ROYAUME DU MAROC

OFFICE MAROCAIN DE LA PROPRIETE INDUSTRIELLE ET COMMERCIALE
---

(19)

المملكة المغربية

المكتب المغربي للملكية الصناعية و التجارية
---

(12) BREVET D'INVENTION

(11) N° de publication :
MA 38026 A1

(51) Cl. internationale :
H04L 12/00

(43) Date de publication :
29.12.2017

(21) N° Dépôt :
38026

(22) Date de Dépôt :
21.04.2015

(71) Demandeur(s) :
UNIVERSITÉ MOHAMMED V DE RABAT, Angle avenue Allal El Fassi et Mfadel Cherkaoui Al Irfane 8007. N.U RABAT (MA)

(72) Inventeur(s) :
SOUIDI Mamoun

(74) Mandataire :
FATIMA ZAOUI

(54) Titre : **METHOE DE DISSIMULATION DE L'INFORMATION PAR CODES CONVOLUTIFS EN STENOGRAPHIE**

MA 38026A1

Abrégé

Notre invention concerne le domaine de la cryptologie, nous introduisons une nouvelle application de la théorie des codes convolutifs en sténographie définis suivant les termes de la théorie des systèmes linéaires pour minimiser la distorsion et augmenter ainsi la performance de la dissimulation.

MA 38026A1

Titre : Méthode de dissimulation de l'information par codes convolutifs en sténographie

Description

Notre invention concerne le domaine de la cryptologie, nous introduisons une nouvelle application de la théorie des codes convolutifs en sténographie définis suivant les termes de la théorie des systèmes linéaires pour minimiser la distorsion et augmenter ainsi la performance de la dissimulation.

Dans l'art antérieur les codes en blocs sont largement utilisés pour la dissimulation d'informations tout en utilisant le processus sténographique. Ces méthodes posent des problèmes de déformation des supports d'informations parfois détectables à l'œil ce qui rend le cryptage défaillant. Notre méthode suggère un protocole steganographique bas´e sur les codes convolutifs définis grâce aux termes de la théorie des systèmes linéaires. Comme définis auparavant, les codes convolutifs peuvent ˆetre donnés par le quadruple (A, B, C, D), et le but ici est de considérer un protocole d'insertion pour un message à l'intérieur d'un message u avec le moins de distorsion possible.

CODES CONVOLUTIFS ET STÉGANOGRAPHIE D'UN POINT DE VUE DE LA THÉORIE DES SYSTÈMES LINÉAIRES.

L. E. Um[1], El M. Souidi[1], H. Jouhari[1], and M. I. García-Planas[2]

[1] Dept. d'Informatique
Université Mohammed V, Morocco
[2] Dept. de Matemàtica Aplicada I
Universitat Politècnica de Catalunya, Spain

Résumé Dans ce travail, nous introduisons une application de la théorie des codes convolutifs en stéganographie. Concrètement nous suggérons un protocole stéganographique basé sur les codes convolutifs definis suivant les termes de la théorie des sytsèmes linéaires.

1 Introduction

Dans ce chapître, nous introduisons une application de la théorie des codes convolutifs en stéganographie. Comme nous le savons déjà, les codes en blocs sont largement utilisés pour la dissimulation d'informations tout en utilisant le processus stéganographique, comme dans [3]. Ici, nous suggérons un protocole stéganographique basé sur les codes convolutifs définis grâce aux termes de la théorie des systémes linéaires. Comme définis auparavant, les codes convolutifs peuvent être donnés par le quadruple (A, B, C, D), et le but ici est de considerer un protocôle d'insertion pour un message m à l'intérieur d'un message u avec le moins de distortion possible.

2 Stéganographie

La stéganographie peut être comparable à la protection de la communication, puisqu'elle est connue comme une technique étant une technique utilisée pour protéger de l'information qui doit être échangée en cachant même son existence, dans des fichiers numériques, que ce soit des photographies, ou bien des vidéogrammes. Connaissant la cryptographie comme la science et la technique derrière la protection de l'information et des messages devant être transmis, l'idée de la stéganographie est de prévenir un observateur mal intentionné de detecter même déjà le besoin pour cette protection au départ, et cela dépend aussi des situations et circonstances, par exemple dans des circonstances où la cryptographie ne peut être utilisée. Parfois, il est aussi possible de combiner ensemble les deux techniques pour protection de la communication et de l'information. L'exemple classique connu pour illustrer l'utilisation d'un schéma stégnographique est le problème d'échange des messages sous la surveillance d'un gardien, [11].

