Introduction aux méthodes de tatouage asymétriques dans le cadre de la protection de copie

Teddy FURON et Pierre DUHAMEL,*†

18 octobre 2001

Résumé

Cet article présente la problématique de la protection de copie des contenus enregistrés sur DVD[1] vidéo. Une analyse montre la valeur à protéger et le niveau de sécurité requis pour cette application. Un système de protection de copie est construit à partir de primitives cryptographiques et d'une technique de tatouage. Une analyse des menaces définit proprement les attaques possibles concernant la fonction de tatouage. Une évaluation de leur complexité définit un niveau de sécurité. Dans le cadre particulier de la protection de copie, le paradigme de tatouage asymétrique fournit un meilleur niveau de sécurité que les techniques classiques à étalement de spectre.

Mots-clé : Protection de copie, tatouage, analyse de menaces, méthodes asymétriques.

Abstract

The concept of asymmetric watermarking schemes is introduced in the framework of copy protection. While the targeted application is detailed, we focus on the role of the watermarking tool in the global copy protection system. A threat analysis underlies the attacks pirates will likely proceed. The estimated complexity of these attacks is the security level that the watermarking technique provides to the the global system. In the copy protection framework, classical spread spectrum techniques are not secure enough. This is the reason why we introduced the concept of asymmetric schemes.

Keywords: Copy protection, watermarking, threat analysis, asymmetric schemes.

*E-mail: teddy.furon@ieee.org, pierre.duhamel@lss.supelec.fr

†Ce travail est financé par THOMSON multimedia.

1. Digital Versatile Disc

1 Introduction

Le tatouage est l'art de cacher un message binaire dans un contenu. Ce domaine scientifique nouveau est né au début des années 90, mais le nombre d'articles sur ce sujet est déjà incroyablement important. Cet intérêt sans cesse grandissant est expliqué par le fait que le tatouage une fonctionnalité désirée dans de nombreuses applications. Mais chacune d'elles impose un cahier des charges spécifique. Celui-ci se détermine généralement en ces termes : non perception, capacité, robustesse, complexité des algorithmes, sécurité. Il est naïf de croire qu'une seule technique de tatouage peut répondre aux différentes spécificités de chaque application. De même, il est vain d'inventer multitudes de techniques de tatouage en ignorant les contraintes du cadre d'utilisation visé. Dans cet article, le cadre applicatif est celui de la protection de copie.

Celle-ci est un sujet différent la protection des droits d'auteur, thème traditionnellement associé au tatouage. Notre cadre applicatif est, en fait, assez mal connu. La première partie de l'article en brosse un présentation plutôt sommaire, mais nécessaire pour bien comprendre les objectifs et contraintes types liés à cette application. Un fait important est que le tatouage ne peut résoudre à lui seul notre problématique. Associé à d'autres primitives cryptographiques, il soutient une architecture globalement sécurisée. Son rôle est une deuxième barrière de défense nécessaire pour protéger les contenus.

La deuxième partie de l'article se concentre sur la fonction de tatouage. De plus, elle n'évoque que des méthodes ou des principes. Elle ne détaille en aucun cas une technique particulière complètement spécifiée. Nous voulons avant tout mettre en avant les problèmes inhérents à la méthode dite à étalement de spectre qui est utilisée dans 70 % des techniques de tatouage. La conclusion est que cette méthode classique de marquage n'est pas appropriée à cette application. En fait, elle procure un niveau de sécurité trop faible. Le paradigme nouveau de la méthode asymétrique décrit à la fin de l'article est plus adapté. Encore une fois, nous insistons sur le fait que ce propos n'est valide que dans le cadre de la protection de copie.

2 La problématique de la protection de copie

Cette section présente les dangers inhérents à la distribution commerciale de contenus et les spécifications pour un système de protection de contenus.

2.1 Distribution commerciale de contenus

Un contenu est une représentation physique d'une œuvre créée par un artiste. Il s'agit d'un morceau de musique, d'une image, d'un film... Codé dans un format numérique, il est représenté par une séquences de mots binaires. Cependant, une œuvre est un concept plus riche englobant des notions

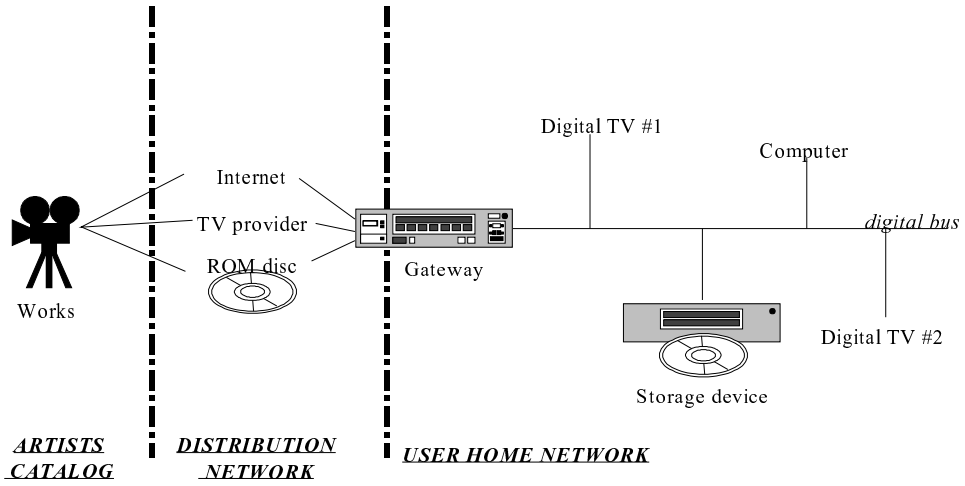

FIG. 1 – *Chaîne de distribution commerciale*

comme un auteur, un titre, ou des ayant droits. Ces notions ne sont plus présentes dans le flux numérique, si bien que le public utilise le contenu sans pour autant les connaître. Le contenu numérique est envoyé à travers une chaîne de distribution à son public : téléchargement de fichiers MP3 via Internet, achat de DVD vidéo, abonnement à des bouquets de chaînes télévisées par décodeur satellite ou câble. La figure 1 illustre ces cas.

La technologie numérique a révolutionné complètement les chaînes de distribution en diminuant de manière radicale les coûts des supports physiques. Cependant, le modèle commercial est resté très classique. Les contenus sont considérés comme des biens à valeur ajoutée. Les chaînes de distribution achètent les droits des œuvres à leur auteur et vendent les contenus aux utilisateurs. La principale menace de ce modèle commercial est la redistribution illégale par des utilisateurs malhonnêtes. Cela prive de leurs revenus non seulement les chaînes de distribution mais aussi les artistes. Le but des systèmes de protection de copie est de palier ce problème. Ils ne défendent que la valeur commerciale des contenus et en aucun cas le nom de l'auteur ou des ayant droits : protection de copie et protection de copyright sont des applications radicalement différentes. Ces systèmes doivent seulement empêcher la création et/ou la distribution des copies illégales (aussi appelées copies pirates). Retrouver l'identité des utilisateurs malhonnêtes n'est pas non plus le but d'un système de protection de copie. Cette fonctionnalité est trop coûteuse car elle implique la saisie de copies illégales, l'identification des pirates et leur poursuite judiciaire. Cette procédure ne peut s'appliquer, en général, que dans la distribution de professionnel à professionnel [6].

Le but du système de protection de copie n'est pas aussi simpliste. Faire des copies de sauvegarde est un droit quand l'utilisateur possède un contenu original, de même qu'enregistrer pour des fins personnelles des émissions té-

lévisées. La tâche du système de protection de copie n'est pas d'empêcher la copie quelqu'elle soit, mais d'assurer le droit légal de copier tout en empêchant la duplication pirate à but lucratif. Ceci est extrêmement difficile : comment distinguer une copie à des fins personnelles d'une copie distribuée illégalement. Dans cet article, nous ne détaillerons qu'une manière grossière de faire cette distinction. Cette approche est partagée par de nombreux comités de travail internationaux (CPTWG [2] et DVD Forum, par exemple). A chaque contenu, un statut de la liste suivante est attribué :

- 'Libre de copier' : pour des contenus sans valeur commerciale tels que les spots publicitaires, les bandes annonces, les films amateurs personnels...
- 'Copiable une fois' : pour les contenus 'originaux' (CD audio, DVD-ROM, émissions télévisées) qu'il est légal de copier pour des fins personnelles.
- 'Non recopiable' : pour les copies de sauvegarde des contenus 'Copiable une fois'. On suppose que la duplication d'une copie de sauvegarde n'est pas un acte légal.
- 'Interdit de copier' : pour les contenus à valeur commerciale trop importante (cas de la télévision à la demande ou 'Pay-TV').

