SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks **00**: 1–16 (2008)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sec.0000

Performance of the Verification for Binary Memoryless Channels [†]

Vladimir B. Balakirsky, A. J. Han Vinck

American University of Armenia, 0019 Yerevan, Armenia
Institute for Experimental Mathematics, 45326 Essen, Germany

Summary

We consider the verification problem when the verifier receives a pair of binary vectors $(\mathbf{a}, \mathbf{b})$ and makes the acceptance or the rejection decision. The acceptance decision has to be made if the vector $\mathbf{b}$ is the result of transmission of the vector $\mathbf{a}$ over a known memoryless channel. An attacker knows the weight of the vector $\mathbf{a}$ and substitutes a vector $\mathbf{b}$ generated by a stationary Bernoulli source having the assigned probability of bit 1, which is unknown to the verifier. We design the verification algorithm with the metric depending on the weight of the vector $\mathbf{a}$. The use of this metric allows us to restrict the possibilities of the attacker in such a way that the best strategy becomes the flipping of a fair coin. The performance of the algorithm is essentially better than the performance of the scheme, which is based on the maximum likelihood decision for the known channel. We also show that the algorithm is a special case of a general verification scheme for an arbitrary memoryless channel. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Verification; Authentication; Decoding; Hypotheses Testing

1. Introduction and statement of the problem for binary memoryless channels

1.1. Subclass of binary channels

Let us consider a binary channel, denoted by V_{ξ_0,ξ_1}, whose conditional probabilities are determined by the matrix

$$\left[\begin{array}{cc} V_{\xi_0,\xi_1}(0|0) & V_{\xi_0,\xi_1}(1|0) \\ V_{\xi_0,\xi_1}(0|1) & V_{\xi_0,\xi_1}(1|1) \end{array}\right] = \left[\begin{array}{cc} 1-\xi_0 & \xi_0 \\ \xi_1 & 1-\xi_1 \end{array}\right],$$

where

$$\xi_0 + \xi_1 \in (0,1). \tag{1}$$

In particular, we will be interested in 3 special cases of the channels created when the probability that bit 0 is corrupted is equal to $\xi \in (0, 1/2)$ and the probability that bit 1 is corrupted takes values in the $[0, 1/2]$ interval:

– an asymmetric channel where 1's are noiselessly transmitted, $V_{\xi,0}$;
– a binary symmetric channel, $V_{\xi,\xi}$;
– an asymmetric channel where 1's are destroyed, $V_{\xi,1/2}$.

If $(\xi_0, \xi_1) = (\eta, 1-\eta)$ for some $\eta \in [0,1]$, then $\xi_0 + \xi_1 = 1$, i.e., the constraint (1) does not hold, and the output bit is independent of the input bit. Thus, transmission of a bit over the $V_{\eta,1-\eta}$ channel is equivalent to the generating of the output bit by a Bernoulli $(1-\eta, \eta)$ source. We will use this equivalence to unify formal considerations.

Given a pair of probabilities (ξ_0, ξ_1) and a pair of binary vectors $(\mathbf{a}, \mathbf{b})$ of length n, the conditional

[*] Correspondence to: E-mail: v_b_balakirsky@rambler.ru
[†] This work was partially supported by the DFG.

probability of receiving the vector $\mathbf{b}$ when the vector $\mathbf{a}$ was sent over the memoryless V_{ξ_0,ξ_1} channel is equal to

$$V_{\xi_0,\xi_1}(\mathbf{b}|\mathbf{a}) = \prod_{t=1}^{n} \begin{cases} 1-\xi_{a_t}, & \text{if } b_t = a_t, \\ \xi_{a_t}, & \text{if } b_t \neq a_t. \end{cases}$$

The probabilistic description of the transmission over the $V_{\eta,1-\eta}$ channel is introduced similarly.

Notice that the following data transmission scheme can be considered as the transformation $V_{\xi,0} \to V_{\xi,1/2}$. Suppose that there is a source generating bit 0 with probability $1-\omega$ and bit 1 with probability ω. Let the generated bit be transmitted over two parallel $V_{\xi,0}$ channels. The probabilities that the received bit is 0 and 1 at the output of any of the channels are equal to

$$\begin{aligned} Q(0) &= (1-\omega)(1-\xi), \\ Q(1) &= (1-\omega)\xi + \omega, \end{aligned}$$

and the probabilities that the pair of received bits is $(1,0)$ and $(0,1)$ are equal to

$$\begin{aligned} Q(1,0) &= (1-\omega)\xi(1-\xi), \\ Q(0,1) &= (1-\omega)(1-\xi)\xi. \end{aligned}$$

Hence,

$$\omega = \frac{\xi - 2\xi^2}{1+\xi-2\xi^2} \Rightarrow \frac{Q(0,1)}{Q(1)} = 1/2.$$

Furthermore

$$\frac{Q(1,0)}{Q(0)} = \xi.$$

The ratios $Q(0,1)/Q(1)$ and $Q(1,0)/Q(0)$ represent the conditional probabilities of bits 0 and 1 at the output of one of the channels when 1 and 0 are bits at the output of another channel. The channel describing the dependence of bits at the outputs of parallel $V_{\xi,0}$ channels is obviously more noisy than the $V_{\xi,0}$ channel. The degradation can be such that the conditions for the transmission of 1's are reverted.

A numerical illustration of the considerations above is as follows,

$$\begin{bmatrix} 3/4 & 1/4 \\ 0 & 1 \end{bmatrix} \overset{\omega=0.9}{\to} \begin{bmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{bmatrix}.$$

The probabilities that the received bit is 0 and 1 at the output of the $V_{1/4,0}$ channel are equal to

$$\begin{aligned} Q(0) &= (8/9)(3/4) = 2/3, \\ Q(1) &= (8/9)(1/4) + (1/9)1 = 1/3. \end{aligned}$$

and the probabilities that the pair of received bits at the outputs of parallel $V_{1/4,0}$ channels are different is equal to

$$Q(1,0) = Q(0,1) = (8/9)(1/4)(3/4) = 1/6.$$

Hence, the conditional probabilities $0 \to 1$ and $1 \to 0$ of the channel, which relates the observations are equal to

$$\frac{1/6}{2/3} = 1/4, \quad \frac{1/6}{1/3} = 1/2.$$

The independent use of the data transmission scheme n times brings the scheme in Figure 1.

The example above is of interest for processing biometric data. Suppose that there are n features and that the t-th component of the vector $\mathbf{c}$ is equal to 1 if and only if the person has the t-th feature, $t = 1, \ldots, n$. Let the vector $\mathbf{c}$ be observed at the enrollment stage as the vector $\mathbf{a}$ in such a way that if the feature is present, then it is definitely registered. The vector $\mathbf{a}$ is stored in the database under the name of the person. If the same person appears at the verification stage, then the vector $\mathbf{c}$ is observed again and the vector $\mathbf{b}$ represents the outcomes of the observations. The pair of vectors $(\mathbf{a}, \mathbf{b})$ is available to the verifier, who analyzes the dependence between these vectors to accept or to reject the identity claim of the person. This dependence is established by the conditional probability of receiving the vector $\mathbf{b}$ given the vector $\mathbf{a}$ when the vector $\mathbf{c}$ is unknown. An available estimate of the conditional probability can be obtained if the verifier assumes that this vector is generated at random because it characterizes “a randomly chosen person”. If the statistics over biometric characteristics of people is such that each feature is registered with the relative frequency ω and the features are independent, the verifier can assume that each bit of the vector $\mathbf{c}$ is independently generated by a Bernoulli source.

From a formal point of view, our example contains a paradox: although all features of the person are noiselessly observed at the enrollment stage, the verifier has to take into account only the features that were not observed while computing the conditional probability of the vector $\mathbf{b}$ received at the verification stage. However, the verification algorithm should not be only based on the obtained conditional probability (an attacker, who presents the all–zero vector to the verifier, always passes through the verification stage with the acceptance decision). Therefore, bits of the vector $\mathbf{b}$ received at positions, where the features were registered at the enrollment stage, are also

DOI: 10.1002/sec

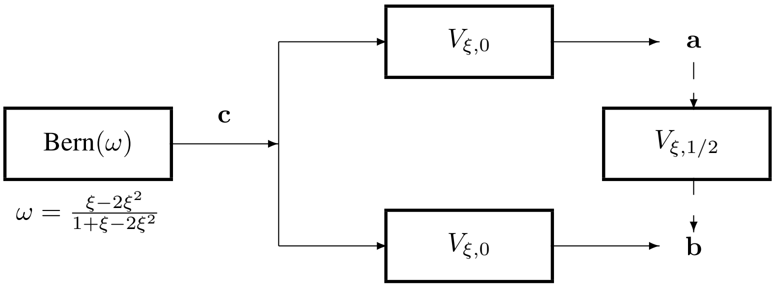

Fig. 1. Creating the channel $V_{\xi,1/2}$ between observations at the outputs of two parallel $V_{\xi,0}$ channels whose input is generated by a Bernoulli source.

important. These considerations are continued in the next subsection.

1.2. A basic verification scheme

Let us consider the scheme in Figure 2 that will be considered as a basic verification scheme for binary memoryless channels. The pair of binary vectors $(\mathbf{a},\mathbf{b})$ is given to the verifier. The output of the verifier is a binary variable taking values “Acceptance” and “Rejection”, abbreviated as Acc and Rej. Let the verification be formalized as a mapping

$$(\mathbf{a},\mathbf{b}) \;\rightarrow\; \text{Decision} = \begin{cases} \text{Acc}, & \text{if } \mathbf{b} \in \mathcal{D}_{\mathbf{a}}, \\ \text{Rej}, & \text{if } \mathbf{b} \notin \mathcal{D}_{\mathbf{a}}, \end{cases}$$

where $\mathcal{D}_{\mathbf{a}}, \mathbf{a} \in \{0,1\}^n$, are the acceptance sets. These sets have to be fixed in advance in such a way that the verifier can reliably distinguish between the following cases.

