Performance Analysis of Identification System Based on Order Statistics List Decoder

Farhadzadeh, Farzad; Voloshynovskyy, Svyatoslav; Koval, Oleksiy

How to cite

FARHADZADEH, Farzad, VOLOSHYNOVSKYY, Svyatoslav, KOVAL, Oleksiy. Performance Analysis of Identification System Based on Order Statistics List Decoder. In: IEEE International Symposium on Information TheoryInformation Theory Proceedings (ISIT). Austin (TX). [s.l.] : Institute of Electrical and Electronics Engineers (IEEE), 2010. p. 2652–2656. doi: 10.1109/ISIT.2010.5513690

This publication URL: https://archive-ouverte.unige.ch/unige:47635
Publication DOI: 10.1109/ISIT.2010.5513690

Performance Analysis of Identification System Based on Order Statistics List Decoder

Farzad Farhadzadeh, Sviatoslav Voloshynovskiy, and Oleksiy Koval
Computer Science Department
University of Geneva
7 route de Drize, CH 1227, Geneva, Switzerland
Email: {Farzad.Farhadzadeh, svolos, Oleksiy.Koval}@unige.ch

***Abstract*—In this work we advocate an approach for the statistical performance analysis of an identification system. The statistical performance analysis is accomplished for the corresponding probability of miss and false acceptance based on the order statistic list decoding framework.**

I. Introduction

In the past decade, the cryptography community developed a set of powerful tools to protect physical objects [1]. Historically, the protection of items is based on technologies that use features that are difficult to duplicate, copy or clone.

The present work addresses an identification problem based on biometrics or Physical Unclonable Functions (PUFs). Both biometrics and PUFs are well-known techniques in forensic applications [1] because of their ability to serve as a unique identifier for many people and objects.

Channel distortions, due to acquisition imperfection, compression *etc.*, impact data and make it noisy. Therefore, the identification system should be able to cope with data variations. The decoders in classical identification setups estimate a unique index for a given query. This makes them relatively sensitive to strong distortions in their input. Another approach, which can be considered as the generalization of the above mentioned one, was firstly proposed by Elias [2] in communication theory and is known as *list decoding*. The main feature of this type of decoding is to produce a fixed list size of the most likely candidates rather than a single one. The result of [2] was generalized by Forney to a variable list size [3]. Using a Neyman-Pearson type optimality criterion it was demonstrated that the proposed decoder guarantees maximal Gallager-type error exponents. In many identification problems, the final sink of information is a human being. This restriction makes variable list size decoding undesirable, due to the high variability of the list size, for very noisy environments the list might be exceedingly long.

As mentioned, these types of decoders have been used in a communication setup, where the decoder estimates the sent message from a fixed codebook. It is also believed that list decoding might bring additional benefits for identification systems that operate in very noisy environments. However, contrary to digital communications, in the identification setup the decoder should determine whether a given query is related to some elements of the database, and if so, which one. Therefore, just using a list decoder is not sufficient to ensure that the estimated indices are really related to the query, i.e., without restricting the probability of false acceptance. To generalize the list decoder to the identification setup, we must add an erasure option to the decoding rule, which means that the decision regions are not exhaustive.

The main contribution of this paper can be summarized as follows. We introduce a new identification setup by using a fixed maximum list size decoder based on an *order statistic list decoder* (OSLD) and analyze its performance versus unique or so-called ordinary decoding for the identification problem. For reasons of computational complexity, privacy and security, it is undesirable for an identification system to retain the biometrics or PUFs in their original form. In most cases, some non-invertible dimensionality reduction transform is applied to produce a template (a.k.a. digital fingerprint) that is quantized at the second stage. We use the random projection transform and binarization for binary template generation [6]. By using i.i.d. Gaussian random projectors, the data and channel statistics change to the Gaussian model and binarization converts them to a binary model. Therefore, we analyze the OSLD probability of miss and false acceptance over two channels, the additive white Gaussian noise (AWGN) and the binary symmetric channel (BSC).

