

SIMULATION DE TATOUAGE NUMÉRIQUE PAR CODES TARDOS



15/03/2020

—
João FREITAS, Augustin LARUELLE



TABLE DES MATIÈRES

1 Le code de Tardos	4
1.1 Notations et modèle	4
1.1.1 Notations	4
1.1.2 <i>Marking assumption</i>	4
1.1.3 Accusation ε -sécuritaire	4
1.2 Algorithme de Tardos	4
1.2.1 Génération des tatouages	5
1.2.2 Calcul du score de Tardos	5
1.2.3 Étude du score de Tardos	5
1.2.4 Accusation selon le critère de Tardos	6
1.3 Performances de l'algorithme	6
1.3.1 Non accusation d'innocents	6
1.3.2 Accusation d'un pirate de la collusion	6
1.3.3 Optimalité de la borne de non accusation d'innocents	6
2 Comparaison des distributions du score des innocents face à celle de la collusion	7
2.1 Description des stratégies	7
2.1.1 Stratégie pile ou face	7
2.1.2 Stratégie par vote majoritaire	7
2.1.3 Stratégie par entrelacement	8
2.2 Étude de la distribution du score d'un innocents	8
2.2.1 Étude empirique	8
2.2.2 Étude théorique	10
2.3 Étude de la distribution du score d'un pirate	11
2.3.1 Étude empirique	11
2.3.2 Étude théorique	13
3 Estimation de la p-valeur du score de Tardos	16
3.1 Monte-Carlo naïf	16
3.2 <i>Adaptive Multilevel Splitting</i>	17
3.2.1 Pseudo-code	17
3.2.2 Résultats empiriques	18
4 Stratégie pour accuser un deuxième pirate	20
4.1 Description des stratégies envisagées	20
4.1.1 Stratégie de somme des scores	20
4.1.2 Stratégie des bits "cachés"	21
4.2 Optimalité du score de Tardos pour l'accusation d'un deuxième pirate	22
4.2.1 Définition du score	22
4.2.2 Notations pour la stratégie de la collusion	22
4.2.3 Calcul de la moyenne du score d'un deuxième pirate	23

INTRODUCTION

Le tatouage numérique permet de cacher un identifiant dans un fichier de manière invisible à l'œil. Une application majeure de ces techniques se trouve dans le monde de la propriété intellectuelle. Chaque destinataire d'une copie d'un document protégé se voit attribué un identifiant. Ainsi, si un utilisateur publie ce document sans en posséder les droits, il peut être identifié en exhibant le tatouage dans le document. Seul celui qui introduit le tatouage connaît la position exacte des bits et donc le pirate ne peut pas brouiller les pistes en masquant son identifiant.

Maintenant, plusieurs pirates peuvent entrer en collusion. Ils peuvent alors détecter les bits qui diffèrent entre leurs deux versions du document et en déduire la position de certains bits du tatouage pour brouiller les pistes. Le but du développement du tatouage numérique est de pouvoir accuser les pirates malgré leur collusion.

1

LE CODE DE TARDOS

Le code de Tardos qui est décrit par ce mathématicien dans son article [5] est une des premières stratégies permettant de mettre en œuvre le tatouage numérique contre les collusions pirates. Ce code est basé sur une approche probabiliste de la programmation.

1.1 NOTATIONS ET MODÈLE

1.1.1 • NOTATIONS

Les notations définies dans cette partie seront conservées pour le reste du document.

Le distributeur est celui qui distribue les tatouages aux n utilisateurs. La longueur du tatouage est notée m .

Nous considérons X la matrice des tatouages. Cette matrice est connue du distributeur de fichier uniquement. X_{ji} représente donc le i-ième bit du tatouage du j-ième utilisateur. Nous avons donc $X \in \{0, 1\}^{n \times m}$.

La collusion pirate est notée $C = \{j_1, \dots, j_c\}$ et est de taille c . Les pirates ont une stratégie ρ telle que $\rho(C) \in \{0, 1\}^m$ est la marque pirate, c'est-à-dire celle qu'ils mettent sur le fichier pirate qu'ils diffusent. Quand le distributeur trouve un fichier pirate, il peut en extraire le tatouage que nous noterons $Y = \rho(C)$.

1.1.2 • MARKING ASSUMPTION

La *marking assumption* est le seul critère que la collusion pirate doit respecter dans l'étude théorique du modèle. Elle se base sur l'idée que les pirates ne peuvent pas repérer seuls la position des bits de tatouage dans le fichier. Ainsi, si tous les pirates ont le même i-ième bits de tatouage alors, quelque soit leur stratégie ρ , la marque pirate conservera se i-ième bit.

$$\forall \rho, C, \forall i \in [|1, m|] \quad X_{j_1 i} = X_{j_2 i} = \dots = X_{j_c i} \Rightarrow \rho(C)_i = X_{j_1 i}$$

1.1.3 • ACCUSATION ε -SÉCURITAIRE

Rappelons que le but du tatouage numérique est de pouvoir accuser la collusion pirate. Cependant il faut à tout prix éviter d'accuser un innocent.

Notons FN l'événement le distributeur n'accuse aucun pirate et FP celui où le distributeur accuse un innocent.

Une stratégie d'accusation σ est dite ε -sécuritaire si $\mathbb{P}(FN \cup FP) \leq \varepsilon$.

1.2 ALGORITHME DE TARDOS

Le code de Tardos est un processus de tatouage aléatoire garantissant les bornes minimales en terme longueur du tatouage par rapport à un niveau de fiabilité voulu. Cet algorithme est proposé en 2003 par Tardos [5].

1.2.1 • GÉNÉRATION DES TATOUAGES

La première étape avant de générer les tatouages est de tirer $p \in [t, 1-t]^m$ vecteur des paramètres qui nous permettront de tirer les bits de tatouage. Chaque p_i est tiré indépendamment selon la loi de densité $f(p) = \frac{Cte}{\sqrt{p(1-p)}} \mathbb{1}_{[t,1-t]}(p)$ avec $Cte = \frac{1}{2[\text{Arcsin}(\sqrt{1-t}) - \text{Arcsin}(\sqrt{t})]}$.

Le t est un hyperparamètre de l'algorithme. Celui-ci permet aux p_i de ne pas s'approcher trop de 1 et de 0 pour que $\frac{p_i}{1-p_i}$ et $\frac{1-p_i}{p_i}$ restent bornés.

Ensuite sont attribués les tatouages. Chaque X_{ji} est tiré indépendamment selon la loi de Bernoulli de paramètre p_i .

Nous remarquerons que ce processus de génération de tatouage peu ne nécessite pas la connaissance à priori du nombre total d'utilisateur n . Ainsi à chaque nouvel utilisateur, il est facile au distributeur d'attribuer un tatouage.

1.2.2 • CALCUL DU SCORE DE TARDOS

Face à une marque pirate Y , il faut être capable de déterminer si un utilisateur est susceptible d'avoir contribué à la collusion qui à permis d'aboutir à cette marque. Pour cela, chaque utilisateur se voit attribuer un score de Tardos $S_j = S(Y, X_{j.}, p)$. Nous noterons que dans son article [5], Tardos n'utilise pas exactement le score ci-dessous mais un score non centré. Cette version recentrée du score a été proposée par Skorić en 2008 [4].

$$S_j = \sum_{i=0}^m U(Y_i, X_{ji}, p_i) \quad \text{avec} \quad \begin{cases} U(1, 1, p) = \sqrt{\frac{1-p}{p}} \\ U(1, 0, p) = -\sqrt{\frac{p}{1-p}} \\ U(0, 1, p) = -\sqrt{\frac{1-p}{p}} \\ U(0, 0, p) = \sqrt{\frac{p}{1-p}} \end{cases}$$

Plus S_j est grand plus l'utilisateur j à tendance à avoir une tête de coupable. En effet, quand les bits de la marque pirate et de l'utilisateur j coïncident alors le terme de score est positif alors que sinon il est négatif. De plus, le facteur p intervient pour rendre cette ajout ou retrait de score plus ou moins important. Par exemple, quand les deux bits sont des 1, si p est proche de 1 alors le facteur de score est presque nul car presque tous les utilisateurs auront des 1 à cet emplacement. Par contre, si p est proche de 0 alors peu d'utilisateurs ont des bits 1 à cet emplacement donc la probabilité que l'utilisateur j soit dans la collusion augmente fortement et le facteur de score pour ce bit est élevé.

1.2.3 • ÉTUDE DU SCORE DE TARDOS

Notons $U_i = U(Y_i, X_{ji}, p_i)$ avec p_i déterministe et un utilisateur j innocent. On rappel que Y_i et X_{ji} sont indépendants et que X_{ji} est distribué selon une loi de Bernoulli de paramètre p_i . Avec la formule de la fonction U , on peut réécrire U_i de la sorte :

$$\begin{aligned} (U_i | Y_i = 1) &= p_i \delta \sqrt{\frac{1-p_i}{p_i}} + (1-p_i) \delta - \sqrt{\frac{p_i}{1-p_i}} \\ (U_i | Y_i = 0) &= p_i \delta - \sqrt{\frac{1-p_i}{p_i}} + (1-p_i) \delta \sqrt{\frac{p_i}{1-p_i}} \end{aligned}$$

Ainsi, $\mathbb{E}(U_i | Y_i = 1) = p_i \sqrt{\frac{1-p_i}{p_i}} + (1-p_i)(-\sqrt{\frac{p_i}{1-p_i}}) = 0$.

De même, $\mathbb{E}(U_i|Y_i = 0) = 0$. Donc $\boxed{\mathbb{E}(U_i) = 0}$

Nous avons bien que le score de Tardos est centré.