2.1 Charactéristiques d'un schéma stéganographique

A schéma stéganographique est caractérisé par quelques conditions nécessaires et composantes qui sont :

i) le choix d'un support de communication
ii) le message à encapsuler
iii) la fonction d'insertion
iv) la fonction d'extraction
v) la clé stéganographique optionelle de management

Les fonctions d'insertion et d'extraction sont comme leurs noms indiquent les fonctions servant à cacher et extraire le message secret ou l'information. Pour la stéganographie numérique comme dans notre cas, le but est de cacher ou d'encapsuler une séquence de bits dans une couverture numérique, tout en respectant certaines conditions essentielles, telles s'assurer que l'objet de couverture ne montre aucun signe perceptible de distortion.
Etant donné que dans le monde numérique, le choix des objets de couverture est tellement large (fichiers graphiques, messages, etc..), mais est tout de même dicté par la nature de l'information à cacher. Pour ce faire, la performance d'une méthode stéganographique peut être estimée sur un objet de couverture principalement grâce à la distortion moyenne et le ratio d'insertion.

Par exemple, une méthode très populaire utilisée en stéganographie numérique est appelée Least Significant Bit(LSB). Elle consiste à cacher de l'information dans un fichier graphique, en remplaçant les bits les moins significatifs de certains pixels spécifiquement selectionés par des bits de message, de telle manière que ceux-ci soient "visuellement imperceptibles".

Definition 1. *Un schéma stéganographique numérique S de type $[n, k]$ sur un alphabet A est une paire de fonctions :*

$$\begin{aligned} emb &: A^n \times A^k \longrightarrow A^n \\ rec &: \quad A^n \longrightarrow A^k \end{aligned}$$

telle que

$$rec(emb(c, m)) = m \text{ pour tout } c \in A^n \text{ and } m \in A^k,$$

avec m étant le message secret, et c le vecteur de couverture.

Le schéma est noté par : (emb, rec).
Nous avons : $c' = emb(c, m)$ and $rec(c') = m$

Le schéma a les paramètres suivants :

a) la longueur de l'objet de couverture n
b) la capacité d'insertion k
c) le rayon d'insertion ρ, défini par :

$$\rho = \max\{d(c, emb(c, m)) \mid c \in A^n, m \in A^k\}.(d : \text{Hamming distance})$$

d) le nombre moyen de changements de couverture R_a, donné par :

$$R_a = \frac{1}{q^{kn}} \sum d(c, emb(c, m))$$

où $\sharp A = q$

Nous avons la proposition suivante :

Proposition 1. *[10] Soit $S = (emb, rec)$ un schema stéganographique de type $[n, k]$ defini sur l'alphabet A. Donc :*

1) la fonction rec est surjective ;

2) pour un $c \in A^n$, la fonction $emb(c, -) : A^k \longrightarrow A^n$ est injective.

En particulier, $k \leq n$.

Démonstration. La conclusion découle de : $rec(emb(c, m)) = m$.

Etant donné que le but du stegoschéma est de sceller autant d'information que possible, avec le minimum de changement possible, nous avons la définition d'un schéma convenable ;

Definition 2. *Un schéma stéganographique $S = (emb, rec)$ est dit convenable si le nombre de changements produits dans la couverture est le nombre minimum possible permis par la fonction de extraction.*

$S = (emb, rec)$ est convenable si et seulement si

$$d(c, emb(c, m)) = d(c, rec^{-1}(m)), \text{ for all } c \in A^n \text{ and } m \in A^k.$$

Nous avons la proposition suivante :

Proposition 2. *Soit $S = (emb, rec)$ un schéma stéganographique de type $[n, k]$ sur A. Il existe un stegoschéma convenable $S^* = (emb^*, rec)$ du même type $[n, k]$ tel que $R_a(S^*) \leq R_a(S)$ (R_a : nombre moyen de changements au cours de l'insertion).*

3 Stéganographie et Codage

Il existe quelques protocoles stéganographiques interessants, déjà définis á partir de la théorie du codage, en considérant le fait que les codes correcteurs d'erreurs sont utilisés pour detecter et corriger les erreurs durant le transfert de données. Si nous considérons par exemple quelques unes des méthodes impliquant l'existence de la matrice de parité, nous pouvons implementer la méthode du codage de syndrome. Considerant un protocole stéganographique faisant partie du domaine de l'échelle des images du gris, inspiré par [9]. Cette approche suggère de diviser le bloc de couverture en blocs de tailles égales.