Nous entendons par copie une duplication du contenu dans son intégralité. Le statut des contenus originaux ne change jamais, si bien qu'il est possible de faire plusieurs copies de sauvegarde. Cependant, il est impossible de dupliquer une copie de sauvegarde. Le système de protection de copie a pour but de limiter l'utilisateur à la première génération de copie. Cela est supposé empêcher l'augmentation exponentielle du nombre de copies illégales, et de la rendre linéaire dans le temps. Ce système de statut est clairement imparfait, mais il a le mérite d'être simple.

2.2 Classification des pirates

Habituellement, en sécurité, les pirates sont classés en trois catégories.

- Piraterie des 'particuliers' : ce sont les particuliers qui n'ont aucun savoir en sécurité. Leur budget est très limité ; ils n'achètent que des appareils grand public. Ils copient de façon occasionnelle pour des amis, souvent même, en ignorant leur inégalité.
- Piraterie de garage : ce sont des personnes ayant de bonnes connaissances en électronique et en informatique. Ils partagent leur savoir faire entre groupes. Leur outil privilégié est un ordinateur bien équipé. Ils

2. Copy Protection Technical Working Group

vendent des copies illégales à travers un réseau de connaissances. Leur but est d'au moins rembourser les frais d'équipement.

- Piraterie mafieuse : ce sont des groupes criminels organisés peu nombreux dans le monde. Leur savoir faire est presque illimité. Ils vendent des millions de copies pirates. Cela représente pour eux un investissement rentable offrant peu de risques.

Il est estimé que chaque classe de pirate produit le même nombre de copies illégales. Seule l'application du pouvoir législatif peut lutter contre la piraterie mafieuse. Les barrières techniques d'un système de protection de copie sont inefficaces. La loi est, en revanche, inapplicable pour prévenir la piraterie des 'particuliers'. Il est impossible de vérifier les copies de tous les particuliers. Seul un système de protection de copie enfoui dans les appareil grand public sera viable. Cela restreint considérablement les ambitions d'un tel système : il doit être peu coûteux et maintenir les honnêtes gens honnêtes[3].

2.3 Conformité

Les chaînes de distribution n'ont pas les moyens de pression pour imposer le système de protection sur tous les appareils. Ainsi, ces derniers seront divisés en deux mondes. Le monde 'conforme' est composé des appareils produits sous une certaine licence de protection de copie. Ils comportent toutes les fonctionnalités du système de protection de copie. Le monde 'non conforme' est l'ensemble de tous les autres appareils : non protégés, 'hackés' ou anciens modèles.

Les producteurs de contenus veulent minimiser les interactions entre ces deux mondes. Les contenus à grande valeur commerciale ne sont disponibles que dans le monde 'conforme'. Ces appareils ne doivent pas pouvoir jouer des contenus piratés provenant du monde 'non conforme'. Ainsi, le consommateur a le choix suivant : acheter un appareil conforme pour jouer des contenus protégés et pour copier quand cela est autorisé, ou acheter un appareil non conforme pour jouer et enregistrer des contenus non protégés ou piratés.

2.4 Architectures de systèmes de protection de copie

Nous avons à disposition plusieurs outils pour construire un système protégeant les contenus. Ce sont surtout des fonctions cryptographiques comme le chiffrement, la signature numérique, l'authentification d'appareils, la fonction de hachage. Il serait trop long de les détailler ici, mais prenons conscience qu'avec ceux-ci, il est possible de bâtir des systèmes presque efficaces gérant les différents statuts de la section 2.1. En effet, il n'existe qu'une seule faille : la copie des contenus en clair.

3. 'Keep honest people honest' est le leitmotiv du CPTWG.

L'utilisateur malhonnête est avant tout un utilisateur qui a acquis le droit de déchiffrer les contenus : achat d'un lecteur DVD conforme, abonnement à un système d'accès conditionnel. Ces contenus multimédia protégés sont destinés à être rendus sous forme analogique. Ces signaux électriques seront échantillonnés puis compressés par l'utilisateur malhonnête. Certes, les contenus en clair ainsi piratés ne seront pas d'aussi bonne qualité, mais le but du pirate est de les vendre à une somme très inférieure au prix normal. Il existe aussi des moyens de prélever le flux numérique en clair entre les circuits de déchiffrement et de décodage source (ou convertisseur numérique analogique). Le problème est encore plus facile sur des ordinateurs personnels [47]. Les constructeurs sont obligés de re-chiffrer le contenu transitant entre les cartes de décodage vidéo et les moniteurs numériques. Ceci n'est pas trivial puisque le contenu est alors en bande de base (i.e. sous la forme de pixel), et donc à un très haut débit.

C'est pourquoi un système de protection de copie utilise aussi une technique de tatouage robuste. La robustesse signifie qu'il est possible de détecter le tatouage même si le contenu a subi des transformations (codage source avec pertes, filtrage, D/A + A/D, ajout de bruit sont des exemples). Dans notre application, le tatouage ne sert qu'à empêcher l'attaque décrite ci-dessus. Un appareil conforme refuse ainsi de jouer un contenu tatoué non chiffré. Nous n'avons pas besoin d'enfouir un message binaire dans le contenu. Nous détectons juste la présence ou l'absence d'une marque.

Mais, l'efficacité du tatouage est relative. Les distributeurs de contenus n'ont pas les moyens d'obliger les constructeurs à incorporer dans tous les appareils un détecteur de tatouage. Ainsi, nul ne peut empêcher les appareils non conformes d'accepter des contenus en clair. La stratégie des distributeurs de contenus est d'utiliser un système de chiffrement non seulement comme accès conditionnel, mais aussi, pour ne distribuer la clé de déchiffrement qu'aux constructeurs garantissant la présence d'un détecteur de tatouage dans leurs produits.

Ce dernier problème explique pourquoi le tatouage n'est toujours pas utilisé sur le marché ou dans un standard. Beaucoup ne sont pas convaincus de l'efficacité réelle de cette deuxième barrière de défense. Le CPTWG, par exemple, délibère depuis quatre ans sur un standard de tatouage vidéo sans aucun résultat probant pour l'instant. Un autre débat est qui, des distributeurs de contenus, des constructeurs ou des utilisateurs, va payer le surcoût engendré.

3 Enjeux de cette étude

3.1 Cahier des charges

La section précédente nous aide à établir le cahier des charges pour une technique de tatouage dans le cadre de la protection de copie :

- aucune dégradation perceptive pour respecter la haute qualité des contenus distribués,
- capacité d'un bit. Le détecteur ne recherche qu'une présence ou absence de marque,
- robustesse aux opérations de conversion analogique / numérique et codage source avec pertes,
- faible complexité de l'algorithme de détection,
- sécurité relative.

Le cinquième critère est assez subjectif. Quelle est la différence entre sécurité et robustesse ? La robustesse mesure l'impact de transformations appliquées, de façon intentionnelle ou non, sur le contenu. Ces transformations sont des traitements classiques comme ceux proposés par un logiciel de retouche d'images. On parle parfois d'attaques aveugles au sens où c'est une tentative désespérée de la part de l'attaquant. La sécurité mesure, quant à elle, l'impact d'un traitement purement intentionnel dédié à enlever la marque. On parle aussi d'attaque malicieuse au sens où l'attaquant connaît parfaitement l'algorithme d'incrustation du tatouage.

La subjectivité du dernier critère vient du fait qu'elle est en contradiction avec l'objectif sans ambition de garder les honnêtes gens honnêtes. Autrement dit, il ne doit pas y avoir de failles complètement triviales débouchant sur une attaque facilement reproductible. Pour rendre notre recherche un peu plus passionnante, nous augmentons l'importance de ce critère. Ainsi, nous cherchons quel est le meilleur niveau de sécurité possible dans ce cadre applicatif.

3.2 Principe de Kerckhoffs

Le principe fondateur de la cryptographie a été établi en 1883 par A. Kerckhoffs [31]. Il stipule que l'inventeur d'une technique de chiffrement doit supposer que l'attaquant connaît tout de l'algorithme excepté un paramètre secret. Ainsi, la sécurité du crypto-système doit reposer uniquement sur la mise au secret de cette clé, l'algorithme étant public. Le principe de Kerckhoffs est une heuristique qui se défend par deux arguments :

- Il existe des algorithmes cryptographiques publics qui n'ont pas été cassés, e.g. RSA, DES.

– Il existe des algorithmes propriétaires (i.e. qui violent le principe de Kerckhoffs) qui ont été cassés. Le livre [45] en cite quelques exemples. Le plus connu est sûrement la machine Enigma. En protection de copie, le 'hack' du CSS[4] est une autre illustration [2].