Étudions maintenant la variance de U_i .

$$\mathbb{E}(U_i^2|Y_i = 1) = p_i \frac{1-p_i}{p_i} + (1-p_i) \frac{p_i}{1-p_i} = 1. \text{ De même, } \mathbb{E}(U_i^2|Y_i = 0) = 1.$$

D'où, $\boxed{Var(U_i) = \mathbb{E}(U_i^2) = 1}$

1.2.4 • ACCUSATION SELON LE CRITÈRE DE TARDOS

Le critère de Tardos [5] fixe un seuil, $Z = 20c \ln(\varepsilon)$, d'accusation des utilisateurs. Si $S_j > Z$ alors l'utilisateur j est accusé. Ce seuil permet notamment d'avoir une stratégie d'accusation ε -sécuritaire, comme nous le verrons dans la partie suivante. Cependant ce seuil n'est pas explicite. En effet, dans un cas pratique, la taille c de la collusion n'est pas connue et donc ce seuil ne peut pas être fixé.

1.3 PERFORMANCES DE L'ALGORITHME

Les preuves de ces performances sont données par Tardos [5].

1.3.1 • NON ACCUSATION D'INNOCENTS

Dans son article, Tardos donne la preuve de l'efficacité de son code. Notamment $\mathbb{P}(S \geq Z) \leq \varepsilon$, ce qui signifie que la probabilité qu'un innocent ait un score supérieur au seuil est inférieure à ε . La seule condition pour que cette borne s'applique est que $m \geq c^2 \log(\frac{n}{\varepsilon})$. En fait, Tardos atteint ici la borne inférieure de la longueur du tatouage pour avoir une accusation ε -sécuritaire. Cette borne inférieure a été établie par Boneh et Shaw [1] en 1998 et jamais atteinte avant le code proposé par Tardos.

1.3.2 • ACCUSATION D'UN PIRATE DE LA COLLUSION

Avec la même condition que le théorème précédent, la probabilité de ne pas accuser de coupable est inférieure à une puissance de ε .

$$\mathbb{P}(FN) \leq \varepsilon^{c/4}$$

Ne pas connaître la taille de la collusion, n'est pas très grave ici. On peut minorer c par 1 et donc garder une puissance fixe de ε . De plus, l'important dans ce processus de tatouage numérique est surtout de ne jamais accuser d'innocents et d'avoir un seuil dissuasif d'accusation pour les pirates.

1.3.3 • OPTIMALITÉ DE LA BORNE DE NON ACCUSATION D'INNOCENTS

Dans son papier [5], Tardos fixe la valeur de tout les paramètres sans nécessairement justifier. Notamment dans la preuve de non accusation d'innocents, il fixe $\alpha = \frac{1}{10c}$. Nous allons voir que ce choix est bien optimal.

Par une inégalité de Chernoff, Tardos montre l'inégalité suivante : $\forall \alpha > 0, \mathbb{P}(S \geq Z) \leq e^{\alpha^2 m - \alpha Z}$. Avec $Z = 20ck$, $m = 100c^2 k$ et $k = \lceil \ln(1/\varepsilon) \rceil$.

Cette borne de Chernoff est optimale pour $\alpha = \frac{Z}{2m} = \frac{1}{10c}$. En effet, $\alpha^2 m - \alpha Z = m \left(\alpha - \frac{Z}{2m} \right)^2 - \frac{Z^2}{4m}$.

2

COMPARAISON DES DISTRIBUTIONS DU SCORE DES INNOCENTS FACE À CELLE DE LA COLLUSION

Nous allons étudier différents types de stratégies et voir quelle est la distribution empirique et théorique des innocents face à ces stratégies. Nous considérerons trois stratégies : pile ou face, par vote majoritaire et par entrelacement, citées en exemple par Teddy Furon [3].

2.1 DESCRIPTION DES STRATÉGIES

2.1.1 • STRATÉGIE PILE OU FACE

Cette stratégie consiste à tirer une pièce équilibrée pour mettre 0 ou 1 à chaque fois qu'un bit de tatouage est repéré. Nous rappelons que la collusion pirate et leur stratégie doit tout de même respecter la *marking assumption*. Ainsi les bits communs à tous les pirates sont conservés.

Nous allons à présent déterminer la loi de la marque pirate d'une collusion qui suit cette stratégie. Rappelons que $\forall i \in [|1, m|]$, $\forall k \in [|1, c|]$, $X_{j_k i} \stackrel{iid}{\sim} \text{Bernoulli}(p_i)$.

Soit $i \in [|1, m|]$,

$$\begin{cases} \mathbb{P}(\forall k \in [|1, c|], X_{j_k i} = 1) = p_i^c \\ \mathbb{P}(\forall k \in [|1, c|], X_{j_k i} = 0) = (1 - p_i)^c \end{cases} \quad \text{d'où,} \quad \begin{cases} \mathbb{P}(Y_i = 1) = p_i^c + \frac{1}{2}(1 - p_i^c - (1 - p_i)^c) \\ \mathbb{P}(Y_i = 0) = (1 - p_i)^c + \frac{1}{2}(1 - p_i^c - (1 - p_i)^c) \end{cases}$$

Ce qui est intéressant pour une stratégie, c'est d'avoir la loi jointe de Y_i et $X_{j_k i}$ (avec $X_{j_k i}$ coupable), car celle-ci nous permet d'étudier la distribution score de Tardos d'un pirate par rapport à celui d'un innocent.

$$\mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \mathbb{P}(X_{j_k i} = 1)\mathbb{P}(Y_i = 1 | X_{j_k i} = 1) = p_i[\frac{1}{2}(1 - p_i^{c-1}) + p_i^{c-1}]$$

Avec des calculs similaires, on trouve :

$$\begin{cases} \mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \frac{p_i}{2}(1 + p_i^{c-1}) \\ \mathbb{P}(X_{j_k i} = 1, Y_i = 0) = \frac{p_i}{2}(1 - p_i^{c-1}) \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 1) = \frac{1-p_i}{2}(1 - (1 - p_i)^{c-1}) \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 0) = \frac{1-p_i}{2}(1 + (1 - p_i)^{c-1}) \end{cases}$$

2.1.2 • STRATÉGIE PAR VOTE MAJORITY

Cette stratégie a pour principe de mettre le bit le plus représenté à chaque position du tatouage. Nous remarquons que cette stratégie vérifie tout de même la *marking assumption* même si les pirates connaissent la position des bits du tatouage.

Nous allons à présent déterminer la loi de la marque pirate d'une collusion qui suit cette stratégie.

Soit $i \in [|1, m|]$,

$$\begin{cases} \mathbb{P}(Y_i = 1) = \sum_{l=\lceil \frac{c}{2} \rceil}^c \binom{c}{l} p_i^l (1 - p_i)^{c-l} \\ \mathbb{P}(Y_i = 0) = \sum_{l=0}^{\lceil \frac{c}{2} \rceil - 1} \binom{c}{l} p_i^l (1 - p_i)^{c-l} \end{cases}$$

Nous allons à présent déterminer la loi jointe de Y_i et $X_{j_k i}$ (avec $X_{j_k i}$ coupable).

$$\mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \mathbb{P}(X_{j_k i} = 1)\mathbb{P}(Y_i = 1 | X_{j_k i} = 1) = p_i \sum_{l=\lceil \frac{c}{2} \rceil}^c \binom{c-1}{l-1} p_i^{l-1} (1-p_i)^{c-l}$$

Avec des calculs similaires, on trouve :

$$\begin{cases} \mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \sum_{l=\lceil \frac{c}{2} \rceil}^c \binom{c-1}{l-1} p_i^l (1-p_i)^{c-l} \\ \mathbb{P}(X_{j_k i} = 1, Y_i = 0) = \sum_{l=0}^{\lceil \frac{c}{2} \rceil - 2} \binom{c-1}{l} p_i^{l+1} (1-p_i)^{c-l-1} \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 1) = \sum_{l=\lceil \frac{c}{2} \rceil}^c \binom{c-1}{l} p_i^l (1-p_i)^{c-l} \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 0) = \sum_{l=0}^{\lceil \frac{c}{2} \rceil - 1} \binom{c-1}{l} p_i^l (1-p_i)^{c-l} \end{cases}$$

2.1.3 • STRATÉGIE PAR ENTRELACEMENT

Cette stratégie consiste à tirer aléatoirement un pirate dans la collusion pour chaque bit. Le pirate choisi imposera ensuite son bit de tatouage. Comme la stratégie par vote majoritaire, cette stratégie vérifie la *marking assumption* quelque soit l'information détenue par la collusion.

Nous allons à présent déterminer la loi de la marque pirate d'une collusion qui suit cette stratégie.

Soit $i \in [|1, m|]$,

$$\begin{cases} \mathbb{P}(Y_i = 1) = \sum_{k=1}^c \mathbb{P}(X_{j_k i} \text{ choisi}) \mathbb{P}(Y_i = 1 | X_{j_k i} \text{ choisi}) = \sum_{k=1}^c \frac{1}{c} \mathbb{P}(X_{j_k i} = 1) = p_i \\ \mathbb{P}(Y_i = 0) = \sum_{k=1}^c \mathbb{P}(X_{j_k i} \text{ choisi}) \mathbb{P}(Y_i = 0 | X_{j_k i} \text{ choisi}) = \sum_{k=1}^c \frac{1}{c} \mathbb{P}(X_{j_k i} = 0) = 1 - p_i \end{cases}$$

Nous allons à présent déterminer la loi jointe de Y_i et $X_{j_k i}$ (avec $X_{j_k i}$ coupable).

$$\mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \sum_{k'=1}^c \mathbb{P}(X_{j_{k'} i} \text{ choisi}) \mathbb{P}(X_{j_k i} = 1, Y_i = 1 | X_{j_{k'} i} \text{ choisi}) = \frac{1}{c} p_i + \frac{c-1}{c} p_i^2$$

Avec des calculs similaires, on trouve :

$$\begin{cases} \mathbb{P}(X_{j_k i} = 1, Y_i = 1) = \frac{1}{c} p_i + \frac{c-1}{c} p_i^2 \\ \mathbb{P}(X_{j_k i} = 1, Y_i = 0) = \frac{c-1}{c} p_i (1-p_i) \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 1) = \frac{c-1}{c} p_i (1-p_i) \\ \mathbb{P}(X_{j_k i} = 0, Y_i = 0) = \frac{1}{c} (1-p_i) + \frac{c-1}{c} (1-p_i)^2 \end{cases}$$