Par exemple, considérons le protocole suivant ayant l'objet de couverture v dont les valeurs de LSB sont données par $v = v_0, v_1, \ldots, v_n$ sur $\mathbb{F}_2^n$, le message m à cacher $m = m_0, m_1, \ldots, m_t$ avec $t < n$ sur $\mathbb{F}_2^t$, le code donné par sa matrice de parité H.

Cacher m dans v produit le stégo objet $r = r_0, r_1, \ldots, r_n$, donné par la relation :

$$m = r \times H^t \tag{1}$$

Afin d'extraire m, c'est la même equation 1 qui est utilisée. Après insertion, quelques uns des bits du bloc de couverture sont modifiés (soit 0 ou 1) ; si nous

En effet, decoder un système pour ce type de système consiste à résoudre le suivant :

$$T_\ell \begin{pmatrix} x(0) \\ u \end{pmatrix} = y. \tag{5}$$

Nous rappelons qu'il est usuel de considerer l'etat initial du système $x(0) = 0$, comme dans notre cas par exemple; ainsi, notre nouvelle matrice d'output-observabilité est réduite :

$$\widehat{T}_{\ell-1} = \begin{pmatrix} D & & & \\ CB & D & & \\ CAB & CB & D & \\ \vdots & & \ddots & \ddots \\ CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB\ D \end{pmatrix} \tag{6}$$

Pour pouvoir réaliser l'insertion, le processus consiste en considerer la matrice d'output-observabilité comme la matrice de controle. Le modèle de stéganographie que nous construisons est inspiré par la représentation (A, B, C, D) du code convolutif. En sachant que nous utilisons cette structure sur les codes convolutifs pour le décodage, étape par étape, la stéganographie joue le rôle et sert de base pour la fonction d'insertion. Dans ce cas, nous décidons d'approcher ce problème de manière séquentielle ce qui signifie que pour chacune des étapes du processus stéganographique, à chacune des étapes $t = 1, \dots, \ell$, le protocole consiste à cacher la séquence message, en altérant légèrement la séquence de couverture avec quelques erreurs, de faon à construire la stégo-séquence. Pour ce faire, nous avons besoin de trouver la meilleure séquence correspondant aux bits á modifier qui minimise la mdification, qui correspond au coset leader de la liste des potentiels "vecteurs erreurs $e(t)$"pour encapsuler $m(t)$ dans $u(t)$ par la formule : $u(t) + e(t)$.

Pour ce qui est de la récuperation du message caché, c'est à ce niveau que la matrice d'output-observabilité apparaît explicitement pour résoudre l'equation correspondant. En effet, le processus d'extraction du message caché consiste en encoder la stégo-sequence. Ce qui représente une analogie de la méthode du protocole stéganographique basé sur le syndrome, pour récuperer le message caché, nous extrayons á chacune des étapes de la stégo-sequence, en utilisant le "bloc of matrices", chacune des parties du message caché.

Definition 3. *Le quasi-syndrome noté par s est la valeur à partir de laquelle on choisit la pertubation estimée e pour l'insertion de m dans u. Il est donné par : $s = He$.*

Il peut être lié à la notion de syndrome telle que défini pour les stégo-codes, basés sur le codage en bloc.

L'algorithme qui suit donne la méthode pour le processus d'insertion.

Algorithme d'insertion pour la fonction d'insertion emb
Entrée : Message m, Séquence de couverture u
Sortie : Stégo-séquence St

if rank D is row maximal **then**
  for t :=0 to ℓ **do**
    $s(t) = m(t) - Du(t) - \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k))$;
    list $sysdTableD = \{e(t)\}$;
    compute $\min_{e \in \text{list}} w(e(t))$;
    pick and store one corresponding $e(t)$;
    $St(t) = u(t) + e(t)$;
  end for
  $St = (St(0), \ldots, St(\ell))$
else
    Choose an adequate code with D corresponding
end if

L'algorithme suivant est utilisé pour extraire le message encapsulé.

Algorithme d'extraction pour la fonction d'extraction rec

$m = \widehat{T}_{\ell-1}\left(u + e\right)$
for t :=0 to ℓ **do**
  $m(t) = Du(t) + De(t) + \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k))$;
end for
$m(t) = (m(0), \ldots, m(\ell))$

Example 2. Dans $\mathbb{F}_2$, considérons le code $C(A, B, C, D)$ défini par :

$$A = \begin{pmatrix} 1\,0\,1\,0 \\ 0\,1\,0\,1 \\ 1\,0\,0\,0 \\ 0\,1\,0\,0 \end{pmatrix}, B = \begin{pmatrix} 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \\ 0\,0\,1 \end{pmatrix}, C = \begin{pmatrix} 1\,1\,0\,1 \\ 1\,0\,1\,1 \end{pmatrix}, D = \begin{pmatrix} 0\,1\,0 \\ 0\,1\,1 \end{pmatrix}$$

Nous observons que le système est output-observable.