Acc: the vector $\mathbf{b}$ is received as a result of transmission of the vector $\mathbf{a}$ over the V_{ξ_0,ξ_1} channel.

Rej: the vector $\mathbf{b}$ is generated by a Bernoulli source with the probability of bit 1 equal to η (equivalently, the vector $\mathbf{b}$ is received as a result of transmission of the vector $\mathbf{a}$ over the $V_{\eta,1-\eta}$ channel).

The vector $\mathbf{a}$ *and the pair* (ξ_0,ξ_1) *are known and* (1) *holds, while the value of parameter* η *is unknown.*

Let us illustrate the difficulties of the problem and consider the case $(\xi_0,\xi_1) = (\xi, 1/2)$. Suppose that the verifier wants to have the false rejection rate at most ε for a given vector $\mathbf{a}$. Due to the structure of the $V_{\xi,1/2}$ channel, there is an obvious (the maximum likelihood) decision: the verifier does not pay attention to the bits of the vector $\mathbf{b}$ received at positions where the vector $\mathbf{a}$ contains ones and only controls the weight of the vector received at positions where the vector $\mathbf{a}$ contains zeroes. Namely, the vector $\mathbf{b}$ has to be included into the set $\mathcal{D}_{\mathbf{a}}$ if and only if the weight does not exceed k_ε, which is chosen from the inequality

$$\sum_{k=0}^{k_\varepsilon} \text{Bin}_\xi(k|n-\text{wt}(\mathbf{a})) \geq 1-\varepsilon,$$

where

$$\text{wt}(\mathbf{a}) = \left|\left\{ t \in \{1,\ldots,n\} : a_t = 1 \right\}\right|$$

denotes the weight of the vector $\mathbf{a}$ and

$$\text{Bin}_\xi(k|\tilde{n}) = \binom{\tilde{n}}{k}(1-\xi)^{\tilde{n}-k}\xi^k$$

for all $\tilde{n}$ and k (in the following considerations, we will assume that $\binom{\tilde{n}}{k} = 0$ if $k \notin \{0,\ldots,\tilde{n}\}$). However, such an assignment includes the all–zero vector in all acceptance sets. If $\eta = 0$, then the attacker presents this vector to the verifier in the Rej case, and he always passes through the verification with the acceptance decision.

To overcome the difficulties, we notice that there are many acceptance sets with the property that the sum of conditional probabilities of receiving their entries, when the vector $\mathbf{a}$ is transmitted over the $V_{\xi,1/2}$ channel, is not less than $1-\varepsilon$. The particular set can be specified by the set $\mathcal{K}$ consisting of pairs of integers, where the first integer belongs to the set

DOI: 10.1002/sec

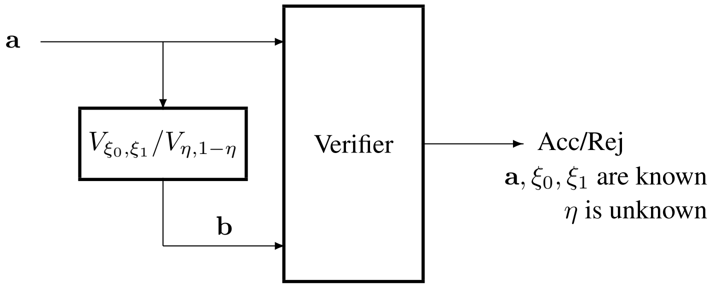
$\mathbf{a}$
$V_{\xi_0,\xi_1}/V_{\eta,1-\eta}$
$\mathbf{b}$
Verifier
Acc/Rej
$\mathbf{a}, \xi_0, \xi_1$ are known
η is unknown

Fig. 2. The structure of a basic verification scheme for binary memoryless channels.

$\{0,\dots,n-\text{wt}(\mathbf{a})\}$ and the second integer belongs to the set $\{0,\dots,\text{wt}(\mathbf{a})\}$. Given a vector $\mathbf{b}$, the verifier computes

$$k_0 = \left|\left\{ t \in \{1,\dots,n\} : (a_t,b_t) = (0,1) \right\}\right|, \quad (2)$$

$$k_1 = \left|\left\{ t \in \{1,\dots,n\} : (a_t,b_t) = (1,0) \right\}\right|, \quad (3)$$

and makes the acceptance decision if and only if $(k_0,k_1) \in \mathcal{K}$. The set $\mathcal{K}$ has to be fixed in such a way that

$$\sum_{(k_0,k_1)\in\mathcal{K}} \text{Bin}_\xi(k_0|n-\text{wt}(\mathbf{a}))\binom{\text{wt}(\mathbf{a})}{k_1}2^{-\text{wt}(\mathbf{a})} \geq 1-\varepsilon.$$

The solution presented in the further considerations is a regular construction of the set $\mathcal{K}$ that can be also used for the $V_{\xi,1/2}$ channel.

1.3. Formal statement of the problem

Let us denote the normalized weight of the vector $\mathbf{a}$ by

$$p_\mathbf{a} = \frac{\text{wt}(\mathbf{a})}{n}. \quad (4)$$

The probabilities of the verification errors, called the false rejection and the false acceptance rates, can be expressed as

$$\text{FRR}_\mathbf{a} = \sum_{\mathbf{b}\notin\mathcal{D}_\mathbf{a}} V_{\xi_0,\xi_1}(\mathbf{b}|\mathbf{a}),$$

$$\text{FAR}_\mathbf{a}(\eta) = \sum_{\mathbf{b}\in\mathcal{D}_\mathbf{a}} V_{\eta,1-\eta}(\mathbf{b}|\mathbf{a}),$$

and the problem is formulated as several requirements.

R$_0$: *Find a regular construction for the acceptance sets* $\mathcal{D}_\mathbf{a}$, $\mathbf{a} \in \{0,1\}^n$.

“A regular construction” is understood as a threshold–type set

$$\mathcal{D}_\mathbf{a} = \left\{ \mathbf{b} \in \{0,1\}^n : m(\mathbf{a},\mathbf{b}) < T_{p_\mathbf{a}} \right\}, \quad (5)$$

where the function $m(\mathbf{a},\mathbf{b})$, called the metric between the vectors $\mathbf{a}$ and $\mathbf{b}$, is defined as an additive extension of the component–wise metric,

$$m(\mathbf{a},\mathbf{b}) = \frac{1}{n}\sum_{t=1}^{n} m(a_t,b_t). \quad (6)$$

The values of the component–wise metric are determined by the matrix

$$\mathbf{M} = \begin{bmatrix} m(0,0) & m(0,1) \\ m(1,0) & m(1,1) \end{bmatrix} = \begin{bmatrix} 0 & m_0 \\ m_1 & 0 \end{bmatrix}. \quad (7)$$

The rules above include the constructions of the acceptance sets according to the Hamming distance between the vectors $\mathbf{a}$ and $\mathbf{b}$ when $m_0 = m_1 = 1$ and to the maximum likelihood decoding for the V_{ξ_0,ξ_1} channel when

$$(m_0,m_1) = \left(\ln\frac{1-\xi_0}{\xi_0}, \ln\frac{1-\xi_1}{\xi_1} \right). \quad (8)$$

This observation will be presented in more details in the following sections.

Notice also that the construction above can be equivalently introduced as an assignment of the set

$$\mathcal{K}_\mathbf{a} = \left\{ (k_0,k_1) : k_0m_0 + k_1m_1 < nT_{p_\mathbf{a}} \right\} \quad (9)$$

DOI: 10.1002/sec

satisfying the condition

$$\sum_{(k_0,k_1)\in\mathcal{K}_{\mathbf{a}}} \mathrm{Bin}_{\xi_0}(k_0|n-\mathrm{wt}(\mathbf{a})) \cdot \mathrm{Bin}_{\xi_1}(k_1|\mathrm{wt}(\mathbf{a})) \geq 1-\varepsilon.$$

Having received the vector $\mathbf{b}$, the verifier computes (k_0, k_1) by (2), (3) and makes the acceptance decision if and only if $(k_0, k_1) \in \mathcal{K}_{\mathbf{a}}$. Therefore the previous considerations represent a special case of the general setup.

R$_R$: *For all vectors* $\mathbf{a} \in \{0,1\}^n$,

$$\mathrm{FRR}_{\mathbf{a}} \leq \varepsilon.$$

This requirement is oriented to practical applications of the verification scheme. For example, the scheme should guarantee a certain false rejection rate for an arbitrary chosen person whose biometric characteristics are expressed by the vector $\mathbf{a}$.

R$_A$: *Given a* $p \in \{0/n, 1/n, \ldots, n/n\}$,

$$\max_{\eta\in[0,1]} \max_{\mathbf{a}\in\{0,1\}^n_{np}} \mathrm{FAR}_{\mathbf{a}}(\eta) \to \min, \qquad (10)$$

where

$$\{0,1\}^n_{np} = \left\{ \mathbf{a} \in \{0,1\}^n : p_{\mathbf{a}} = p \right\}$$

denotes the set of binary vectors of weight np.

In the minimization problem stated above, we assume that the attacker knows the weight of the vector $\mathbf{a}$, which is equal to np. By the memoryless assumptions and the restriction of the metric between any pair of vectors to an additive extension of the component–wise metric, the value of $\mathrm{FAR}_{\mathbf{a}}(\eta)$ is the same for all vectors $\mathbf{a} \in \{0,1\}^n_{np}$. Therefore, instead of taking the maximum in (10), we can equivalently require the minimum value of $\mathrm{FAR}_{\mathbf{a}}(\eta)$ for any vector $\mathbf{a}$ having the weight np.