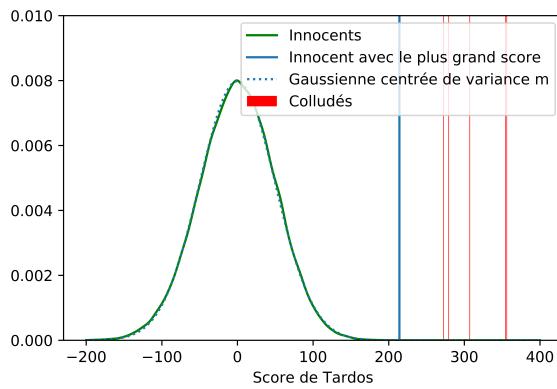
2.2 ÉTUDE DE LA DISTRIBUTION DU SCORE D'UN INNOCENTS

2.2.1 • ÉTUDE EMPIRIQUE

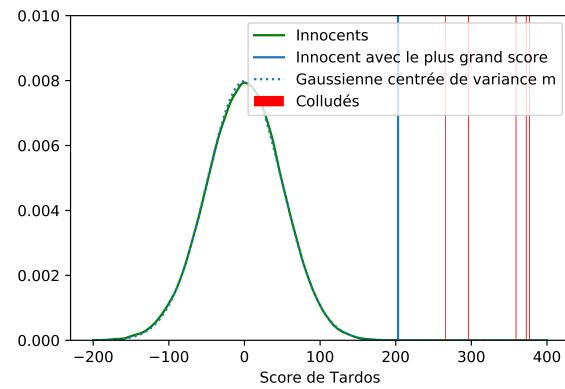
Chacune des stratégies mentionnées ci-dessus va être testée sur un jeu de données générées aléatoirement. Nous avons pris les valeurs suivantes pour les paramètres du problème :

- $n = 10^5$
- $m = 2,5 \cdot 10^3$
- $c = 5$
- $t = \frac{1}{300}$

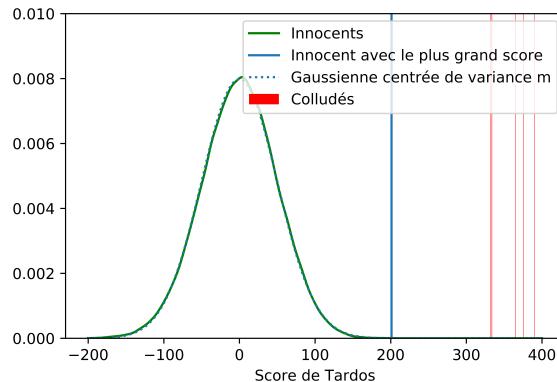
Avec ce choix de paramètres nous avons un code ε -sécuritaire, avec $\varepsilon \geq n \exp(-\frac{m}{5c^2}) = 2,01 \cdot 10^{-4}$. Le facteur 5 dans l'exponentielle vient de la constante devant la minoration de la longueur du tatouage. Cette constante a été étudiée [3] et vaut $\frac{\pi^2}{2} < 5$ pour $c \rightarrow +\infty$.



(a) Stratégie pile ou face



(b) Stratégie par vote majoritaire



(c) Stratégie par entrelacement

FIGURE 1 – Distributions empiriques des scores des innocents face à différentes stratégies de collusion

On remarque dans cette étude empirique, dont les résultats sont exhibés en figure 1, que la distribution du score des innocents semble suivre une loi gaussienne centrée de variance m . Cette conjecture s'applique aux trois stratégies pirates considérées ici. Dans la partie suivante, nous montrerons que ce résultat limite est vrai quelque soit la stratégie pirate. De plus, il est important de noter que le score de Tardos discrimine bien les innocents des pirates. En effet, quand la taille du tatouage est respectée, l'innocent avec le score le plus élevé a presque toujours un score plus faible que n'importe lequel des colludés.

2.2.2 • ÉTUDE THÉORIQUE

Nous allons étudier la convergence de la distribution du score d'un innocent vers une gaussienne centrée de variance m , et ce quelle que soit la stratégie de la collusion.

Pour ce faire nous utiliserons la condition de Lyapunov pour la généralisation du théorème central limite. Ce critère s'applique à une suite de variables indépendantes $(X_n)_{1 \leq n \leq 0}$ définies sur un même espace de probabilités d'espérance finie μ_n et d'écart-type fini σ_n . Notons $s_n^2 = \sum_{i=0}^n \sigma_i^2$ et

$$Z_n = \frac{1}{s_n} \sum_{i=0}^n (X_i - \mu_i). \text{ Avec ces notations la condition de Lyapunov donne :}$$

$$\exists \delta > 0, \frac{1}{s_n^{2+\delta}} \sum_{i=0}^n \mathbb{E}(|X_i - \mu_i|^{2+\delta}) \xrightarrow[n \rightarrow \infty]{\text{--}} 0 \Rightarrow Z_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Notons $U_i = U(Y_i, X_{ji}, p_i)$ avec p_i déterministe et un utilisateur j innocent. On rappel que Y_i et X_{ji} sont indépendants et que X_{ji} est distribué selon une loi de Bernoulli de paramètre p_i .

Soit $i \in [|1, m|]$,

$$\mathbb{E}(U_i^4 \mid Y_i = 1) = p_i \frac{(1-p_i)^2}{p_i^2} + (1-p_i) \frac{p_i^2}{(1-p_i)^2} = \frac{(1-p_i)^2}{p_i} + \frac{p_i^2}{1-p_i}.$$

$$\text{De même, } \mathbb{E}(U_i^4 \mid Y_i = 0) = \frac{(1-p_i)^2}{p_i} + \frac{p_i^2}{1-p_i}.$$

$$\text{D'où, } \mathbb{E}(U_i^4) = \frac{(1-p_i)^2}{p_i} + \frac{p_i^2}{1-p_i}.$$

Or $p_i \in [t, 1-t]$, donc $\mathbb{E}(U_i^4) \leq \frac{(1-t)^2}{t} + \frac{t^2}{1-t} = M$. On obtient ainsi $\frac{1}{m^2} \sum_{i=0}^m \mathbb{E}(U_i^4) \leq \frac{M}{m} \xrightarrow[m \rightarrow \infty]{\text{--}} 0$.

Le critère de Lyapunov est donc vérifié pour $\frac{S}{\sqrt{m}}$ pour $\delta = 2$. En effet, rappelons que les U_i sont indépendants, centrés et réduits. On en déduit donc

$$\frac{S}{\sqrt{m}} \xrightarrow[m \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

Notons que forcer $p_i \in [t, 1-t]$ est crucial pour ce critère de convergence.

$$\text{Sinon, } \frac{1}{m} \sum_{i=0}^m \mathbb{E}(U_i^4) = \frac{1}{m} \sum_{i=0}^m \frac{(1-p_i)^2}{p_i} + \frac{p_i^2}{1-p_i} \xrightarrow[m \rightarrow \infty]{\text{LFGN}} \mathbb{E}(f(Z)),$$

$$\text{avec } Z \text{ de loi à densité } p \mapsto \frac{1}{\pi \sqrt{p(1-p)}} \text{ et } f : x \mapsto \frac{(1-x)^2}{x} + \frac{x^2}{1-x}.$$

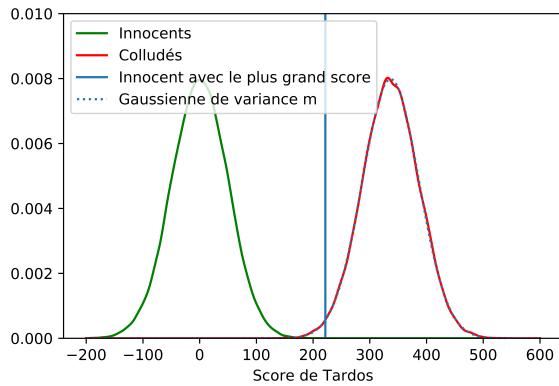
Puis, par symétrie par rapport à $\frac{1}{2}$, on obtient $\frac{1}{m} \sum_{i=0}^m \mathbb{E}(U_i^4) \xrightarrow[m \rightarrow \infty]{\text{--}} 2\mathbb{E}(\tilde{f}(Z))$, avec $\tilde{f} : x \mapsto \frac{(1-x)^2}{x}$.

Or $x \mapsto \frac{(1-x)^2}{\pi x \sqrt{x(1-x)}}$ n'est pas intégrable en 0, soit $\mathbb{E}(\tilde{f}(Z)) = \infty$, ce qui ne permet plus d'avoir le critère de Lyapunov.

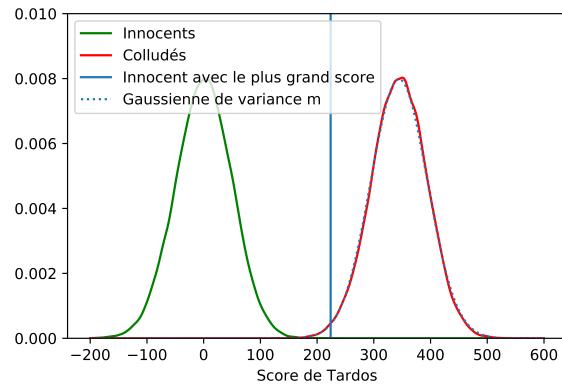
2.3 ÉTUDE DE LA DISTRIBUTION DU SCORE D'UN PIRATE

2.3.1 • ÉTUDE EMPIRIQUE

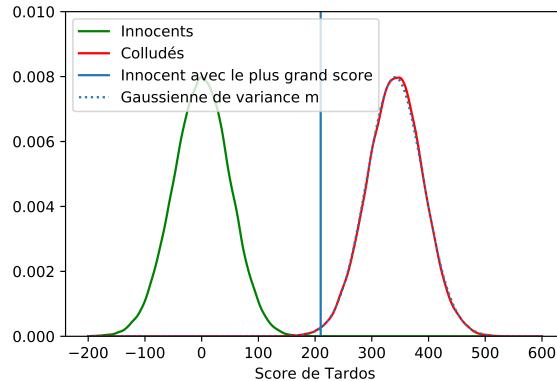
Pour cette première partie de l'étude, les paramètres de l'étude empirique de la distribution du score des innocents ont été conservés. Pour obtenir, une distribution sur les scores des pirates, $N = 10^4$ collusions ont été tirées uniformément.