Considérons que le message á cacher soit $m = (m(0), m(1), m(2))$, et que l'opération soit réalisée section par section, à chacune des étapes, dans une séquence de couverture notée par u.
Let's consider $m = (10,\ 00,\ 01)$

Considérons que connaissant chacune des séquences d'entrées u, nous essaierons de trouver les caractéristiques de chacune des séquences e des bits à modifier qui seront ajoutées à u, afin d'encapsuler le message m. Le quasi-syndrome correspodant est noté par s.

Considerons la matrice de décodage pour les codes convolutifs donnée par :

$$\widehat{T}_{\ell-1} = \begin{pmatrix} D & & & \\ CB & D & & \\ CAB & CB & D & \\ \vdots & & \ddots & \ddots \\ CA^{\ell-1}B & CA^{\ell-2}B & \ldots & CB\ D \end{pmatrix}$$

A chacune des étapes : $0, 1, \ldots, \ell$, nous essaierons d'évaluer notre valeur erreur e.

Nous avons déjà : $m = \widehat{T}_{\ell-1}\left(u + e\right)$

Pour les entrées, nous avons m and u.

Tout d'abord, évaluons le sous-ensemble des erreurs potentielles e; comme formule génerale, m et u sont donnés par : $m(t) = \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k)) + D(u(t) + e(t))$; ce qui signifie que :

$$De(t) = s(t) = m(t) - \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k)) - Du(t)$$

Parcourant tous les cas possibles de syndromes (ou de séquences de couverture), nous établissons le sous-ensemble des erreurs e.

s	e
00	000
	100
10	011
	111
01	001
	101
11	010
	110

Figure 1. Syndrome Table 2

Ainsi, á chacune des étapes, il existe toujours une séquence e qui peut être utilisée pour l'insertion tel que : $w(e) \leq 2$.

Travaillons avec $u = (111\ 010\ 001)$ avec $m = (10,\ 00,\ 01)$.

A l' étape $t = 0$, nous avons : $D(u(0) + e(0)) = m(0)$
Ainsi, $s = (00)$, et $e(0) = (000)$

A l' étape $t = 1$, nous avons : $D(u(1) + e(1)) = m(1) - CB(u(0) + e(0))$
Ainsi, $s = (0\,0)$, nous choisissons $e(0) = (0\,0\,0)$ and $e(1) = (0\,0\,0)$

A l' étape $t = 2$, nous avons : $D(u(2) + e(2)) = m(2) - CAB(u(0) + e(0)) - CB(u(1) + e(1))$
Ainsi, $s = (1\,0)$, et $e(2) = (0\,1\,1)$.

Ensuite, pour $u = (1\,1\,1,\ 0\,1\,0,\ 0\,0\,1)$, nous pouvons cacher $m = (10,\ 00,\ 01)$ avec la séquence : $e = (0\,0\,0,\ 0\,0\,0,\ 0\,1\,1)$.

Pour cette opération, nous cachons 6 bits dans une séquence de couverture de 9-length en changeant 2 bits.

Fonctions d'insertion et d'extraction Dans le but de définir notre code convolutif, pour le stégoschéma correspondant, en fonction de notre protocole, la condition nécessaire suivante est réquise.

Proposition 3. *Soit (A, B, C, D) une représentation d'un code convolutif $\mathcal{C}$, avec $A \in M_\delta(\mathbb{F})$, $B \in M_{\delta\times k}(\mathbb{F})$, $C \in M_{p\times\delta}(\mathbb{F})$, $D \in M_{p\times k}(\mathbb{F})$ (avec $D \neq 0$ et $p = n - k$). Soit $p < k$.*
Une condition nécessaire pour construire un stégoschéma à partir de $\mathcal{C}$ est que le rang de D ($\operatorname{rank} D$) soit maximal par lignes.

Conditions sur les sous-sections modifiées de la couverture Ici, nous avons certaines conditions s'appliquant à la modification de la séquence de couverture pendant l'insertion.