Remark. The vector $\mathbf{b}$ is the value of a randomly chosen vector $B^n = (B_1, \ldots, B_n)$ whose probability distributions in the Acc and in the Rej cases are defined as

$$\begin{aligned} \Pr\{\, B^n = \mathbf{b} \,\} &= V_{\xi_0,\xi_1}(\mathbf{b}|\mathbf{a}), && (11)\\ \Pr\{\, B^n = \mathbf{b} \,|\, \eta \,\} &= V_{\eta,1-\eta}(\mathbf{b}|\mathbf{a}). && (12) \end{aligned}$$

The problem can be viewed as the discrimination between these distributions on the basis of the received vector $\mathbf{b}$ when η is unknown. Notice that if the vector $\mathbf{a}$ contains both zeroes and ones, then $V_{\xi_0,\xi_1}(\mathbf{b}|\mathbf{a})$ specifies the non–stationary memoryless distribution over the set $\{0,1\}^n$, which is known to the verifier. At the same time, the channel $V_{\eta,1-\eta}(\mathbf{b}|\mathbf{a})$ determines a stationary memoryless distribution and one can get an arbitrary distribution by changing η. Therefore, the problem is as follows: *discriminate between the known non–stationary memoryless distribution and an arbitrary stationary memoryless distribution over the set* $\{0,1\}^n$ *on the basis of the received vector* $\mathbf{b}$.

1.4. Summary of our contributions and plan of the paper

The problem under considerations is one of the basic problems of cryptography (see, for example [1], [2]) where the vector $\mathbf{a}$ carries information, like “I am a legitimate user (a plane belonging to your air forces)” or “I am user A”, while an attacker tries to substitute a vector that is considered as a noisy version of the vector $\mathbf{a}$. It also appears as synchronization problem when the receiver scans the space to find the arrival time of a certain signal.

We address the problem under the following constraints:

- ⊙ the vector $\mathbf{a}$ is fixed and there are no probability distributions over possible input vectors;
- ⊙ the verifier receives a noisy version of the vector $\mathbf{a}$ in the Acc case, and the channel is fixed;
- ⊙ the only possibility, which is given to a designer of the system, is an assignment of the verification (decoding) algorithm;
- ⊙ the verification algorithm should have the complexity, which is a linear function of the observation length, and it has to provide a certain performance for any vector $\mathbf{a}$ having the same weight;
- ⊙ the attacker knows both the weight of the vector $\mathbf{a}$ and the verification algorithm; his possibilities are restricted by the selection of the probability of bit 1 at the output of a Bernoulli source, which is independently used n times to generate the substitution vector.

An idea, developed in the paper, is an assignment of the component–wise metric depending on the whole vector $\mathbf{a}$. Then the metric associated with the pair of input vectors, which is defined by (6), being an additive extension of the component–wise metric, becomes a memory containing function. We show that such an assignment can be fixed in a way that the

DOI: 10.1002/sec

best strategy of the attacker is a fair coin tossing strategy where the substitution vector is generated by a Bernoulli source with probabilities of bits 0 and 1 equal to $1/2$. Both the false rejection and the false acceptance rates are exponentially decreasing functions of n for all channels satisfying (1).

The competitor of our verification algorithm, which we call the $p_{\mathbf{a}}$-verification scheme, is the ML-verification scheme when the acceptance set is defined as the set of the most likely vectors at the output of the channel, given the input vector $\mathbf{a}$. We already showed that the ML-verification scheme cannot be used for the $V_{\xi,1/2}$ channel. This fact could be explained by the asymmetric structure of the channel, and one can expect that the ML-verification scheme should be recommended for the $V_{\xi,\xi}$ channel where it is reduced to the computation of the Hamming distance between the received vectors and its comparison with the threshold. However, we will show that the performance of such a scheme is essentially worse than the performance of the $p_{\mathbf{a}}$-verification scheme. The comparison of schemes under considerations for the $V_{\xi,\xi}$ channel is given in details, since it clearly demonstrates the points where the schemes differ.

Simple analytic formulas for the false rejection and the false acceptance rates can be derived if the binomial-type probability distributions of the metric is approximated by Gaussian distributions. We use this possibility and show that the accuracy of the numerical results is good enough even when n is rather small.

Finally, we show that the $p_{\mathbf{a}}$-verification scheme for binary memoryless channels is derived from a general construction of the metric for an arbitrary discrete memoryless channels. This construction can be also extended to the channels with continuous output alphabet and to the channels with memory.

The verification problem in different variants was the subject of the most of the cryptography papers, but we do not know publications where the authors seriously considered the verification algorithm in the following sense. The ML-verification scheme can be viewed as an obvious solution to the problem. The point that the scheme can be modified in such a way that the complexity is reduced is clear from communication theory point of view, but we think that the point that it can be modified to restrict the possibilities of the attacker is our contribution. Therefore we only cite [3] where an example of the construction of acceptance sets for asymmetric channels was given and [4] where the fact that the ML-verification scheme has a poor performance even for binary symmetric channels was noticed.

2. The $p_{\mathbf{a}}$-verification scheme

2.1. Introduction of the metric

Let us introduce the component–wise metric as follows.

⊙ *Both* m_0, m_1 *are fixed depending on* $p_{\mathbf{a}}$. *Specifically,*

$$(m_0, m_1) = (p_{\mathbf{a}}, 1 - p_{\mathbf{a}}). \tag{13}$$

Notice that the dependence of m_0, m_1 on $p_{\mathbf{a}}$ means that the value of the metric $m(\mathbf{a}, \mathbf{b})$ also depends on $p_{\mathbf{a}}$, and (6) specifies a memory containing function, which can be presented as

$$m(\mathbf{a}, \mathbf{b}) = \frac{1}{n} \sum_{t=1}^{n} \begin{cases} 0, & \text{if } a_t = b_t, \\ p_{\mathbf{a}}, & \text{if } (a_t, b_t) = (0, 1), \\ 1 - p_{\mathbf{a}}, & \text{if } (a_t, b_t) = (1, 0). \end{cases} \tag{14}$$

The verification scheme where the metric is assigned by (14) will be referred to as the $p_{\mathbf{a}}$-verification scheme.

Two examples of computing the values of the metric are given below. If

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} 00011 \\ 10101 \end{bmatrix},$$

then

$$m(\mathbf{a}, \mathbf{b}) = \frac{2/5 + 0 + 2/5 + 3/5 + 0}{5}.$$

If

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} 00010 \\ 10101 \end{bmatrix},$$

then

$$m(\mathbf{a}, \mathbf{b}) = \frac{1/5 + 0 + 1/5 + 4/5 + 1/5}{5}.$$

Both examples contain the same pairs of bits at the first 4 positions. However, their contributions to the metric are different.

An implementation of the $p_{\mathbf{a}}$-verification scheme is as follows. There are $n+1$ matrices $\mathbf{M}$ and $n+1$ values of the threshold stored in the memory. Having received the vector $\mathbf{a}$, the verifier finds $p_{\mathbf{a}}$, reads the corresponding matrix $\mathbf{M}$ and the threshold $T_{p_{\mathbf{a}}}$, computes $m(\mathbf{a}, \mathbf{b})$ for the vector $\mathbf{b}$ using (14), and checks whether it is less than the threshold or not.

DOI: 10.1002/sec

2.2. The expected values and the variances of the metric

Let us denote

$$\sigma_p^2 = p(1-p)$$

for all $p \in [0,1]$ and notice that σ_p^2 is the variance of the random variable generated by a Bernoulli source having the probability of bit 1 equal to p.

Since the particular vector $\mathbf{b}$ is the value of a random variable B^n, the metric $m(\mathbf{a},\mathbf{b})$ is the value of a random variable $m(\mathbf{a},B^n)$. The expected value and the variance of this random variable are its basic characteristics. The analytic expressions for these characteristics are given below for the distributions that specify the Acc and the Rej cases. We also refer to Figure 3 where an example of the probability density functions of the Gaussian approximations of the probability distributions of the metric are presented; these approximations will be introduced in the following sections.

Proposition 1. *Let* $\mathsf{E}, \mathsf{E}_\eta$ *and* $\mathrm{Var}, \mathrm{Var}_\eta$ *denote the expected values and the variances of the metric* $m(\mathbf{a},\mathbf{b})$ *computed according to the probability distributions* (11), (12)*, respectively. Then*

$$\begin{aligned}&(\mathsf{E}, \mathrm{Var})\\ &= \left((\xi_0+\xi_1)\sigma_{p_\mathbf{a}}^2,\ (p_\mathbf{a}\sigma_{\xi_0}^2 + (1-p_\mathbf{a})\sigma_{\xi_1}^2)\frac{\sigma_{p_\mathbf{a}}^2}{n} \right)\end{aligned}$$

and

$$(\mathsf{E}_\eta, \mathrm{Var}_\eta) = \left(\sigma_{p_\mathbf{a}}^2,\ \sigma_\eta^2 \frac{\sigma_{p_\mathbf{a}}^2}{n} \right).$$

In particular, if $\xi_0 = \xi_1 = \xi$*, then*

$$(\mathsf{E}, \mathrm{Var}) = \left(2\xi\sigma_{p_\mathbf{a}}^2,\ \sigma_\xi^2 \frac{\sigma_{p_\mathbf{a}}^2}{n} \right), \qquad (15)$$

$$(\mathsf{E}_\eta, \mathrm{Var}_\eta) = \left(\sigma_{p_\mathbf{a}}^2,\ \sigma_\eta^2 \frac{\sigma_{p_\mathbf{a}}^2}{n} \right). \qquad (16)$$

Proof: The random variable $m(\mathbf{a},B^n)$ can be expressed as

$$m(\mathbf{a},B^n) = \frac{1}{n}\sum_{t=1}^{n} m(a_t,B_t),$$

where $m(a_1,B_1),\ldots,m(a_n,B_n)$ are independent random variables. Therefore,

$$n\mathsf{E}[\,m(\mathbf{a},B^n)\,] = \sum_{t=1}^{n} \mathsf{E}[\,m(a_t,B_t)\,]$$

and

$$n^2\mathrm{Var}[\,m(\mathbf{a},B^n)\,] = \sum_{t=1}^{n} \mathrm{Var}[\,m(a_t,B_t)\,].$$

Suppose that, for all $t = 1,\ldots,n$, there are $d_t \neq 0$ and $\alpha_t \in [0,1]$ such that

$$\Pr\{m(a_t,B_t) = d\} = \begin{cases} 0, & \text{if } d \notin \{0,d_t\},\\ 1-\alpha_t, & \text{if } d = 0,\\ \alpha_t, & \text{if } d = d_t. \end{cases}$$

Then

$$\mathsf{E}[\,m(a_t,B_t)\,] = d_t\alpha_t$$

and

$$\mathrm{Var}[\,m(a_t,B_t)\,] = d_t^2\alpha_t - (d_t\alpha_t)^2 = d_t^2\sigma_{\alpha_t}^2.$$

In the Acc case, there are

⊙ $n(1-p_\mathbf{a})$ indices t with $(d_t,\alpha_t) = (p_\mathbf{a},\xi_0)$;
⊙ $np_\mathbf{a}$ indices t with $(d_t,\alpha_t) = (1-p_\mathbf{a},\xi_1)$.

