STATISTICAL ANALYSIS OF A LINEAR ALGEBRA
ASYMMETRIC WATERMARKING SCHEME

G. Boato\textsuperscript{1}, F. G. B. De Natale\textsuperscript{1}, C. Fontanari\textsuperscript{2}, and F. Pérez-González\textsuperscript{3}

\textsuperscript{1}Dept. of Information and Communication Technology, University of Trento, Italy
\textsuperscript{2}Dept. of Mathematics, School of Information Technologies, Politecnico di Torino, Italy
\textsuperscript{3}Dept. of Signal Theory and Communication, University of Vigo, Spain
boato@dit.unitn.it; denatale@ing.unitn.it; claudio.fontanari@polito.it; fperez@tsc.uvigo.es

ABSTRACT
We introduce a novel asymmetric watermarking scheme, involving a private key for embedding and a public key for detection, and we detail its statistical analysis, relying on Neyman-Pearson criterion. The proposed scheme solves part of the problems connected to previous watermarking approaches based on linear algebra. In particular, special attention is paid at reducing the side information required at the detector, as well as at achieving higher robustness by emphasizing the contribution of the watermark in the detection phase.

Index Terms—asymmetric watermarking, statistical analysis, linear algebra

1. INTRODUCTION
Digital watermarking and cryptography represent two different approaches to achieve security in telecommunications. Recently, the scientific community has started to investigate their connections, becoming increasingly aware of the potential value and effectiveness of hybrid solutions. In particular, the analogy with public key cryptography suggests to consider asymmetric watermarking, involving a private key for embedding and a public key for detection (see [1], [2] and [3] for a detailed survey and a critical discussion).

Starting from [4], we are exploiting linear algebra tools for designing asymmetric watermarking schemes. We stress that our approach substantially improves previous ones. Indeed, the eigenvector watermarking scheme introduced in [5] has been defeated by an effective attack (see [6], Section 4.4) and the weakness of the method presented in [7] has been demonstrated in [8]. On the other hand, in the scheme proposed in [9] the watermark cannot be chosen arbitrarily, but it turns out to be heavily dependent on the host image (see in particular statement c) of the Theorem on p. 787, which shows that the watermark is forced to be a suitable multiple of a sequence deterministically extracted from the original image). As a consequence, the method of [9] is appropriate just for copyright protection, where only one key is assigned to each image, but definitely not for fingerprinting, where every recipient is identified by its own key. On the contrary, our approach is suitable also for fingerprinting, allowing the insertion into any image of different watermarking sequences (even sequentially into the same image, see [10]).

In this paper, a substantial step forward is introduced, consisting of two main improvements. First, we reduce the side information required by the detector, and second we enforce the robustness of the method by introducing a new parameter $\beta$ which emphasizes the contribution of the watermark in the detection phase. These two aspects are fundamental for the viability of the proposed method in practical applications. A detailed statistical analysis of the watermarking scheme is provided under the standard assumption (for analytical purposes, see [11]) that the cover image is i.i.d. zero mean Gaussian distributed.

Correspondingly, Section 1 provides a formal description of the method, while Section 2 is devoted to its statistical analysis. Finally, Section 3 collects some numerical results and concluding remarks.

2. WATERMARKING SCHEME
We are going to construct an asymmetric watermarking procedure suitable for digital fingerprinting. Let $V$ be a feature space of dimension $d$ (for instance, the space $\mathbb{R}^d$ corresponding to the entries in the top left corner of the DCT of a digital image), fix an original $\phi \in V$ and let $\{u_1, \ldots, u_n\}$ be an ordered set of users. For each $i = 1, \ldots, n$, the user $u_i$ is associated with a secret signature $s_i \in V$ such that $\{\phi, s_1, \ldots, s_n\}$ is an orthogonal set of vectors.