Proposition 4. *Soit (A, B, C, D) une représentation d'un code convolutif $\mathcal{C}$ pour un schéma stéganographique S. Ainsi, à chaque étape t de la séquence de convolution, la séquence à modifier e introduite pour l'insertion de m dans u peut être donné par la formule :*

$$De(t) = m(t) - Du(t) - \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k))$$

Démonstration. Etant donné que l'insertion de m est donnée par : $m = \widehat{T}_{\ell-1}\left(u + e\right)$, et ayant : $\widehat{T}_{\ell-1} = \begin{pmatrix} D & & & \\ CB & D & & \\ CAB & CB & D & \\ \vdots & & \ddots & \ddots \\ CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB\ D \end{pmatrix}$, nous pouvons déduire le résultat.

Borne classique d'imperceptibilité

Proposition 5. *Soit (A, B, C, D) une représentation d'un code convolutif $\mathcal{C}$ pour un schéma stéganographique S. Considérons l'insertion de m dans u, avec la séquence d'erreur e. Supposons que D aît toutes ses colonnes non nulles et distinctes.*
Ainsi, pour chacune des étapes t de la séquence de convolution : $\exists e(t)$ tel que $w(e(t)) \leq \tau - 1$.

Démonstration. Considerant $\widehat{T}_{\ell-1}$ la matrice de contrôle, l'insertion de m est donnée par : $m = \widehat{T}_{\ell-1}\left(u + e\right)$;

à chaque étape t, nous avons : $De(t) = s(t)$.

Considérons D_j les colonnes de D ; sachant que τ est le nombre minimal de colonnes linéairement-dépendantes de D, pour tout t, pour $s(t) \neq 0$, il existe n columns D_j tel que : $\sum_{j=1}^{n} D_j = s(t) = De(t) \neq 0$, avec $n \leq \tau - 1$;
par conséquent, $w(e(t)) = n \leq \tau - 1$; et pour $s(t) = 0$, il y a toujours $e(t) = 0$ qui vérifie : $w(e(t)) = 0 \leq \tau - 1$.

À partir de cela, nous pouvons déduire le résultat.

Cette proposition découle de la précédente.

Proposition 6. *Considérons un schéma stéganographique S donné par un code convolutif (A, B, C, D), et les fonctions emb et rec. Supposons que D aît toutes ses colonnes non nulles et distinctes. Ainsi, à l'intérieur d'une séquence de couverture de longueur ℓk , nous pouvons encapsuler au maximum un message de longueur ℓp en modifiant au maximum $\ell(\tau - 1)$ bits*

Example 3. Dans $\mathbb{F}_2$, considérons le code $\mathcal{C}(A, B, C, D)$ défini par :

$$A = \begin{pmatrix} 1\,0\,1\,1\,0\,1 \\ 0\,0\,1\,0\,0\,1 \\ 1\,1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,1\,1 \\ 1\,1\,0\,0\,0\,1 \\ 0\,0\,0\,0\,1\,0 \end{pmatrix}, B = \begin{pmatrix} 1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,0 \\ 1\,0\,0\,1\,0 \\ 0\,0\,0\,0\,1 \\ 1\,1\,1\,0\,1 \\ 0\,1\,1\,0\,0 \end{pmatrix}$$

$$C = \begin{pmatrix} 1\,0\,0\,1\,1\,0 \\ 0\,1\,0\,0\,1\,1 \\ 0\,0\,1\,0\,1\,1 \end{pmatrix}, D = \begin{pmatrix} 1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,0 \\ 1\,0\,1\,0\,1 \end{pmatrix}$$

Nous observons que le système est output-observable.

Considérons le message à encapsuler $m = (m(0), m(1), m(2))$, et cela étant fait avec chacune de ces sections, à chaque étape, dans une séquence de couverture dénotée par u.

Considérons $m = (111,\ 100,\ 001)$

Nous considérons que connaissant chaque séquence entrée u, nous essayerons de trouver les charactéristiques de chaque séquence e des bits à modifier qui

ont été ajoutés à u, dans le but d'encapsuler le message m. Le quasi-syndrome correspondant est dénoté par s.

Considérons la matrice de décodage pour les codes convolutifs donnée par :

$$\widehat{T}_{\ell-1} = \begin{pmatrix} D & & & \\ CB & D & & \\ CAB & CB & D & \\ \vdots & & \ddots & \ddots \\ CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB\ D \end{pmatrix}$$

À chaque étape ici : $0, 1, \dots, \ell$, nous essaierons d' évaluer notre valeur erreur e

Nous avons déjà : $m = \widehat{T}_{\ell-1}\,(u + e)$

Comme entrées, nous avons m et u.