(a) Stratégie pile ou face



(b) Stratégie par vote majoritaire



(c) Stratégie par entrelacement

FIGURE 2 – Distributions empiriques des scores des innocents et des colludés avec différentes stratégies

On remarque dans cette étude empirique, dont les résultats sont exhibés en figure 2, que la distribution du score des pirates semble suivre une loi gaussienne de variance m . Cette conjecture s'applique aux trois stratégies pirates considérées ici. Dans la partie suivante nous montrerons que ce résultat limite est vrai quelque soit la stratégie pirate. Cependant la moyenne et variance de cette distribution normale dépend de la stratégie employée.

Dans cette deuxième partie de l'étude empirique des scores des pirates, les paramètres précédents sont conservés à l'exception de m qui varie. Comme on peut remarquer sur les deux études précédentes les trois stratégies testées n'influent pas beaucoup sur les résultats. Ceci nous permet de faire cette étude empirique avec une seule stratégie, celle par vote majoritaire.

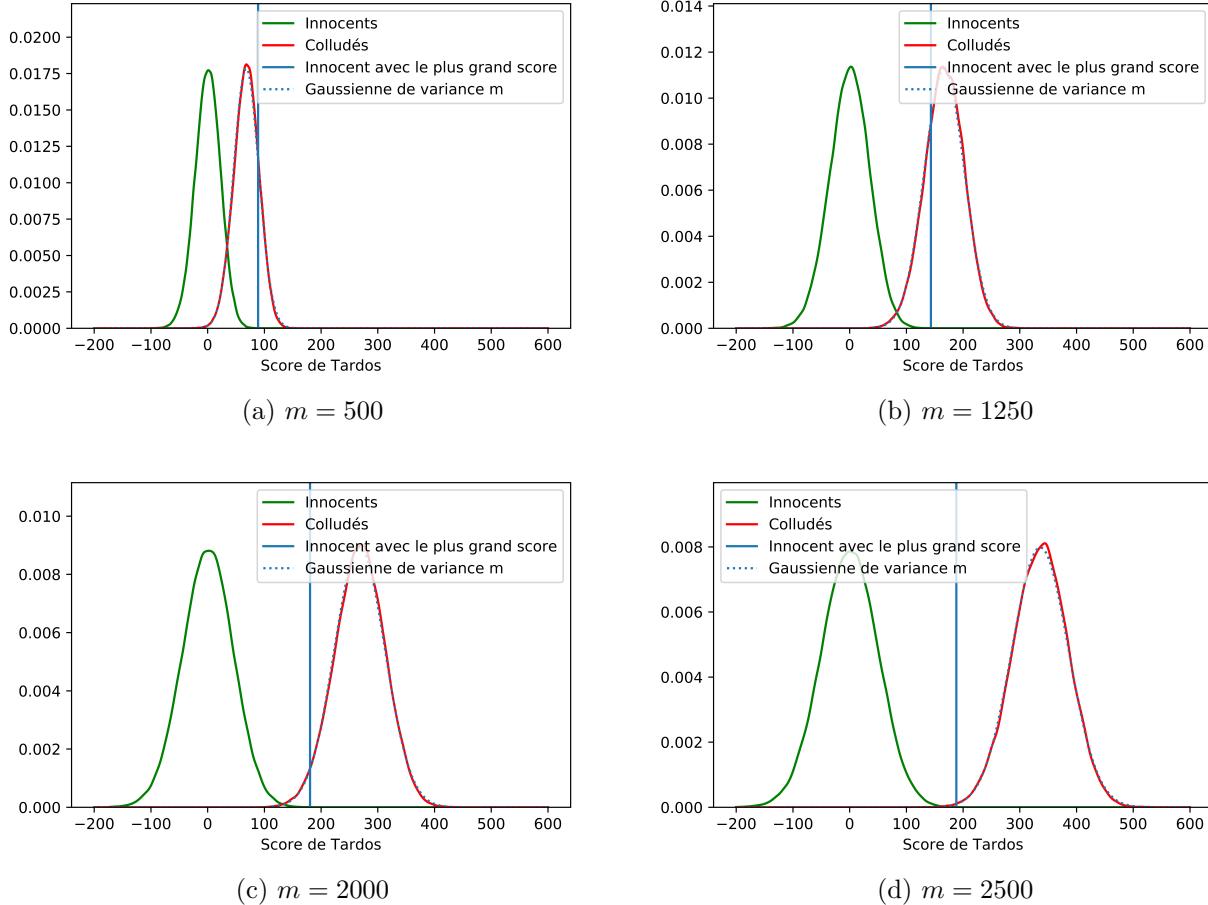


FIGURE 3 – Distributions empiriques des scores des innocents et des pirates avec différentes tailles m de tatouage

Les résultats montrés en figure 3 mettent en avant le rôle de la longueur du tatouage m . En effet, quand cette dernière est trop petite ($m = 500$ dans notre cas) les gaussiennes du score des innocents et des pirates ont une importante intersection et il devient difficile d'accuser un coupable. Notamment, l'innocent avec le score le plus élevé est au dessus de la moyenne de la gaussienne des pirates. Plus m grandit et plus les deux gaussiennes se séparent pour n'avoir qu'une très faible intersection pour $m = 2500$, qui respecte la borne inférieure donnée par Tardos.

Une troisième partie de cette étude se concentre sur la variation de la taille c de la collusion. Rappelons les paramètres qui on servit à tracer les courbes de la figure 4 :

- $n = 10^5$
- $m = 2, 5.10^3$
- $t = \frac{1}{300}$

Les résultats de cette étude sont sans appel, la séparation des distributions des scores de Tardos des pirates et des innocents est très liée au nombre de pirates c dans la collusion. En effet, en triplant la taille de la collusion, on passe de pirates très peu cachés, avec le plus grand score d'innocent dans la queue de la distribution pour $c = 5$, à des pirates presque indétectables, avec une moyenne deux fois

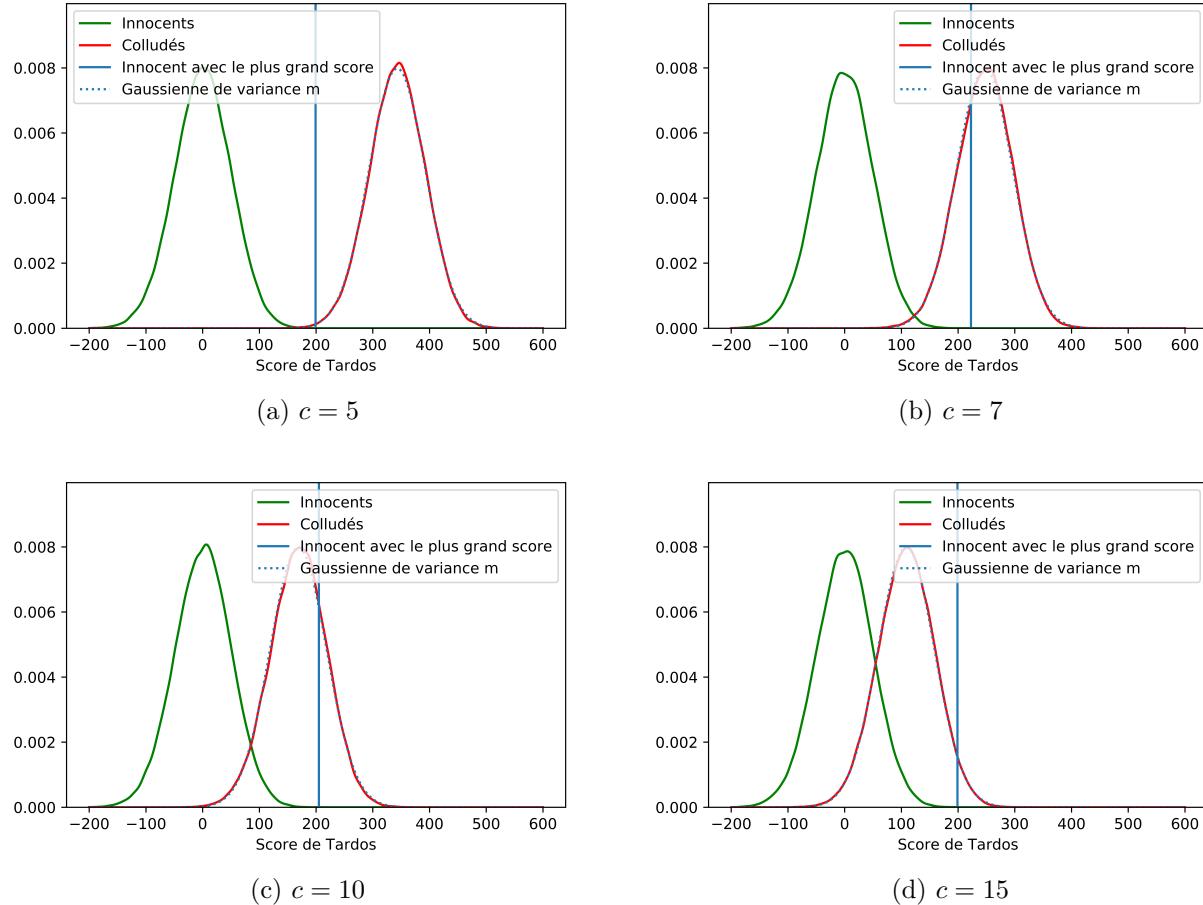


FIGURE 4 – Distributions empiriques des scores des innocents et des pirates avec différentes tailles c de la collusion

plus faible que le plus grand score d'un innocent. Rappelons que le code de Tardos fonctionne sous certaines conditions, notamment que $m \geq kc^2 \log(\frac{n}{\varepsilon})$ avec k une constante dont Furon [3] rappelle les différentes bornes.

