JANIS: JUST ANOTHER N-ORDER SIDE-INFORMED WATERMARKING SCHEME

T. Furon[*], *B. Macq*,

Université Catholique de Louvain
Dept. TELE
Louvain-la-neuve, Belgium

N. Hurley, G. Silvestre[†].

University College of Dublin
Computer Science Dpt.
Dublin, Ireland

ABSTRACT

This paper deals with some detection issues of watermark signals. We propose an easy way to implement an informed watermarking embedder whatever the detection function. This method shows that a linear detection function is not suitable for side information. This is the reason why we build a family of non-linear functions named JANIS. Used with a side-informed embedder, its performance is much better than the classical spread spectrum method.

1. INTRODUCTION

We model the watermarking problem as follows. From an original content C_o, an extraction function measures N features ordered in a vector $\mathbf{r_o}$. This vector is modified by a mixing function to create a watermarked vector $\mathbf{r_w} = F(\mathbf{r_o}, g\mathbf{w})$. $\mathbf{w}$ is the watermark signal whose variance is set to one and g is the embedding strength. Usually, mixing functions are additive (Eq. (1)) or proportional (Eq. (2)):

$$\mathbf{r_w} = \mathbf{r_o} + g\mathbf{w} \quad (1)$$

$$\mathbf{r_w} = \mathbf{r_o} \star (\mathbf{1} + g\mathbf{w}) \quad (2)$$

where $\star$ is the product component by component. The inverse extraction function completes the embedding stage creating a content C_w whose features vector is $\mathbf{r_w}$.

In this paper, we assume that the components of original vectors are i.i.d. and distributed as $\mathcal{N}(0, \sigma^2_{r_o})$. This represents the reality for some watermarking techniques, but it is also the simplest framework to carry out statistic studies. Especially, it is widely believed that, with the Gaussian assumption, the spread spectrum (SS) method is the optimum scheme where the detector is a correlator. This is proven in section 2.

The goal of the authors is to show that this belief is not true. The rationale is the following one. First, the watermarking channel is not a Shannon channel because the embedder knows what noise will corrupt the transmission: it is the original content modelled by $\mathbf{r_o}$. It is obvious that the embedder should take advantage of this knowledge. Section 3 shows how this can be implemented easily in a watermarking scheme. But it turns out that correlator receivers are not suitable for such a strategy. This is the reason why we build a family of detection functions called JANIS (for 'Just Another N-order side-Inform Scheme'). These receivers are dedicated to fully take benefit of the side information. Theoretical and experimental results show that JANIS is much more efficient than the SS method.

[*]Thanks to Certimark European Project.

[†]The three authors collaborate thanks to Ulysse France-Ireland funds.

2. THE CLASSICAL VIEW

The watermark detector is a device receiving unknown contents whose features vectors are $\mathbf{r_u}$. Some have been watermarked (hypothesis H_1, vectors $\mathbf{r_w}$), others are in their original form (hypothesis H_0, vectors $\mathbf{r_o}$). Its output is a binary variable $\check{d}$ which equals 1 (0) if the content is considered as watermarked (resp. not watermarked). This yields detection errors. P_{fa} is the probability of false alarm that $\check{d} = 1$ when H_0 is true. P_{md} is the probability of a misdetection that $\check{d} = 0$ whereas the content was watermarked (H_1 is true).

2.1. Neyman-Pearson strategy

The first thing to set when building a watermark detector is to choose the right strategy. We list some alternatives.

To minimise the Bayesian risk. Denote

$$c = P_{fa}C_{fa}P(\mathsf{H}_0) + P_{md}C_{md}P(\mathsf{H}_1)$$

the bayesian risk where $\{C_{fa}, C_{md}\}$ are the costs of the detection errors and $\{P(\mathsf{H}_0), P(\mathsf{H}_1)\}$ are a priori hypothesis probability. This strategy is not suitable for watermarking as these a priori probabilities are not known in practice.

To minimise the maximum risk. Denote

$$c^\star = P_{fa}C_{fa}P^\star(\mathsf{H}_0) + P_{md}C_{md}P^\star(\mathsf{H}_1)$$

the maximum risk where $\{P^\star(\mathsf{H}_0), P^\star(\mathsf{H}_1)\}$ are probability of the worse case, i.e. maximising the Bayesian cost.