The outline of this paper is as follows. In section II, we introduce the identification setup. In section III, we analyze the error events related to our setup. The simulation results and conclusions are presented in section IV and V.

Notations: We use capital letters X to denote scalar random variables and $\mathbf{X}$ to denote vector random variables. Corresponding small letters x and $\mathbf{x}$ denote the realizations of scalar and vector random variables. All vectors without sign tilde are assumed to be of the length N and with the sign tilde of length L. We use $h(\cdot)$ and $H(\cdot)$ to denote differential entropy, entropy and binary entropy. $\mathcal{N}(\mu, \sigma_X^2)$ stands for the Gaussian distribution with mean μ and variance σ_X^2. $\mathcal{B}(N, p)$ denotes the Binomial distribution with N trials and probability of success p. $\Phi(\cdot)$ denotes cumulative distribution function (CDF) for a $\mathcal{N}(0, 1)$ random variable and $\mathrm{Q}(\cdot)$ designates Q-function, $\mathrm{Q}(v) = 1 - \Phi(v)$. $\|\cdot\|$ denotes Euclidean vector norm. $V_{(r:M)}$ stands for the r-th order statistics of M i.i.d. random variables.

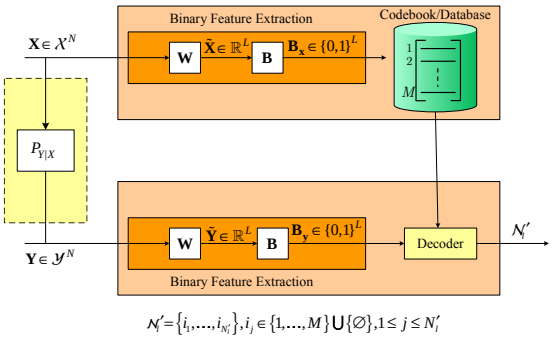


Fig. 1. Identification problem based on binary templates.

II. Identification Setup

The identification setup under analysis is shown in Fig. 1. The *Codebook/Database* is generated by recording biometrics or PUFs of each item to be identified $\mathbf{x}(m) \in \mathcal{X}^N, m = 1, \ldots, M$, during the enrollment stage. It is assumed that these biometrics or PUFs are drawn independently from a common stationary distribution $p_X(x)$. The receiver observes a noisy version $\mathbf{Y}$ of the biometric or PUFs of a given person or item, where the probabilistic mismatch between $\mathbf{X}$ and $\mathbf{Y}$ is modeled by $p(\mathbf{y}|\mathbf{x}) = \prod_{n=0}^{N-1} p(y[n]|x[n])$. The second step of enrollment is to reduce the dimensionality from N to L. The reduction is accomplished by applying random projectors [6]; we use an approximation of a so-called *orthoprojector* $\mathbf{W} \in \mathbb{R}^{L\times N}$, where each element $w_{ij} \sim \mathcal{N}(0, \frac{1}{N})$ with $1 \leq i \leq L, 1 \leq j \leq N$. A dimensionality reduction step is applied to produce $\tilde{\mathbf{X}} = \mathbf{WX}$ and $\tilde{\mathbf{Y}} = \mathbf{WY}$. L-length binary data is derived from the projected data by taking the sign of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$, producing binary templates, $\mathbf{B_x}$ and $\mathbf{B_y}$.

The decoding process in the identification setup is accomplished in two steps. At the first step, the primary candidates are chosen by the OSLD. At the second step, a threshold is applied to all candidates, and the candidates which satisfy the constraint remain on the list. The OSLD decoding procedure can be summarized as follows[1]:

1) The likelihood functions, $p(\mathbf{y}|\mathbf{x}(m)), 1 \leq m \leq M$, for all database entries are evaluated.
2) The computed likelihood functions are sorted in ascending order.
3) The N_l indices with the largest likelihood functions are chosen which form a set $\mathcal{N}_l$. Parameter N_l is referred as a list size.
4) The final output set of the decoder is defined as:

$$\mathcal{N}_l' = \{m \in \mathcal{N}_l : p(\mathbf{y}|\mathbf{x}(m)) \geq e^{N\gamma}\} \tag{1}$$

where the parameter γ controls the number of final candidates.