Therefore

$$n\mathsf{E} = n(1-p_\mathbf{a})p_\mathbf{a}\xi_0 + np_\mathbf{a}(1-p_\mathbf{a})\xi_1$$

and

$$n^2\mathrm{Var} = n(1-p_\mathbf{a})p_\mathbf{a}^2\sigma_{\xi_0}^2 + np_\mathbf{a}(1-p_\mathbf{a})^2\sigma_{\xi_1}^2.$$

In the Rej case, there are

⊙ $n(1-p_\mathbf{a})$ indices t with $(d_t,\alpha_t) = (p_\mathbf{a},\eta)$;
⊙ $np_\mathbf{a}$ indices t with $(d_t,\alpha_t) = (1-p_\mathbf{a},1-\eta)$.

Therefore

$$n\mathsf{E}_\eta = n(1-p_\mathbf{a})p_\mathbf{a}\eta + np_\mathbf{a}(1-p_\mathbf{a})(1-\eta)$$

and

$$n^2\mathrm{Var}_\eta = n(1-p_\mathbf{a})p_\mathbf{a}^2\sigma_\eta^2 + np_\mathbf{a}(1-p_\mathbf{a})^2\sigma_{1-\eta}^2.$$

Simple algebraic manipulations with the expressions above reduce these expressions to the forms given in the claim.

Discussion. We will assume that $p_\mathbf{a} = p$.

1. *The expected value of the metric in the* Rej *case does not depend on* η.
   If $\eta = 0$, then the attacker presents the all–zero vector, and each position t where $a_t = 1$ gives the contribution to the metric equal to $(1-p)/n$. Since pn is the number of these

DOI: 10.1002/sec

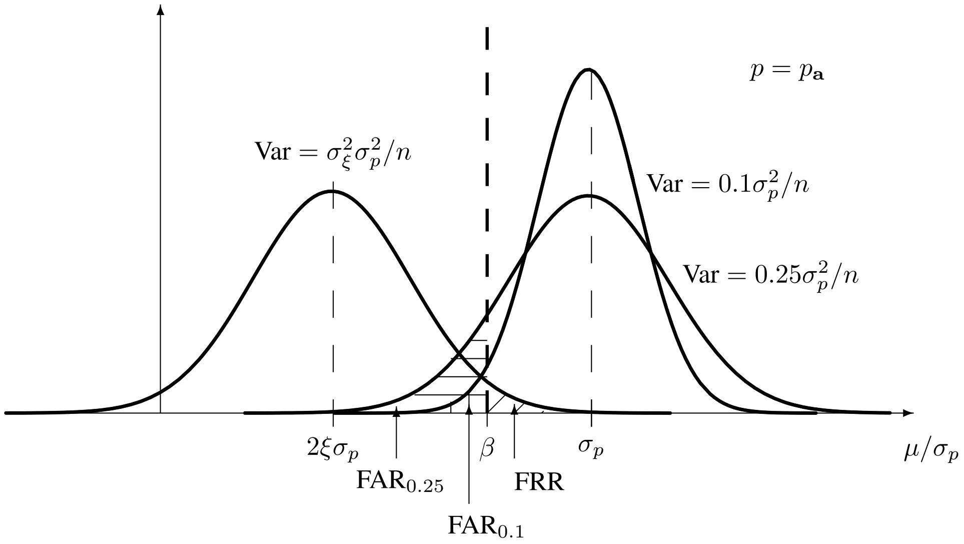

Fig. 3. Example of Gaussian approximations to the probability distributions of the metric for the $p_{\mathbf{a}}$- verification scheme.

positions, the metric is equal to $p(1-p) = \sigma_p^2$. Similarly, if $\eta = 1$, then the attacker presents the all–one vector, and each position t where $a_t = 0$ gives the contribution to the metric equal to p/n. Since $(1-p)n$ is the number of these positions, the metric is also equal to σ_p^2. If $\eta \in (0,1)$, then the attacker randomizes over the vectors that can be presented, and he receives the constant σ_p^2 as the expected value of the metric.

2. *The variance of the metric in the* Rej *case is maximized when $\eta = 1/2$.*
   If the threshold T_p is greater than σ_p^2, then the attacker, who presents either the all–zero or the all–one vector, always passes through the verification with the acceptance decision. If $T_p < \sigma_p^2$, then the attacker has to choose $\eta \in (0,1)$, since he would always get the rejection decision otherwise. At least when the probability distribution of the metric is Gaussian, the maximum deviation from the expected value is attained when the attacker maximizes the variance of the metric. As a result we come to the claim that *the $p_{\mathbf{a}}$-verification scheme has a perfect algorithmic secrecy in the* Rej *case*, meaning the following considerations. For any cryptographic scheme, including our verification scheme, there is a blind attacker, who presents a vector generated by flipping a fair coin with probabilities of 0's and 1's equal to $1/2$. It turns out that if the attacker is informed about the weight of the vector $\mathbf{a}$ and the verification algorithm, then he cannot include this knowledge into the attack and it does not change the fair coin tossing strategy. Notice that the notion of the perfect algorithmic secrecy differs from the conventional notions of perfectness, which are based on the measurement of the conditional entropy of the input to the scheme given the leaking information; in our case, the probability distribution over the input vectors is not given and we cannot use this criterion in each case.
3. *The expected value of the metric in the* Acc *case is less than the expected values of the metric in the* Rej *case.*
4. *If $p \notin \{0,1\}$, then the variances of the metric in all cases our considerations vanish with the length as the function $1/n$.*
   By the previous observation, one can assign a threshold T_p between E and E_η. As the variances of the metric vanish with the length n, by Chebyshev inequality, both the false acceptance and the false rejection rates also vanish. Moreover, by the known upper

bounds on the erf-function, the Gaussian approximations for the probability distributions of the metric lead to the statement that *the false acceptance and the false rejection rates vanish exponentially fast when n increase and we approach the performance of a data processing scheme, which uses the repetition code.* These points will be presented in details in the following sections.

5. *The ratios of the expected value of the metric and the square root of the variance are proportional to $\sigma_p\sqrt{n}$.*
   The false acceptance and the false rejection rates are determined by the exponent of the square of this ratio taken with the negative sign, as it will be shown in the following sections. Hence, the performance essentially depends on p and the smallest probabilities of error are reached when $p = 1/2$.

3. The ML-verification scheme

We will analyze the metric

$$\tilde{m}(\mathbf{a},\mathbf{b}) = \frac{1}{n}\sum_{t=1}^{n}\begin{cases} -\ln(1-\xi_{a_t}), & \text{if } b_t = a_t, \\ -\ln\xi_{a_t}, & \text{if } b_t \neq a_t \end{cases}$$

corresponding to the maximum likelihood decoding for the V_{ξ_0,ξ_1} channel. One can easily check that

$$\tilde{m}(\mathbf{a},\mathbf{b}) = -(1-p_{\mathbf{a}})\ln(1-\xi_0) - p_{\mathbf{a}}\ln(1-\xi_1) + m^*(\mathbf{a},\mathbf{b}), \tag{17}$$

where

$$m^*(\mathbf{a},\mathbf{b}) = \frac{1}{n}\sum_{t=1}^{n}\begin{cases} 0, & \text{if } a_t = b_t, \\ \mu_{\xi_0}, & \text{if } (a_t,b_t) = (0,1), \\ \mu_{\xi_1}, & \text{if } (a_t,b_t) = (1,0), \end{cases} \tag{18}$$

and

$$\mu_{\xi_0} = \ln\frac{1-\xi_0}{\xi_0}, \quad \mu_{\xi_1} = \ln\frac{1-\xi_1}{\xi_1}. \tag{19}$$

The first two terms in the expression at the right-hand side of (17) do not depend on $\mathbf{b}$, and the comparison of the metric $\tilde{m}(\mathbf{a},\mathbf{b})$ with some threshold T^* is equivalent to the comparison of the metric $m^*(\mathbf{a},\mathbf{b})$ with the threshold

$$T^*_{p_{\mathbf{a}}} = T^* + (1-p_{\mathbf{a}})\ln(1-\xi_0) + p_{\mathbf{a}}\ln(1-\xi_1), \tag{20}$$

i.e., the acceptance set is defined as

$$\mathcal{D}^*_{\mathbf{a}} = \left\{ \mathbf{b} \in \{0,1\}^n : m^*(\mathbf{a},\mathbf{b}) < T^*_{p_{\mathbf{a}}} \right\}. \tag{21}$$

Therefore the general assignment of the metric by (6)–(8) includes "the maximum likelihood assignment" of the acceptance sets as a special case when $(m_0,m_1) = (\mu_0,\mu_1)$ and only the value of the threshold depends on $p_{\mathbf{a}}$ (notice that if $\xi_0 = \xi_1$, then the value of the threshold does not depend on $p_{\mathbf{a}}$). The verification scheme, where the metric between the vectors $\mathbf{a}$ and $\mathbf{b}$ is specified by (18), will be referred to as the ML-verification scheme.