This strategy is also hard to implement as defining a cost for each error is not easy in practice.

To maximise the detection power while bounding the probability of false alarm above . Denote $P_p = 1 - P_{md}$ the detection power. The Neyman-Pearson strategy is to maximise P_p while $P_{fa} \leq P_{sl}$ (P_{sl} is the significance level). It reflects correctly what watermarkers are doing in practice: e.g. P_{sl} values appear in the calls for proposal of CPTWG[1] and SDMI[2].

According to Neyman-Pearson theorems, the best test is then based on the following sufficient statistic [1]:

$$d_{NP} = \log \frac{p_{R_w}(\mathbf{r_u})}{p_{R_o}(\mathbf{r_u})} \tag{3}$$

In the watermarking framework, the mixing function is usually inversible ($\mathbf{r_o} = F^{-1}(\mathbf{r_w}, g\mathbf{w})$) so that $p_{R_w}(\mathbf{r_u}) = p_{R_o}(F^{-1}(\mathbf{r_w}, g\mathbf{w}))/J$ where J is the determinant of the Jacobian matrix of $F^{-1}(.)$. Finally, the tested statistic is compared to a threshold Thr:

$$\check{d} = \begin{cases} 1 & \text{if } d_{NP} = \log \frac{p_{R_o}(F^{-1}(\mathbf{r_u}, g\mathbf{w}))}{J p_{R_o}(\mathbf{r_u})} > Thr \\ 0 & \text{if } d_{NP} \leq Thr \end{cases} \tag{4}$$

Thr is set so that $E\{d_{NP} > Thr | \mathsf{H}_0\} = P_{fa} \leq P_{sl}$.

2.2. Locally most powerful test

The problem is not correctly modelled since g is not constant in practice. There are contents that perceptually bear a relatively strong embedding distortion, whereas others are extremely sensitive and support only small values of g. The hypothesis are then :

$$\mathsf{H}_0 : g = 0 \quad \text{versus} \quad \mathsf{H}_1 : g > 0 \tag{5}$$

The detection is no longer a simple hypothesis test but a one-sided test. The behaviour the Neyman-Pearson test depends then on the pdf of $\mathbf{r_o}$. As stated in Eq. (4), the value of g, which is, a priori, unknown from the detector, is necessary to measure d_{NP}. For some very few pdf, this is not important as the Neyman-Pearson test is always the best $\forall g > 0$, i.e. it is *uniformly most powerful*. Yet, in the general case, the test is only *locally most powerful*. Small values of g are selected because the embedding strength is small in practice and because contents are hardly distinguishable in that case. Making a Taylor expansion of Eq. (4), a new tested statistic is defined as:

$$d_{LMP} = \left. \frac{\partial}{\partial g} d_{NP}(g) \right|_{g=0} \tag{6}$$

The derivative of the log-likelihood in 0 usually replaces d_{NP} in tests detecting weak signals.

[1] Copy Protection Technical Working Group

[2] Secure Digital Music Initiative

2.3. Examples

Suppose the mixing function is additive (Eq.(1)) and that the features are i.i.d. Then,

$$d_{LMP} = \sum_{i=0}^{N-1} w[i] \left(-\frac{p'_{R_o}(r_u[i])}{p_{R_o}(r_u[i])} \right) \tag{7}$$

Denote $f_{NL}(x) = -p'_{R_o}(x)/p_{R_o}(x)$. The test is a non-linear correlator as sketched in Fig. 1. If r_o is a gaussian r.v., then $f_{NL}(x) = 2x/\sigma^2_{r_o}$ and Eq.(7) is the classical correlation statistic.