To investigate performances of the decoders, we should consider a composite hypothesis test where:

$\mathcal{H}_0$: The query $\mathbf{Y}$ is unrelated to any database entry,

$\mathcal{H}_m$: The query $\mathbf{Y}$ is related to the m^{th} entry of database.

The performance of the decoder is evaluated by:

- the Probability of miss (a related query is incorrectly rejected, i.e., not in the final list $\mathcal{N}_l'$);
- the Probability of false acceptance (an unrelated query is incorrectly accepted).

III. Error Events

Before considering error events, we will consider *Order Statistics* which will be used in the computation of the probability of errors. We suppose that $V(1), V(2), \ldots, V(M)$ are M i.i.d. random variables, each with the cumulative distribution function (CDF) $F(v)$. Let $F_{(r:M)}(v)$ denote the CDF of the r-th order statistics $V_{(r:M)}$, which corresponds to the r-th position of $v_{(1:M)} \leq v_{(2:M)} \leq \ldots \leq v_{(r:M)} \leq \ldots \leq v_{(M:M)}$ for a specific outcome. The CDF of the r-th order statistic $V_{(r:M)}$ is given by [5]:

$$\begin{aligned} F_{(r:M)}(v) &= \Pr\left\{V_{(r:M)} \leq v\right\} \\ &= \Pr\{\text{at least } r \text{ of } V_i \text{ are less or equal to } v\} \\ &= \sum_{i=r}^{M} \binom{M}{i} F^i(v)[1-F(v)]^{M-i} \end{aligned} \tag{2}$$

since the term in the summand is the binomial probability that *exactly* i of $V(1), V(2), \ldots, V(M)$ are less than equal to v.

In the following subsections, we consider the probability of miss and probability of false acceptance for the AWGN channel and the BSC.

A. Probability of Miss

Once the list of primary candidates is selected by the OSLD, the final candidates are extracted by applying the threshold to their likelihoods. A miss event occurs when the query related to the database entry does not belong to the list of final candidates. The probability of miss, P_M, is given by:

$$\begin{aligned} P_M = 1 - \sum_{m=1}^{M} &\Pr\{\mathbf{x}(m) \in \mathcal{N}_l \text{ and } p(\mathbf{y}|\mathbf{x}(m)) \geq e^{N\gamma}|\mathcal{H}_m\} \times \\ &\Pr\{\mathcal{H}_m\}, \end{aligned} \tag{3}$$

where $\mathcal{N}_l$ is the primary list of candidates. As the entries of the database are identically distributed and equally likely to be queried, the overall probability of miss does not depend on the particular index and hence:

$$P_M = 1 - \Pr\{\mathbf{x}(1) \in \mathcal{N}_l \text{ and } p(\mathbf{y}|\mathbf{x}(1)) \geq e^{N\gamma}|\mathcal{H}_1\}. \tag{4}$$

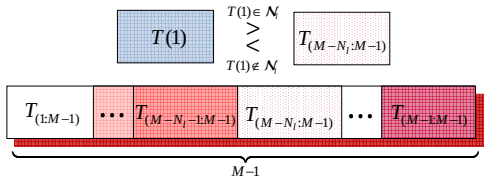
The miss event is considered for the reduced dimensionality data over the AWGN channel and binary data over the BSC.

1) The Additive White Gaussian Noise Channel: We assume that the decoder input is generated by the following additive channel[2]:

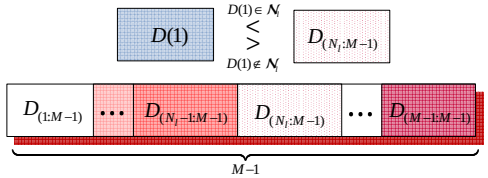
$$\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} + \tilde{\mathbf{Z}}, \tag{5}$$

[1]The low-complexity identification of OSLD decoding based on the concept of bit reliability is given in [7].