Tout d'abord, évaluons le sous-ensemble des potentielles erreurs e ; comme formule générale, m et u sont donnés par : $m(t) = \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k)) + D(u(t) + e(t))$; d'où :

$$De(t) = s(t) = m(t) - \sum_{k=0}^{t-1} CA^{t-1-k}B(u(k) + e(k)) - Du(t)$$

En parcourant tous les cas possibles de syndromes (ou séquences de couverture), nous avons l'ensemble des erreurs e.

Par conséquent, à chacune des étapes, il y a toujours une séquence e qui peut être utilisée pour encapsuler telle que : $w(e) \leq \tau - 1$ ($\tau = 3$).

Travaillons avec u = (11011 01000 10101) ayant m = (111, 100, 001).

À l' étape $t = 0$, nous avons : $D(u(0) + e(0)) = m(0)$
Ainsi, $s = (1\,0\,1)$, et $e(0) = (1\,0\,0\,0\,0)$

À l' étape $t = 1$, nous avons : $D(u(1) + e(1)) = m(1) - CB(u(0) + e(0))$
Ainsi, $s = (0\,0\,0)$, nous avons $e(0) = (1\,0\,0\,0\,0)$ et $e(1) = (0\,0\,0\,0\,0)$

À l' étape $t = 2$, nous avons : $D(u(2) + e(2)) = m(2) - CAB(u(0) + e(0)) - CB(u(1) + e(1))$
Ainsi, $s = (0\,0\,0)$, et $e(2) = (0\,0\,0\,0\,0)$.

Ainsi, pour u = (11011, 01000, 10101), nous pouvons encapsuler m = (111, 100, 001) avec la séquence de modification : $e = (1\,0\,0\,0\,0,\ 0\,0\,0\,0\,0,\ 0\,0\,0\,0\,0)$.

Pour cette opération, nous avons encapsulé 9 bits dans une séquence de couverture de longueur 15-length en changeant uniquement 1 bit.

s	e (w < τ)	e (w $\geq \tau$)
000	00000	00111
		11110
		11001
001	00001	11111
	11000	
	00110	
010	00010	11100
	00101	11011
100	01000	01111
	10001	
011	00011	11010
	00100	11101
101	01001	01110
	10000	10111
110	01010	01101
	10100	10011
111	01100	01011
	10010	10101

Figure 2. Syndrome Table 3

Interêts et bénéfices comparés à la stéganographie basée sur les codes en bloc

Example 4. Considerons un protocole stéganographique basé sur un code en bloc donné par sa matrice de transfert : $H = \begin{pmatrix} 1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,0 \\ 1\,0\,1\,0\,1 \end{pmatrix}$

Considerons un ensemble de messages m à encapsuler dans un ensemble de sequences de couverture u donnés par : $m(0) = (111)$ dans $u(0) = (11011)$, $m(1) = (100)$ dans $u(1) = (01000)$, and $m(2) = (001)$ dans $u(2) = 10101$ l'un après l'autre, avec ce processus stéganographique.

Etant donné la table de syndrome :

s	e (w $< \tau$)	e (w $\geq \tau$)
000	00000	00111
		11110
		11001
001	00001	11111
	11000	
	00110	
010	00010	11100
	00101	11011
100	01000	01111
	10001	
011	00011	11010
	00100	11101
101	01001	01110
	10000	10111
110	01010	01101
	10100	10011
111	01100	01011
	10010	10101

Figure 3. Syndrome Table 4

La séquence de modification est donnée par : e.
L'insertion de m est donnée par : $m(X) = H(u(X) + e(X))$.

Ainsi, $m(0) = H(u(0) + e(0))$; ce qui signifie que : $He(0) = m(0) - Hu(0) = s(0)$; $s(0) = (1\,0\,1)$, et nous choisissons $e(0) = (1\,0\,0\,0\,0)$

Tout comme pour $e(1)$, nous avons $De(1) = m(1) - Hu(1) = s(1)$; $s(1) = (000)$, et nous avons $e(1) = (00000)$

Il en est de même pour $e(2)$, nous avons $De(2) = m(2) - Hu(2) = s(2)$; $s(2) = (110)$, et nous prenons $e(2) = (01010)$

Pour la séquence mise tout ensemble, nous avons :
Pour $u = (11011,\ 01000,\ 10101)$, nous pouvons cacher $m = (111,\ 100,\ 001)$ avec la sequence à rajouter : $e = (10000,\ 00000,\ 01010)$.

Travaillant avec le code convolutif, comme précedemment, nous avons : pour la même séquence $u = (11011,\ 01000,\ 10101)$, nous pouvons encapsuler $m = (111,\ 100,\ 001)$ avec la séquence de modification : $e = (10000,\ 00000,\ 00000)$.