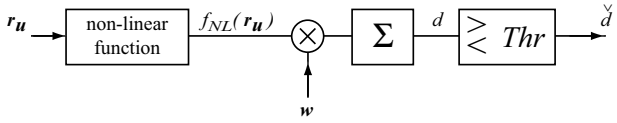

Fig. 1. Structure of a non-linear correlator

Suppose the mixing function is proportional (Eq. (2)) and the features are i.i.d. Then,

$$d_{LMP} = \sum_{i=0}^{N-1} w[i] \left(-r_u[i] \frac{p'_{R_o}(r_u[i])}{p_{R_o}(r_u[i])} \right) - w[i] \tag{8}$$

Assuming $\mathbf{w}$ is centered then the last term can be forgotten. Denote $f_{NL}(x) = -x p'_{R_o}(x)/p_{R_o}(x)$ and retrieve the non-linear correlator of figure 1. If r_o follows a Weibull distribution parameterised by β, then $f_{NL}(x) = x^\beta$. We find the test proposed by J. Oostveen and *al.* [2].

3. A NEW STRATEGY

Somehow, the basic strategy in the watermarking field up to now is to fix the embedding scheme and to try to optimise the detection stage as we have seen in the last section. The basic idea of this article is to do things the other way around: we fix a detection function $d \equiv D(\mathbf{r_u})$ and try to optimise the embedding stage. Actually, $D(.)$ depends of a secret key k, but to simplify notation we did not write it down.

3.1. Side Information

We present an intuitive way to maximise the detectability of the presence of a watermark. The embedding function is assumed to be additive. As the components of the watermark signal are very small, we make a Taylor development to the first order of the detection function:

$$D(\mathbf{r_w}) = D(\mathbf{r_o} + g\mathbf{w}) \sim D(\mathbf{r_o}) + g\mathbf{w}^T \nabla D(\mathbf{r_o}) \tag{9}$$

We design $D(.)$ such that $E\{D(\mathbf{r_o})\} = 0$ as usually done in watermarking. Then, to distinguish easily the hypothesis

H_1 from H_0, the goal is to maximise the value of $D(\mathbf{r_w})$. For this purpose, we create $\mathbf{w}$ as follows:

$$\mathbf{w} = K\nabla D(\mathbf{r_o}) \tag{10}$$

where K is a scalar normalising the power of $\mathbf{w}$ to one.

3.2. Comparison

M. Costa introduced in 1983 a new communication paradigm caracterised by a channel state known at the encoder only [3], which is, in watermarking, the original content component $\mathbf{r_o}$. To maximise the *channel capacity*, the watermark signal $\mathbf{w}$ depends on this state contrary to the independent creation of $\mathbf{w}$ as done classically in Shannon paradigm. This notion was applied to watermarking by B. Chen [4], J. Eggers and *al.* [5]. On the other hand, I. Cox and *al.* also thought about side information to increase the *robustness* of watermarking techniques [6].

Our study is clearly inspired by these previous works. But, we provide a way to optimise (to the first order) whatever the detection function is. Especially, we show that some detection functions give better results than others: e.g., linear functions do not provide any enhancement. Improvements of the article [6] are not so impressive because their detection function is not sufficiently non-linear.

Others advantages stems from the dependence of $\mathbf{w}$ on $\mathbf{r_o}$. For copyright protection applications, it is a good point that the watermark signal is dedicated to one content to avoid the 'copy attack'. For video contents, it is dangerous to always add the same watermark signal as a pirate could estimate it averaging frames. It is also dangerous to add a watermark signal that is evolving in time too fast in low-motion videos as a pirate could erase it with a time low-pass filter. Thanks to its dependence, our watermark signal follows the original content time evolution.

4. JANIS

JANIS is a family of detection functions $\{D_n(.)\}$. The larger the integer n is, the less linear $D_n(.)$ is.

4.1. Correlator

To explain why we create the JANIS scheme, we show that the classical correlator function used in almost all symmetric watermarking techniques is not suitable for side information. The detection function is just a correlation with a secret signal $\mathbf{a}$: $D(\mathbf{r_u}) = {\mathbf{r_u}}^T\mathbf{a}/N$. The gradient is then fixed $\nabla D(\mathbf{r_o}) = \mathbf{a}/N$. It means that the watermark equals the secret signal $\mathbf{a}$ whatever the state of the channel, i.e. no attention is paid to the value of $\mathbf{r_o}$.