[2]This channel results from the random projections with possible diagonalization [6].



(a) AWGN Channel.



(b) BSC.

Fig. 2. The illustration of the event that the first entry of database related to the query is not on the primary list.

where $\tilde{\mathbf{X}} \sim \mathcal{N}(\mathbf{0}, \frac{\xi}{L}\mathbf{I}_L), \xi = \frac{L}{N}\|\mathbf{x}\|^2$ and $\tilde{\mathbf{Z}} \sim \mathcal{N}(\mathbf{0}, \sigma_Z^2\mathbf{I}_L)$. Since

$$\begin{aligned} T(m) &= \sum_{n=0}^{L-1} \tilde{y}[n]\tilde{x}(m)[n] - \frac{1}{2}\sum_{n=0}^{L-1} \tilde{x}^2(m)[n] \\ &= \sum_{n=0}^{L-1} \tilde{y}[n]\tilde{x}(m)[n] - \frac{1}{2}\xi \end{aligned} \tag{6}$$

is a sufficient statistic, the first event in (3) occurs iff $T(m)$ is among the N_l largest of $\{T(1), T(2), \ldots, T(M)\}$. Fig. 2a illustrates this given the fact that the query is related to the first entry. Therefore the probability of miss can be stated as,

$$\begin{aligned} P_M &= 1 - \Pr\{T_{(M-N_l:M-1)} < T(1) \text{ and } T(1) \geq \lambda | \mathcal{H}_1\} \\ &= 1 - \int_{\lambda}^{\infty} \Pr\{T_{(M-N_l:M-1)} < t | \mathcal{H}_1, T(1) = t\} p_{T(1)}(t)\, \mathrm{d}t, \end{aligned} \tag{7}$$

where from (1) and (6), $\lambda = \frac{\xi + L\sigma_Z^2}{2} + \sigma_Z^2 L\left(\frac{1}{2}\ln(2\pi\sigma_Z^2) + \gamma\right)$, $p_{T(1)(t)}$ denotes the PDF of $T(1)$. For equiprobable, equal energy and orthogonal $\tilde{\mathbf{x}}(m)$, conditioned on $\mathcal{H}_1$:

$$T(m) \sim \begin{cases} \mathcal{N}(\frac{1}{2}\xi, \sigma_Z^2\xi), & \text{for } m = 1, \\ \mathcal{N}(-\frac{1}{2}\xi, \sigma_Z^2\xi), & \text{for } m \neq 1. \end{cases} \tag{8}$$

From (2), the CDF of the $(M - N_l)^{th}$ order statistics of the i.i.d random variables $T(m), m \neq 1$, is given by:

$$\begin{aligned} &\Pr\{T_{(M-N_l:M-1)} < t\} = F_{(M-N_l:M-1)}(t) = \\ &\sum_{p=M-N_l}^{M-1} \binom{M-1}{p} \Phi^p\left(\frac{t + \frac{1}{2}\xi}{\sqrt{\xi\sigma_Z^2}}\right)\left[\mathrm{Q}\left(\frac{t + \frac{1}{2}\xi}{\sqrt{\xi\sigma_Z^2}}\right)\right]^{(M-1)-p} \end{aligned} \tag{9}$$

From (7), (8), (9) and letting $u \triangleq (t + \frac{1}{2}\xi)/\sqrt{\xi\sigma_Z^2}$, the probability of miss over the AWGN channel can be expressed as[3]:

$$\begin{aligned} P_M = 1 - \Bigg\{ &\sum_{p=M-N_l}^{M-1} \binom{M-1}{p} \\ &\times \int_{\frac{\lambda + \frac{1}{2}\xi}{\sqrt{\xi\sigma_Z^2}}}^{\infty} \Phi^p(u)\mathrm{Q}^{(M-1)-p}(u) \\ &\times \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(u - \sqrt{\frac{\xi}{\sigma_Z^2}}\right)^2\right] \mathrm{d}u \Bigg\}. \end{aligned} \tag{10}$$

