Document Authentication Using Graphical Codes: Impacts of the Channel Model

Anh Thu Phan Ho, Bao An Hoang Mai, Wadih Sawaya, Patrick Bas

Document Authentication Using Graphical Codes: Impacts of the Channel Model

Anh Thu Phan Ho
LAGIS UMR 8219 CNRS
Inst. Telecom, Telecom Lille1
59000 Villeneuve d'Ascq, FR
phanho@telecom-lille1.eu

Bao An Mai Hoang
LAGIS UMR 8219 CNRS
Telecom Lille1
59000 Villeneuve d'Ascq, FR
maihoang@telecom-lille1.eu

Wadih Sawaya
LAGIS UMR 8219 CNRS
Inst. Telecom, Telecom-Lille1
59000 Villeneuve d'Ascq, FR
wadih.sawaya@telecom-lille1.eu

Patrick Bas
LAGIS UMR 8219 CNRS
Ecole Centrale de Lille
59651 Villeneuve d'Ascq, FR
patrick.bas@ec-lille.fr

ABSTRACT

This paper proposes to investigate the impact of the channel model for authentication systems based on codes that are corrupted by a physically unclonable noise such as the one emitted by a printing process. The core of such a system is the comparison for the receiver between an original binary code, an original corrupted code and a copy of the original code. We analyze two strategies, depending on whether or not the receiver use a binary version of its observation to perform its authentication test. By deriving the optimal test within a Neyman-Pearson setup, a theoretical analysis shows that a thresholding of the code induces a loss of performance. This study also highlights the fact that the probability of the type I and type II errors can be better approximated, by several orders of magnitude, computing Chernoff bounds instead of the Gaussian approximation. Finally we evaluate the impact of an uncertainty for the receiver on the opponent channel and show that the authentication is still possible whenever the receiver can observe forged codes and uses them to estimate the parameters of the model.

Categories and Subject Descriptors

K.6.5.0 [**Management of Computing and Information Systems**]: Computer Security and Protection—*Authentication*
; H.1.1 [**Models and Principles**]: Systems and Information Theory—*Information theory*

IH&MMSec'13, June 17–19, 2013, Montpellier, France.

Keywords

Authentication, Statistical Analysis, Hypothesis testing, Binary thresholding

1. INTRODUCTION

Authentication of physical products such as documents, goods, drugs, jewels, is a major concern in a world of global exchanges. According to the Organization for Economic Co-operation and Development (OECD), international trade in counterfeit and pirated goods reached more than US $250 billion in 2009 [10], additionally the World Health Organization in 2005 claimed that nearly 25% of medicines in developing countries are forgeries [9].

One way to perform authentication of physical products is to rely on the stochastic structure of the material that composes the product. Authentication can be performed for example by recording the random patterns of the fiber of a paper [6], but such a system is practically heavy to deploy since each product needs to be linked to its high definition capture stored in a database. Another solution is to rely on the degradation induced by the interaction between the product and a physical process such as printing, marking, embossing, carving ... Because of both the defaults of the physical process and the stochastic nature of the mater, this interaction can be considered as a Physically Unclonable Function (PUF) [12] that cannot be reproduced by the forger and can consequently be used to perform authentication. In [5], the authors measure the degradation of the inks within printed color-tiles, and use discrepancy between the statistics of the authentic and print-and-scan tiles to perform authentication. Other marking techniques can also be used, in [11] the authors propose to characterize the random profiles of laser marks on materials such as metals (the technique is called LPUF for Laser-written PUF) and to use them as authentication features.

We study in this paper an authentication system which uses the fact that a printing process at very high resolution can be seen as a stochastic process due to the nature of different elements such as the paper fibers, the ink heterogeneity,

or the dot addressability of the printer. Such an authentication system has been proposed by Picard et al. [8, 7] and uses 2D pseudo random binary codes that are printed at the native resolution of the printer (2400 dpi on a standard offset printer or 812 dpi on digital HP Indigo printer). The whole system is depicted on Fig. 1: once printed on a package to be authenticated, the degraded code can be scanned and thresholded by an opponent (the forger). Note that at this stage the thresholding is necessary because the industrial printers can only print dots, e.g. binary versions of the scanned code. The opponent will produce a printed copy of the original code to manufacture his forgery and the receiver will compare the scanned (and potentially post-processed) version of the original code with the scanned (and potentially post-processed) version of the copied code in order to perform authentication. One advantage of this system over previously cited ones is that it is easy to deploy since the authentication process needs only a scan of the graphical code under scrutiny and the seed used to generate the original one, no fingerprint database is required.