We restrict ourselves to the formulas describing the performance of the ML-verification scheme for binary symmetric channels. The proof of the statement below is similar to the proof of Proposition 1.

Proposition 2. *Let* $\mathsf{E}^*, \mathsf{E}^*_\eta$ *and* $\mathsf{Var}^*, \mathsf{Var}^*_\eta$ *denote the expected values and the variances of the metric* $m^*(\mathbf{a},\mathbf{b})$ *computed according to the probability distributions* (11), (12)*, respectively. If* $\xi_0 = \xi_1 = \xi$*, then*

$$(\mathsf{E}^*, \mathsf{Var}^*) = \left(\xi\mu_\xi,\, \sigma_\xi^2\frac{\mu_\xi^2}{n}\right), \tag{22}$$

$$(\mathsf{E}^*_\eta, \mathsf{Var}^*_\eta) = \left((p*\eta)\mu_\xi,\, \sigma_\eta^2\frac{\mu_\xi^2}{n}\right). \tag{23}$$

Like for the $p_{\mathbf{a}}$-verification scheme, the variances Var and Var* vanish when n increases. However, this property holds for all $p \in [0,1]$. The expected value of the metric in the Rej case depends on η. Moreover, if $p*\eta < \xi$, then it is less than the expected value of the metric in the Acc case. If η increases and stays less than $1/2$, then $p*\eta$ also increases, i.e., "the Rej case becomes farther from the Acc case". However, the variance of the metric also increases, and the increase of η can result in a smaller false acceptance rate. This effect is discussed in the next section where we conclude that there is an optimum value of η meaning the maximization of the false acceptance rate for a given p. Therefore, the ML-verification scheme does not have the perfect algorithmic property. Notice also that the ratio of the expected value of the metric and the square root of the variance does not depend on μ_ξ. In other words, the scheme has the same performance for any positive number substituted for μ_ξ. In particular, substitution of the value 1 gives an equivalent scheme where the decision is made on the basis of the Hamming distance between the input vectors to the verifier, and the formulas (22), (23) become

$$(\mathsf{E}^*, \mathsf{Var}^*) = \left(\xi,\, \sigma_\xi^2\frac{1}{n}\right), \tag{24}$$

$$(\mathsf{E}^*_\eta, \mathsf{Var}^*_\eta) = \left(p*\eta,\, \sigma_\eta^2\frac{1}{n}\right). \tag{25}$$

DOI: 10.1002/sec

If $p * \eta < \xi$, then the definition of the acceptance sets by (21) has to be changed in such a way that the acceptance decision is made when $m^*(\mathbf{a},\mathbf{b})$ *is greater than the threshold* $T^*_{p_\mathbf{a}}$. Otherwise, we really obtain the false acceptance rate equal to 1, and one can complain that the scheme should not be called “the ML-scheme” (this situation corresponds to the algorithm when the verifier has to make the acceptance decision when the Hamming distance between the input vectors is large enough). We do not consider this case in the following discussion.

4. Analysis of the verification schemes

4.1. Exact formulas for the false rejection and the false acceptance rates for the $p_\mathbf{a}$-verification scheme

Let the vector $\mathbf{a}$ be fixed in such a way that $p_\mathbf{a} = p$ and let $p = 1/\ell$, where $\ell > 1$ is an integer. We will also assume that n/ℓ is an integer and omit the subscript $\mathbf{a}$ for formal brevity.

If k_0, k_1 are determined by (2), (3) for a given vector $\mathbf{b}$, then

$$n\ell m(\mathbf{a},\mathbf{b}) = k,$$

where

$$k = k_0 + (\ell-1)k_1$$

is an integer. Since $k_0 \le n(1-1/\ell)$ and $k_1 \le n/\ell$, the maximum value of k is equal to

$$k_{\max} = n(1-1/\ell) + (\ell-1)n/\ell = \frac{2n(\ell-1)}{\ell}.$$

By (9), (13),

$$\begin{aligned}\text{FRR} &\\ = &\sum_{k=T}^{k_{\max}} \sum_{k_1=0}^{n/\ell} \text{Bin}_{\xi_0}\Big(k-(\ell-1)k_1 \,\Big|\, \frac{n(\ell-1)}{\ell}\Big)\\ &\cdot \text{Bin}_{\xi_1}\Big(k_1 \,\Big|\, \frac{n}{\ell}\Big),\end{aligned}$$

$$\begin{aligned}\text{FAR}(\eta) &\\ = &\sum_{k=0}^{T-1} \sum_{k_1=0}^{n/\ell} \text{Bin}_{\eta}\Big(k-(\ell-1)k_1 \,\Big|\, \frac{n(\ell-1)}{\ell}\Big)\\ &\cdot \text{Bin}_{1-\eta}\Big(k_1 \,\Big|\, \frac{n}{\ell}\Big).\end{aligned}$$

Table I. Some values of the false rejection and the false acceptance rates attained by the $p_\mathbf{a}$-verification scheme when $n = 129$ and $p_\mathbf{a} = 1/3$.

$(\xi_0,\xi_1) = (1/4,0)$				
T	FRR	FAR$(1/2)$	FAR$(1/3)$	FAR(0)
47	$3.1\cdot10^{-9}$	$2.5\cdot10^{-7}$	$1.4\cdot10^{-7}$	0
44	$1.7\cdot10^{-7}$	$3.3\cdot10^{-8}$	$1.8\cdot10^{-8}$	0
41	$4.6\cdot10^{-6}$	$8.6\cdot10^{-9}$	$4.9\cdot10^{-9}$	0
38	$8.3\cdot10^{-5}$	$6.4\cdot10^{-9}$	$3.7\cdot10^{-9}$	0
35	$1.0\cdot10^{-3}$	$6.2\cdot10^{-9}$	$3.6\cdot10^{-9}$	0
$(\xi_0,\xi_1) = (1/4,1/4)$				
T	FRR	FAR$(1/2)$	FAR$(1/3)$	FAR(0)
76	$5.7\cdot10^{-6}$	$9.6\cdot10^{-2}$	$8.4\cdot10^{-2}$	0
70	$1.5\cdot10^{-4}$	$2.0\cdot10^{-2}$	$1.5\cdot10^{-2}$	0
64	$2.4\cdot10^{-3}$	$2.4\cdot10^{-3}$	$1.7\cdot10^{-3}$	0
58	$2.1\cdot10^{-2}$	$1.7\cdot10^{-4}$	$1.0\cdot10^{-4}$	0
52	$1.1\cdot10^{-1}$	$6.3\cdot10^{-6}$	$3.6\cdot10^{-6}$	0
$(\xi_0,\xi_1) = (1/4,1/2)$				
T	FRR	FAR$(1/2)$	FAR$(1/3)$	FAR(0)
88	$1.4\cdot10^{-3}$	$5.7\cdot10^{-1}$	$5.7\cdot10^{-1}$	1
82	$1.4\cdot10^{-2}$	$2.9\cdot10^{-1}$	$2.7\cdot10^{-1}$	0
76	$7.7\cdot10^{-2}$	$9.6\cdot10^{-2}$	$8.4\cdot10^{-2}$	0
70	$2.6\cdot10^{-1}$	$2.0\cdot10^{-2}$	$1.6\cdot10^{-2}$	0
64	$5.5\cdot10^{-1}$	$2.4\cdot10^{-3}$	$1.7\cdot10^{-3}$	0

The sums above can be easily computed. Some numerical results for $(n,p) = (129,1/3)$ and three channels, $V_{\xi,0}$, $V_{\xi,\xi}$, $V_{\xi,1/2}$, where $\xi = 1/4$, are presented in Table I. If $T < n(\ell-1)/\ell = 129 \cdot 2/3 = 86$, then the false acceptance rate is maximized when $\eta = 1/2$. In this case, the presentation of the all–zero vector, corresponding to $\eta = 0$, gives the false acceptance rate equal to 0. Otherwise, if $T \ge 86$, then $\eta = 0$ results in the false acceptance rate equal to 1. We also present data for $\eta = p = 1/3$ to illustrate the point that it is not the optimum assignment. One can conclude that the $p_\mathbf{a}$-verification scheme has a good performance even when the channel is very noisy.

Let us illustrate the difference between the $p_\mathbf{a}$- and the ML-verification schemes using the notation of this section. Let K_0 and K_1 denote the random variables defined as the number of ones in the prefix of length $n(\ell-1)/\ell$ and in the suffix of length n/ℓ of a randomly chosen vector of length n. Then

$$\text{Met}^* = K_0 + K_1$$

is the random variable defined as the Hamming distance between the vector $\mathbf{a}$ consisting of $n(\ell-1)/\ell$ zeroes followed by n/ℓ ones. Furthermore, denote

$$\text{Met} = K_0 + (\ell-1)K_1. \qquad (26)$$

The values of the random variable Met* are used in the ML-verification scheme and the values of the

DOI: 10.1002/sec

random variable Met are used in the $p_{\mathbf{a}}$-verification scheme. One can easily see that the expected values of the metrics and the variances are given by (15), (16) and (24), (25) when $\sigma^2_{p_{\mathbf{a}}} = (\ell-1)/\ell^2$ up to the normalization factors $n\ell$ for the expected values and $n^2\ell^2$ for the variances. One can also obtain these quantities directly. In particular,

$$\begin{aligned}
\mathsf{E}_\eta[\mathrm{Met}] &= \frac{n(\ell-1)}{\ell}\eta + \frac{n}{\ell}(\ell-1)(1-\eta)\\
&= \frac{n(\ell-1)}{\ell},\\
\mathsf{E}^*_\eta[\mathrm{Met}^*] &= \frac{n(\ell-1)}{\ell}\eta + \frac{n}{\ell}(1-\eta)\\
&= ((\ell-2)\eta+1)\frac{n}{\ell},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}_\eta[\mathrm{Met}] &= \frac{n(\ell-1)}{\ell}\sigma^2_\eta + \frac{n}{\ell}(\ell-1)^2\sigma^2_\eta\\
&= \sigma^2_\eta n(\ell-1),\\
\mathrm{Var}^*_\eta[\mathrm{Met}^*] &= \frac{n(\ell-1)}{\ell}\sigma^2_\eta + \frac{n}{\ell}\sigma^2_\eta\\
&= \sigma^2_\eta n.
\end{aligned}$$

The idea of introducing the metric by (26) can be viewed as the restriction of the attacker's possibilities. In this case, the expected value of the metric does not depend on η, and the attacker can only change the variance of the metric by choosing a certain η. Moreover, the variance is maximized when $\eta = 1/2$, leading to the fair coin tossing strategy and making the attacker more predictable. The influence of η on the performance of the ML-verification scheme, is very high. We will demonstrate this point via comparison of the performances of the schemes for binary symmetric channels and conclude that the best choice of η gives the false acceptance rate for the ML-verification scheme always greater than the false acceptance rate for the $p_{\mathbf{a}}$-verification scheme.