2) Binary Symmetric Channel: After dimensionality reduction and binarization, we have binary data with the length L, where $L < N$. In order to evaluate the probability of miss, we consider it over the BSC with a crossover probability of P_b. For any $\mathbf{b}_{\mathbf{x}}(m), \mathbf{b}_{\mathbf{y}} \in \{0,1\}^L$, the likelihood function

$$p(\mathbf{b}_{\mathbf{y}} | \mathbf{b}_{\mathbf{x}}(m)) = P_b^{d(m)}(1 - P_b)^{L - d(m)} \tag{11}$$

is a decreasing function of the Hamming distance $d(m) \triangleq d_H(\mathbf{b}_{\mathbf{y}}, \mathbf{b}_{\mathbf{x}}(m))$ for $0 \leq P_b \leq 0.5$. Therefore, the first event in (4) occurs, if $D(m)$ is not among the N_l smallest $\{D(1), D(2), \ldots, D(M)\}$. Similarly to the AWGN channel case, for a given query related to $\mathcal{H}_1$ the first event occurs when $D(1)$ of the related query is on the primary list (Fig. 2b). Therefore, the probability of miss can be stated as:

$$\begin{aligned} P_M &= 1 - \Pr\{D_{(N_l:M-1)} > D(1) \text{ and } D(1) \leq \eta | \mathcal{H}_1\} \\ &= 1 - \sum_{d=0}^{\eta} \Pr\{D_{(N_l:M-1)} > d | \mathcal{H}_1, D(1) = d\} p_{D(1)}(d), \end{aligned} \tag{12}$$

where from (1) and (11), $\eta = L\frac{\gamma - \ln(1-P_b)}{\ln(P_b/(1-P_b))}$ and $p_{D(1)}(d)$ denotes the PMF of $D(1)$. The dimensionality reduction and binarization change the statistics of the database generated by $p_{\mathbf{X}}(\mathbf{x})$ to the Binomial distribution, i.e., $\mathbf{B}_{\mathbf{x}} \sim \mathcal{B}(L, 1/2)$ for $p_{\mathbf{X}}(\mathbf{x}) = \mathcal{N}(\mathbf{0}, \sigma_X^2\mathbf{I}_N)$. Conditioned on $\mathcal{H}_1$, the sufficient statistics can be expressed as follows:

$$D(m) \sim \begin{cases} \mathcal{B}(L, P_b), \text{for } m = 1, \\ \mathcal{B}(L, \frac{1}{2}), \text{ for } m \neq 1. \end{cases} \tag{13}$$

From (2), the CDF of the N_l^{th} order statistics of the i.i.d. random variables $D(m), m \neq 1$ is given by:

$$\begin{aligned} &\Pr\{D_{(N_l:M-1)} < d\} = F_{(N_l:M-1)}(d) = \\ &\sum_{p=N_l}^{M-1} \binom{M-1}{p} S(d)^p (1 - S(d))^{(M-1)-p}, \end{aligned} \tag{14}$$

where $S(d) \triangleq \left(\frac{1}{2}\right)^L \sum_{x=0}^{d} \binom{L}{x}$. From (12), (13) and (14), the miss probability over the BSC is given by:

$$\begin{aligned} P_M = 1 - \Bigg\{ &\sum_{d=0}^{\eta} \binom{L}{d} P_b^d (1 - P_b)^{L-d} \\ &\times \sum_{p=0}^{N_l - 1} \binom{M-1}{p} S(d)^p (1 - S(d))^{(M-1)-p} \Bigg\}. \end{aligned} \tag{15}$$

[3]It should be pointed out that for $N_l = 1$ and $\lambda = -\infty$, (10) coincides with the error probability of the ML decoder [4], page 121.