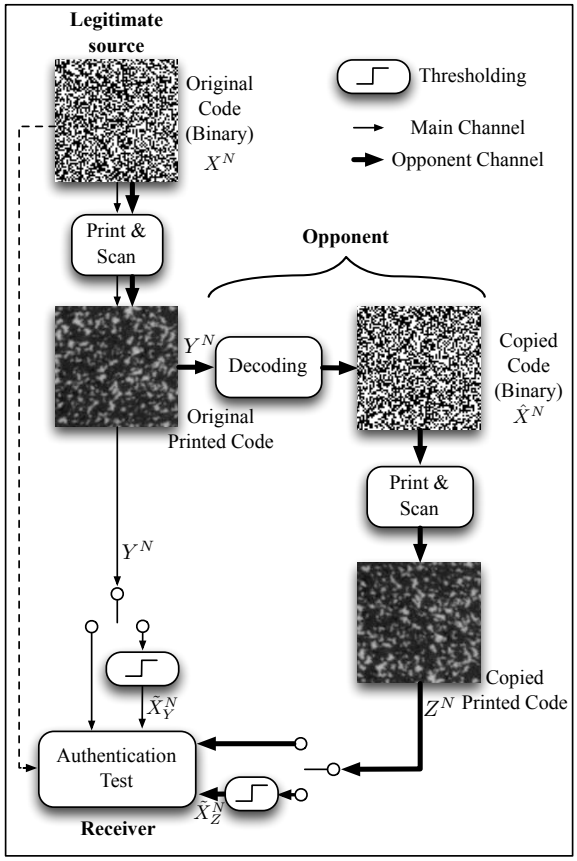

Figure 1: Principle of authentication using graphical codes.

The security of this system solely relies on the use of a PUF, i.e. the impossibility for the opponent to accurately estimate the original binary code. Different security analysis have already been performed w.r.t. this authentication system, or to very similar ones. In [1], the authors have studied the impact of multiple printed observations of same graphical codes and the authors have shown that the power of the noise due to the printing process can be reduced in this particular setup. In [3], the authors use machine learning tools in order to try to infer the original code from an observation of the printed code, their study shows that the estimation accuracy can be increased without recovering perfectly the original code. In [2], the authors consider the security analysis in the rather similar setup of passive fingerprinting using binary fingerprints under informed attacks (the channel between the original code and the copied code is assumed to be a Binary Symmetric Channel), they show that the security increase with the code length and they propose a practical threshold when type I error (original detected as a forgery) and type II error (forgery detected as an original) are equal.

The goal of this paper is to analyze what are the different strategies for the receiver with respect to the post-processing step. We assume that the strategy of the opponent is fixed and that the copied binary code suffers a binary input binary output channel. We show that it is in the receiver's interest to process directly the scanned grayscale code instead of a binary version and we evaluate the impact of the Gaussian approximation of the test with respect to its asymptotic expression. We also investigate the impact of the estimation of the opponent printing channel over the authentication performances.

2. AUTHENTICATION CHANNEL

2.1 Notations

We designate sets by calligraphic font e.g. $\mathcal{X}$ and random variables (RV) ranging over these sets by the same italic capitals e.g. X. The cardinality of the set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. The sequence of N variables $(X_1, X_2, ..., X_N)$ is denoted X^N.

2.2 The setup

The authentication sequence is a binary sequence X^N chosen at random from the message set $\mathcal{X}^N$, and is shared secretly with the legitimate receiver. In our authentication model, X^N is published as a noisy version Y^N, taking values in the set of points $\mathcal{V}^{\mathcal{N}}$ (see Fig. 1). An opponent may observe Y^N and, naturally, tries to retrieve the original authentication sequence. He obtains an estimated sequence $\hat{X}^N$ and prints it to forge a fake sequence Z^N hoping that it will be accepted by the receiver as coming from the legitimate source. The receiver observes then a sequence $\mathcal{O}^N$ which may be one of the two possible sequences Y^N or Z^N, and has to decide whether it comes from the legitimate source or not.

The authentication model may then be viewed as a secret communication problem involving two channels $\mathcal{X} \to (\mathcal{Y}, \mathcal{Z})$, where unreliable communication is desired for one of them (the opponent channel), and perfect communication for the second one (the main channel). We define the main channel as the channel between the legitimate source and the receiver, and the opponent channel as the channel between the legitimate source and the receiver but passing through the counterfeiter channel (see Fig. 1).

2.3 Channel modeling

Let $\mathrm{P}_{V/X}$ be the generic transition matrix modeling the whole physical processes used here, precisely printing and scanning devices. The entries of this matrix are conditional probabilities $P_{V/X}(v/x)$ relating the input alphabet $\mathcal{X}$ and

the output alphabet $\mathcal{V}$ of the whole processes. In practical and realistic situations, $\mathcal{X}$ is a binary alphabet standing with black (0) and white (1) elements of a digital code and the channel output set $\mathcal{V}$ stands for the set of gray level values with cardinality K (for printed and scanned images, $K = 256$). Transition matrix $\mathrm{P}_{V/X}$ may be any discrete distribution over the set $\mathcal{V}$. In our global authentication model, the two channels $\mathcal{X} \to (\mathcal{Y}, \mathcal{Z})$ are considered being discrete and memoryless with conditional probability distribution $P_{YZ/X}(y, z/x)$. The marginal channels $\mathrm{P}_{Y/X}$ and $\mathrm{P}_{Z/X}$ constitute the transition probability matrices of the main channel and the opponent channel respectively. While $\mathrm{P}_{Y/X} = \mathrm{P}_{V/X}$, $\mathrm{P}_{Z/X}$ depends on the opponent processing. We aim here at expressing this marginal distribution considering that the opponent tries to restore the original sequence before publishing his fraudulent sequence Z^N.