Notre thème de recherche peut donc se résumer en une question : Le principe de Kerckhoffs est-il valide en tatouage pour protection de copie ? Contrairement à la cryptographie, le tatouage est une science trop jeune pour se targuer d'expériences positives : il n'y a pas de techniques de tatouage publiques et sures en protection de copie (à la connaissance des auteurs). De plus, des techniques propriétaires ont été cassées (cf. challenge SDMI[5] [41]).

Les enjeux de cette question pourtant simple sont extrêmement importants. Analysons les deux réponses :

– Le principe de Kerckhoffs n'est pas viable en tatouage. Chacun développera dans son laboratoire une technique de tatouage, dont la sécurité résidera sur le secret de l'algorithme (« Secrecy by Obscurity »). La recherche académique qui en publiant, par définition, rend ses résultats publics, n'a donc plus lieu d'être dans ce domaine. De même, il est risqué de déposer des brevets. Le niveau de sécurité des techniques sera assez faible : comment garder un secret plus de deux ans dans des applications grand public ? Le leitmotiv « Garder les honnêtes gens honnêtes» sera la seule justification possible de l'emploi d'une technique aussi peu sure. La technique adoptée ne sera pas la meilleure mais avant tout la moins complexe produisant ainsi une sécurité aussi éphémère que le secret de son algorithme.

– Le principe de Kerckhoffs est viable. C'est un crédit pour la recherche académique. Les experts en la matière peuvent analyser des techniques publiques pour évaluer leur niveau de sécurité. Leurs résultats seront connus de tous. La technique choisie sera celle produisant le meilleur score selon les critères cités ci-dessus. Ces performances seront garanties car approuvées ou démontrées publiquement.

La communauté tatouage estime plus ou moins que le principe de Kerckhoffs est valable pour des certaines applications . Elle a acquis une certaine expérience lui permettant ou bien de mesurer objectivement un niveau de sécurité (cf. le critère de Cachin en stéganographie [4]) ou bien d'énoncer certains principes pour éviter des attaques malicieuses (cf. la 'dead-lock attack' [42] ou la 'copy attack' [34] en protection de copyright). Le problème reste cependant entier en protection de copie.

4. Content Scrambling System, algorithme de chiffrement des DVD vidéo actuels
5. Secure Digital Music Initiative

4 Technique de tatouage

Le but de cette section est de décrire une modélisation des deux processus de tatouage que sont l'incrustation et la détection. C'est aussi la première occasion d'introduire des notations qui seront maintes fois reprises par la suite. Elles sont en parties issues de l'article [17]. Cette modélisation est surtout le reflet d'une certaine approche du tatouage qui n'est pas partagée par tous.

4.1 Incrustation

L'espace des contenus est noté $\mathcal{C}$ de dimension $\dim(\mathcal{C})$. A partir d'un contenu original C_o de $\mathcal{C}$, l'incrustation fabrique un contenu tatoué C_w de la manière suivante.

Tout d'abord, une fonction d'extraction $X(.)$ crée un vecteur $\mathbf{r_o}$ de taille N : $\mathbf{r_o} = X(C_o)$. L'espace des vecteurs extraits est noté $\mathcal{W}$ et appelé l'espace de tatouage. Il est en général isomorphe à $\mathbb{R}^N$. La fonction d'extraction sélectionnera, par exemple, certains coefficients d'une transformée linéaire du contenu et les classera dans un vecteur. La section 4.5 propose plusieurs critères de choix de la fonction d'extraction.

Puis, une fonction de multiplexage $F(.,.)$ mélange le vecteur extrait et un signal de tatouage $\mathbf{w}$. Le résultat est un vecteur tatoué $\mathbf{r_w} = F(\mathbf{r_o}, \mathbf{w})$.On note $D(\mathbf{r_w}, \mathbf{w})$ la vraisemblance que le signal de tatouage soit présent dans le vecteur extrait. La fonction de multiplexage ne modifie cependant que légèrement le vecteur extrait pour des raisons perceptives. Cette contrainte est modélisée par une distance $\|\mathbf{r_w} - \mathbf{r_o}\|_{\mathcal{W}} < r_{\mathcal{W}}$. Ainsi, le but de la fonction de multiplexage est de trouver le meilleur vecteur tatoué $\mathbf{r_w}$:

$$\mathbf{r_w} = \arg \max_{\|\mathbf{r}-\mathbf{r_o}\|_{\mathcal{W}} < r_{\mathcal{W}}} D(\mathbf{r}, \mathbf{w}) \tag{1}$$

Cette méthode décrite dans [17, 36] est rarement utilisée. La maximisation d'un critère sous contrainte est trop complexe. En pratique, la fonction de multiplexage est clairement sous-optimale. Elle se réduit à un multiplexage basique :

$$\mathbf{r_w} = \mathbf{r_o} + g\mathbf{w} \tag{2}$$

où g est un scalaire réglant la puissance du signal de tatouage.

La fonction d'extraction $X(.)$ ne peut clairement pas être inversée puisque $\dim(\mathcal{C}) > N$. La fonction d'extraction inverse $Y(.,.)$ modifie le contenu C_o à l'aide du vecteur tatoué. Le contenu obtenu $C_w = Y(\mathbf{r_w}, C_o)$ est perceptiblement identique à C_o tout en ayant $\mathbf{r_w}$ pour vecteur extrait.

4.2 Détection

Le processus de détection reçoit le contenu C_u. Sa fonction est de décider si ce contenu est protégé, i.e. s'il a été tatoué. Le résultat binaire de la détec-

tion est donc un booléen $\check{d}(C_u)$: $\check{d}(C_u) = 1$ si le contenu est déclaré protégé, $\check{d}(C_u) = 0$ si le contenu est déclaré non protégé. L'algorithme employé $\check{d}(.)$ est appelé règle de décision.

La détection est envisagée en trois étapes. En premier lieu, le vecteur extrait est calculé comme à l'incrustation : $\mathbf{r_u} = X(C_u)$. Il est ensuite comparé au vecteur de tatouage. On nomme résultat scalaire du détecteur $d = D(\mathbf{r_u}, \mathbf{w})$. Finalement, le résultat binaire est la comparaison du résultat scalaire à un seuil positif Thr : $\check{d}(C_u) = (d > Thr)$.

Cela revient à partitionner l'espace de tatouage $\mathcal{W}$ en deux régions. La région critique $\mathcal{R}(\mathbf{w})$ est la région des vecteurs extraits considérés comme protégés. Sa définition est la suivante :

$$\mathcal{R}(\mathbf{w}) = \{\mathbf{r} \in \mathcal{W} | D(\mathbf{r}, \mathbf{w}) > Thr\} \tag{3}$$

La région correspondant aux vecteurs non protégés est $\mathcal{W} \setminus \mathcal{R}$.

4.3 Probabilités de fausse détection

Le détecteur ne sait pas si le contenu reçu a été tatoué par l'incrustation. Il y a donc deux hypothèses exclusives :

- H_1 : Le contenu reçu a été tatoué.
- H_0 : Le contenu reçu n'a pas été tatoué.

Il y a deux probabilités de fausse détection. Premièrement, la probabilité de fausse alarme P_{fa} est la probabilité que le détecteur considère un contenu non tatoué comme protégé : $P_{fa} = \mathbf{Prob}(\mathbf{r_u} \in \mathcal{R}(\mathbf{w}) | \mathsf{H}_0)$. Deuxièmement, la probabilité de raté P_{md} est la probabilité que le détecteur ne considère pas un contenu tatoué comme protégé : $P_{md} = \mathbf{Prob}(\mathbf{r_u} \notin \mathcal{R}(\mathbf{w}) | \mathsf{H}_1)$. On préfère parler de la puissance de la détection $P_p = 1 - P_{md}$.

4.4 Commentaires

Cette modélisation a l'air anodin, mais en fait, elle reflète une approche 'télécommunication' du tatouage. Le tatouage est une *superposition* d'un signal à un contenu : le signal de tatouage est transmis de l'incrustation vers le détecteur à travers un canal représenté par le contenu. Le vecteur extrait du contenu original est un bruit qui pollue cette communication. C'est un canal souvent modélisé comme un canal à bruit additif blanc Gaussien. La détection suit l'architecture des récepteurs de télécommunications. Une première partie démodule le signal reçu (fonction d'extraction et calcul du résultat scalaire). La seconde partie est la prise de décision quant au symbole émis.