2.3.2 • ÉTUDE THÉORIQUE

Nous allons étudier la convergence de la distribution du score des pirates vers une gaussienne dont la moyenne et la variance dépendent de la stratégie pirate employée. Les notations de cette partie reprennent celles introduites par Furon [3].

Notons $U_i = U_i(Y_i, X_{ji}, p_i)$ avec p_i déterministe et j un utilisateur pirate.

Notons $\Sigma_i = \sum_{k=1}^c X_{j_k i}$, la somme des i -ème bits de tatouage des pirates.

Notons $\rho_{c,\sigma} = \mathbb{P}(Y = 1 \mid \Sigma = \sigma)$. On peut supposer que cette valeur ne dépend pas de la position i du bit de tatouage car le code de Tardos s'applique de manière indépendante à chaque bit de tatouage, les pirates ne gagnent rien en faisant évoluer leur stratégie selon les bits de tatouage. Remarquons que $(\rho_{c,\sigma})_{0 \leq \sigma \leq c}$ décrit entièrement la stratégie pirate ρ d'une collusion de taille c .

Étudions $\Pi(p) = \mathbb{P}(Y = 1 \mid p)$.

$$\Pi(p) = \sum_{\sigma=0}^c \mathbb{P}(Y = 1 \mid \Sigma = \sigma) \mathbb{P}(\Sigma = \sigma) = \sum_{\sigma=0}^c \rho_{c,\sigma} \binom{c}{\sigma} p^\sigma (1-p)^{c-\sigma}$$

On obtient notamment que Π est un polynôme de degré au plus c et que $\forall p \in [t, 1-t], p^c \leq \Pi(p) \leq 1 - (1-p)^c$.

Avant de se lancer dans les calculs d'espérances, il nous faut introduire

$$\Pi_x(p) = \mathbb{P}(Y = y \mid X_{ji} = x, p_i).$$

$$\Pi_x(p) = \sum_{\sigma=x}^{c-1+x} \rho_{c,\sigma} \binom{c-1}{\sigma-x} p^{\sigma-x} (1-p)^{c-1-\sigma+x}$$

Par conditionnement, on obtient $\Pi = p\Pi_1 + (1-p)\Pi_0$.

$$\text{Un calcul algébrique donne : } \begin{cases} \Pi_1(p) = \Pi(p) + \frac{1-p}{c} \Pi' \\ \Pi_0(p) = \Pi(p) - \frac{p}{c} \Pi' \end{cases}$$

Calculons l'espérance de U_i .

$$\mathbb{E}(U_i) = \sum_{x,y \in \{0,1\}} U(x, y, p_i) \mathbb{P}(Y = y \mid X_{ji} = x, p_i) \mathbb{P}(X_{ji} = x \mid p_i) = \\ U(0, 0, p_i)[1 - \Pi_0(p_i)][1 - p_i] + U(0, 1, p_i)[\Pi_0(p_i)][1 - p_i] + U(1, 0, p_i)[1 - \Pi_1(p_i)][p_i] + U(1, 1, p_i)[\Pi_1(p_i)][p_i] =$$

$$\sqrt{\frac{p_i}{1-p_i}}[1 - p_i][1 - \Pi_0(p_i)] - \sqrt{\frac{p_i}{1-p_i}}[1 - p_i]\Pi_0(p_i) - \sqrt{\frac{1-p_i}{p_i}}[p_i][1 - \Pi_1(p_i)] + \sqrt{\frac{1-p_i}{p_i}}[p_i]\Pi_1(p_i) = \\ 2\sqrt{p_i(1-p_i)}[\Pi_1(p_i) - \Pi_0(p_i)] = \frac{2}{c}\sqrt{p_i(1-p_i)}\Pi'(p_i)$$

Notons $\mu : p \mapsto \frac{2}{c}\sqrt{p(1-p)}\Pi'(p)$, alors $\boxed{\mathbb{E}(U_i) = \mu(p_i)}$.

Pour la suite nous noterons $\bar{\mu} = \mathbb{E}[\mu(X)]$ où X est une variable aléatoire réelle de densité

$$f(p) = \frac{Cte}{\sqrt{p(1-p)}} \mathbb{1}_{[t, 1-t]}(p) \text{ avec } Cte = \frac{1}{2[\text{Arcsin}(\sqrt{1-t}) - \text{Arcsin}(\sqrt{t})]}.$$

$$\bar{\mu} = \int_t^{1-t} \mu(p) f(p) dp = \frac{2Cte}{c} \int_t^{1-t} \Pi'(p) dp = \frac{2Cte}{c} (\Pi(1-t) - \Pi(t)) \xrightarrow[t \rightarrow 0]{} \frac{2}{\pi c}$$

Calculons l'espérance de U_i^2 .

$$\mathbb{E}(U_i^2) = \sum_{x,y \in \{0,1\}} U(x, y, p_i)^2 \mathbb{P}(Y = y \mid X_{ji} = x, p_i) \mathbb{P}(X_{ji} = x \mid p_i) =$$

$$p_i[1 - \Pi_0(p_i)] + p_i\Pi_0(p_i) + [1 - p_i][1 - \Pi_1(p_i)] + [1 - p_i]\Pi_1(p_i) = 1$$

Notons $\sigma^2 : p \mapsto 1 - \frac{4}{c^2}p(1-p)\Pi'^2(p)$, alors $\boxed{\text{Var}(U_i) = \sigma^2(p_i)}$.

Étudions $s_m^2 = \sum_{i=1}^m \sigma^2(p_i)$. Soit X une variable aléatoire réelle de densité $f(p) = \frac{Cte}{\sqrt{p(1-p)}} \mathbb{1}_{[t, 1-t]}(p)$

avec $Cte = \frac{1}{2[\text{Arcsin}(\sqrt{1-t}) - \text{Arcsin}(\sqrt{t})]}$. Comme $\sigma^2 f$ est continue sur $[t, 1-t]$ compact, alors

$$\mathbb{E}[\sigma^2(X)] < +\infty.$$

Rappelons que les p_i sont indépendant identiquement distribués selon la loi de X . La loi forte des grands nombres donne $\frac{1}{m} \sum_{i=1}^m \sigma^2(p_i) \xrightarrow{p.s.} \mathbb{E}[\sigma^2(X)] \stackrel{\text{noté}}{=} \bar{\sigma}^2$.

On a donc $\boxed{s_m^4 \sim \bar{\sigma}^4 m^2}$.

Nous allons utiliser le critère de Lyapunov pour $\delta = 2$, afin de montrer la convergence de $\frac{S}{\sqrt{m}} = \frac{\sum_{i=1}^m U_i}{\sqrt{m}}$ vers une gaussienne. Pour ce faire, il nous faut étudier $\mathbb{E}[(U_i - \mu_i)^4]$.

$$\mathbb{E}[(U_i - \mu_i)^4] = \sum_{x,y \in \{0,1\}} U(x,y,p_i)^4 \mathbb{P}(Y = y \mid X_{ji} = x, p_i) \mathbb{P}(X_{ji} = x \mid p_i) =$$

$$\left[\sqrt{\frac{p_i}{1-p_i}} - \frac{2}{c} \sqrt{p_i(1-p_i)} \Pi'(p_i) \right]^4 [1 - \Pi_0(p_i)][1 - p_i] + \left[\sqrt{\frac{p_i}{1-p_i}} + \frac{2}{c} \sqrt{p_i(1-p_i)} \Pi'(p_i) \right]^4 [\Pi_0(p_i)][1 - p_i] +$$

$$\left[\sqrt{\frac{1-p_i}{p_i}} + \frac{2}{c} \sqrt{p_i(1-p_i)} \Pi'(p_i) \right]^4 [1 - \Pi_1(p_i)][p_i] + \left[\sqrt{\frac{1-p_i}{p_i}} - \frac{2}{c} \sqrt{p_i(1-p_i)} \Pi'(p_i) \right]^4 [\Pi_1(p_i)][p_i]$$

On déduit de cette expression que $\mathbb{E}[(U_i - \mu(p_i))^4]$ est un fonction continue en p sur $[t, 1-t]$ compact. Donc $\exists M > 0$, $\forall i$, $\mathbb{E}[(U_i - \mu(p_i))^4] \leq M$.

Puis $\frac{1}{s_m^{2+\delta}} \sum_{i=1}^m \mathbb{E}[(U_i - \mu(p_i))^{2+\delta}] \leq \frac{mM}{s_m^4} \xrightarrow[m \rightarrow +\infty]{} 0$. Le critère de Lyapunov est vérifié, donc

$$\frac{1}{s_m} \sum_{i=1}^m [(U_i - \mu(p_i))] \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Enfin, le théorème de Slutsky donne
$$\boxed{\frac{S - m\bar{\mu}}{\sqrt{m\bar{\sigma}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)}.$$

3

ESTIMATION DE LA P-VALEUR DU SCORE DE TARDOS

Comme le seuil de Tardos n'est pas explicite dans un cas pratique, il faut estimer la probabilité qu'un utilisateur j avec un score S_j soit innocent. Pour cela, on calcule la p-valeur du score Tardos. On cherche donc dans cette partie à estimer $q_x = \mathbb{P}(S \geq x | p, Y)$. Nous remarquons que cette valeur dépend de la marque pirate trouvée et doit donc être calculée à chaque nouveau fichier pirate.