When performing a detection to obtain an estimated sequence $\hat{X}^N$ of the original code, the opponent undergoes errors. These errors are evaluated with probabilities $P_{e,W}$ when confusing an original white dot with a black and $P_{e,B}$ when confusing an original black dot with a white. This distinction is due to the fact that the channel distribution $P_{V/X}$ of the physical devices is arbitrary and not necessarily symmetric. Let $\mathcal{D}_{\mathcal{W}}$ and $\mathcal{D}^{\mathrm{c}}_{\mathcal{W}}$ be respectively the optimal decision regions for decoding white or black, obtained after using classical maximum likelihood decoding. As the opponent observes Y^N (we assume that all the physical processes involved are identical for the main channel and the opponent channel), the decision regions will be defined as:

$$\mathcal{D}_{\mathcal{W}} = \left\{ v \in \mathcal{V} :\ P_{V/X}(v/X = 1) > P_{V/X}(v/X = 0) \right\}. \quad (1)$$

Recalling that $\mathrm{P}_{Y/X} = \mathrm{P}_{V/X}$, error probabilities $P_{e,W}$ and $P_{e,B}$ are equal to:

$$P_{e,W} = \sum_{v \in \mathcal{D}^{\mathrm{c}}_{\mathcal{W}}} P_{Y/X}(v/X = 1), \quad (2)$$

$$P_{e,B} = \sum_{v \in \mathcal{D}_{\mathcal{W}}} P_{Y/X}(v/X = 0). \quad (3)$$

The channel $X \to \hat{X}$ can be modeled as a Binary Input Binary Output channel (BIBO) with transition probability matrix $\mathrm{P}_{\hat{X}/X}$:

$$\begin{bmatrix} 1 - P_{e,B} & P_{e,B} \\ P_{e,W} & 1 - P_{e,W} \end{bmatrix} \quad (4)$$

As we can see in Fig. 1, the opponent channel $\mathcal{X} \to Z$ is a physically degraded version of the main channel. Thus, $X \to \hat{X} \to Z$ forms a Markov chain with the relation $P_{\hat{X}Z/X}(\hat{x}, z/x) = P_{\hat{X}/X}(\hat{x}/x) P_{Z/\hat{X}}(z/\hat{x})$. Components of the marginal channel matrix $\mathrm{P}_{Z/X}$ are:

$$\begin{aligned} P_{Z/X}(Z = v/x) &= \sum_{\hat{x}=0,1} P_{\hat{X}Z/X}(\hat{x},\ Z = v/x) \\ &= \sum_{\hat{x}=0,1} P_{\hat{X}/X}(\hat{x}/x) P_{Z/\hat{X}}(Z = v/\hat{x}). \end{aligned} \quad (5)$$

If we assume that the physical processes are identical for the main channel and the opponent channel ($\mathrm{P}_{Z/\hat{X}} = \mathrm{P}_{Y/X} = \mathrm{P}_{V/X}$) the components of the marginal channel matrix $\mathrm{P}_{Z/X}$ will be expressed as:

$$\begin{aligned} P_{Z/X}(Z = v/X = 0) = \quad & (1 - P_{e,B}) P_{V/X}(v/X = 0) \\ & + P_{e,B} P_{V/X}(v/X = 1), \end{aligned} \quad (6)$$

$$\begin{aligned} P_{Z/X}(Z = v/X = 1) = \quad & (1 - P_{e,W}) P_{V/X}(v/X = 1) \\ & + P_{e,W} P_{V/X}(v/X = 0). \end{aligned} \quad (7)$$

2.4 Receiver's strategies

Two strategies are possible for the receiver.

2.4.1 *Binary thresholding:*

As a first strategy the receiver decodes the observed sequence $\mathcal{O}^N$ using maximum likelihood criterion and restores a binary version $\tilde{X}^N$ of the original message X^N. Error probabilities in the main channel, i.e. when $\mathcal{O}^N = Y^N$, are the same as (2) and (3). In the opponent channel, i.e. when $\mathcal{O}^N = Z^N$, these probabilities are:

$$\tilde{P}_{e,W} = \sum_{v \in \mathcal{D}^{\mathrm{c}}_{\mathcal{W}}} P_{Z/X}(v/X = 1), \quad (8)$$

$$\begin{aligned} \tilde{P}_{e,W} \quad = \quad & \sum_{v \in \mathcal{D}^{\mathrm{c}}_{\mathcal{W}}} (1 - P_{e,W}) P_{V/X}(v/X = 1) \\ + \quad & P_{e,W} P_{V/X}(v/X = 0). \end{aligned}$$

Finally we have:

$$\tilde{P}_{e,W} = (1 - P_{e,W}) P_{e,W} + P_{e,W}(1 - P_{e,B}). \quad (9)$$

The same development yields:

$$\tilde{P}_{e,B} = (1 - P_{e,B}) P_{e,B} + P_{e,B}(1 - P_{e,W}). \quad (10)$$

For this first strategy, the opponent channel may be viewed as the cascade of two binary input/binary output channels:

$$\begin{bmatrix} 1 - \tilde{P}_{e,B} & \tilde{P}_{e,B} \\ \tilde{P}_{e,W} & 1 - \tilde{P}_{e,W} \end{bmatrix} =$$

$$\begin{bmatrix} 1 - P_{e,B} & P_{e,B} \\ P_{e,W} & 1 - P_{e,W} \end{bmatrix} \times \begin{bmatrix} 1 - P_{e,B} & P_{e,B} \\ P_{e,W} & 1 - P_{e,W} \end{bmatrix}. \quad (11)$$

When the channel distribution $P_{V/X}$ is symmetric, we have $P_{e,W} = P_{e,B} = p$, and expressions (9) and (10) are unified giving $\tilde{p} = 2p(1 - p)$. We recognize here the cross over probability of two cascaded binary symmetric channels with cross probability p. As we will see in the next section, the test that the receiver will perform to decide whether the observed decoded sequence $\hat{X}^N$ comes from the legitimate source or not is tantamount to counting the number of errors in this case.