Avec cette approche 'télécommunication', il n'est pas surprenant de constater que dans 70 % des techniques, la vraisemblance $D(.,.)$ est basée sur une corrélation.

$$D(\mathbf{r_u}, \mathbf{w}) = \frac{\mathbf{r_u}^T \mathbf{w}}{\|\mathbf{r_u}\| \|\mathbf{w}\|} \tag{4}$$

On retrouve le très classique récepteur à filtre adapté (ou récepteur par corrélation) [23]. De même, le vecteur $\mathbf{w}$ est une séquence pseudo-aléatoire, comme dans les systèmes de télécommunication numérique à étalement de spectre DSSS[6] : l'information à transmettre, ici un bit indiquant la présence de la marque, a été codée sur les N composantes du vecteur $\mathbf{r_w}$. Cependant, à la différence de communications CDMA[7], le vecteur $\mathbf{w}$ ne sera pas une séquence pseudo-aléatoire classique comme une m-sequence ou une séquence de Gold [23]. Il constituera la clé secrète de la technique. N'oublions pas que l'étalement de spectre est une technique inventée pendant la seconde guerre mondiale[8] qui n'a été déclassée pour application commerciale qu'en 1985. Dans le cadre des communications militaires, le fait que la porteuse $\mathbf{w}$ reste un secret confère les avantages suivants :

- cacher la présence du signal et la localisation de l'émetteur. On parle de faible probabilité d'interception. La comparaison avec les contraintes perceptives en tatouage est évidente.
- combattre les interférences dues au brouillage intentionnel, à la présence d'autres communications ou à la sélectivité du canal (e.g. à cause des trajets multiples). La comparaison avec les contraintes de robustesse est aussi évidente.
- atteindre un certain niveau de sécurité puisque la porteuse n'est pas publique. Les récepteurs ne la connaissant pas ne peuvent pas décoder, émettre ou modifier le message, ce qui amène un certain niveau de sécurité.

Cependant, une confusion demeure quant à la sécurité apportée par l'étalement de spectre. Comme l'écrit T. Kalker [30], l'étalement de spectre apporte une solution pour transmettre un message à cacher dans un contenu : il s'agit de la couche physique d'une transmission. Cependant, la sécurité n'est absolument pas acquise par l'étalement de spectre (même s'il est un principe de base en télécommunications militaires).

4.5 Choix du canal

Le choix technique le plus critique est celui de la fonction d'extraction. Autrement dit, il s'agit de choisir sur quel canal le signal de tatouage sera

6. Direct Sequence Spread Spectrum
7. Code Division Multiple Access
8. système SIGSALY [28]

émis. Ce thème est le sujet de nombreux articles. Nous détaillons par la suite quelques critères souvent évoqués.

Principe de Cox Le premier article de référence est celui de I. Cox et *al.* [16] où il est stipulé que la fonction d'extraction doit sélectionner les caractéristiques du contenu les plus importantes d'un point de vue perceptif. Ces caractéristiques étant extrêmement sensibles, on utilisera une technique à large étalement de spectre pour modifier de façon infime chaque composante du vecteur extrait. Cependant, grâce à leur importance perceptive, elles sont aussi relativement invariantes aux attaques du pirate tant qu'une certaine qualité de contenu est préservée. La technique de tatouage sera robuste parce que la fonction d'extraction est robuste. Cette stratégie est à notre avis trop utopique pour être mise en œuvre.

Faible compléxité C'est parfois une contrainte de faible complexité qui motive le dessin de la fonction $X(.)$. Cette stratégie ne cherche pas à enfouir un tatouage d'énergie maximale. Elle utilise un modèle de masquage simple garantissant l'imperceptibilité du tatouage. L'énergie de ce dernier est faible mais suffisante pour l'application visée.

C'est le cas de la technique baptisée 'JAWS' (Just Another Watermarking System) de protection de copie vidéo [35]. L'espace d'insertion $\mathcal{W}$ est le domaine $\mathcal{C}$. Pour construire une technique la plus universelle possible, i.e. qui ne soit pas attaché à un format d'image précis, les auteurs ont choisi de tatouer directement la luminance des pixels.

Avec le même soucis de faible complexité, certains ont préféré, au contraire, tatouer (détecter) lors du codage (resp. décodage) source. Ainsi la technique de tatouage est dépendante d'un codec. Elle partage avec lui certaines ressources. La fonction d'extraction sera de fait la transformée utilisée dans le codec. On ne compte plus les techniques de tatouage basée sur la DCT[9] 8x8 (cf. l'état de l'art de G. Langelaar et *al.* [12]) ou les transformées en ondelettes utilisées dans JPEG2000 [38] ou MPEG4 [3]. Alors que la théorie du codage source fournit des arguments pour choisir cette transformée (e.g., maximiser le gain de codage), rien nous montre qu'elle soit adaptée pour jouer aussi le rôle de la fonction d'extraction. Les inventeurs de 'JAWS' nous rappellent même que si la tranformée est linéaire, alors incruster et détecter un tatouage peut se faire dans le domaine $\mathcal{C}$ ou $\mathcal{W}$ indifféremment [27] (la corrélation et la norme Euclidienne dans $\mathcal{C}$ s'écriront comme une corrélation et une norme Euclidienne dans $\mathcal{W}$).

Cependant, si on juge que les attaques du pirate ne seront que des compressions à faible débit, il est intéressant de choisir le même domaine 'transformée' pour modéliser ces attaques. R. Wolfgang et *al.* tatouent des images

9. Discrete Cosine Transform

fixes dans le domaine DCT ou DWT[10] et les attaquent par compression JPEG ou SPIHT [40]. Leur conclusion est qu'il est plus robuste d'incruster le signal de tatouage dans le domaine où le contenu sera quantifié. J. Eggers modélise l'incrustation du signal de tatouage suivie d'une quantification due à la compression du contenu tatoué par une quantification avec 'dithering' [20]. Cette modélisation lui permet d'atteindre de bonnes performances en robustesse contre une compression JPEG.

Robustesse Le principe de Cox tire sa robustesse de la fonction d'extraction. Or, certaines attaques sont des transformations géométriques. Elles ne diminuent pas la puissance du tatouage mais déplacent celui-ci à des endroits où le détecteur ne le cherchera pas. Ce sont des attaques par désynchronisation. Une des réponses de la communauté tatouage est de prendre des fonctions d'extraction invariantes : les modules de la DFT[11] sont invariants par translations cycliques [33], les modules de la transformée de Fourier-Mellin sont invariants par rotation, ré-échantillonnage uniforme et translation [21, 5]. Or, ces coefficients ne sont pas des 'descripteurs perceptiblement importants', comme le principe de Cox le réclame. La thèse de S. Derrode explique très bien que l'information principale se trouve dans la phase de ces coefficients et non dans leur module [43]. Rares sont ceux d'ailleurs qui utilisent leurs phases en tatouage [22, 15]. Ces fonctions d'extraction invariantes sont en complète contradiction avec le principe de Cox. Le plus surprenant est qu'elles résistent bien à des attaques par compressions aux transformées à fort gain de codage. M. Ramkumar, dans le troisième chapitre de sa thèse [37], explique que ces fonctions d'extraction invariantes sont un 'trou', i.e. une zone non exploitée, des algorithmes de compression standard.

Détectabilité La fonction d'extraction peut être choisie afin de faciliter la tâche du détecteur. Il est préférable que les composantes des vecteurs extraits soient, par exemple, indépendantes. Les performances du détecteur sont relativement meilleures et le calcul de la vraisemblance que le contenu soit tatoué est indéniablement plus facile. Les arguments sont ici purement statistiques et sont connus de la théorie des tests d'hypothèses.

5 Failles de sécurité

Avec le principe de l'étalement de spectre couplé à une technique d'incrustation par transformée, on peut couvrir bon nombre des articles proposées, y compris pour la protection de copie comme [35, 13, 18]. Cependant, dans ce cadre, cette méthode n'est pas sûre.

10. Discrete Wavelet Transform
11. Discrete Fourier Transform

5.1 Attaques sur contenus

Dans nombre d'applications, les clés secrètes et les messages à cacher changent régulièrement car ils sont fonctions de paramètres comme un nom d'auteur, une date ou un résumé cryptographique du contenu. Par conséquent, il existe peu de contenus tatoués à l'identique. Or, dans notre cas, nous nous trouvons dans la situation opposée. Tous les contenus sont tatoués de la même façon. Autrement dit, il existe une quantité quasi infinie de vecteurs extraits tatoués $\mathbf{r_w}$ contenant le même signal de tatouage $\mathbf{w}$, considéré comme la clé secrète. Ce cas extrême où il n'y a qu'une seule clé et qu'un seul message à cacher est typique de la protection de copie.