Lorsque comparé au cas convolutif, nous modifions 2 bits de plus.

Example 5. Considerons un protocole steganographique basé sur un code en bloc donné par sa matrice de transfert : $H = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

Considerons un ensemble de messages m à encapsuler dans un ensemble de séquences de couverture u données par : $m(0) = (10)$ dans $u(0) = (111)$, $m(1) = (00)$ dans $u(1) = (010)$, et $m(2) = (01)$ dans $u(2) = (001)$ l'un après l'autre, avec ce processus steganographique.

Etant donné la table des syndromes :

s	e
00	000
	100
10	011
	111
01	001
	101
11	010
	110

Figure 4. Syndrome Table 5

La séquence de modification à chacune des étapes est donnée par : e.
Encapsuler m est donné par : $m(X) = H(u(X) + e(X))$.

Ainsi, $m(0) = H(u(0) + e(0))$; ce qui signifie que : $He(0) = m(0) - Hu(0) = s(0)$; $s(0) = (0\,0)$, et nous choisissons $e(0) = (0\,0\,0)$

Tout comme pour $e(1)$, nous avons $De(1) = m(1) - Hu(1) = s(1)$; $s(1) = (1\,1)$, et nous avons $e(1) = (0\,1\,0)$

Tout comme pour $e(2)$, nous avons $De(2) = m(2) - Hu(2) = s(2)$; $s(2) = (0\,0)$, et nous choisissons $e(2) = (0\,0\,0)$

Pour toute la séquence complète, nous avons :
pour $u = (1\,1\,1,\ 0\,1\,0,\ 0\,0\,1)$, nous pouvons encapsuler $m = (10,\ 00,\ 01)$ avec la séquence de modification : $e = (0\,0\,0,\ 0\,1\,0,\ 0\,0\,0)$

Dans le cas des codes convolutionnels, comme plus tôt, nous avons :
pour le même $u = (1\,1\,1,\ 0\,1\,0,\ 0\,0\,1)$, nous pouvons encapsuler $m = (10,\ 00,\ 01)$ avec la séquence de modification : $e = (0\,0\,0,\ 0\,0\,0,\ 0\,1\,1)$

Lorsque comparé au cas convolutionnel, nous changeons 1 bit de moins.

Example 6. Considérons un protocole stéganographique basé sur un code en bloc donné par sa matrice de transfert : $H = \begin{pmatrix} 1\,0\,0\,1\,0\,1\,1 \\ 0\,1\,0\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1 \end{pmatrix}$

Considerons un ensemble de messages m à encapsuler dans un ensemble de séquences de couverture u donné par : $m(0) = (110)$ dans $u(0) = (1100110)$, $m(1) = (010)$ dans $u(1) = (0010011)$, and $m(2) = (011)$ dans $u(2) = (1001010)$ l'un aprés l'autre, avec ce processus steganographique.

Etant donné la table des syndromes :

La séquence à chaque étape donnée par : e.

Insérer m est donné par : $m(X) = H(u(X) + e(X))$.

Ainsi, $m(0) = H(u(0) + e(0))$; ce qui signifie que : $He(0) = m(0) - Hu(0) = s(0)$; $s(0) = (1\,0\,0)$, et nous choisissons $e(0) = (1\,0\,0\,0\,0\,0\,0)$

Tout comme pour $e(1)$, nous avons $De(1) = m(1) - Hu(1) = s(1)$; $s(1) = (0\,0\,1)$, et nous avons $e(1) = (0\,0\,1\,0\,0\,0\,0)$

Tout comme pour $e(2)$, nous avons $De(2) = m(2) - Hu(2) = s(2)$; $s(2) = (1\,1\,0)$, et nous avons $e(2) = (0\,0\,0\,1\,0\,0\,0)$

Pour la séquence au complet, nous avons :
pour $u = (1\,1\,0\,0\,1\,1\,0,\ 0\,0\,1\,0\,0\,1\,1,\ 1\,0\,0\,1\,0\,1\,0)$, nous pouvons encapsuler $m = (110,\ 010,\ 011)$ avec la séquence de modification : $e = (1\,0\,0\,0\,0\,0\,0,\ 0\,0\,0\,1\,0\,0\,0\,0,\ 0\,0\,0\,1\,0\,0\,0)$

s	e (w < τ)
000	0000000
001	0010000
010	0100000
100	1000000
011	0000100
	0110000
101	0000001
	1010000
110	0001000
	1100000
111	0000010
	0011000