Considerations of this section are oriented to the scaling of the metric in such a way that the result is an integer, which simplifies the calculations. If the vector $\mathbf{a}$ has weight w and $(n-w)/w$ is an integer, then the more general form of (26) is

$$\mathrm{Met} = K_0 + \frac{n-w}{w}K_1.$$

If $(n-w)/w$ is not an integer, then we can always find a $c \le w$ such that

$$\mathrm{Met} = c\Big(K_0 + \frac{n-w}{w}K_1\Big)$$

is the integer. In particular, this is true for $c = w$, but the metric takes $\approx n^2$ values in this case, which can cause some computational difficulties. Nevertheless, calculations with Gaussian approximations of the probability distributions of the metric do not cause any difficulties and lead to short analytic formulas for the false rejection and the false acceptance rates, and the accuracy is high enough even when n is small enough. These points are considered in the next subsection.

We also include the following observation, which exposes the ideas of the approach. If w is small, then an attacker can easier guess the positions where the vector $\mathbf{a}$ contains zeroes than the positions where it contains ones. The number of disagreements at positions where the vector $\mathbf{a}$ contains ones is taken to the metric with the coefficient, which is greater than 1, to make the possibilities of the attacker, who wants to minimize the metric, more uniform. Moreover, an assignment of the coefficient by the ratio $(n-w)/w$ makes them exactly uniform.

4.2. Gaussian approximations

The Gaussian probability density function with the mean $\overline{\mu}$ and the variance σ^2 will be denoted by

$$\mathrm{Gaus}(\mu|\overline{\mu},\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp\Big\{-\frac{(\mu-\overline{\mu})^2}{2\sigma^2}\Big\}.$$

We will also use the erf-function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-z^2}\,dz.$$

The fact that the binomial probability distribution converges to the Gaussian probability density function when the length of observations increases is well–known (see, for example [5]). This fact is widely used in information theory to analyze the asymptotics of decoding schemes that use the ML decision rules [6]. In our case, we write

$$\begin{aligned}
\Pr\{\, m(\mathbf{a},B^n) = \mu \,\} &\to \mathrm{Gaus}(\mu|\mathsf{E},\mathrm{Var}),\\
\Pr\{\, m(\mathbf{a},B^n) = \mu \,|\, \eta \,\} &\to \mathrm{Gaus}(\mu|\mathsf{E}_\eta,\mathrm{Var}_\eta),
\end{aligned}$$

where the distributions at the left–hand sides are defined in (11), (12). We do not discuss the accuracy of the approximations in details because of the following reasons: for fixed parameters, the false rejection and the false accepted rates can be computed exactly, as it was demonstrated earlier; some numerical results on these rates will be given and one can compare the exact values and their approximations; the Gaussian

DOI: 10.1002/sec

Table II. Some values of the false rejection and the false acceptance rates and their Gaussian approximations attained by the $p_{\mathbf{a}}$-verification scheme when $n = 129$ and $p_{\mathbf{a}} = 1/3$.

$(\xi_0, \xi_1) = (1/4, 0)$						
T	FRR	$\mathrm{F\hat{R}R}$	FAR$(1/2)$	$\mathrm{F\hat{A}R}(1/2)$	FAR$(1/3)$	$\mathrm{F\hat{A}R}(1/3)$
47	$3.1 \cdot 10^{-9}$	$1.1 \cdot 10^{-10}$	$2.5 \cdot 10^{-7}$	$6.0 \cdot 10^{-7}$	$1.4 \cdot 10^{-7}$	$1.3 \cdot 10^{-7}$
41	$4.6 \cdot 10^{-6}$	$6.0 \cdot 10^{-6}$	$8.6 \cdot 10^{-9}$	$1.1 \cdot 10^{-9}$	$4.9 \cdot 10^{-9}$	$1.4 \cdot 10^{-9}$
35	$1.0 \cdot 10^{-3}$	$3.9 \cdot 10^{-4}$	$6.2 \cdot 10^{-9}$	$1.1 \cdot 10^{-10}$	$3.6 \cdot 10^{-9}$	$8.2 \cdot 10^{-12}$

$(\xi_0, \xi_1) = (1/4, 1/4)$						
T	FRR	$\mathrm{F\hat{R}R}$	FAR$(1/2)$	$\mathrm{F\hat{A}R}(1/2)$	FAR$(1/3)$	$\mathrm{F\hat{A}R}(1/3)$
76	$5.7 \cdot 10^{-6}$	$1.0 \cdot 10^{-6}$	$9.6 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$	$9.3 \cdot 10^{-2}$
64	$2.4 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$	$3.1 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
52	$1.1 \cdot 10^{-1}$	$9.8 \cdot 10^{-2}$	$6.3 \cdot 10^{-6}$	$1.2 \cdot 10^{-5}$	$3.6 \cdot 10^{-6}$	$3.6 \cdot 10^{-6}$

$(\xi_0, \xi_1) = (1/4, 1/2)$						
T	FRR	$\mathrm{F\hat{R}R}$	FAR$(1/2)$	$\mathrm{F\hat{A}R}(1/2)$	FAR$(1/3)$	$\mathrm{F\hat{A}R}(1/3)$
88	$1.4 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$	$5.7 \cdot 10^{-1}$	$6.0 \cdot 10^{-1}$	$5.7 \cdot 10^{-1}$	$6.0 \cdot 10^{-1}$
76	$7.7 \cdot 10^{-2}$	$6.7 \cdot 10^{-2}$	$9.6 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$	$9.3 \cdot 10^{-2}$
64	$5.5 \cdot 10^{-1}$	$5.3 \cdot 10^{-1}$	$2.4 \cdot 10^{-3}$	$3.1 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$

approximations give simple analytic formulas, and we can arrive at some conclusions that are important to understand the behavior of the verification schemes.

If

$$T_p = \beta\sigma_p$$

is the threshold included into the definition of the acceptance set for the vector $\mathbf{a}$ with $p_{\mathbf{a}} = p$ (see (5)), then the Gaussian approximations of the false rejection and the false acceptance rates for the $p_{\mathbf{a}}$-verification scheme are expressed as

$$\begin{aligned}
\mathrm{F\hat{R}R} &= \int_{\beta\sigma_p}^{+\infty} \mathrm{Gaus}(\mu|\mathsf{E}, \mathsf{Var})\, d\mu \\
&= \frac{1}{2} - \frac{1}{2}\mathrm{erf}\Big(\frac{\beta - (\xi_0 + \xi_1)\sigma_p}{\sqrt{p\sigma_{\xi_0}^2 + (1-p)\sigma_{\xi_1}^2}} \sqrt{n/2} \Big), \\
\mathrm{F\hat{A}R}(\eta) &= \int_{-\infty}^{\beta\sigma_p} \mathrm{Gaus}(\mu|\mathsf{E}_\eta, \mathsf{Var}_\eta)\, d\mu \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\Big(\frac{\beta - \sigma_p}{\sigma_\eta} \sqrt{n/2} \Big).
\end{aligned}$$

If $\beta \in ((\xi_0 + \xi_1)\sigma_p, \sigma_p)$, then the use of the well-known estimates of the values of erf-function [7] gives the convergence

$$\begin{aligned}
\frac{-\ln \mathrm{F\hat{R}R}}{n} &\to \frac{(\beta - (\xi_0 + \xi_1)\sigma_p)^2}{2(p\sigma_{\xi_0}^2 + (1-p)\sigma_{\xi_1}^2)}, \\
\frac{-\ln \mathrm{F\hat{A}R}(\eta)}{n} &\to \frac{(\beta - \sigma_p)^2}{2\sigma_\eta^2}.
\end{aligned}$$

as $n \to \infty$, and we conclude that *both the false rejection and the false acceptance rates tend to* 0 *exponentially fast when the observation length increases.*

The accuracy of the false rejection and the false acceptance rates obtained using the Gaussian approximations is high even when the channel is very noisy and the length of observations is small enough, as it is illustrated by the data in Table II.