2.1. Watermark embedding
For each user $u_i$, the following algorithm is implemented:

1) \textbf{Watermark setting}: set $\psi_i := \alpha s_i$ with $0 < \alpha << 1$ in order to meet the usual imperceptibility requirement

2) \textbf{Watermark insertion}: watermark the copy of $\phi$ assigned to user $u_i$ by setting $\phi_i = \phi + \psi_i$.
3) **Definition of auxiliary matrices:** choose a \( d \times d \) orthogonal matrix \( M_i \) and let \( \phi = M_iv_i \) and \( \psi = M_iw_i \) (i.e. \( v_i \), resp. \( w_i \), are the coordinates of \( \phi_i \), resp. \( \psi_i \), in the basis given by the columns of \( M_i \)). Fix a real parameter \( \beta \geq 1 \), which controls the impact of the watermark energy in the detection phase, and let \( b_i := \|\phi_i + \beta w_i\| \) and complete it to an orthonormal basis \((b_1,b_2,\ldots,b_d)\) of \( \mathbb{R}^d \) such that \( <b_1,b_2> = <v_1,w_1> \) (where \( < \ldots > \) denotes linear span). If \( N_i \) is the matrix with \( b_i \) as the \( t \)-th column \((t = 1, \ldots, d)\) and \( \Pi_i = (e_1|e_2| \ldots |e_{i-1}|e_i|e_{i+1} \ldots |e_d) \) is the matrix that permutes the \( i \)-th coordinate vector \( e_i \) and \( e_1 \), then

\[
A_i = \Pi_i \Delta N_i^T \tag{1}
\]

where

\[
\Delta = \begin{pmatrix}
1 & 0 & \ldots \\
0 & S & 0 & \ldots \\
0 & 0 & K & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \ldots & K & 0 & \ldots \\
0 & \ldots & 0 & \ldots & K
\end{pmatrix}
\]

with \( K >> 1 \) an integer appearing \( k \) times and \( 0 < S << K \) a secret real parameter.

4) **Public key releasing:** release to the public the detection key \( D_i = \|v_i + \beta w_i\|A_iM_i^T \). Since \( A_i \) is not invertible, the secrets \( \phi \) and \( \psi \) cannot be reconstructed from \( D_i \).

With this definition of the public key the side information needed by the detector is substantially reduced in comparison with previous approaches [4,10].

2.2. **Watermark detection**

Let now \( \phi_e \) be an extracted feature. The watermark detection is accomplished by the decision function

\[
\delta_i(\phi_e) = \begin{cases} 
1 & \text{if } |\text{sim}(e_i,D_i\phi_e)| \geq T \\
0 & \text{otherwise}
\end{cases} \tag{2}
\]

where \( T = T(\beta) \) is a suitable threshold and

\[
\text{sim}(e_i,D_i\phi_e) = \frac{e_i^T D_i \phi_e}{\| D_i \phi_e \|} \tag{3}
\]

Definitions (2) and (3) for the detector are motivated by the following fact:

**Proposition 1.** For every \( i = 1, \ldots, n \), we have

\[
\begin{align*}
\text{sim}(e_i, D_i \phi_e) &= \frac{\|v_i\|^2 + \|w_i\|^2}{\|D_i \phi_e\|} \\
\text{sim}(e_i, D_i \phi) &= \frac{\|v_i\|^2}{\|D_i \phi\|}
\end{align*}
\]

In particular, if \( \beta >> 0 \) then there exists a decision threshold \( T = T(\beta) \) with \( 0 < T << 1 \) such that \( \delta_i(\phi_e) = 1 \) and \( \delta_i(\phi) = 0 \).

The introduction of parameter \( \beta \), which regulates the impact of the watermark energy at the detector, is therefore crucial for the aim of increasing the robustness of the method.

3. **STATISTICAL ANALYSIS**

For the sake of simplicity we consider the case \( n = 1 \), so that \( A \) is as in (1) with \( \Pi_1 \) the identity matrix. Let now \( \phi_e \) be an image to be tested. We have two alternative hypotheses:

\[
\begin{align*}
H_0 & : \phi_e \text{ does not contain } w \\
H_1 & : \phi_e \text{ contains } w
\end{align*}
\]

The detection aims at defining a test of hypothesis \( H_1 \) versus the alternative \( H_0 \). The Neyman-Pearson detection criterion (see for instance [12]) maximizes the probability of watermark detection \( P_d \) subject to a constraint on the maximum false detection probability \( P_f \). The threshold \( T \) in (2) is calculated in order to satisfy \( P(|\text{sim}(e_1,D_\phi_e)| \geq T|H_0) = P_f \).

In order to apply the Neyman-Pearson criterion it is necessary to know the statistical models of the watermark, the host image and the noise introduced by the attacks.

3.1. **Hypothesis \( H_0 \)**

We are interested in determining the probability of false alarm \( P_f \) where the similarity detector \( |\text{sim}(e_1,D_\phi_e)| \) is used with input \( \phi_e \), a non watermarked image. For mathematical tractability, we assume that vector \( \phi_e \) is i.i.d. zero-mean Gaussian with variance \( \sigma_{\phi_e}^2 \).