3.1 MONTE-CARLO NAÏF

L'approche Monte Carlo naïve de l'estimation de ce paramètre se base uniquement sur la loi forte des grands nombres. L'estimateur obtenu pour N simulations est le suivant :

$$\hat{p}_s = \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{S(X_k) \geq s} \quad \text{avec} \quad X_k \stackrel{iid}{\sim} [\text{Bernoulli}(p_i)]_{1 \leq i \leq m}$$

Cependant cette technique nécessite un nombre moyen de point d'échantillonage inversement proportionnelle à la p-valeur à trouver pour avoir un unique point supérieur au score limite. De plus, le Théorème Centrale Limite ne nous garantie une convergence qu'en $\frac{1}{\sqrt{N}}$. Ainsi, quand on veut estimer des p-valeurs de l'ordre de 10^{-9} , il faut faire 10^{18} simulations pour avoir un résultat fiable. Cette quantité de simulations n'est pas implémentable informatiquement.

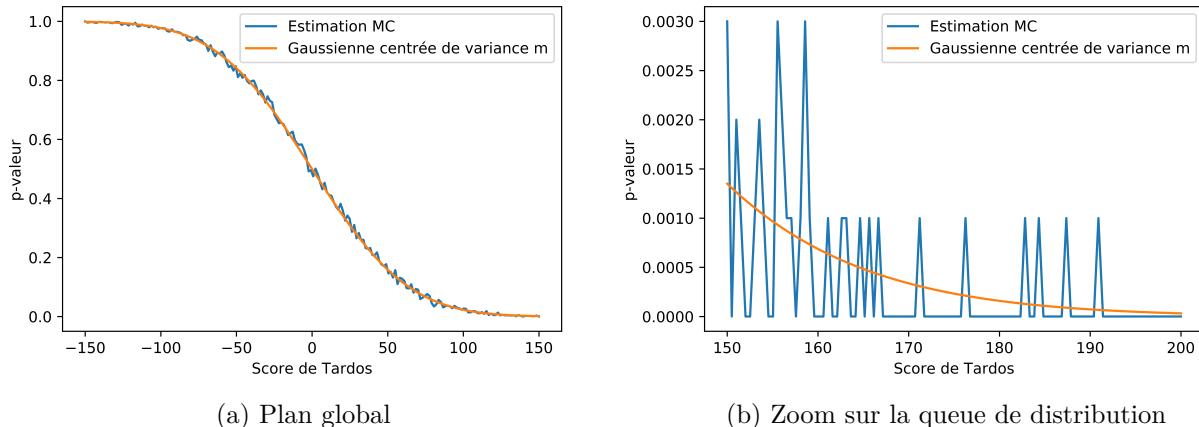


FIGURE 5 – Calcul de la p-valeur du score de Tardos par Monte-Carlo naïf

On peut voir une illustration de ce phénomène sur la figure 5. Ces courbes ont été obtenues avec les paramètres suivants : $N = 1000$ et $m = 2500$. On observe que cet algorithme fonctionne bien pour la majorité de la distribution, comme on peut le voir sur le plan global. Cependant, en queue de distribution, notre estimateur n'a même plus un comportement continu. Celui-ci fait des sauts pour des cas contingents et peut varier du simple au triple pour une variation de score de 1. De plus, l'estimateur se stabilise à 0 pour $s \geq 192$ alors que la probabilité attendue est de l'ordre de 10^{-4} , ce qui n'est donc pas satisfaisant dans notre contexte où l'on veut estimer une très faible p-valeur.

3.2 ADAPTIVE MULTILEVEL SPLITTING

Pour pouvoir calculer avec précision la p-valeur d'événements rares, nous avons fait appel à la méthode d'*Adaptive Multilevel Splitting*, notamment présentée par Cérou [2]. L'idée derrière cette méthode est de simuler la distribution d'origine de manière séquentielle. À chaque étape les échantillons loin de l'objectif de score sont clonés sur des points plus proches puis de les décorrélés par un processus de sauts aléatoires.

3.2.1 • PSEUDO-CODE

Algorithme 1 : Adaptive Multilevel Splitting

```

1 Hyperparamètres :  $N_{simulation}$ ,  $J$ ,  $T$ 
2 Données :  $p$ ,  $y$ ,  $S_{lim}$ 
3  $X \leftarrow$  matrice de tatouages de  $N_{simulation}$  innocents tirés grâce à  $p$ 
4  $S \leftarrow$  score de Tardos de  $X$ 
5 Trier  $X$  selon le score des tatouages associés
6 Trier  $S$ 
7  $Seuil \leftarrow$  médiane( $S$ )
8  $Compteur \leftarrow 0$ 
9 tant que  $Seuil < S_{lim}$  faire
10    $Compteur \leftarrow Compteur + 1$ 
11   pour chaque  $j \in [|0, N_{simulation}/2 - 1|]$  faire
12      $X[j] \leftarrow X[N_{simulation}/2 + j]$ 
13      $S[j] \leftarrow S[N_{simulation}/2 + j]$ 
14   pour chaque  $k \in [|1, T|]$  faire
15      $I \leftarrow J$  indices tiré aléatoirement sans remise entre 0 et  $m - 1$ 
16     pour chaque  $j \in [|0, N_{simulation}/2|]$  faire
17        $\tilde{X} \leftarrow$  copie de  $X[j]$ 
18       pour chaque  $i \in I$  faire
19          $\tilde{X}[i] \leftarrow \text{Bernoulli}(p[i])$ 
20        $\tilde{S} \leftarrow$  score de Tardos de  $\tilde{X}$ 
21       si  $\tilde{S} \geq Seuil$  alors
22          $X[j] \leftarrow \tilde{X}$ 
23          $S[j] \leftarrow \tilde{S}$ 
24    $Seuil \leftarrow$  médiane( $S$ )
25  $q \leftarrow$  proportion de valeurs de  $S$  supérieures à  $S_{lim}$ 
26 Renvoyer :  $q \left( \frac{1}{2} \right)^{Compteur-1}$ 
```

La méthode d'*Adaptive Multilevel Splitting* suit le procédé décrit par l'algorithme 1. Le principe de base consiste à générer, à chaque étape i , la distribution $S \mid S \geq Seuil_i$ en choisissant $Seuil_i$ de

telle sorte que $\mathbb{P}(S \geq Seuil_i \mid S \geq Seuil_{i-1}) = \frac{1}{2}$. Pour générer cette distribution c'est un algorithme de type Metropolis-Hastings qui est utilisé.

Notons que dans cet algorithme les hyperparamètres sont très importants et un bon résultat dépend entièrement de leurs réglages.

En effet, si $N_{simulation}$ est trop faible, nous n'entrerons pas dans une zone de confiance à chaque étape et l'erreur s'accumulera. Généralement, ce paramètre est fixé selon la puissance de calcul.

De même, si T est trop faible, alors les points clonés ne seront pas suffisamment décorrélatés de leur origine pour prétendre représenter la distribution. En effet, ce paramètre vient de l'algorithme de Metropolis-Hastings qui nous assure la convergence de la distribution empirique vers la distribution théorique, mesure stationnaire de la chaîne de Markov créée. Ainsi trop peu d'itérations ne nous permet plus d'affirmer que nous sommes proche de la distribution théorique.

Enfin, J a un double rôle. Ce paramètre ne peut être ni trop grand ni trop faible. En effet, un J trop élevé induirait beaucoup de rejet par le critère de Metropolis-Hastings et donc finalement peu de sauts et peu de décorrélation. À l'opposé, le saut proposé par un J petit sera presque tout le temps accepté mais ne génère presque pas de décorrélation.

3.2.2 • RÉSULTATS EMPIRIQUES

Les estimations de p-valeurs de cette étude ont été réalisées avec les hyperparamètres suivants :

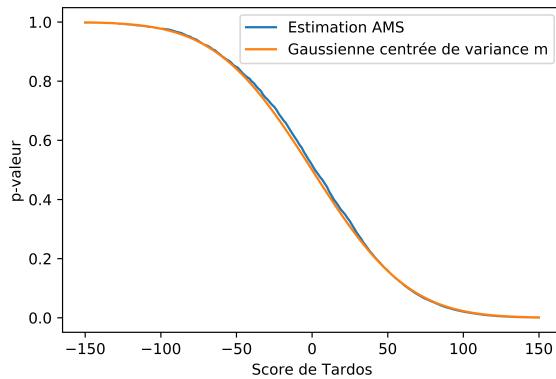
- $N_{simulation} = 3000$,
- $J = 5$,
- $T = 40$.

On remarque, tout d'abord, que l'estimation est beaucoup plus régulière, surtout en fin de distribution. En effet, cela est prévisible parce que le modèle n'est calculé qu'une seule fois et complète les étapes en fonctions des résultats précédents. Ce qui est plus important c'est que notre distribution ne devient pas stationnaire de valeur nulle pour des événements rares avec des scores supérieurs à 180, contrairement à ce que nous renvoyait l'estimateur Monte-Carlo naïf.

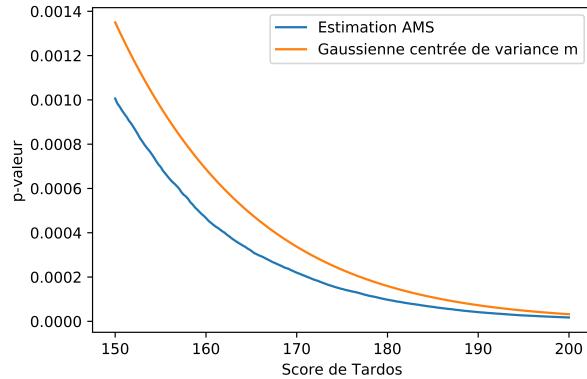
On observe tout de même un écart entre l'estimation AMS et la gaussienne centrée de variance m vers la queue de la distribution, alors que le score des innocents est censé tendre vers cette loi. En effet la pente du logarithme de la p-valeur semble être légèrement plus négative que celle de la gaussienne. Plusieurs facteurs peuvent expliquer cet écart. Premièrement, la loi du score d'un innocent à un support fini et donc la différence entre la gaussienne centrée de variance m est d'autant plus importante qu'on arrive dans des queues de la distribution. Sachant que sur la représentation de la p-valeur ces écarts se somment, contrairement à une représentation de densité comme on a pu en observer dans les parties précédentes. Deuxièmement, la distribution gaussienne est une limite quand $m \rightarrow \infty$ sur le score normalisé $\frac{S}{\sqrt{m}}$. Ceci peut nous laisser penser que la limite n'est pas encore atteinte mais aussi que des termes négligeables devant \sqrt{m} peuvent encore avoir des effets.