2.4.2 *Grey level observations:*

In the second strategy, the receiver performs his test directly on the received sequence $\mathcal{O}^N$ without any given decoding. We will see in the next section that this strategy is better than the previous one for authentication.

3. HYPOTHESIS TESTING

As the observed sequence may come from the legitimate receiver or from a counterfeiter, the receiver considers two hypothesis H_0 and H_1 corresponding respectively to each of the former cases. This problem is formulated by the fact that the observed sequence may be described by two probabilities, say Q_0 and Q_1. A decision rule will assign one of

the two hypothesis for each possible observed sequence and the observed sequence space will then be partitioned into two regions $\mathcal{H}_0$ and $\mathcal{H}_1$. Accepting hypothesis H_0 while it is actually a fake (H_1 is true) leads to an error of type II having probability β. Rejecting hypothesis H_0 while actually the observed sequence comes from the legitimate source (H_0 is true) leads to an error of type I with probability α. An optimal decision rule will be given by the Neyman Pearson criterion. The eponymous theorem states that under the constraint $\alpha \leq \alpha^*$, β is minimized when the choice of H_0 is done if only if the following log-likelihood test is verified:

$$\log \frac{Q_0(v^N)}{Q_1(v^N)} \geq \gamma, \tag{12}$$

where γ is a threshold verifying the constraint $\alpha \leq \alpha^*$.

3.1 Binary thresholding:

In the first strategy, the final observed data is $\tilde{X}^N$ and the original sequence X^N is a side information containing two types of data ("0" and "1"). The distribution of each component ($\tilde{X}_i$, X_i) of the sequence ($\tilde{X}^N$, X^N) is the same for each of these types. We derive now the probabilities that describe $\tilde{X}^N$ for each of the two possible hypothesis. Under hypothesis H_j, $j \in \{0, 1\}$, these probabilities are expressed conditionally to the known original code:

$$\begin{aligned} P(\tilde{X}^N = \tilde{x}^N / X^N = x^N, H_j) &= \prod_{i/X_i=0}^{N_B} P(\tilde{x}_i / X_i = 0, H_j) \\ &\times \prod_{i/X_i=1}^{N_W} P(\tilde{x}_i / X_i = 1, H_j), \end{aligned}$$

where N_B and N_W are respectively the number of black and white components in the original code.

- Under hypothesis H_0 the channel $X \to \hat{X}$ has distributions given by (2) and (3) and we have:

$$\begin{aligned} P\left(\tilde{x}^N / x^N, H_0\right) &= (P_{e,B})^{n_{e,B}} (1 - P_{e,B})^{N_B - n_{e,B}} \\ &\times (P_{e,W})^{n_{e,W}} (1 - P_{e,W})^{N_W - n_{e,W}}, \end{aligned}$$

where $n_{e,B}$ and $n_{e,W}$ are the number of errors ($\tilde{x}_i \neq x_i$) when black is decoded into white and when white is decoded into black respectively.

- Under hypothesis H_1, the channel $X \to \hat{X}$ has distributions given by (9) and (10) and we have:

$$\begin{aligned} P\left(\tilde{x}^N / x^N H_1\right) &= (\tilde{P}_{e,B})^{n_{e,B}} (1 - \tilde{P}_{e,B})^{N_B - n_{e,B}} \\ &\times (\tilde{P}_{e,W})^{n_{e,W}} (1 - \tilde{P}_{e,W})^{N_W - n_{e,W}}. \end{aligned}$$

Applying now the Neyman Pearson criterion (12) the test is expressed as:

$$L_1 = \log \frac{P\left(\tilde{X}^N = \tilde{x}^N / X^N = x^N, H_1\right)}{P\left(\tilde{X}^N = \tilde{x}^N / X^N = x^N, H_0\right)} \underset{H0}{\overset{H1}{\gtrless}} \gamma, \tag{13}$$

$$\begin{aligned} L_1 &= n_{e,B} \log \left(\frac{\tilde{P}_{e,B}(1 - P_{e,B})}{P_{e,B}(1 - \tilde{P}_{e,B})} \right) \\ &+ n_{e,W} \log \left(\frac{\tilde{P}_{e,W}(1 - P_{e,W})}{P_{e,W}(1 - \tilde{P}_{e,W})} \right) \underset{H0}{\overset{H1}{\gtrless}} \lambda_1, \end{aligned} \tag{14}$$

where $\lambda_1 = \gamma - N_B \log\left(\frac{1-\tilde{P}_B}{1-P_B}\right) - N_W \log\left(\frac{1-\tilde{P}_W}{1-P_W}\right)$. For symmetric channels, this expression is simplified by

$$n_{e,B} + n_{e,W} \underset{H0}{\overset{H1}{\gtrless}} \lambda_1'. \tag{15}$$

This expression of the test has the practical advantage to only count the number of errors in order to perform the authentication task without even knowing the opponent channel, but at a cost of a loss of optimality.