B. Probability of False Acceptance

The main reason to consider the probability of false acceptance is to show the reliability of the decoding process with respect to various attacking strategies. There are different scenarios to investigate the reliability of the decoder in identification setups:

- The attacker has no access to the PDF of database generation, $p_X(x)$.
- The attacker has access to the PDF of database generation, $p_X(x)$.
- The database entries are partially known by the attacker.
- The database entries are totally known by the attacker.

After the Gaussian random projections, reduced database entries have the Gaussian distribution; therefore, the first and second scenarios coincide. In this paper, we consider the scenario in which the PDF is fully known by the attacker. Then, blindly generated codewords that follow the PDF are sent to the decoder. For this scenario, the probability of false acceptance can be defined as:

$$P_{FA} = \Pr\{\mathcal{N}_l' \neq \varnothing | \mathcal{H}_0\}. \tag{16}$$

In the following subsections, the false acceptance event is considered over the AWGN channel and the BSC.

1) Additive White Gaussian Noise Channel: For a given query, which is unrelated to any entry of the database, the correlation between the query and all entries is computed and sorted in ascending order. The following events are defined:

$$E_{T_{(i:M)}} = \{T_{(i:M)} \geq \lambda | \mathcal{H}_0\}, \tag{17}$$

where $M - N_l + 1 \leq i \leq M$ and $T(m) \sim \mathcal{N}(-\frac{1}{2}\xi, \xi\sigma_z^2)$ for large L. From (16) and (17), the probability of false acceptance can be defined as:

$$\begin{aligned} P_{FA} &= \Pr\{\cup_{i=M-N_l+1}^{M} E_{T_{(i:M)}} | \mathcal{H}_0\} \\ &= 1 - \Pr\{\cap_{i=M-N_l+1}^{M} E^c_{T_{(i:M)}} | \mathcal{H}_0\} \\ &\overset{(a)}{=} 1 - \Pr\{E^c_{T_{(M:M)}} | \mathcal{H}_0\} = \Pr\{E_{T_{(M:M)}} | \mathcal{H}_0\}, \end{aligned} \tag{18}$$

where $E^c_{T_{(M:M)}}$ is the complement of $E_{T_{(M:M)}}$, and (a) follows from the fact that as $E^c_{T_{(M:M)}}$ occurs the rest of the events will certainly occur. Then the probability of false acceptance can be derived as in [5]:

$$\begin{aligned} P_{FA} &= \Pr\left\{ \max_{1 \leq m \leq M} T(m) \geq \lambda | \mathcal{H}_0 \right\} \\ &\cong 1 - \Phi^M \left(\frac{\lambda + \frac{1}{2}\xi}{\sqrt{\xi \sigma_Z^2}} \right). \end{aligned} \tag{19}$$

The interesting thing to note is that the probability of false acceptance is independent of the primary list size N_l, under the considered attacking scenario.

2) Binary Symmetric Channel: The same as for the AWGN channel, we define the following events:

$$E_{D_{(i:M)}} = \{D_{(i:M)} \leq \eta | \mathcal{H}_0\}, \tag{20}$$

where $1 \leq i \leq N_l$, $D(m) \sim \mathcal{B}(L, \frac{1}{2})$ and $E_{D_{(i:M)}}$ is the event that the i^{th} ascending ranked Hamming distance between the query and an entry of the database is smaller than the threshold. The probability of false acceptance is found as:

$$\begin{aligned} P_{FA} &= \Pr\{\cup_{i=1}^{N_l} E_{D_{(i:M)}} | \mathcal{H}_0\} \\ &= 1 - \Pr\{\cap_{i=1}^{N_l} E^c_{D_{(i:M)}} | \mathcal{H}_0\} \\ &\overset{(a)}{=} 1 - \Pr\{E^c_{D_{(1:M)}} | \mathcal{H}_0\} = \Pr\{E_{D_{(1:M)}} | \mathcal{H}_0\} \end{aligned} \tag{21}$$