Quant au réglage des hyperparamètres, on obtient la même estimation pour $T = 1000$, ce qui signifie que les points clonés sont suffisamment décorrélatés pour $T = 40$. Le J est donc aussi bien réglé puisque la décorrélation à lieu.

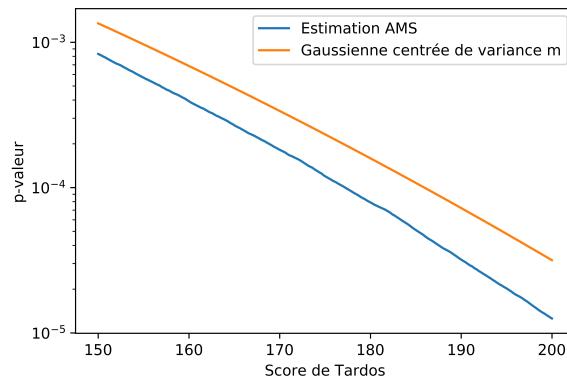
Cette estimation de la p-valeur par *Adaptive Multilevel Splitting* est donc préférable à une approche Monte-Carlo naïve mais aussi à une réduction à la distribution gaussienne limite.



(a) Plan global



(b) Zoom sur la queue de distribution



(c) Zoom sur la queue de distribution avec une échelle logarithmique

FIGURE 6 – Calcul de la p-valeur du score de Tardos par AMS

4

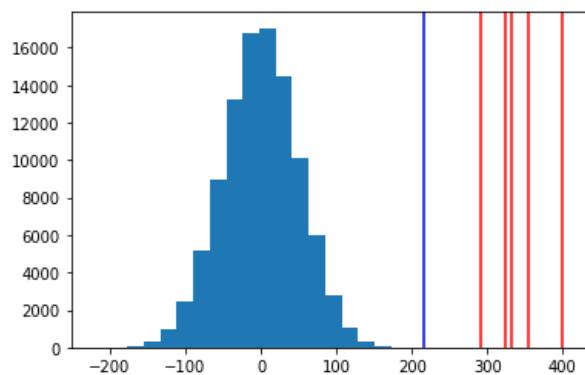
STRATÉGIE POUR ACCUSER UN DEUXIÈME PIRATE

Dans cette partie on suppose qu'un pirate est connu, il a par exemple reconnu les faits ou la probabilité d'innocence est suffisamment faible pour l'inculper. Le but est de tirer partie de cette information pour accuser un deuxième pirate, s'il y en a un et éventuellement le reste de la collusion. Pour ce faire, nous avons essayé deux fonctions de score différentes, toutes les deux indépendantes de la stratégie de la collusion.

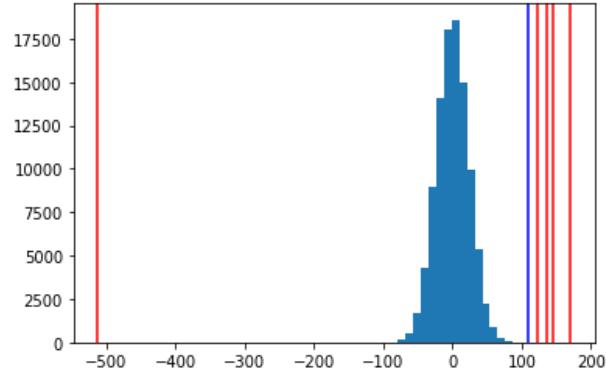
4.1 DESCRIPTION DES STRATÉGIES ENVISAGÉES

4.1.1 • STRATÉGIE DE SOMME DES SCORES

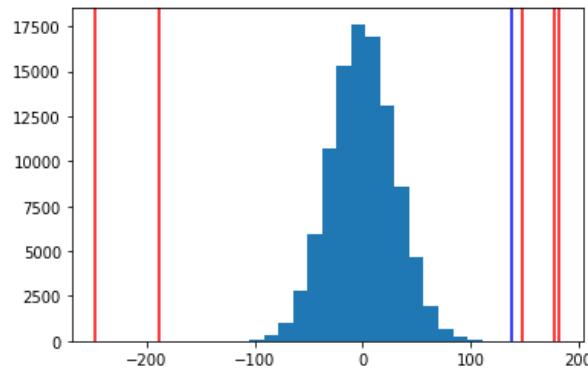
Pour la première stratégie, nous avons proposé un score équidistant du premier accusé et de la marque pirate. Ce score se définit comme suit : $\frac{1}{\sqrt{2}}[S(Y, X_{j.}, p) + S(X^c, X_{j.}, p)]$. Ce score garde les propriétés d'être centré et réduit pour chaque bit de tatouage. Celui-ci pénaliserait pareillement les 0 et les 1, hors de la *marking assumption*.



(a) Score de Tardos initiaux



(b) Score après accusation d'un pirate



(c) Score après accusation de deux pirates

FIGURE 7 – Répartition des nouveaux scores pour des accusations successives

Après avoir fait des essais nous nous sommes rendus compte que ce score cachait bien certains coupables et pouvait inculper des innocents. En effet, des innocents pouvaient avoir un score relatif à l'accusé plus grand qu'un autre coupable. Ainsi, par addition, un innocent pouvait devenir accusable.

Les résultats empiriques de la figure 7 ont été réalisés avec les paramètres suivants :

- $n = 10^5$
- $m = 2,5 \cdot 10^3$
- $t = \frac{1}{300}$

La figure 7 montre que certains pirates ont maintenant des scores très négatifs et sont donc peu accusables.

4.1.2 • STRATÉGIE DES BITS "CACHÉS"

La deuxième stratégie à laquelle nous avons pensé était de calculer le score seulement sur les bits où les codes des accusés et la marque pirate coïncidaient. Ces bits sont encore cachés pour la collusion et vérifie la *marking assumption*.

Là, nous avons eu un autre problème, parce que ce score cachait la pénalisation des pirates qui avaient la même valeur pour un bit, c'est-à-dire, il n'était plus pénalisés parce que ce bit n'était pas comptabilisé.

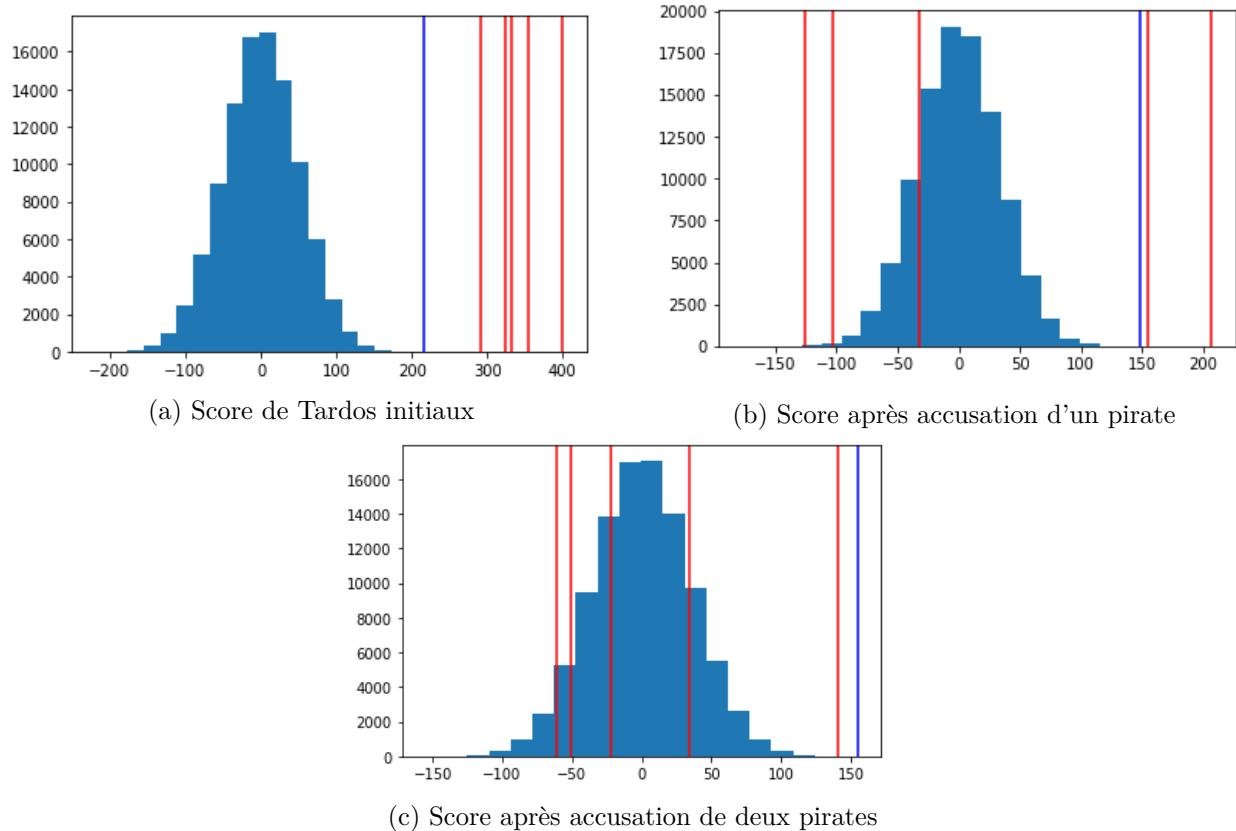


FIGURE 8 – Répartition des nouveaux scores pour des accusations successives

Les résultats empiriques de la figure 8 ont été réalisés avec les paramètres suivants :

- $n = 10^5$
- $m = 2,5 \cdot 10^3$
- $t = \frac{1}{300}$

La figure 8 montre notamment que le troisième utilisateur à être accusé est en fait un innocent, ce qui ne convient pas.