3.2 Grey level observations:

In the second strategy, the observed data is $\mathcal{O}^N$. Here again, the distribution of each component ($\mathcal{O}_i$, X_i) of the sequence ($\mathcal{O}^N$, X^N) is the same for each type of data of X. The Neyman Pearson test is expressed as:

$$L_2 = \log \frac{P(\mathcal{O}^N = v^N / X^N = x^N, H_1)}{P(\mathcal{O}^N = v^N / X^N = x^N, H_0)} \underset{H0}{\overset{H1}{\gtrless}} \lambda_2, \tag{16}$$

which can be developed as

$$\begin{aligned} L_2 &= \sum_{i/X_i=1}^{N_W} \log \frac{P_{Z/X}(\mathcal{O}_i = v / X_i = 1)}{P_{Y/X}(\mathcal{O}_i = v / X_i = 1)} \\ &+ \sum_{i/X_i=0}^{N_B} \log \frac{P_{Z/X}(\mathcal{O}_i = v / X_i = 0)}{P_{Y/X}(\mathcal{O}_i = v / X_i = 0)} \underset{H0}{\overset{H1}{\gtrless}} \lambda_2, \end{aligned} \tag{17}$$

$$\begin{aligned} L_2 = &\sum_{i/X_i=1}^{N_W} \log \left(1 - P_{e,W} + P_{e,W} \frac{P_{V/X}(\mathcal{O}_i/0)}{P_{V/X}(\mathcal{O}_i/1)}\right) + \\ &\sum_{i/X_i=0}^{N_B} \log \left(1 - P_{e,B} + P_{e,B} \frac{P_{V/X}(\mathcal{O}_i/1)}{P_{V/X}(\mathcal{O}_i/0)}\right) \underset{H0}{\overset{H1}{\gtrless}} \lambda_2. \end{aligned} \tag{18}$$

Note that here the expressions of the channel models $P_{V/X}(\mathcal{O}_i/X_i)$ are required in order to perform the optimal test.

3.3 Performance of hypothesis testing

3.3.1 The Gaussian approximation

In the previous section we have expressed the Neyman-Pearson test for the two proposed strategies resumed by (14) and (18). These tests may then be practically performed on the observed sequence in order to make a decision about its authenticity. We aim now at expressing the error probabilities of type I and II, and comparing the two possible strategies described previously. Let $m = 1, 2$ be the index denoting the strategy, a straightforward calculation gives

$$\alpha_m = \sum_{l > \lambda_m} P_{L_m}(l / H_0), \tag{19}$$

$$\beta_m = \sum_{l < \lambda_m} P_{L_m}(l / H_1). \tag{20}$$

As the length N of the sequence is generally large, we use the central limit theorem to study the distributions P_{L_m}, $m = 1, 2$.

For the binary thresholding strategy, the observed sequence is $\hat{X}^N$. In (14) $n_{e,W}$ and $n_{e,B}$ are binomial random variables,

with parameters depending on the source of the observed sequence, i.e. if it comes from the legitimate source or from the counterfeiter. Let N_x and $P_{e,x}$ stand respectively for the number data of type x in the original code and the cross over probabilities of the BIBO channels (4) and (11). When N is large enough, the binomial random variables are approximated with a Gaussian distribution. We have:

$$n_{e,x} \sim \mathcal{N}(N_x P_{e,x},\ N_x P_{e,x}(1-P_{e,x})). \tag{21}$$

One can obviously now deduce the parameters of the normal approximation describing the log-likelihood L_1.

For the second strategy, i.e. when the receiver tests directly the observed gray level sequence, the log-likelihood L_2 Eq. (18) may be expressed as two sums of i.i.d. and becomes:

$$L_2 \quad = \quad \sum_{i/X_i=1}^{N_W} \ell(\mathcal{O}_i;\ 1) + \sum_{i/X_i=0}^{N_B} \ell(\mathcal{O}_i;\ 0) \underset{H0}{\overset{H1}{\gtrless}} \lambda_2, \tag{22}$$

where $\ell(v;\ x)$ is a function $\ell :\ \mathcal{V} \to \mathbb{R}$ with parameter $x = 0,\ 1$ and having some distribution with mean and variance equal to:

$$\mu_x = E[\ell(V;\ x) \mid H_j] = \sum_{v \in \mathcal{V}} \ell(v,\ x) P_{V/X}(v/x), \tag{23}$$

and

$$\mathrm{var}[\ell(V;\ x) \mid H_j] = \sum_{v \in \mathcal{V}} (\ell(v,\ x) - \mu_x)^2 P_{V/X}(v/x), \tag{24}$$

with $P_{V/X} = P_{Y/X}$ (resp. $P_{V/X} = P_{Z/X}$) for $j = 0$ (resp. 1) . The central limit theorem is then used again for the distribution of L_2 to compute the type I and type II error probabilities.

3.3.2 Asymptotic expression

One important problem is the fact that the Gaussian approximation proposed previously provides inaccurate error probabilities when the threshold λ_m in (19) and (20) is far from the mean of the random variable L_m. Chernoff bound and asymptotic expression are preferred in this context as very small error probabilities of type I and II may be desired [4]. Given a real number s the Chernoff bound on type I and II errors may be expressed for $m = 1,\ 2$ as:

$$\alpha_m = \Pr(L_m \geq \lambda_m) \leq e^{-s\lambda_m} g_{L_m}(s) \text{ for any } s > 0, \tag{25}$$

$$\beta_m = \Pr(L_m \leq \lambda_m) \leq e^{-s\lambda_m} g_{L_m}(s) \text{ for any } s < 0, \tag{26}$$

where the function $g_{L_m}(s)$ is the moment generating function of L_m defined as:

$$g_{L_m}(s) = E_{L_m}\left[e^{sL_m}\right]. \tag{27}$$

These bounds are significant for λ_m far from $E[L_m]$, namely when bounding the tails of a distribution. The tightest bound is obtained by finding the value of s that provides the minimum of the RHS of (25) and (26), i.e. the minimum of $e^{-s\lambda_m} g_{L_m}(s)$. Taking the derivative, the value s that provides the tightest bound is such that[1]:

$$\lambda_m = \frac{\frac{\mathrm{d}g_{L_m}(s)}{\mathrm{d}s}}{g_{L_m}(s)} = \frac{\mathrm{d}}{\mathrm{d}s} \ln g_{L_m}(s). \tag{28}$$