where $E^c_{D_{(i:M)}}$ is the complement of $E_{D_{(i:M)}}$, and (a) follows from the fact that if the event $E^c_{D_{(1:M)}}$ occurs the rest of the events will certainly occur. Then the probability of false acceptance can be derived [5]:

$$\begin{aligned} P_{FA} &= \Pr\left\{ \min_{1 \leq m \leq M} D(m) \leq \eta | \mathcal{H}_0 \right\} \\ &= 1 - \left[1 - \left(\frac{1}{2}\right)^L \sum_{x=0}^{\eta} \binom{L}{x} \right]^M. \end{aligned} \tag{22}$$

Similarly to the AWGN channel case, the probability of false acceptance is independent of the primary list size.

IV. Simulation Results

The proposed decoder performance is evaluated by using databases of synthetic data with different sizes that are independently and identically normally distributed, i.e., $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$. The Signal-to-Noise Ratio (SNR) is defined as $\text{SNR} = 10\log_{10}\frac{\sigma_X^2}{\sigma_Z^2}, \sigma_X^2 = 1$.

Fig. 3 and Fig. 4 confirm the fact that the empirical and analytical probability of miss and false acceptance over the AWGN channel and the BSC coincide with each other as far as the precision of computer simulation allows.

From (10), (19) and (15), (22) the probability of miss and false acceptance over the AWGN channel and the BSC are computed and *receiver operating characteristic* (ROC) curves are shown for different SNRs, database sizes M and primary list sizes N_l.

Fig. 5 shows that if the SNR is high, the performance for the unique decoder and the OSLD are the same. However, in low SNR scenarios, the OSLD helps improve the performance while the probability of correct detection $P_D = 1 - P_M$ is not close to one. Fig. 6 shows the impact of N_l and M on the identifier performance. Increasing the primary list size N_l improves the identifier performance but after a certain value of N_l it does not change anymore, and increasing the database size M decreases the performance.

Fig. 7 and Fig. 8 confirm the above conclusions for the BSC case.

V. Conclusion

We have analyzed the identification setup based on the OSLD framework. In light of this framework, we have investigated the OSLD performance by deriving analytical equations for the probability of miss and false acceptance.

Simulation results show that on the one hand, the OSLD can only improve the identifier performance in very low SNR scenarios. On the other hand, this improvement is restricted by a certain range of list sizes. The obtained results can

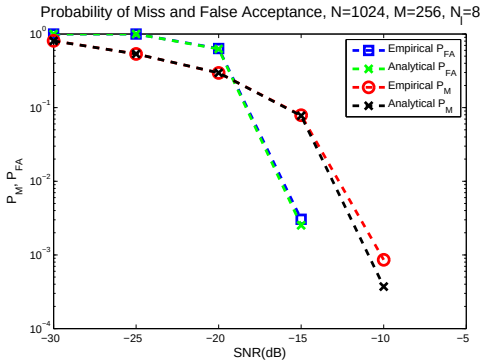


Fig. 3. The probability of miss and false acceptance over the AWGN channel.

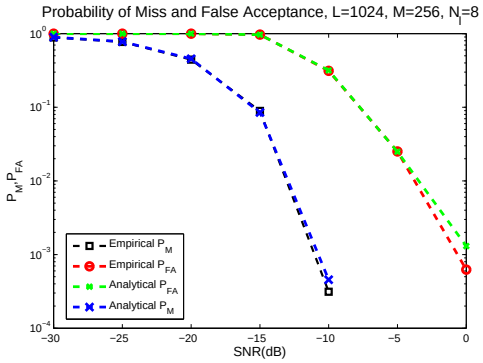


Fig. 4. The probability of miss and false acceptance over the BSC.

