= \Pr[\bigcup_{a\in\mathcal{A}_\epsilon^{(J)}}\bigcup_{w=1}^{M} d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) \le \gamma L|H_0] \\ &\le \sum_{a\in\mathcal{A}_\epsilon^{(J)}}\sum_{w=1}^{M} \Pr[d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) \le \gamma L|H_0] \\ &= M|\mathcal{A}_\epsilon^{(J)}|\Pr[d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) \le \gamma L|H_0] \\ &\le |\mathcal{A}_\epsilon^{(J)}|2^{-L(1-H_2(\gamma)-R_{id})}, \\ &= 2^{-L(1-H_2(\gamma)-R_{id}-\frac{1}{L}\log_2|\mathcal{A}_\epsilon^{(J)}|)}, \end{aligned} \tag{23}$$

where the first inequality is due to the union bound and the second one follows from the application of the Chernoff bound on the tail of binomial distribution $\mathcal{B}(L, 0.5)$.[11]

Similarly, one can bound the probability of incorrect identification P_{ic} in the following way:

$$\begin{aligned} P_{ic} &= \Pr[\bigcup_{a\in\mathcal{A}_\epsilon^{(J)}} d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) > \gamma L \cup \bigcup_{s\neq w}^{M} d_H(b^L_{\mathbf{y}_a}(s), b^L_{\mathbf{z}_a}) \le \gamma L|H_w] \\ &\le \sum_{a\in\mathcal{A}_\epsilon^{(J)}} \Pr[d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) > \gamma L|H_w] \\ &\quad + \sum_{a\in\mathcal{A}_\epsilon^{(J)}}\sum_{s\neq w}^{M} \Pr[d_H(b^L_{\mathbf{y}_a}(s), b^L_{\mathbf{z}_a}) \le \gamma L|H_w] \\ &= |\mathcal{A}_\epsilon^{(J)}|\Pr[d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) > \gamma L|H_m] \\ &\quad +(M-1)|\mathcal{A}_\epsilon^{(J)}|\Pr[d_H(b^L_{\mathbf{y}_a}(w), b^L_{\mathbf{z}_a}) \le \gamma L|H_m] \\ &\le |\mathcal{A}_\epsilon^{(J)}|2^{-LD(\gamma||P_{b_e})} + |\mathcal{A}_\epsilon^{(J)}|2^{-L(1-H_2(\gamma)-R_{id})} \\ &\le 2^{-L(D(\gamma||P_{b_e})-\frac{1}{L}\log_2(|\mathcal{A}_\epsilon^{(J)}|))} \\ &+ 2^{-L(1-H_2(\gamma)-R_{id}-\frac{1}{L}\log_2(|\mathcal{A}_\epsilon^{(J)}|))}. \end{aligned} \tag{24}$$

where $D(\gamma||P_{b_e}) = \gamma\log_2\frac{\gamma}{P_{b_e}} + (1-\gamma)\log_2\frac{1-\gamma}{1-P_{b_e}}$ is the divergence and both inequalities are justified by the same arguments as in case of (23).

Finally, taking into account (8), one can rewrite (23) and (24) as follows:

$$P_f = 2^{-L(1-H_2(\gamma)-R_{id}-\frac{J(H(A)+\epsilon)}{L})}; \tag{25}$$

$$P_{ic} = 2^{-L(D(\gamma||P_{b_e})-\frac{J(H(A)\epsilon)}{L}))} + 2^{-L(1-H_2(\gamma)-R_{id}-\frac{J(H(A)+\epsilon)}{L}))}. \tag{26}$$

Therefore, the average probability of error for equally likely hypothesis case is given by:

$$P_e = 0.5P_f + 0.5P_{ic} \tag{27}$$

that is a function of γ. It is easy to show that (27) attains its minimum for $\gamma = \gamma_{opt}$, where

$$\gamma_{opt} = \frac{1 - R + \log_2(1-P_{b_e}) - 1/L}{\log_2\left(\frac{1-P_{b_e}}{P_{b_e}}\right)}, \tag{28}$$

that coincides with the result obtained in.[11] Therefore, although the probabilities of error are impacted by desynchronization in a fixed length multimedia object identification, the system design (threshold selection) remains unchanged with respect to the case when identification is performed over DMC channels. Furthermore, one can claim asymptotic invariance of the considered protocol to the class of desynchronizations of interest (for $L \to \infty$).