4.2 OPTIMALITÉ DU SCORE DE TARDOS POUR L'ACCUSATION D'UN DEUXIÈME PIRATE

Les études empiriques menées dans les sous-parties précédentes nous conduisent à penser que l'on apprend pas beaucoup d'information en connaissant un membre de la collusion. Le score de Tardos est toujours le score le plus fiable pour la distribution des p_i considérée pour accuser un deuxième pirate. Nous allons maintenant démontrer ce résultat.

4.2.1 • DÉFINITION DU SCORE

Le score considéré prend en compte le tatouage d'un coupable X^c , la marque pirate Y et la distribution des p_i .

$$\text{Ainsi, } S = \sum_{i=0}^m U_i \text{ avec } U_i = U(X_i, X_i^c, Y_i, p_i).$$

Pour garder des résultats exploitables, cette fonction de score doit être centrée pour les innocents ce qui nous conduit à la condition suivante :

$$\forall X_i^c, Y_i \in \{0, 1\}, \forall p_i, p_i U(1, X_i^c, Y_i, p_i) + (1 - p_i) U(0, X_i^c, Y_i, p_i) = 0 \quad (1)$$

De plus, nous noterons $\tilde{U}(p) = U(0, 1, 0, p) - U(0, 1, 1, p) = U(0, 0, 0, p) - U(0, 0, 1, p)$. L'égalité de ces deux quantités provient de la structure de symétrie du score. En effet, les bits 1 et 0 doivent jouer des rôles équivalents dans un score optimal. Nous remarquons que dans le cas du score défini par Tardos et réduit par Skorić $\tilde{U}(p) = 2\sqrt{\frac{p}{1-p}}$.

4.2.2 • NOTATIONS POUR LA STRATÉGIE DE LA COLLUSION

Nous reprendrons les notations définies pour l'étude théorique de la distribution du score d'un pirate. C'est à dire :

- $\Pi(p) = \mathbb{P}(Y = 1 | p)$
- $\Pi_x(p) = \mathbb{P}(Y = 1 | X_i^c = x, p)$

Dans la continuité nous noterons :

$$\begin{aligned} \Pi_{x_1, x_2}(p) &= \mathbb{P}(Y = 1 | X_i^{c_1} = x_1, X_i^{c_2} = x_2, p) \\ \Pi_{x_1, x_2}(p) &= \sum_{\sigma=x_1+x_2}^{c-2+x_1+x_2} \binom{c-2}{\sigma - x_1 - x_2} p^{\sigma - x_1 - x_2} (1-p)^{c-2-\sigma+x_1+x_2} \end{aligned}$$

Avec ces notations et des calculs élémentaires nous obtenons les relations suivantes :

$$\Pi_{1,0} - \Pi_{0,0} = \frac{\Pi'_0}{c-1} \quad (2)$$

$$\Pi_{1,1} - \Pi_{1,0} = \frac{\Pi'_1}{c-1} \quad (3)$$

$$p\Pi'_1 + (1-p)\Pi'_0 = \frac{c-1}{c}\Pi' \quad (4)$$

4.2.3 • CALCUL DE LA MOYENNE DU SCORE D'UN DEUXIÈME PIRATE

Prenons $U_i = U(\tilde{X}, X_i^c, Y_i, p_i)$ pour \tilde{X} pirate.

Calculons l'espérance de U_i .

$$\begin{aligned} \mathbb{E}(U_i) &= \sum_{x,x^c,y \in \{0,1\}} U(x, x^c, y, p_i) \mathbb{P}(Y = y \mid \tilde{X} = x, X_i^c = x^c, p_i) \mathbb{P}(\tilde{X} = x, X_i^c = x^c \mid p_i) = \\ &\sum_{x,x^c,y \in \{0,1\}} U(x, x^c, y, p_i) \mathbb{P}(Y = y \mid \tilde{X} = x, X_i^c = x^c, p_i) \mathbb{P}(\tilde{X} = x \mid p_i) \mathbb{P}(X_i^c = x^c \mid p_i) = \\ &p_i^2 \Pi_{1,1}(p_i) U(1, 1, 1, p_i) + p_i^2 [1 - \Pi_{1,1}(p_i)] U(1, 1, 0, p_i) + p_i(1 - p_i) \Pi_{1,0}(p_i) U(1, 0, 1, p_i) + \\ &p_i(1 - p_i) [1 - \Pi_{1,0}(p_i)] U(1, 0, 0, p_i) + p_i(1 - p_i) \Pi_{1,0}(p_i) U(0, 1, 1, p_i) + p_i(1 - p_i) [1 - \Pi_{1,0}(p_i)] U(0, 1, 0, p_i) + \\ &(1 - p_i)^2 \Pi_{0,0}(p_i) U(0, 0, 1, p_i) + (1 - p_i)^2 [1 - \Pi_{0,0}(p_i)] U(0, 0, 0, p_i) \end{aligned}$$

En utilisant 1, on obtient :

$$\mathbb{E}(U_i) = p(1-p)[U(0, 1, 1) - U(0, 1, 0)][\Pi_{1,0} - \Pi_{1,1}] + (1-p)^2[U(0, 0, 1) - U(0, 0, 0)][\Pi_{0,0} - \Pi_{1,0}]$$

Puis en utilisant 3, 2 et 4, on a :

$$\mathbb{E}(U_i) = \frac{1-p_i}{c-1} \tilde{U}(p_i) [p_i \Pi'_1(p_i) + (1-p_i) \Pi'_0(p_i)] = \frac{1-p_i}{c} \tilde{U}(p_i) \Pi'(p_i)$$

Ainsi, par la loi forte des grands nombres, en notant \tilde{S} le score d'un deuxième pirate, on a :

$$\mathbb{E}\left(\frac{\tilde{S}}{m}\right) \xrightarrow[m \rightarrow \infty]{LFGN} \int_t^{1-t} \frac{1-p}{c} \tilde{U}(p) \Pi'(p) f(p) dp = \frac{Cte}{c} \int_t^{1-t} \tilde{U}(p) \sqrt{\frac{1-p}{p}} \Pi'(p) dp$$

$$\text{avec } f(p) = \frac{Cte}{\sqrt{p(1-p)}} \mathbb{1}_{[t,1-t]}(p) \text{ et } Cte = \frac{1}{2[\text{Arcsin}(\sqrt{1-t}) - \text{Arcsin}(\sqrt{t})]}.$$

Cette dernière valeur peut s'interpréter comme un jeu à somme nulle dans lequel les pirates choisissent pour stratégie Π et le distributeur \tilde{U} . Cependant le distributeur joue en premier et dévoile sa stratégie avant que les pirates jouent. En effet, on peut supposer que les articles sur ce sujet et les avancées de la recherche sont publics. Face à ce constat, la stratégie pirate peut exploiter les variations de $p \mapsto \tilde{U}(p) \sqrt{\frac{1-p}{p}}$ pour optimiser leur réponse. Une stratégie d'équilibre de Nash est donc pour le distributeur de rendre les pirates indifférents, c'est à dire que $\tilde{U}(p) \sqrt{\frac{1-p}{p}}$ doit être constant. Puis $\tilde{U}(p)$ est proportionnel à $\sqrt{\frac{p}{1-p}}$ et on retrouve le score de Tardos initial.

Ainsi conserver le score de Tardos est une bonne stratégie pour continuer d'accuser des pirates en connaissant l'un d'entre eux.

CONCLUSION

Pour conclure, le tatouage numérique par score de Tardos est performant. Cette technique permet d'atteindre la borne inférieure en terme de longueur du tatouage nécessaire pour avoir un algorithme ϵ -sécuritaire pour une collusion de taille inférieure à c .

Cette approche probabiliste, avec le choix de la distribution des p_i et de la fonction de score, fait tendre les distributions de score des coupables et des innocents vers des gaussiennes. Peut importe la stratégie pirate la distribution du score des innocents se rapproche d'une loi normale centrée de variance m . La moyenne du score des coupables tends vers $\frac{2m}{\pi c}$ peut importe la stratégie envisagée quand t est suffisamment faible. Quant à la variance, elle dépend de la stratégie choisie.

Pour utiliser le code de Tardos en pratique et accuser un membre de la collusion lors de la découverte d'un fichier pirate contenant une marque pirate, il faut être capable d'estimer la p-valeur du score de Tardos car la taille c de la collusion est inconnue. Pour se faire, l'approche par *Adaptive Multilevel Splitting* est beaucoup plus performante qu'une approche par Monte Carlo naïf, surtout dans la queue de la distribution qui concentre des événements rares.

C'est alors posée la question de l'accusation d'un deuxième pirate, en supposant le premier confirmé est certain. Dans ce cas encore, le score de Tardos sans modification ni prise en compte de l'information donnée par le premier pirate est un bon moyen de discriminer innocent et pirate. Ce score est notamment le seul score centré à renvoyer la même moyenne pour le score des colludés quelque soit la stratégie. D'autres scores pourraient être envisagés sur le principe d'un compromis biais variance.

RÉFÉRENCES

- [1] D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Transactions of Information Theory*, 44, 1998.
- [2] F. Cérou and al. Adaptive multilevel splitting : Historical perspective and recent results. *Chaos, American Institute of Physics*, 2019.
- [3] T. Furon. *Sécurité Multimédia - Partie 1 : Authentification et Insertion de Données Cachées*, chapter 6. Traçage de traîtres. ISTE Editions, 2019.
- [4] B. Skorić and al. Tardos fingerprinting is better than we thought. 2008.
- [5] G. Tardos. Optimal probabilistic fingerprint codes. *STOC'03*, 2003.