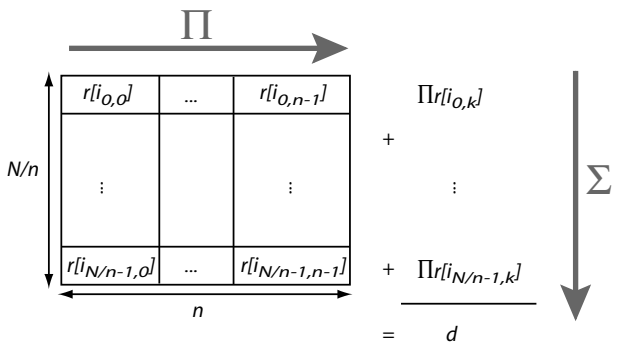

Fig. 2. JANIS detection function

4.2. n-order statistic

Let us describe the $D_n(.)$ function. First, components of $\mathbf{r_u}$ are multiplied by components of a secret vector $\mathbf{a} \in \{-1, 1\}^N$: $\mathbf{r} = \mathbf{r_u} \star \mathbf{a}$. Then its components are ordered according a secret arrangement in a $n \times N/n$ matrix ($n \geq 1$ and we suppose N/n is an integer). Its elements are multiplied each other in a row and then added to yield d. This is sketched figure 2. d is the following polynomial function.

$$d = D_n(\mathbf{r_u}) = \frac{1}{N}\sum_{k=0}^{N/n-1}\prod_{j=0}^{n-1} a[i_{j,k}]r_u[i_{j,k}] \tag{11}$$

As each component appears in only one monomial term composing d, its gradient is easy to calculate:

$$\nabla D(\mathbf{r_o})[i_{l,k}] = \frac{a[i_{l,k}]}{N}\prod_{j=0,j\neq l}^{n-1} a[i_{j,k}]r_o[i_{j,k}] \tag{12}$$

Notice that for $n = 1$, we are back with a classical SS as dealt in subsection 4.1.

4.3. Efficiency

If d is Gaussian distributed, then the power of the test is:

$$P_p = Q\left(\frac{\sigma_{d|\mathsf{H}_0}}{\sigma_{d|\mathsf{H}_1}}Q^{-1}(1 - P_{fa}) - \varepsilon\right) \tag{13}$$

where $Q(.)$ is the cdf of $\mathcal{N}(0, 1)$ and the efficiency $\varepsilon = (\mu_{d|\mathsf{H}_1} - \mu_{d|\mathsf{H}_0})/\sigma_{d|\mathsf{H}_1}$. Larger efficiencies give more powerful tests. Equations (15) and (16) provide the mean and variance of d under both hypothesis. $M(k)$ is k^{th} moment of $\mathcal{N}(0, 1)$ and $G = g^2/\sigma^2_{r_o}$. This leads to the efficiency $\varepsilon[n]$ of $D_n(.)$ which equals to the first order:

$$\varepsilon[n] = \sqrt{nGN} + o(\sqrt{G}) \tag{14}$$

This efficiency, better than SS one by a factor $\sqrt{n}$, has never been reached before in the watermarking field!

$$\mu_{d|\mathsf{H}_0}[n] \;=\; 0 \;\;;\;\; {\sigma_{d|\mathsf{H}_0}}^2[n] = \frac{\sigma_r^{2n}}{nN} \;\;;\;\; \mu_{d|\mathsf{H}_1}[n] = \frac{\sigma_r^n}{n}\sum_{s=0}^{n} \mathrm{C}_n^s \sqrt{G^s} M(s+1)^{n-s} M(s-1)^s \tag{15}$$

$${\sigma_{d|\mathsf{H}_1}}^2[n] \;=\; \frac{\sigma_r^{2n}}{nN}\left(\sum_{s=0}^{n} \mathrm{C}_n^s G^s M(2(s+1))^{n-s} M(2(s-1))^s + (2\sqrt{G})^n M(n)^n\right) - \frac{n}{N}{\mu_{d|\mathsf{H}_1}}^2 \tag{16}$$