Notice that if $\xi_0 = \xi_1 = \xi$, then

$$\mathrm{F\hat{R}R} = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\Big(\frac{\beta - 2\xi\sigma_p}{\sigma_\xi} \sqrt{n/2} \Big). \tag{27}$$

Let us restrict ourselves to the analysis of the ML-verification scheme for $\xi_0 = \xi_1 = \xi$. We write

$$\begin{aligned}
\Pr\{\, m^*(\mathbf{a}, B^n) = \mu \,\} &\to \mathrm{Gaus}(\mu|\mathsf{E}^*, \mathsf{Var}^*), \\
\Pr\{\, m^*(\mathbf{a}, B^n) = \mu \,|\, \eta \,\} &\to \mathrm{Gaus}(\mu|\mathsf{E}_\eta^*, \mathsf{Var}_\eta^*).
\end{aligned}$$

If

$$T^* = \beta^*\mu_\xi$$

is the threshold included into the definition of the acceptance set for the vector $\mathbf{a}$ (see (21) and notice that the value of the threshold does not depend on $p_{\mathbf{a}}$ when $\xi_0 = \xi_1 = \xi$, as it follows from (20)), then the Gaussian approximations of the false rejection and the false acceptance rates for the ML-verification scheme

DOI: 10.1002/sec

are expressed as

$$\begin{aligned} \hat{\mathrm{FRR}}^{*} &= \int_{\beta^{*}\mu_{\xi}}^{+\infty} \mathrm{Gaus}(\mu|\mathrm{E}^{*},\mathrm{Var}^{*})\,d\mu \\ &= \frac{1}{2} - \frac{1}{2}\mathrm{erf}\Big(\frac{\beta^{*}-\xi}{\sigma_{\xi}}\sqrt{n/2}\Big), \qquad (28) \\ \hat{\mathrm{FAR}}^{*}(\eta) &= \int_{-\infty}^{\beta^{*}\mu_{\xi}} \mathrm{Gaus}(\mu|\mathrm{E}^{*}_{\eta},\mathrm{Var}^{*}_{\eta})\,d\mu \\ &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\Big(\frac{\beta^{*}-p*\eta}{\sigma_{\eta}}\sqrt{n/2}\Big) (29) \end{aligned}$$

4.3. Comparison of the performances of the $p_{\mathbf{a}}$- and the ML-verification schemes for binary symmetric channels

Suppose that $\xi_0 = \xi_1 = \xi$ and compare the performances of the $p_{\mathbf{a}}$- and the ML-verification schemes by choosing the value of parameter β^* in such a way that the false rejection rates for the both schemes coincide. Then

$$\frac{\beta^{*}-\xi}{\sigma_{\xi}} = \frac{\beta-2\xi\sigma_{p}}{\sigma_{\xi}}$$

or

$$\beta^{*} = \beta + \xi(1-2\sigma_{p}), \qquad (30)$$

as it follows from (27), (28).

First of all, notice that *the* ML-*verification scheme is better protected against the blind attack*: if $\eta = 1/2$, then $p*\eta = 1/2$ and

$$\begin{aligned} \hat{\mathrm{FAR}}(1/2) &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\Big(\frac{\beta-\sigma_{p}}{1/2}\sqrt{n/2}\Big), \\ \hat{\mathrm{FAR}}^{*}(1/2) &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\Big(\frac{\beta^{*}-1/2}{1/2}\sqrt{n/2}\Big). \end{aligned}$$

However,

$$\begin{aligned} \xi < 1/2 \ &\Rightarrow\ \xi(1-2\sigma_{p}) < 1/2-\sigma_{p} \\ &\Rightarrow\ \beta+\xi(1-2\sigma_{p})-1/2 < \beta-\sigma_{p} \\ &\Rightarrow\ \beta^{*}-1/2 < \beta-\sigma_{p} \\ &\Rightarrow\ \hat{\mathrm{FAR}}^{*}(1/2) < \hat{\mathrm{FAR}}(1/2). \end{aligned}$$

While the blind attack is the best strategy for the $p_{\mathbf{a}}$-verification scheme, this is not true for the ML-verification scheme, and the attacker can essentially increase the false acceptance rate by choosing $\eta = \eta^*$, where

$$\eta^{*} = \arg\max_{\eta\in[0,1]} \frac{\beta^{*}-p*\eta}{\sigma_{\eta}}.$$

In particular, the numerical example given in Table III for $n = 129$ and $p = 1/3$ shows that the false acceptance rate is increased by 5 orders of magnitude. However, *this increase can be attained when the attacker knows the value of p, because η^* is the function of p.*

Notice that if $\beta^* > p$, then $\eta^* = 0$ brings the false acceptance rate equal to 1, i.e., $\hat{\mathrm{FAR}}^{*}(0) = 1$.

Without loss of generality, we can assume that $p \le 1/2$ and continue the analysis for the case when

$$\beta + \xi(1-2\sigma_{p}) < p. \qquad (31)$$

Notice that the sum at the left–hand side is less than $p*\eta$ for all $\eta \in [0,1]$.

One can easily see that $p*\eta = p + (1-2p)\eta$ and that η^* is determined by the equation

$$\frac{\partial}{\partial\eta}\frac{\beta^{*}-p-(1-2p)\eta}{\sqrt{\eta(1-\eta)}}\Big|_{\eta=\eta^{*}} = 0.$$

Thus,

$$-2(1-2p)\eta^{*}(1-\eta^{*}) = (\beta^{*}-p-(1-2p)\eta^{*})(1-2\eta^{*})$$

or

$$\eta^{*} = \frac{p-\beta^{*}}{1-2\beta^{*}}.$$

Furthermore,

$$\begin{aligned} \frac{(\beta^{*}-p*\eta^{*})^{2}}{\sigma^{2}_{\eta^{*}}} &= 4(p-\beta^{*})(1-p-\beta^{*}) \\ &= \frac{(\beta^{*}-\sigma_{p})^{2}}{1/4} - \frac{(1-2\sigma_{p})\beta^{*}}{1/4}. \end{aligned}$$

Let us denote

$$f_{\beta}(\xi) = \frac{(\beta^{*}-p*\eta^{*})^{2}}{\sigma^{2}_{\eta^{*}}} - \frac{(\beta-\sigma_{p})^{2}}{1/4}.$$

Then, by (30),

$$\begin{aligned} \frac{f_{\beta}(\xi)}{4} &= (\beta+\xi(1-2\sigma_{p})-\sigma_{p})^{2} \\ &\quad - (1-2\sigma_{p})(\beta+\xi(1-2\sigma_{p})) \\ &\quad - (\beta-\sigma_{p})^{2}, \end{aligned}$$

which results in the expression

$$\frac{f_{\beta}(\xi)}{4(1-2\sigma_{p})} = (1-2\sigma_{p})\xi^{2} - (1-2\beta)\xi - \beta$$

Since $f_\beta(\xi)$ is a convex down function of ξ, the maximum is attained when $\xi = 0$ or $\xi = \xi^*$, where

$$\xi^{*} = \frac{p-\beta}{1-2\sigma_{p}}$$

DOI: 10.1002/sec

is the maximum possible value of ξ satisfying (31). We write

$$f_\beta(0) = -4(1-2\sigma_p)\beta < 0$$

and

$$\begin{aligned}\frac{f_\beta(\xi^*)}{4(1-2\sigma_p)} &= \frac{(p-\beta)^2}{1-2\sigma_p} - \frac{(p-\beta)(1-2\beta)}{1-2\sigma_p} - \beta \\ &= -\frac{(1-p+\beta)(p-\beta)}{1-2\sigma_p} - \beta < 0,\end{aligned}$$

i.e., $f_\beta(\xi) < 0$ and

$$\frac{(\beta^* - p * \eta^*)^2}{\sigma^2_{\eta^*}} < \frac{(\beta-\sigma_p)^2}{1/4}.$$

Hence,

$$\frac{p*\eta^* - \beta^*}{\sigma_{\eta^*}} < \frac{\sigma_p - \beta}{1/2}$$

or

$$\frac{\beta^* - p*\eta^*}{\sigma_{\eta^*}} > \frac{\beta - \sigma_p}{1/2},$$

which implies

$$\hat{\text{FAR}}^*(\eta^*) > \hat{\text{FAR}}(1/2).$$

We conclude that *for any binary symmetric channel, the $p_{\mathbf{a}}$-verification scheme brings a better performance than the* ML*-verification scheme if the value of the threshold $\beta\sigma_p$ is chosen in such a way that* (31) *holds.*

Examples of the functions $\text{Gaus}(\mu \,|\, (p*\eta)\mu_\xi, \sigma^2_\eta\mu^2_\xi/n)$ are given in Figure 4 to illustrate the point that there is an optimum value of η maximizing the size of the region below the curve when the argument is less than a fixed β^*.

5. Derivation of the $p_{\mathbf{a}}$-verification scheme as a special case of the general construction for arbitrary memoryless channels

One can consider our assignment of the metric for the $p_{\mathbf{a}}$-verification scheme as a kind of a trick oriented to binary channels. We will present our construction for arbitrary discrete memoryless channels and derive the metric for binary channels as a special case.

Suppose that the channel is specified by conditional probabilities

$$(\,V(b|a),\, b \in \mathcal{B}, a \in \mathcal{A}\,),$$

where $\mathcal{A}$ and $\mathcal{B}$ are finite sets. To avoid technical difficulties, we will assume that all conditional probabilities are positive.

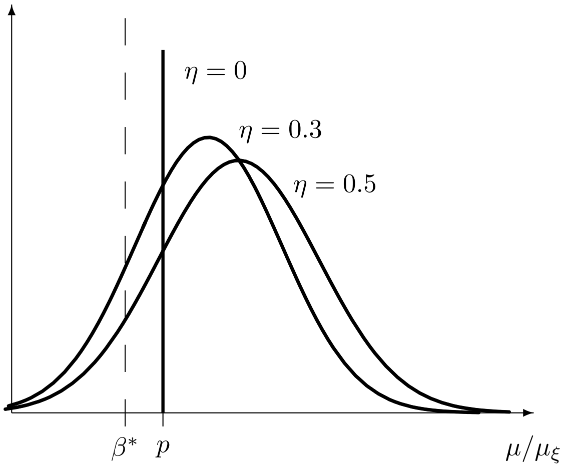

Fig. 4. Examples of the functions $\text{Gaus}(\mu \,|\, (p*\eta)\mu_\xi, \sigma^2_\eta\mu^2_\xi/n)$.