Etant donné la fonction de mixage de l'Eq. (2), l'attaquant n'a plus qu'à moyenner un grand nombre de vecteurs extraits pour estimer la clé secrète.

$$\widehat{\mathbf{w}} = \frac{1}{T}\sum_{i=0}^{T-1} \mathbf{r_{w,i}} = E\{\mathbf{r_o}\} + g\mathbf{w} \tag{5}$$

L'espérance $E\{\mathbf{r_o}\}$ est une donnée publique estimable par tous. Une fois le signal de tatouage estimé, l'attaquant peut facilement modifier $\mathbf{r_w}$ pour construire un contenu piraté C_p. Par exemple, une simple orthogonalisation de Graham Schmidt suffit :

$$\mathbf{r_p} = X(C_p) = \mathbf{r_w} - \frac{\mathbf{r_w}^T.\widehat{\mathbf{w}}}{||\widehat{\mathbf{w}}||}\widehat{\mathbf{w}} \tag{6}$$

Comme, par hypothèse, la qualité du contenu n'a pas trop souffert lors de l'incrustation du tatouage, cette attaque ne la dégrade que très peu. En pratique, elle produit des contenus piratés de meilleure qualité qu'une attaque aveugle par compression.

Nous avons expliqué cette technique en nous appuyant sur des modèles très simplifiés de tatouage. En réalité, les choses sont plus complexes notamment à cause de modèle de perception humaine utilisé lors de l'incrustation du tatouage. La fonction de mixage est donc plus complexe, mais elle reste suffisamment linéaire pour que l'attaque décrite soit une menace réaliste. Il suffit que l'estimée $\widehat{\mathbf{w}}$ soit suffisamment colinéaire à $\mathbf{w}$ pour que $D(\mathbf{r_p}, \mathbf{w}) < Thr$. C'est ainsi que S. Craver et J. Stern ont cassé l'une des techniques de tatouage proposées par le SDMI [41].

Une autre menace est une attaque à paire de contenu original/tatoué. S'il vient à l'idée des distributeurs de contenus de protéger des œuvres déjà disponibles sur le marché, l'attaquant n'aura plus qu'à faire la différence des vecteurs extraits pour estimer le signal de tatouage.

Cette analyse nous prouve que le signal de tatouage $\mathbf{w}$ ne peut pas jouer le rôle d'une clé secrète car il est facilement estimable dans le cadre de la protection de copie. Dans l'optique de l'étalement de spectre, la solution à ce type d'attaque est bien connue des télécommunications militaires ou des

cryptographes : il faut utiliser une clé $\mathbf{w}$ aussi longue que possible [28] ou, en théorie suivant le théorème cryptographique de Shannon [44], utiliser une clé aussi longue que le contenu à tatouer. Ainsi, l'estimation de l'Eq. (5) n'est plus possible. Si l'attaquant a réussi à estimer le signal de tatouage sur un contenu particulier, il ne possédera qu'une partie de la clé secrète, ce qui ne remet pas en cause la sécurité de tout le système. Malheureusement, ceci est impossible avec la technique de tatouage à étalement de spectre : il n'y a aucun moyen de synchroniser le processus d'incrustation et celui de détection dans les appareils des utilisateurs ; de plus, cela demande une mémoire infinie pour stocker une clé secrète aussi longue.

5.2 Attaque sur détecteur

Des attaques plus sophistiquées sont aussi possibles, notamment si la détection est basée sur un processus linéaire. Comme chaque appareil conforme possède un détecteur sous la forme d'une boîte noire scellée, le pirate peut questionner ce dernier comme un oracle autant de fois que nécessaire. En étudiant les réponses à des entrées choisies, l'attaquant essaie d'estimer la clé secrète $\mathbf{w}$.

Supposons que le résultat de la détection soit une corrélation de $\mathbf{r_w}$ avec $\mathbf{w}$ comme dans l'Eq. (4), c'est donc une projection de $\mathbf{r_w}$ sur la direction indiquée par $\mathbf{w}$. En testant N fois le détecteur avec des vecteurs de la base canonique $\{\mathbf{e_k}\}$ de l'espace de 'tatouage', le pirate retrouve parfaitement le secret :

$$\widehat{\mathbf{w}} = \sum_{k=0}^{N-1} D(\mathbf{e_k}, \mathbf{w})\mathbf{e_k} \tag{7}$$

Cette attaque n'est pas réalisable en pratique puisque le pirate n'a pas accès à la corrélation mais au résultat de sa comparaison avec le seuil Thr. Cependant, T. Kalker a montré dans l'article [29] que ce résultat binaire donne assez d'information pour estimer $\mathbf{w}$ en $O(N)$ détections.

6 Méthodes asymétriques

Jusqu'en 1999, toutes les techniques de tatouage étaient symétriques (aussi appelées 'à clé privée'). La symétrie signifie que le processus de détection utilise les mêmes paramètres secrets que le processus d'incrustation. Ceci a été illustré au chapitre précédent où le signal de tatouage $\mathbf{w}$ était sensé jouer le rôle de clé secrète. Nous avons vu que la connaissance de cette clé aide grandement l'adversaire à pirater des contenus.

Pour palier ce problème, nous proposons la notion d'asymétrie. L'idée est d'inventer un schéma de tatouage où le processus de détection ne vérifie pas si un signal de référence est caché dans un contenu. Mais, il vérifie si le contenu à une propriété statistique spéciale due à la présence du signal

de tatouage. Cette propriété ne peut être attendue naturellement des contenus non tatoués. Une quantité plus ou moins grande de signaux de tatouage donnent au contenu une telle propriété, si bien qu'une recherche exhaustive n'est clairement pas envisageable pour le pirate. Cela donne la possibilité de changer de signal de tatouage si celui-ci est dévoilé, ce qui confère un renouvellement du processus d'incrustation sans aucun changement du côté de la détection. Une autre stratégie est de ne jamais utiliser le même signal de tatouage. Ainsi, si des contenus apportent de l'information sur le signal de tatouage incrusté, cela ne livre à l'utilisateur malhonnête qu'un indice pour pirater ces contenus. La protection des autres contenus n'est pas compromise comme les signaux de tatouage sont statistiquement indépendants. Le système global reste donc toujours sûr.

Plusieurs méthodes asymétriques, inventées de façon indépendante, ont été proposées. R. Van Schyndel et A. Tirkel ont présenté leur idée dans l'article [39] qui a été analysée [19] et améliorée [26] par J. Eggers et B. Girod. La proposition de J. Smith et C. Dodge est publiée dans les actes du troisième workshop 'Information Hiding' [46]. Elle a été redécouverte par G. Sylvestre et N. Hursley [14] d'une part, et J. Stern et J.-P. Tillich [25, 24] avec une approche plus cryptographique, d'autre part. Enfin, les auteurs ont exposé leur méthode au même workshop [9], ainsi qu'une application aux images fixes [11] et à l'audio [10]. Nous avons récemment prouvé qu'en fait toutes ces méthodes, en apparence très différentes, sont basées sur une formulation mathématique de l'algorithme de détection commune [7]. Nous présentons dans la suite de l'article la méthode exposée dans l'article [9].

6.1 Algorithmes

Dans cette méthode asymétrique, le détecteur ne compare pas le vecteur extrait $\mathbf{r_u}$ à un signal de tatouage spécifique $\mathbf{w}$, mais il vérifie si $\mathbf{r_u}$ a une propriété statistique due à la présence de $\mathbf{w}$. Nous décrivons d'abord l'algorithme d'incrustation.

Tout d'abord, un vecteur $\mathbf{v}$ dont les composantes sont i.i.d. représentant un processus aléatoire centré blanc Gaussien de variance unité, est convolué par le filtre h. Le vecteur résultant est alors entrelacé. Cette entrelaceur agit comme une permutation pseudo-aléatoire $\pi(.)$ des composantes du vecteur. Les vecteurs résultants de l'entrelacement seront accentués avec le symbole tilde : $\widetilde{r_o}[n] = r_o[\pi(n)]$ et $r_o[n] = \widetilde{r_o}[\pi^{-1}(n)] \quad \forall n \in \{0, \ldots, N-1\}$. Finalement, le vecteur de tatouage est défini par : $\forall \in \{0, \ldots, N-1\} \quad w[n] = (h \otimes \mathbf{v})[\pi(n)]$. La formule d'incrustation est donc :

$$r_w[n] = r_o[n] + g[n].(h \otimes \mathbf{v})[\pi(n)] \quad \forall \in \{0, \ldots, N-1\}$$

Le filtre normalisé h et le signal $\mathbf{v}$ sont les paramètres secrets du processus d'incrustation. Noter que n'importe quel signal $\mathbf{v}$ pseudo-aléatoire centré blanc Gaussien convient. La détection n'a pas besoin de ces paramètres. Elle

doit connaître le désentrelaceur (c'est à dire la permutation inverse $\pi^{-1}(.)$) et le module de la réponse fréquentielle du filtre h. Cet ensemble de paramètres $\{\pi^{-1}(.), |H(f)|\}$ caractérisent la propriété statistique attendu : le spectre de la séquence entrelacée $\widetilde{\mathbf{r_u}}$ à la forme de $|H(f)|^2$. Un simple test décide à quelle hypothèse (H_0 ou H_1) le contenu inconnu C_u appartient le plus vraisemblablement.