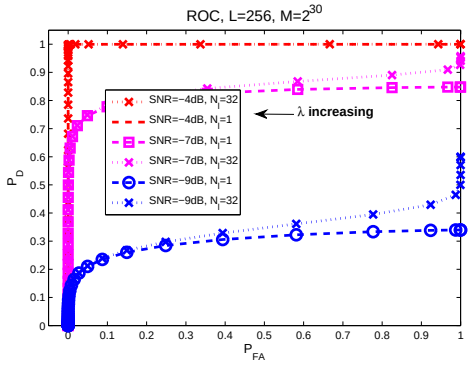


Fig. 5. The OSLD performance for the Gaussian setup and various SNRs.

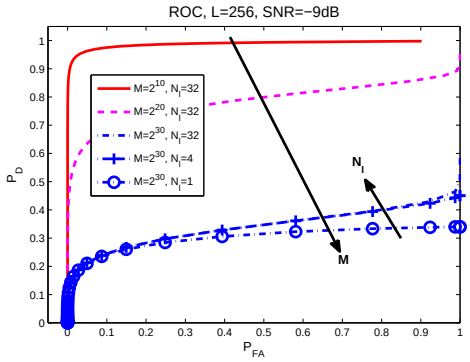


Fig. 6. The OSLD performance for the Gaussian setup.

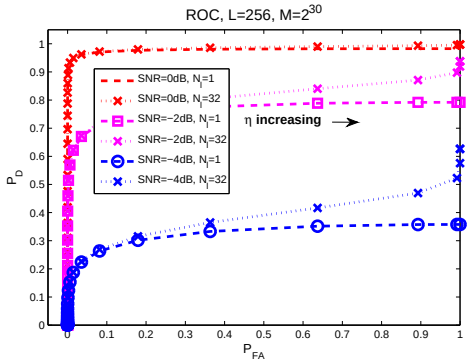


Fig. 7. The OSLD performance for the binary setup and various SNRs.

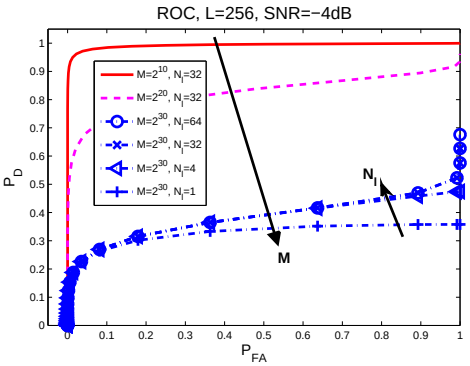


Fig. 8. The OSLD performance for the binary setup.

be of interest for security and content-based retrieval system analysis.

ACKNOWLEDGMENT

The authors would like to thank Prof. E. Telatar and SIP group members F. Beekhof and T. Holotyak for stimulating discussions. This paper was partially supported by SNF projects 200021-119770.

REFERENCES

[1] P. Tuyls, B. Skoric, and T. Kevenaar, *Security with noisy data: On Private Biometrics, Secure Key Storage and Anti-Counterfeiting*, Springer, 2007.

[2] P. Elias, *List decodeing for noisy channels*, Tech. Rept. 335, Research Labratoary of Electronics, M.I.T, 1955.

[3] G. D. Forney, Jr, *Exponential error bounds for erasure, list and decision feedback schemes*, IEEETrans Inf. Theory, vol. IT-14, no. 2, pp. 206-220, Mar. 1968.

[4] S. M. Kay, *Fundamental of Statitical Signal Processing, Detection Theory*, vol. 2, pp. 121, Prentice-Hall, 1998.

[5] H. A. David, H. N. Nagaraja, *Order Statistics*, 3rd ed, pp. 9, Wiley-Interscience, 2003.

[6] S. Voloshynovskiy, O. Koval, F. Beekhof, and T. Pun, *Conception and limits of robust perceptual hashing: toward side information assisted hash functions*, SPIE Photonics West, San Jose, USA, 2009.

[7] F. Beekhof, S. Voloshynovskiy, O. Koval, and T. Holotyak, *Fast Identification Algorithms for Forensic Applications*, IEEE International Workshop on Information Forensics and Security, London, 2009.