[1](one can show that $e^{-s\lambda_m} g_{L_m}(s)$ is convex)

Reminding that L_m is a sum of N independant random variables, asymptotic analysis in probability theory (when N is large enough) shows that bounds similar to (25) and (26) are much more appropriate for estimating α_m and β_m than the Gaussian approximation. To make this more clear, we will introduce the semi-invariant moment generating function after an acute observation of the identity (28). The semi-invariant moment generating function of L_m is $\mu_{L_m}(s) = \ln g_{L_m}(s)$. This function has many interesting properties that ease the extraction of an asymptotic expression for (25) and (26) [4]. For instance, this function is additive for the sum of independant random variables, which yields for example for $m = 2$:

$$\mu_{L_2}(s) = \sum_{i/X_i=1}^{N_W} \mu_{\ell_{i/1}}(s) + \sum_{i/X_i=0}^{N_B} \mu_{\ell_{i/0}}(s), \tag{29}$$

where $\mu_{\ell_{i/x}}(s)$ is the semi-invariant moment generating function of the random variable $\ell_{i/x} = \ell(\mathcal{O}_i;\ x)$. In addition, the s optimizing the bound and obtained from (28) may be driven from the sum of the derivatives:

$$\lambda_m = \sum_{i/X_i=1}^{N_W} \mu'_{\ell_{i/1}}(s) + \sum_{i/X_i=0}^{N_B} \mu'_{\ell_{i/0}}(s). \tag{30}$$

Chernoff bounds on type I and II errors (25) and (26) may then be expresses as:

$$\begin{aligned} \alpha_m \quad &= \quad \Pr(L_m \geq \lambda_m) \\ &\leq \quad \exp\left[\sum_{i/X_i=1}^{N_W} \left(\mu_{\ell_{i/1}}(s) - s\mu'_{\ell_{i/1}}(s)\right) \right. \\ &+ \quad \left. \sum_{i/X_i=0}^{N_B} \left(\mu_{\ell_{i/0}}(s) - s\mu'_{\ell_{i/1}}(s)\right) \right] \text{ for any } s > 0, \end{aligned} \tag{31}$$

and

$$\begin{aligned} \beta_m \quad &= \quad \Pr(L_m \leq \lambda_m) \\ &\leq \quad \exp\left[\sum_{i/X_i=1}^{N_W} \left(\mu_{\ell_{i/1}}(s) - s\mu'_{\ell_{i/1}}(s)\right) \right. \\ &+ \quad \left. \sum_{i/X_i=0}^{N_B} \left(\mu_{\ell_{i/0}}(s) - s\mu'_{\ell_{i/1}}(s)\right) \right] \text{ for any } s < 0. \end{aligned} \tag{32}$$

The distribution of each component $(\mathcal{O}_i, X_i)$ of the sequence $(\mathcal{O}^N,\ X^N)$ is the same for each type of data of X, and $\mu_{\ell_{i/x}}(s) = \mu_{\ell/x}(s)$ is independent from i for a given type of data $x = 0,\ 1$. The RHS in (31) and (32) can be simplified as:

$$\exp\left[N_W \left(\mu_{\ell/1}(s) - s\mu'_{\ell/1}(s)\right) + N_B \left(\mu_{\ell/0}(s) - s\mu'_{\ell/0}(s)\right)\right]. \tag{33}$$

The asymptotic expression is evaluated (see [4], Appendix 5A) for the sum of i.i.d and for large N we have (for $N_B \approx N_W \approx N/2$), for $s > 0$:

$$\begin{aligned} \alpha_m \quad &= \quad \Pr(L_m \geq \lambda_m) \\ &\underset{N\to\infty}{\to} \quad \frac{1}{|s|\sqrt{N\pi\mu''_\ell(s)}} \exp\left\{\tfrac{N}{2}\left[\mu_\ell(s) - s\mu'_\ell(s)\right]\right\}. \end{aligned} \tag{34}$$

and for $s < 0$:

$$\begin{aligned}\beta_m &= \Pr(L_m \le \lambda_m) \\ &\underset{N\to\infty}{\to} \frac{1}{|s|\sqrt{N\pi\mu''_\ell(s)}} \exp\left\{\frac{N}{2}\left[\mu_\ell(s) - s\mu'_\ell(s)\right]\right\}.\end{aligned} \tag{35}$$

where $\mu_\ell(s) = \mu_{\ell/0}(s) + \mu_{\ell/1}(s)$, $\mu'_\ell(s) = \mu'_{\ell/0}(s) + \mu'_{\ell/1}(s)$, and $\mu''_\ell(s) = \mu''_{\ell/0}(s) + \mu''_{\ell/1}(s)$ is the second derivative of the semi invariant moment generating function of random variable $\ell(v;\, x)$ defined by:

$$\begin{aligned}\ell(v;\,0) &= \log\left(1 - P_{e,W} + P_{e,W}\frac{P_{V/X}(v/0)}{P_{V/X}(v/1)}\right), \\ \ell(v;\,1) &= \log\left(1 - P_{e,B} + P_{e,B}\frac{P_{V/X}(v/1)}{P_{V/X}(v/0)}\right).\end{aligned}$$

Fig. 2 illustrates the gap between the estimation of α and β using the Gaussian approximation and the asymptotic expression. The Monte-Carlo simulations confirm the fact that the derived Chernoff bounds are tight.