Notice that since \( D = \|v + \beta w\|AM^T \), with \( M \) an orthogonal matrix, it follows that \( D_\phi_e \) is Gaussian distributed with covariance matrix \( \|v + \beta w\|^2 \Delta^2 \). Therefore, \( e_i^T D_\phi_e \) is Gaussian with zero mean and variance \( \sigma_{\phi_e}^2 \|v + \beta w\|^2 \). On the other hand, the term \( |D_\phi_e| \) has the same distribution as the random variable \( \|v + \beta w\|((x_1^2 + S^2x_2^2 + K^2\sum_{i=3}^{k+2}x_i^2)^{1/2} \), where \( K \) is the number of elements equal to \( K \) in the diagonal of \( A \) and \( x_i \sim \mathcal{N}(0,\sigma_{\phi_e}^2) \) for every \( i \). Since \( K >> \max(1,s) \), we can consider instead the approximation \( \|v + \beta w\|K(\sum_{i=3}^{k+2}x_i^2)^{1/2} \). Notice that this random variable is independent of \( e_1^T D_\phi_e \). This fact will be very useful in the computation of \( P_f \).

Moreover, we can normalize by \( \|v + \beta w\|\sigma_{\phi_e} \) both the numerator and the denominator of the similarity function to arrive at the following approximately equivalent problem:

\[
P_f = Pr \left\{ \frac{x_1}{K(\sum_{i=3}^{k+2}x_i^2)^{1/2}} > T \right\}
\]

where \( T \) is the threshold and \( x_i \sim \mathcal{N}(0,1) \) for all \( i \). Equivalently, we can write

\[
P_f = Pr \left\{ x_1 > TK \left( \sum_{i=3}^{k+2}x_i^2 \right)^{1/2} \right\}
\]
The term \( R = \sum_{k+2}^{i=3} x_i^2 \) follows a generalized Rayleigh distribution with pdf \( P_R(r) = \frac{r^{k-1} e^{-r/2}}{2 \Gamma(k/2)} \), \( r \geq 0 \), \( k \geq 1 \) where \( \Gamma() \) denotes the Gamma function. Then, in order to compute \( P_f \) we can fix a value of \( x_1 \) and determine the probability that \( R \) is smaller than \( \frac{r}{x_1} \). Finally, we must average the result with respect to the pdf of \( x_1 \). Hence

\[
P_f = \int_{-\infty}^{+\infty} P_r \left( R < \frac{x_1}{TK} \right) f_{x_1}(x_1) dx_1
\]

The cumulative distribution function of \( R \) for the case \( k \) even, i.e. \( k = 2m \), is \( P_{R}(r) = 1 - e^{-r/2} \sum_{n=0}^{m-1} \frac{r^{n+1}}{(2\pi)^{n+1/2}} \). Hence

\[
P_f = 1 - \sum_{n=0}^{m-1} n! \frac{1}{n!} \int_{-\infty}^{+\infty} e^{-r/2} \sum_{n=0}^{m-1} \frac{r^{n+1}}{(2\pi)^{n+1/2}} dx_1
\]

\[
P_f = 1 - \sum_{n=0}^{m-1} \frac{TKT (n + 1/2)}{(2\pi)^{n+1/2}}
\]

\[
P_f = 1 - \sum_{n=0}^{m-1} \frac{TK (2n - 1))}{(2\pi)^{n+1/2}}
\]

where

\[
n!! \begin{cases} n \times (n-2) \ldots 5 \times 3 \times 1 & \text{n odd} \\ n \times (n-2) \ldots 6 \times 4 \times 2 & \text{n even} \\ 1 & \text{n = -1, 0} \end{cases}
\]

Notice that \( P_f \) depends solely on the product \( TK \). This means that for a fixed \( P_f \) we just have to make the product \( TK \) constant (as in the experimental results reported in Fig. 1).