• H_0: Le vecteur extrait $\mathbf{r_u}$ n'est pas tatoué, donc il ne partage pas la propriété statistique particulière. Grâce à l'action supposée idéale de la permutation pseudo-aléatoire $\pi^{-1}(.)$, les composantes de $\widetilde{\mathbf{r_u}}$ sont supposées représenter un processus blanc et stationnaire, si bien que son spectre $S_0(f)$ est constant, en espérance : $S_0(f) = \sigma^2_{r_u} + \mu^2_{r_u}.\delta(f)$ où μ_{r_u} et $\sigma^2_{r_u}$ sont la moyenne et la variance des composantes de $\mathbf{r_u}$.

• H_1: Le vecteur extrait $\mathbf{r_u}$ a été tatoué. Supposons que $\widetilde{\mathbf{g}}$, $\widetilde{\mathbf{w}}$ et $\widetilde{\mathbf{r_o}}$ sont des vecteurs aléatoires statistiquement indépendants et stationnaires, la relation suivante donne :

$$\begin{aligned} \varphi_{\widetilde{r_u}}[l] &= E\{\widetilde{r_u}[m].\widetilde{r_u}[m+l]\} = \varphi_{\widetilde{r_o}}[l] + \varphi_{\widetilde{g}}[l].\varphi_{\widetilde{w}}[l] \\ \Phi_{\widetilde{r_u}}(f) &= S_1(f) = \Phi_{\widetilde{r_o}}(f) + \Phi_{\widetilde{g}}(f) \otimes \Phi_{\widetilde{w}}(f) \end{aligned}$$

$E\{.\}$ est l'espérance mathématique, $\varphi_{\widetilde{r_o}}[.]$ la fonction de corrélation de la séquence $\{\widetilde{r_o}[m]\}$ et $\Phi_{\widetilde{r_o}}(f)$ sa transformée de Fourrier, qui est la densité spectrale de puissance. Nous supposons de plus que $\widetilde{\mathbf{r_o}}$ et $\widetilde{\mathbf{g}}$ sont des processus aléatoires blancs. Comme $\widetilde{\mathbf{w}} = h \otimes \mathbf{v}$, $\Phi_{\widetilde{w}}(f) = |H(f)|^2$. Le filtre h est normalisé si bien que $\int |H(f)|^2 \, df = 1$. Finalement, la densité spectrale de puissance espérée dans le cas H_1, est :

$$\begin{aligned} S_1(f) &= \sigma^2_{r_o} + \sigma^2_g + \mu^2_{r_o}\delta(f) + \mu^2_g \, |H(f)|^2 \\ &= \mu^2_{r_u}.\delta(f) + \sigma^2_{r_u} + \mu^2_g(|H(f)|^2 - 1) \end{aligned}$$

Ainsi, dans ce cas, l'estimation du spectre $S_1(f)$ du vecteur extrait centré $\widetilde{\mathbf{r_u}}$ est de la forme de $|H(f)|^2$.

La région critique $R'(\{\pi(.), |H(f)|^2\})$ est l'ensemble des vecteurs extraits de l'espace de tatouage partageant la propriété statistique particulière. Elle dépend seulement des paramètres $\{\pi(.), |H(f)|^2\}$, et elle peut être définie par un test d'hypothèses en analyse spectrale basée sur un critère de maximum de vraisemblance :

$$R'(\{\pi(.), |H(f)|^2\}) = \{\mathbf{r_u} | U_N(\widetilde{\mathbf{r_u}}, S_0) - U_N(\widetilde{\mathbf{r_u}}, S_1) \geq Thr'\} \qquad (8)$$

où Thr' est un seuil positif dépendant de la probabilité de fausse alarme fixée dans le cahier des charges et $U_N(\widetilde{\mathbf{r_u}}, S_i)$ est la partie principale de la vraisemblance de Whittle que le spectre du processus aléatoire $\widetilde{\mathbf{r_u}}$ correspond avec la densité spectrale de puissance S_i. Son expression simplifiée est la suivante :

$$U_N(\widetilde{\mathbf{r_u}}, S_i) = 2N \int_{-\frac{1}{2}}^{\frac{1}{2}} (\log(S_i(f)) + \frac{I_N(f)}{S_i(f)}) df$$

où $I_N(f)$ est le périodogramme de la séquence $\widetilde{\mathbf{r_u}}$:

$$I_N(f) = \left| \sum_{k=0}^{N-1} \widetilde{r_u}[k].e^{2\pi i k f} \right|^2 \quad \forall f \in]-\frac{1}{2}, \frac{1}{2}]$$

6.2 Performances

6.2.1 Versatilité

Le signal **w** enfoui dans le contenu est un bruit Gaussien blanc grâce à l'action de l'entrelaceur $\pi(.)$. Ainsi, cette méthode s'adapte à toute technique de tatouage à étalement de spectre. Seule la manière de créer le signal de tatouage a changé. Au lieu de choisir sans cesse le même vecteur, on tire, pour chaque contenu, un vecteur pseudo-aléatoire que l'on filtre puis entrelace.

6.2.2 Puissance de détection

Le cahier des charges des systèmes de protection impose en général un niveau α de fausse alarme maximum. Le seuil Thr' doit être calculé de sorte que $P_{fa} < \alpha$. On reconnaît ici un test d'hypothèse suivant la stratégie de Neyman-Pearson. Le test est alors d'autant plus efficace que sa puissance P_p est grande (cf. 4.3). P_p est en fait une fonction croissante du paramètre d'efficacité e :

$$e = \frac{\mu_{d|\mathrm{H}_1} - \mu_{d|\mathrm{H}_0}}{\sigma_{d|\mathrm{H}_1} + \sigma_{d|\mathrm{H}_0}} \tag{9}$$

Dans le cas d'une technique de tatouage symétrique à étalement de spectre sur des vecteurs extraits Gaussiens et blancs, un calcul classique montre que $e_s \propto \sqrt{GN}$ où G est le rapport de puissance tatouage à original $G = \frac{g^2}{\sigma^2_{r_o}} \ll 1$. En revanche, les méthodes asymétriques ne produisent qu'une efficacité $e_a \propto G\sqrt{N}$. Par conséquent, les méthodes asymétriques sont bien moins efficaces que les techniques symétriques. Pour palier cet inconvénient, la seule solution est d'accroître N. Comme le rapport de puissance G est de l'ordre de -20dB il faut détecter la présence du tatouage sur des vecteurs environ 10 fois plus grands, ce qui amène de nombreuses difficultés sur la complexité, la taille mémoire et le temps de réponse nécessaires aux détecteurs asymétriques.

6.2.3 Robustesse

Ce paragraphe détaille la robustesse contre des transformations classiques de contenus. Le problème est de comparer la robustesse de la méthode de tatouage asymétrique à celle de la méthode symétrique, pour une technique donnée, c'est à dire pour une fonction d'extraction $X(.)$ fixée. Nous avons choisi, comme exemple, une technique de tatouage d'images fixes très robuste détaillée dans l'article [1]. Les modules d'un ensemble des coefficients