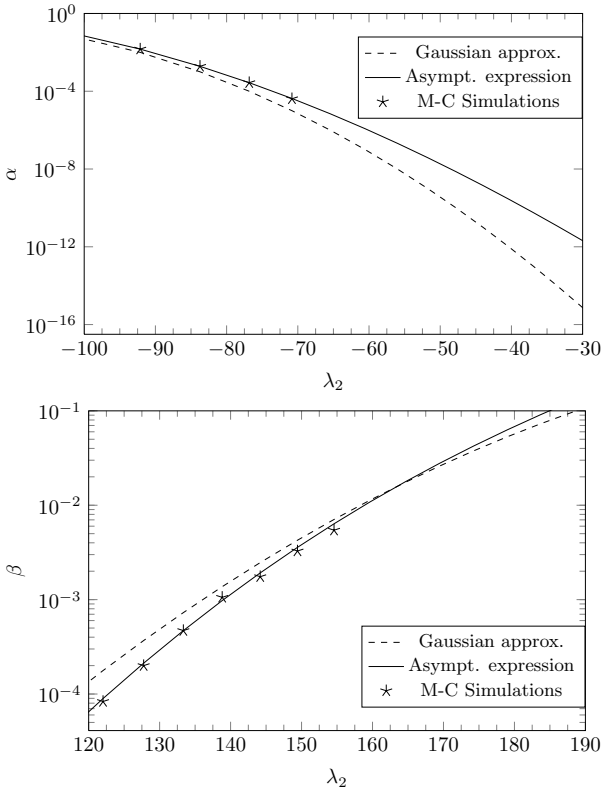

Figure 2: Comparison between the Gaussian approximation, the asymptotic expression and Monte-Carlo simulations (10^6 trials) for the second strategy, $N = 2000$, $\sigma = 50$.

3.4 Comparison between the two strategies

In this setup and without loss of generality, we assume that the print and scan channel is modeled by a discreet non-symmetric and memoryless channel with binary input alphabet $\mathcal{X}$ and grey level outputs $\mathcal{V}$, generated from a normalized discrete Gaussian distribution $P_{V/x}(v/x)$. For $x = 0,\ 1$:

$$P_{V/x}(v/x) = \frac{\exp(-(v-\mu_x)^2/2\sigma^2)}{\sum\limits_{v\in\mathcal{V}} \exp(-(v-\mu_x)^2/2\sigma^2)}. \tag{36}$$

Fig. 3 compares the Receiver Operating Characteristic (ROC) curves associated with the two different strategies, and the impact of the Gaussian approximation. We can notice that the gap between the two strategies is important, this is not a surprise since the binary thresholding removes information about the forged code Y, yet this has a practical impact because one practitioner can be tempted to use the weighted bit error rate given in (15) as an authentication score for its easy implementation.

Moreover, as we will see in the next section, the plain scan of the graphical code can be used whenever the receiver needs to estimate the opponent's channel.

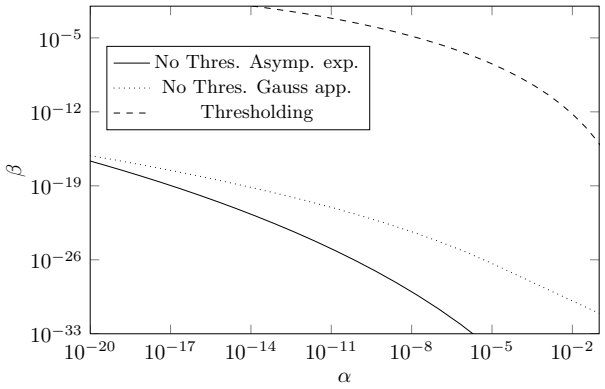

Figure 3: ROC curves for the two different strategies ($N = 2000$, $\sigma = 52$).

4. IMPACT OF THE ESTIMATION OF THE PRINT AND SCAN CHANNEL

The previous scenarios assume that the receiver has a full knowledge of the print and scan channel. Here we assume that the receiver has also to estimate the opponent channel before performing authentication. From the estimated parameters, the receiver will compute a threshold and a test according to a Neyman-Pearson strategy. Depending of the number of observations N_o, the estimated model and test will decrease the performance of the authentication system.

We consider now that the opponent uses different printing device. According to (6) and (7), the parameters to be estimated are $P_{e,W}$, $P_{e,B}$, μ_0, μ_1 and σ. We use the classical Expectation Maximization (E.M.) algorithm combined with the Newton's method to solve the maximization step.

Fig. 4 shows the authentication performances using model estimation for $N_o = 2000$ observed symbols. We can notice that the performance (and the estimated parameters) are very close to an exact knowledge of the model. This analysis shows also that if the receiver has some assumptions of the opponent channel and enough observations, he should perform model estimation instead of using the thresholding strategy. Fig. 5 shows the importance of model estimation when comparing it to a blind authentication test when the receiver assumes that both the opponent channel and his channel are identical.