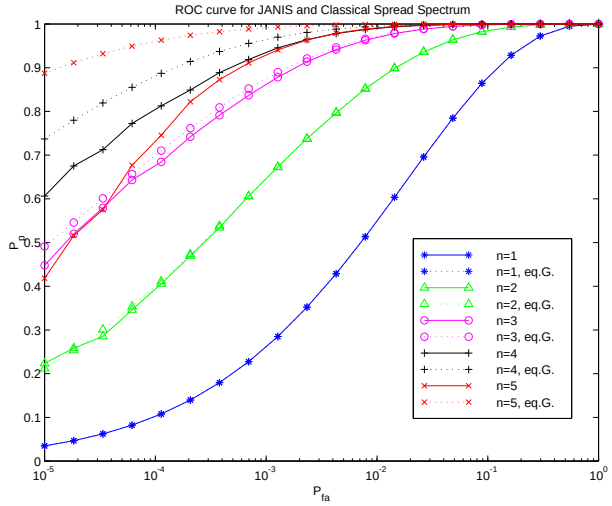

Fig. 3. experimental (solid lines) and theoretical (dotted lines) JANIS ROC curves for $G = -26$dB

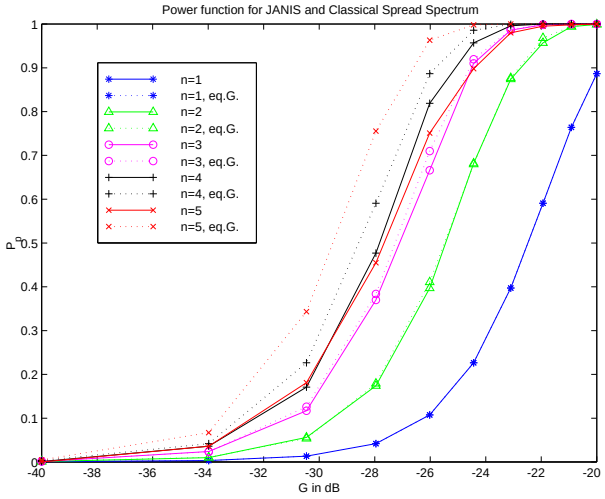

Fig. 4. experimental (solid lines) and theoretical (dotted lines) JANIS power functions for $P_{sl} = 10^{-4}$

4.4. Experimental results

Yet, d is not Gaussian distributed if $n > 1$. Indeed, the smaller n is, the closer to a Gaussian distribution is the pdf of d. We draw the experimental ROC curves $P_p = P_p(n, P_{fa})$ for $1 \leq n \leq 5$ and compare them to the theoretical functions given by Eq. (13). Experiments were done with $N = 2400$ and $g = 0.05$. Figures 3 and 4 show how JANIS function with $n > 1$ are much more efficient than classical SS ($n = 1$). For example, at $P_{sl} = 10^{-4}$ and $G = -26$dB, $P_p[1] = 0.1$ whereas $P_p[4] = 0.75$! In the same way, at $P_{sl} = 10^{-4}$ and $P_p = 0.5$, there is a loss of 6dB between JANIS $n = 4$ and SS! As foreseen, experimental results leave theoretical calculus as n goes larger.

Experimental results on a large images data base will be provided in the final version. The font size will be 9 point.

5. CONCLUSION

The authors try to renew the classical representation of watermarking, proposing a scheme focusing on side information. The JANIS family is extremely promising as it yields very good detection performances but it also has some interesting properties useful for some dedicated applications.

6. REFERENCES

[1] E. Lehmann, *Testing statistical hypothesis*, J. Wiley & Sons, 1986.

[2] Job Oostven, T.Kalker, and J.-P. Linnartz, "Optimal detection of multiplicative watermarks," in *Proc. of the European Signal Processing Conference*, Tampere, Finland, Sept. 2000, EUSIPCO.

[3] M.H.M. Costa, "Writing on dirty paper," *IEEE Trans. on Information theory*, vol. 29, no. 3, May 1983.

[4] B. Chen, *Design and analysis of digital watermarking, information embedding, and data hiding systems*, Ph.D. thesis, Massachusetts Institute of Technology, 2000.

[5] J. Eggers, J. Su, and B. Girod, "Robustness of a blind image watermarking scheme," in *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, Sept. 2000, IEEE.

[6] M. Miller, I. Cox, and J. Bloom, "Informed embedding: exploiting image and detector information during watermark insertion," in *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, Sept. 2000, IEEE.