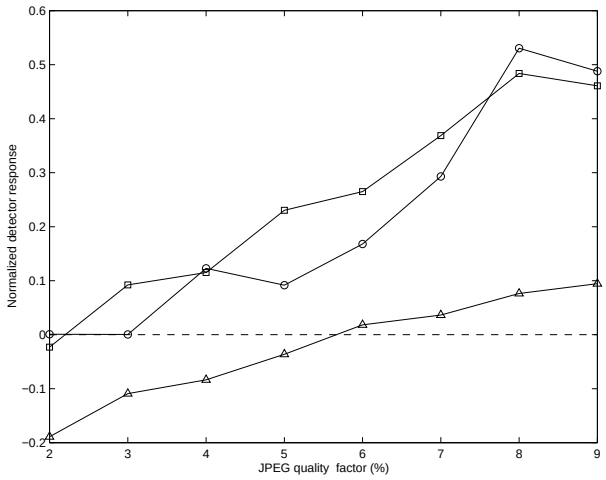

FIG. 2 – *Robustesse contre un codage avec pertes JPEG pour trois détecteurs différents :* ∘ *asymétrique,* △ *symétrique par corrélation, et* □ *symétrique optimum.*

de transformée de Fourrier discrète de l'image sont modifiées proportionnellement : $\mathbf{r_w} = \mathbf{r_o} \star (1 + \gamma \mathbf{w})$ où γ est un paramètre fixant l'amplitude de l'incrustation du tatouage. L'attaque considérée est par exemple une compression JPEG d'un facteur de qualité Q compris entre 0 et 100 %. La réponse du détecteur est normalisé pour qu'à $Q = 100$ %, sa réponse soit de 1. La figure 2 trace le résultat moyen sur un grand nombre d'images pour un facteur de qualité compris entre 0 et 30 %. La détection asymétrique est correcte jusqu'à $Q = 10$ %, alors que la technique symétrique est robuste jusqu'à $Q = 5$ % pour un détecteur basé sur une corrélation et $Q = 3$ % avec un détecteur à maximum de vraisemblance donné dans [1]. Cet exemple montre que la version asymétrique rend la technique de tatouage légèrement moins efficace. Pour retrouver une robustesse comparable, il faut augmenter la taille N des vecteurs extraits.

Une autre attaque est de désynchroniser les entrelaceurs à l'incrustation et à la détection. La détection ne pourra pas retrouver le vecteur $h \otimes \mathbf{v}$ et conclura que le contenu n'est pas tatoué. Dans notre exemple, la technique est robuste à une translation de l'image de quelques pixels. Une rotation lui est en revanche fatale. Cependant, cette remarque est valide pour toutes les techniques de tatouages. Beaucoup de travaux académiques étudient des techniques de synchronisation ou des vecteurs d'extraction invariants pour certaines transformées.

6.2.4 Sécurité

Le rôle de l'entrelaceur est très important. Du côté de l'incrustation, il blanchit le bruit coloré ($h \otimes \mathbf{v}$) avant d'être ajouté au vecteur extrait $\mathbf{r_o}$. Par conséquent, si une technique peut incruster un bruit blanc Gaussien dans un contenu (comme la plupart des techniques dites à étalement de spectre), il est très facile de la rendre asymétrique. Le désentrelaceur blanchit la part du signal provenant du contenu original, ce qui est fondamental pour le bon fonctionnement de la détection. Mais, il joue aussi un rôle essentiel quant à la sécurité de la méthode. En fait, il cache au pirate ce qu'est exactement la propriété statistique attendue des contenus tatoués. Sans sa connaissance, le pirate n'a pas accès au domaine où le cœur de la détection a lieu (les calculs de vraisemblance de l'Eq. (8)). Ainsi, il lui est impossible prédire l'impact de son attaque.

Evaluer le niveau de sécurité de cette méthode asymétrique revient à calculer la complexité nécessaire pour estimer la permutation $\pi(.)$. Supposons que le pirate dispose d'une paire de contenus original/tatoué. Il estime le signal de tatouage $\mathbf{w}$ par la différence $\mathbf{r_w} - \mathbf{r_o}$. L'adversaire doit maintenant trouver la permutation à partir de ce signal. Une possibilité est d'essayer toutes les permutations possibles et de s'arrêter lorsque le signal permuté est coloré. Une difficulté est que l'on ne connaît pas le deuxième paramètre secret $|H(f)|^2$, et par conséquent le spectre espéré S_1. Le pirate prendra à la place un test de sphéricité comme celui de Drouiche et Fay [8], qui détermine la vraisemblance qu'un signal soit coloré. Cependant, il existe $N!$ permutations possibles. Donnons un ordre de grandeur : Pour $N = 2048$, en utilisant l'approximation de Stirling, $N! \sim 2^{19000}$, ce qui est considérablement plus grand que 2^{300}, nombre de particules de particules dans l'univers !

Cette évaluation du niveau de sécurité est trop optimiste. Le pirate sait bien que la permutation π^{-1} a de très bonnes propriétés de décorrélation. De plus, il ne souhaite pas la trouver exactement ; une permutation suffisamment proche convient. Pour une évaluation correcte de la complexité, nous sommes obligés, à l'instar de A. Kerckhoffs, de dévoiler l'algorithme qui a donné naissance à la permutation π. C'est un générateur de permutations pseudo-aléatoires dont l'entrée est un mot de L bits ($L << N$), comme on en trouve dans le livre de D. Knuth [32]. Parcourir l'espace des permutations possibles revient maintenant à essayer ses 2^L éléments. Supposons qu'une génération de permutation suivie d'un test de sphéricité dure une 1μs. L'espérance du temps nécessaire pour tomber sur la bonne permutation est d'environ 2^{L-46} années. Cette analyse ne prend pas en compte la loi de Moore, ni la possibilité de paralléliser l'attaque. Elle montre cependant que le niveau de sécurité est nettement supérieur à celui de la méthode symétrique.

Nous pouvons montrer que l'attaque par oracle est de même plus difficile à réaliser avec la méthode asymétrique. La différence majeure est que le détecteur n'est pas linéaire contrairement à ce qui est supposé dans l'attaque

Eq. (7). Sa complexité est en $O(N^2)$. Certes, ce n'est pas un niveau de sécurité recommandable en cryptographie (i.e. exponentiel), mais notons que le cahier des charges des techniques de taoutage en protection de copie stipule que le détecteur donne une prise de décision toutes les 10 secondes. Ainsi pour $N = 2048$, N essais prennent 6 heures alors que N^2 essais prennent un an et demi.

Un autre point critique est d'analyser les menaces pour le système si le pirate découvre, par 'reverse engineering' de l'implémentation, les paramètres $\{\pi^{-1}(.), |H(f)|\}$. La réponse n'est pas évidente. Le pirate a maintenant accès au vecteur extrait $\mathbf{r_w}$ et sa version permutée $\widetilde{\mathbf{r_w}}$. Il peut créer un vecteur permuté $\widetilde{\mathbf{r_p}}$ que le détecteur considérera comme non tatoué [11]. Mais il n'est pas capable de prédire l'impact visuel de cette attaque. De la même manière, il est capable d'utiliser un modèle perceptif dans l'espace de 'tatouage' pour créer un contenu de bonne qualité, mais il ne peut pas prédire l'impact de ce modèle sur la sortie du détecteur. Ceci est dû à l'entrelaceur qui empêche d'avoir accès dans un même domaine au modèle perceptif et à la formule de détection. Il est très difficile de créer un contenu piraté et de bonne qualité avec la technique décrite dans la section 6.2.3. Les auteurs recherchent des attaques plus efficaces.

7 Conclusion

Le cadre de la protection de copie est complexe. Dans certains contextes, des systèmes utilisant cryptographie et tatouage de manière complémentaire sont efficaces. Cependant, les techniques de tatouage classiques requièrent l'utilisation d'une clé sécrète dont la confidentialité assure la sécurité du système global. Or, les attaques possibles en protection de copie permettent de dévoiler cette clé. C'est pourquoi nous proposons un nouveau schéma de tatouage dit asymétrique. Celui-ci procure au système un niveau de sécurité supérieur, le prix a payé étant une augmentation de la taille des vecteurs extraits.

À travers cet article nous défendons aussi notre conviction que l'application visée conditionne complètement la technique de tatouage. De plus, le fait d'avoir un niveau de sécurité d'un ordre de grandeur supérieur donne du crédit à l'application du principe de Kerckhoffs en tatouage pour protection de copie.

Références

[1] A. de Rosa, M. Barni, F. Bartolini, V. Cappelini, et A. Piva. «Optimum decoding of non-additive full frame DFT watermarks». Dans A. Pfitzmann, éditeur, *Proc. of the third Int. Workshop on Information*

Hiding, pages 159–171, Dresden, Germany, septembre 1999. Springer Verlag.

[2] A. Patrizio. « DVD privacy: It can be done ». http://www.wired.com/news/technology/ 1,1282,32249,00.html.

[3] A. Piva, R. Caldelli, et A. de Rosa. « A DWT-based object watermarking system for MPEG-4 ». Dans *Proc. of Int. Conf. on Image Processing*, volume 3, pages 5–8, Vancouver, Canada, septembre 2000. IEEE.

[4] C. Cachin. «An information-theoretic model for steganography». Dans D. Aucsmith, éditeur, *Proc. of the second Int. Workshop on Information Hiding*, volume 1525 de *L.N.C.S.*, pages 306–318, Portland, Oregon, U.S.A., avril 1998. Springer Verlag.