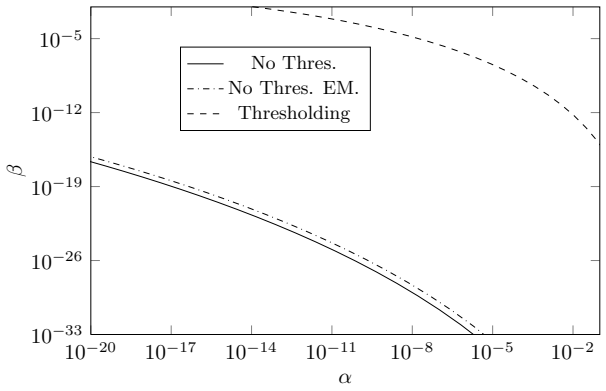

Figure 4: Authentication performance using model estimation with the EM algorithm ($N = 2000$, $N_o = 2000$, $\sigma = 52$, $\mu_0 = 50$, $\mu_1 = 150$). The asymptotic expression is used to derive the error probabilities.

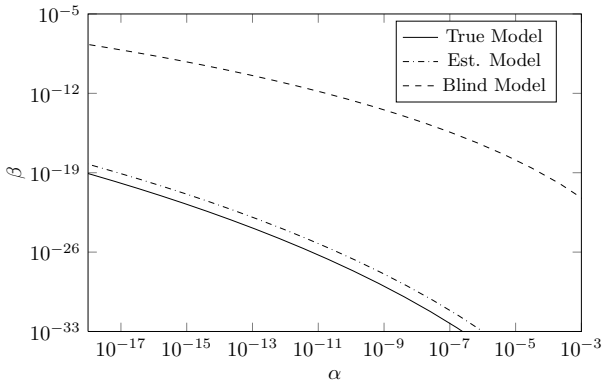

Figure 5: ROC curves comparing different knowledges about the channel while the opponent uses a different printing process ($\sigma = 40$, $\mu_0 = 40$, $\mu_1 = 160$). "True model": the receiver knows exactly this model, "Blind model": the receiver uses his printing process model as the opponent model, "Est. model": the receiver estimates the opponent channel using $N_o = 2000$ observations.

5. CONCLUSIONS AND PERSPECTIVES

This papers brings numerous conclusions on the authentication using binary codes corrupted by a manufacturing stochastic noise:

- The nature of the receiver's input is of upmost importance and thresholding is a bad strategy with respect to getting an accurate version of the genuine or forged code, except if the system requires it.
- The Gaussian approximations used to compute the ROC of the authentication system are not valuable anymore for very low type I or type II errors. Chernoff bounds have to be used instead.
- If the opponent's print and scan channel remains unknown for the receiver, he can use estimation techniques such as the E.M algorithm in order to estimate the channel.
- The proposed methodology is not impacted by the nature of the noise, and can be applied for different memoryless channels that are more realistic for modeling the printing process.

Our futures works plan to address the potential benefits for authentication of structured codes such as error-correcting codes.

6. ACKNOWLEDGEMENTS

This work was partly supported by the National French project ANR-10-CORD-019 "Estampille".

7. REFERENCES

[1] C. Baras, F. Cayre, et al. 2D bar-codes for authentication: A security approach. *Proceedings of EUSIPCO 2012*, 2012.

[2] F. Beekhof, S. Voloshynovskiy, and F. Farhadzadeh. Content authentication and identification under informed attacks. In *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*, pages 133–138. IEEE, 2012.

[3] M. Diong, P. Bas, C. Pelle, and W. Sawaya. Document authentication using 2D codes: Maximizing the decoding performance using statistical inference. In *Communications and Multimedia Security*, pages 39–54. Springer, 2012.

[4] R.G. Gallager. *Information theory and reliable communication*, volume 15. Wiley, 1968.

[5] M.D. Gaubatz, S.J. Simske, and S. Gibson. Distortion metrics for predicting authentication functionality of printed security deterrents. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1489–1492. IEEE, 2009.

[6] T. Haist and H.J. Tiziani. Optical detection of random features for high security applications. *Optics communications*, 147(1-3):173–179, 1998.

[7] J. Picard, C. Vielhauer, and N. Thorwirth. Towards fraud-proof id documents using multiple data hiding technologies and biometrics. *SPIE Proceedings–Electronic Imaging, Security and Watermarking of Multimedia Contents VI*, pages 123–234, 2004.

[8] J. Picard and J. Zhao. Improved techniques for detecting, analyzing, and using visible authentication patterns, July 28 2005. WO Patent WO/2005/067,586.

[9] WCO Press. Global congress addresses international counterfeits threat immediate action required to combat threat to finance/health, 2005. "`http://www.wcoomd.org/en/media/newsroom/2005/november`".

[10] WCO Press. Counterfeiting and piracy endangers global economic recovery, say global congress leaders, 2009. "`http://www.wipo.int/pressroom/en/articles/2009/article_0054.html`".

[11] S.S. Shariati, F.X. Standaert, L. Jacques, B. Macq, M.A. Salhi, and P. Antoine. Random profiles of laser marks. In *Proceedings of the 31st WIC Symposium on Information Theory in the Benelux*, 2010.

[12] G.E. Suh and S. Devadas. Physical unclonable functions for device authentication and secret key generation. In *Proceedings of the 44th annual Design Automation Conference*, pages 9–14. ACM, 2007.