3.2. Hypothesis \( H_1 \)

Let us consider now hypothesis \( H_1 \) and assume that \( \phi_e = \phi_1 + n \), where the components of \( n \) are i.i.d. \( \mathcal{N}(0, \sigma_n^2) \). We have \( e_0^T D\phi_e = \|v\|^2 + \beta \|w\|^2 + n_1' \) where \( n_1' \sim \mathcal{N}(0, \|v + \beta w\|^2 \sigma_n^2) \). On the other hand,

\[
D\phi_e = (\|v + \beta w\|^2 c_1 + n_1', \|v + \beta w\|^2 s c_2 + n_2', \ldots)
\]

where \( e_1^2 + e_2^2 = \|v + \beta w\|^2 \) and \( n_1' \sim \mathcal{N}(0, \|v + \beta w\|^2 \sigma_n^2) \) for all \( i 
\]

By assuming \( K \ll \text{max}(1, s) \), we can write \( \|D\phi_e\|^2 = K^2 \sum_{k+2}^{i=3} n_i'^2 \) and \( e_0^T D\phi_e \) and \( \|D\phi_e\|^2 \) can be regarded as being approximately independent. Now the probability of correct detection \( P_d \) is

\[
P_d = P_r \{ (\|v\|^2 + \beta \|w\|^2 + n_1'^2) > T^2 (K^2 \sum_{i=3}^{k+2} n_i'^2) \}
\]

If we assume \( \|v + \beta w\|^2 \gg n_1' \), then we can make the following simplification:

\[
P_d = P_r \{ (\|v\|^2 + \beta \|w\|^2)^2 + 2 (\|v\|^2 + \beta \|w\|^2) n_1' > (k+2 \sum_{i=3}^{k+2} n_i'^2) \}
\]

If \( n_1'' = \frac{n_1'}{\|v + \beta w\|^2 \sigma_n} \), then \( n_1'' \sim \mathcal{N}(0, 1) \) for all \( i \) and we can write

\[
P_d = P_r \{ \mu_z + \sigma_z n_1'' > (k+2 \sum_{i=3}^{k+2} n_i''^2) \}
\]

3.3. Hypothesis \( H_1 \) with the original as input

Let us calculate now the probability of deciding that \( H_1 \) holds when the detector is given the original. The derivations are almost identical to the previous case with the difference that now \( e_0^T D\phi_e = (\|v\|^2 + n_1' + c_1^2 + c_2^2 = \|v\|^2 \) in (4). The probability of false positive when the detector is given the noisy original \( P_f' \) is

\[
P_f' = P_r \{ (\|v\|^2 + n_1') > T^2 (K^2 \sum_{i=3}^{k+2} n_i'^2) \}
\]

and assuming \( \|v\|^2 \gg n_1' \) we have

\[
P_f' = P_r \{ (\|v\|^2)^2 > T^2 (K^2 \sum_{i=3}^{k+2} n_i'^2) \}
\]

If \( n_1'' = \frac{n_1'}{\|v + \beta w\|^2 \sigma_n} \), then \( n_1'' \sim \mathcal{N}(0, 1) \) for all \( i \) and we can write

\[
P_f' = P_r \{ \mu_z + \sigma_z n_1'' > (k+2 \sum_{i=3}^{k+2} n_i''^2) \}
\]
where $\mu_z = \frac{\|v\|^4}{\|v+\beta w\| \sigma_n^2 K^2}$ and $\sigma_z = \frac{2\|v\|^2}{\|v+\beta w\| \sigma_n^2 K^2}$.

Exactly as above, we finally obtain

$$\mu_z = \frac{(DW R)^2 W N R d}{(DW R + \beta^2) T^2 K^2}, \quad \sigma_z = \frac{2\mu_z}{K T}$$

$$P'_f = \frac{1}{\sqrt{2\pi} \sigma_z} \int_0^\infty F_S(s) e^{-\frac{(s-\mu_z)^2}{2\sigma_z^2}} ds$$

![ROC curves for $DNR = 30dB$](image)

### 4. CONCLUSIONS

Performance of the Newman-Pearson criterion can be evaluated through the ROC (Receiver Operating Characteristic) curves, which plots $P_d$ against $P_f$ and against $P'_f$ (see Fig. 1), and depend on the Document-to-Noise Ratio ($DNR$) $\sigma_\phi^2/\sigma_n^2$.

In both cases $\beta = 5000$ and the product $kd$, which represents the amount of information at the detector side, is kept constant. Hence we see that we can reach an arbitrarily high detection probability and simultaneously a false positive probability bounded under a fixed threshold, thus demonstrating the effectiveness of our method.

Future work will be devoted to its statistical evaluation on more realistic model of the host data.

### 5. REFERENCES