[5] C. Lin, M. Wu, J. Bloom, I. Cox, et M. Miller. «Rotation, scale, and translation resilient public watermarking for images». Dans P.W. Wong et E. Delp, éditeurs, *Security and Watermarking of Multimedia Contents II*, pages 90–98, San Jose, Cal., USA, janvier 2000. SPIE.

[6] C. Simon et B. Macq. « ACTS Project AC019 TALISMAN: Tracing Authors' rights by Labelling Image Services and Monitoring Access Network». http://www.tele.ucl.ac.be/TALISMAN.

[7] F. Furon, I. Venturini, et P. Duhamel. «Unified approach». Dans P.W. Wong et E. Delp, éditeurs, *Security and Watermarking of Multimedia Contents III*, San Jose, Cal., USA, 2001. SPIE.

[8] G. Fay. « *Théorèmes limite pour les fonctionnelles de périodogramme*». PhD thesis, Ecole Nationale Supérieure des Télécommunications, 2000.

[9] T. Furon et P. Duhamel. « An Asymmetric Public Detection Watermarking Technique». Dans A. Pfitzmann, éditeur, *Proc. of the third Int. Workshop on Information Hiding*, pages 88–100, Dresden, Germany, septembre 1999. Springer Verlag.

[10] T. Furon et P. Duhamel. « Audio asymmetric watermarking technique». Dans *Proc. of Int. Conf. on Audio, Speech and Signal Processing*, Istanbul, Turkey, juin 2000. IEEE.

[11] T. Furon et P. Duhamel. «Robustness of an asymmetric technique». Dans *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, septembre 2000. IEEE.

[12] G. Langelaar, I. Setyaiwan, et R. Lagendijk. « Watermarking Digital Image and Video Data». *Signal Processing Magazine*, 17(5), septembre 2000.

[13] G. Rhoads. « Method and apparatus for robust information coding ». patent US 5,748,783, mai 1995. filing date.

[14] G. Silvestre, N. hurley, G. Hanau, et W. Dowling. « Informed Audio Watermarking using Digtial Chaotic Signals ». Dans *Proc. of Int. Conf. on Acoustics, Speech and Signal Processing*, Salt-Lake City, USA, mai 2001. IEEE.

[15] G. Silvestre et W. Dowling. « Embedding data in digital images using CDMA techniques ». Dans *Proc. of Int. Conf. on Image Processing*, volume 1, pages 589–592, Vancouver, Canada, septembre 2000. IEEE.

[16] I. Cox, J. Kilian, T. Leighton, et T. Shamoon. « secure spread spectrum watermarking for multimedia ». *Transaction on image processing*, 6(12):1673–1687, décembre 1997.

[17] I. Cox, M. Miller, et A. McKellips. « Watermarking as communication with side information ». *Proc. of the IEEE*, 87(7):1127–1141, juillet 1999.

[18] I. Cox, M. Miller, T. Kazuyoshi, et W. Yutaka. « Digital data watermarking ». patent EP 0 840 513 A2, novembre 1997. filing date.

[19] J. Eggers et B. Girod. « Robustness of public key watermarking schemes ». Dans V^3D^2 *Watermarking Workshop*, Erlangen, Germany, octobre 1999.

[20] J. Eggers et B. Girod. « Quantization effect on digital watermark ». *Elsevier Signal Processing*, 2001.

[21] J. O'Ruanaidh et T. Pun. « Rotation, Translation and Scale Invariant Spread Sprectrum Digital Image Watermarking ». *Signal Processing, Special issue on copyrigth protection and control*, 66(3), 1998.

[22] J. O'Ruanaidh et T. Pun. « A secure robust digital image watermarking ». Dans *SPIE Electronic Imaging: Processing, printing and publishing in colour*, mai 1998.

[23] J. Proakis. *Digital Communications*. Electrical and computer engineering. McGraw Hill, 3rd édition, 1996.

[24] J. Stern. « *Contribution à la théorie de la protection de l'information* ». PhD thesis, Université de Paris XI, Orsay, Laboratoire de Recherche en Informatique, mars 2001.

[25] J. stern et J.-P. Tillich. « Automatic Detection of a watermarked document using a private key ». To be published in Proc. of the fourth Int. Workshop on Information Hiding., 2001.

[26] J.Eggers, J.Su, et B.Girod. « Public key watermarking by eigenvectors of linear transforms ». Dans *Proc. of the European Signal Processing Conference*, Tampere, Finland, septembre 2000. EUSIPCO.

[27] J.P. Linnartz, G. Depovere, et T. Kalker. « On the design of a watermarking system: considerations and rationales ». Dans A. Pfitzmann, éditeur, *Proc. of the third Int. Workshop on Information Hiding*, pages 253–269, Dresden, Germany, septembre 1999. Springer Verlag.

[28] D. Kahn. « Cryptology and the origins of spread spectrum ». *IEEE spectrum*, pages 70–80, septembre 1984.

[29] T. Kalker. « A security risk for publicly available watermark detectors ». Dans *Benelux Information Theory Symposium*, mai 1998. Veldhoven, The Netherlands.

[30] T. Kalker. « Considerations on watermarking seccurity ». Dans *Proc of the IEEE MMSP'01 conference*, Cannes, France, octobre 2001.

[31] A. Kerckhoffs. « La cryptographie militaire ». *Journal des sciences militaires*, 9:5–38, janvier 1883.

[32] D. Knuth. *The art of computer programming.* Computer Science and Information Processing. Addison-Wesley, 1981.

[33] M. Barni, F. Bartolini, A. De Rosa, et A. Piva. « A new decoder for the optimum recovery of non-additive watermarks ». *IEEE Transactions on Image Processing*, 2001.

[34] M. Kutter, S. Voloshynovskiy, et A. Herrigel. « Watermark copy attack ». Dans P.W. Wong et E. Delp, éditeurs, *Security and Watermarking of Multimedia Contents II*, volume 3971, San Jose, Cal., USA, janvier 2000. SPIE Proceedings.

[35] M. Maes, T. Kalker, J.-P. Linnartz, Joop Talstra, Geert Depovere, et Jaap Haitsma. « Digital watermarking for DVD video copy protection ». *Signal Processing Magazine*, 17(5), septembre 2000.

[36] M. Miller, I. Cox, et J. Bloom. « Informed embedding: exploiting image and detector information during watermark insertion ». Dans *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, septembre 2000. IEEE.

[37] M. Ramkumar. « *Data hiding in multimedia - Theory and applications* ». PhD thesis, University Heights, Newark, New Jersey Institute of Technology, décembre 1999.

[38] R. Grosbois et T. Ebrahimi. « Watermarking in the JPEG 2000 domain». Dans *Proc of the IEEE MMSP'01 conference*, Cannes, France, octobre 2001.

[39] R. Van Schyndel, A. Tirkel, et I. Svalbe. « Key independent watermark detection». Dans *Int. Conf. on Multimedia Computing and Systems*, volume 1, Florence, Italy, juin 1999.

[40] R. Wolfgang, C. Podilchuk, et E. Delp. «The effect of matching watermark and compression transforms in compressed color images ». Dans *Proc. of the Int. Conf. on Image Processing*, volume 1, pages 440–443, Chicago, Ill., USA, octobre 1998.

[41] S. Craver et J. Stern. «Lessons learned from SDMI». Dans *Proc of the IEEE MMSP'01 conference*, Cannes, France, octobre 2001.

[42] S. Craver, N. Memon, B.-L. Yeo, et M. M.Yeung. « Resolving rightful ownership with invisible watermarking techniques: limitations, attacks, and implications ». *IEEE Journal of selected areas in communications*, 16(4), mai 1998. Special issue on copyright and privacy protection.

[43] S. Derrode. « *Représentation de formes planes à niveaux de gris différents approximations de Fourier-Mellin analytique en vue de d'indexation de base d'images*». PhD thesis, Université de Rennes I, Groupe de recherche Images et Formes, décembre 1999.

[44] C.E. Shannon. « Communication theory of secrecy systems ». *Bell system technical journal*, 28:656–715, octobre 1949.

[45] S. Singh. *The code book.* Fourth Estate Limited, 1999. Histoire des codes secrets, publié chez JC Lattès.

[46] J. Smith et C. Dodge. « Developments in steganography ». Dans A. Pfitzmann, éditeur, *Proc. of the third Int. Workshop on Information Hiding*, pages 77–87, Dresden, Germany, septembre 1999. Springer Verlag.

[47] Widevine technologies. «How to steal streaming content». Rapport Technique, http://www.widevine.com/papers/how_to_steal.pdf, 2001.