5. CONCLUSIONS AND FUTURE RESEARCH PERSPECTIVES

In this paper we considered the problem of fingerprint based multimedia object identification over channels with desynchronizations according to the asymmetric identification model. We assumed that desynchronizations impact only the positions of the elements of a multimedia object relevant data and does not introduce any extra dependence into an i.i.d. input as well as do not modify its dimensionality. We analyzed the achievable rate in a theoretical setup with infinitely long codewords that can be attained in such a protocol and concluded that it asymptotically coincides with identification capacity over the DMC. Furthermore, in order to justify the impact of desynchronizations on the performance of such a system in a more realistic scenario with finite length codewords, we investigated its particular implementation based on random projections and binarization using fundamentals of information decision making theory. We confirmed that the presence of the considered class of desynchronizations as a part of identification channel model leads to probabilities of error increase as a function of randomness of the desynchronization parameters. However, this impact is not reflected in the system design meaning that the decision threshold of the BDD coincides with the one used in identification over the DMC.

As a possible extension of the obtained result we see the information-theoretic analysis of desynchronization resilient object identification based on List Decoding and of protocols with privacy amplification.

ACKNOWLEDGMENT

This paper was partially supported by SNF project 200021-119770 and 200021-132337 and CRADA project.

—-

REFERENCES

1. J. Haitsma, T. Kalker, and J. Oostveen, "Robust audio hashing for content identification," in *International Workshop on Content-Based Multimedia Indexing*, pp. 117–125, (Brescia, Italy), September 2001.
2. F. Lefebvre and B. Macq, "Rash : RAdon Soft Hash algorithm," in *Proceedings of EUSIPCO - European Signal Processing Conference*, (Toulouse, France), 2002.
3. F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *Proc. 2003 IEEE Int. Symp. Inform. Theory*, p. 82, (Yokohama, Japan), June 29 - July 4 2003.
4. O. Koval, S. Voloshynovskiy, F. Beekhof, and T. Pun, "Decision-theoretic consideration of robust perceptual hashing: link to practical algorithms," in *WaCha2007, /Third WAVILA Challenge/*, (Saint Malo, France), June 15th 2007.
5. S. Voloshynovskiy, O. Koval, F. Beekhof, and T. Pun, "Robust perceptual hashing as classification problem: decision-theoretic and practical considerations," in *Proceedings of the IEEE 2007 International Workshop on Multimedia Signal Processing*, (Chania, Crete, Greece), October 1–3 2007.
6. A. L. Varna, A. Swaminathan, and M. Wu, "A decision theoretic framework for analyzing hash-based content identification systems," in *ACM Digital Rights Management Workshop*, pp. 67–76, Oct. 2008.
7. O. Koval, S. Voloshynovskiy, F. Farhadzadeh, T. Holotyak and F. Beekhof, "Information-theoretic analysis of desynchronization invariant object identification," in *Proc. ICASSP 2011*, (Prague, CR), May 2010.
8. P. Tuyls, B. Skoric, and T. K. (Eds.), *Security with Noisy Data: On Private Biometrics, Secure Key Storage and Anti-Counterfeiting*, Springer, 2007.
9. T. Cover and J. Thomas, *Elements of Information Theory*, Wiley and Sons, New York, 1991.
10. F. Willems, "On the capacity of a biometrical identification system," in *In: Proc. of the 2003 IEEE Int. Symp. on Inf. Theory*, pp. 8–2, 2003.
11. S. Voloshynovskiy, O. Koval, F. Beekhof, F. Farhadzadeh, and T. Holotyak, "Information-theoretical analysis of private content identification," in *Proc. IEEE Information Theory Workshop (ITW 2010)*, (Dublin, Ireland), August,30 - September, 3 2010.