Given a vector $\mathbf{a} \in \mathcal{A}^n$, introduce the empirical probability distribution over the set $\mathcal{A}$ as

$$P_{\mathbf{a}} = (\,P_{\mathbf{a}}(a),\, a \in \mathcal{A}\,),$$

where

$$P_{\mathbf{a}}(a) = \frac{1}{n}\left|\left\{\, t \in \{1,\ldots,n\} : a_t = a \right\}\right|$$

for all $a \in \mathcal{A}$. Furthermore, let

$$Q_{\mathbf{a}}(b) = \sum_a P_{\mathbf{a}}(a)V(b|a).$$

Let us assign a probability distribution

$$(\Phi(b),\, b \in \mathcal{B})$$

in such a way that

$$\sum_a P_{\mathbf{a}}(a) \ln \frac{V(b|a)}{\Phi(b)} = \text{Const}, \quad \text{for all } b \in \mathcal{B}.$$

Then

$$\sum_a P_{\mathbf{a}}(a) \ln V(b|a) - \ln \Phi(b) = \text{Const}$$

and

$$\Phi(b) = \exp\{-\text{Const}\} \cdot \exp\Big\{ \sum_a P_{\mathbf{a}}(a) \ln V(b|a) \Big\}.$$

The condition

$$\sum_{b'} \Phi(b') = 1$$

Table III. The false acceptance rates attained by the $p_{\mathbf{a}}$- and the ML-verification schemes when the false rejection rates for these schemes are equal to each other when $n = 129$, $p_{\mathbf{a}} = 1/3$, and $\xi = 1/4$.

β	β^*	FRR	FAR	$\text{FAR}^*(\eta^*)$	η^*	$\text{FAR}^*(1/2)$
0.236	0.250	$5.0 \cdot 10^{-1}$	$4.3 \cdot 10^{-8}$	$1.2 \cdot 10^{-5}$	0.167	$1.2 \cdot 10^{-10}$
0.271	0.285	$1.8 \cdot 10^{-1}$	$2.7 \cdot 10^{-6}$	$1.1 \cdot 10^{-3}$	0.112	$1.6 \cdot 10^{-8}$
0.306	0.321	$3.2 \cdot 10^{-2}$	$8.9 \cdot 10^{-5}$	$6.7 \cdot 10^{-2}$	0.035	$1.2 \cdot 10^{-6}$
0.342	0.356	$2.7 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	1	0	$4.4 \cdot 10^{-5}$
0.377	0.391	$1.0 \cdot 10^{-4}$	$1.6 \cdot 10^{-2}$	1	0	$9.2 \cdot 10^{-4}$
0.412	0.427	$1.8 \cdot 10^{-6}$	$9.0 \cdot 10^{-2}$	1	0	$1.0 \cdot 10^{-2}$
0.448	0.462	$1.3 \cdot 10^{-8}$	$3.0 \cdot 10^{-1}$	1	0	$6.6 \cdot 10^{-2}$

implies

$$\exp\{\text{Const}\} = \sum_{b'} \exp\Big\{ \sum_a P_{\mathbf{a}}(a) \ln V(b'|a) \Big\}.$$

Hence $\Phi(b) = \Phi_{\mathbf{a}}(b)$ for all $b \in \mathcal{B}$, where

$$\Phi_{\mathbf{a}}(b) = \frac{\exp\{ \sum_a P_{\mathbf{a}}(a) \ln V(b|a)\}}{\sum_{b'} \exp\{ \sum_a P_{\mathbf{a}}(a) \ln V(b'|a)\}}. \qquad (32)$$

A general construction of the verification scheme. *For each vector* $\mathbf{a}$ *specifying the empirical probability distribution* $P_{\mathbf{a}}$ *over the input alphabet, let the component–wise metric be assigned as*

$$\lambda_{\mathbf{a}}(a,b) = -\ln \frac{V(b|a)}{\Phi_{\mathbf{a}}(b)} \qquad (33)$$

for all $(a,b) \in \mathcal{A} \times \mathcal{B}$. *Define the metric as*

$$\Lambda(\mathbf{a},\mathbf{b}) = \frac{1}{n} \sum_{t=1}^{n} \lambda_{\mathbf{a}}(a,b).$$

Introduce a threshold $\Lambda^*_{\mathbf{a}}$ *and define the acceptance set as*

$$\mathcal{D}_{\mathbf{a}} = \Big\{ \mathbf{b} \in \mathcal{B}^n : \Lambda(\mathbf{a},\mathbf{b}) < \Lambda^*_{\mathbf{a}} \Big\}.$$

If $\mathcal{A} = \mathcal{B} = \{0,1\}$, then $P_{\mathbf{a}} = (1 - p_{\mathbf{a}}, p_{\mathbf{a}})$ and

$$V(b|a) = \begin{cases} 1 - \xi_a, & \text{if } b = a, \\ \xi_a, & \text{if } b \neq a. \end{cases}$$

Hence, if $p_{\mathbf{a}} = p$, then

$$\begin{aligned} \Phi_{\mathbf{a}}(0) &= \frac{(1-\xi_0)^{1-p}\xi_1^p}{S_p}, \\ \Phi_{\mathbf{a}}(1) &= \frac{\xi_0^{1-p}(1-\xi_1)^p}{S_p}, \end{aligned}$$

where

$$S_p = (1-\xi_0)^{1-p}\xi_1^p + \xi_0^{1-p}(1-\xi_1)^p.$$

Notice that

$$\ln \frac{\Phi_{\mathbf{a}}(1)}{\Phi_{\mathbf{a}}(0)} = -(1-p)\mu_{\xi_0} + p\mu_{\xi_1}, \qquad (34)$$

where μ_{ξ_0} and μ_{ξ_1} are defined in (19). Furthermore, by (33),

$$\begin{aligned} \lambda_{\mathbf{a}}(0,0) &= -\ln \frac{1-\xi_0}{\Phi_{\mathbf{a}}(0)}, \\ \lambda_{\mathbf{a}}(0,1) &= -\ln \frac{\xi_0}{\Phi_{\mathbf{a}}(1)}, \\ \lambda_{\mathbf{a}}(1,0) &= -\ln \frac{\xi_1}{\Phi_{\mathbf{a}}(0)}, \\ \lambda_{\mathbf{a}}(1,1) &= -\ln \frac{1-\xi_1}{\Phi_{\mathbf{a}}(1)}, \end{aligned}$$

and

$$\begin{aligned} \lambda_{\mathbf{a}}(0,1) - \lambda_{\mathbf{a}}(0,0) &= -\ln \frac{\xi_0}{\Phi_{\mathbf{a}}(1)} + \ln \frac{1-\xi_0}{\Phi_{\mathbf{a}}(0)}, \\ \lambda_{\mathbf{a}}(1,0) - \lambda_{\mathbf{a}}(1,1) &= -\ln \frac{\xi_1}{\Phi_{\mathbf{a}}(0)} + \ln \frac{1-\xi_1}{\Phi_{\mathbf{a}}(1)}. \end{aligned}$$

Using (34), one can easily check that these equalities imply

$$\begin{aligned} \lambda_{\mathbf{a}}(0,1) - \lambda_{\mathbf{a}}(0,0) &= (\mu_{\xi_0} + \mu_{\xi_1})p_{\mathbf{a}}, \\ \lambda_{\mathbf{a}}(1,0) - \lambda_{\mathbf{a}}(1,1) &= (\mu_{\xi_0} + \mu_{\xi_1})(1 - p_{\mathbf{a}}). \end{aligned}$$

Similarly to (17), we write

$$\begin{aligned} \Lambda(\mathbf{a},\mathbf{b}) &= -(1-p_{\mathbf{a}})\ln(1-\xi_0) - p_{\mathbf{a}}\ln(1-\xi_1) \\ &\quad + (\mu_{\xi_0} + \mu_{\xi_1})m(\mathbf{a},\mathbf{b}), \end{aligned}$$

where $m(\mathbf{a}, \mathbf{b})$ is defined in (14). The influence of the constants that do not depend on the vector $\mathbf{b}$ can be included into the threshold in such a way that

$$\begin{aligned} T_{p_{\mathbf{a}}} &= \frac{1}{\mu_{\xi_0} + \mu_{\xi_1}} \\ &\quad \cdot \Big(\Lambda_{\mathbf{a}}^* + (1 - p_{\mathbf{a}}) \ln(1 - \xi_0) + p_{\mathbf{a}} \ln(1 - \xi_1) \Big), \end{aligned}$$

and we derive the $p_{\mathbf{a}}$-verification scheme as a special case of the construction of this section.

6. Conclusion

The performance of the verification algorithm that can be attained by a certain assignment of the metric was surprising to us. In the present correspondence, we focus on the analysis of binary memoryless channels when the substitution vectors are generated by a stationary Bernoulli source. The further extensions are planed for future papers.

References

[1]. W. Diffie and M. Hellman, “New directions in cryptography”, *IEEE Trans. Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[2]. B. Schneier, *Applied Cryptography*. NY: Addison-Wesley, 1996.

[3]. V. B. Balakirsky and A. J. Han Vinck, “Example of using an information theory approach to verification of biometric data”, Int. Workshop *Applications of Information Theory, Coding and Security*, Yerevan, Armenia, April 14–16, pp. 55–58, 2010.

[4]. A. R. Ghazaryan, “Performance of biometric verification for binary data and binary symmetric observation channels”, Int. Workshop *Applications of Information Theory, Coding and Security*, Yerevan, Armenia, April 14–16, pp. 51–54, 2010.

[5]. A. Papoulis, *Probability, Random variables and Stochastic Processes.* NY: McGraw Hill, 1984.

[6]. S. Kullback, *Information Theory and Statistics.* NY: Dover, 1968.

[7]. M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions.* NY: Dover, 1972.

DOI: 10.1002/sec