Figure 5. Syndrome Table 6

Considerons plutôt le code convolutif, donné par :

$$A = \begin{pmatrix} 1\,0\,0\,1\,1\,0\,1\,1 \\ 0\,0\,1\,0\,0\,1\,1\,0 \\ 1\,1\,1\,0\,1\,1\,0\,1 \\ 0\,0\,1\,1\,0\,0\,1\,1 \\ 1\,0\,1\,1\,0\,1\,0\,0 \\ 0\,1\,0\,0\,0\,0\,1\,0 \\ 0\,0\,1\,0\,0\,1\,0\,1 \\ 1\,0\,0\,1\,1\,0\,0\,1 \end{pmatrix}, B = \begin{pmatrix} 1\,0\,0\,1\,1\,0\,1 \\ 0\,1\,0\,0\,0\,1\,1 \\ 1\,1\,0\,1\,0\,0\,0 \\ 0\,1\,1\,1\,1\,1\,1 \\ 1\,0\,1\,1\,0\,0\,1 \\ 1\,0\,0\,0\,1\,0\,0 \\ 0\,1\,0\,0\,0\,1\,0 \\ 0\,0\,0\,1\,0\,0\,0 \end{pmatrix}, C = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,0\,0\,0\,1\,0\,0 \end{pmatrix}$$

$$D = \begin{pmatrix} 1\,0\,0\,1\,0\,1\,1 \\ 0\,1\,0\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1 \end{pmatrix}$$

A l' étape $t = 0$, nous avons : $D(u(0) + e(0)) = m(0)$
$s(0) = (1\,0\,0)$, et nous choisissons $e(0) = (1\,0\,0\,0\,0\,0\,0)$

A l' étape $t = 1$, nous avons : $D(u(1) + e(1)) = m(1) - CB(u(0) + e(0))$
Ainsi, $s = (0\,1\,1)$, nous choisissons $e(0) = (1\,0\,0\,0\,0\,0\,0)$ et $e(1) = (0\,0\,0\,0\,1\,0\,0)$

A l' étape $t = 2$, nous avons : $D(u(2) + e(2)) = m(2) - CAB(u(0) + e(0)) - CB(u(1) + e(1))$
Ainsi, $s = (0\,1\,1)$, et $e(2) = (0\,0\,0\,0\,1\,0\,0)$.

Pour le même $u = (1\,1\,0\,0\,1\,1\,0,\ 0\,0\,1\,0\,0\,1\,1,\ 1\,0\,0\,1\,0\,1\,0)$, nous pouvons encapsuler $m = (110,\ 010,\ 011)$ avec la séquence de modification : $e = (1\,0\,0\,0\,0\,0\,0,\ 0\,0\,0\,0\,1\,0\,0,\ 0\,0\,0\,0\,1\,0\,0)$

Lorsque comparé au cas convolutif, nous changeons 3 bits, tout comme pour le cas des codes en bloc.

Si nous essayons d'évaluer en termes de relation entre tous les vecteurs erreurs, regardons de pr`s ce que nous obtenons à chacune des étapes
À l'étape $t = 0$, $De(0) = m(0) - Du(0) < \tau$;
À l'étape $t = 1$, $(CB\ D)\begin{pmatrix} e(0) \\ e(1) \end{pmatrix} = m(1) - (CB\ D)\begin{pmatrix} u(0) \\ u(1) \end{pmatrix}$; d'où : $\begin{pmatrix} e(0) \\ e(1) \end{pmatrix} = X_{(CB\ D)}\left(m(1) - (CB\ D)\begin{pmatrix} u(0) \\ u(1) \end{pmatrix}\right)$
d'où : $w(e(0)) + w(e(1)) = C_1$, C_1 étant une constante ;
À l'étape $t = 2$, $(CAB\ CB\ D)\begin{pmatrix} e(0) \\ e(1) \\ e(2) \end{pmatrix} = m(2) - (CAB\ CB\ D)\begin{pmatrix} u(0) \\ u(1) \\ u(2) \end{pmatrix}$; d'où : $\begin{pmatrix} e(0) \\ e(1) \\ e(2) \end{pmatrix} = X_{(CAB\ CB\ D)}\left(m(2) - (CAB\ CB\ D)\begin{pmatrix} u(0) \\ u(1) \\ u(2) \end{pmatrix}\right)$
d'où : $w(e(0)) + w(e(1)) + w(e(2)) = C_2$, C_2 étant une constante ;

En utilisant le même procédé, nous pouvons mettre en relation, les poids de toutes les séquences à partir de $t = 0$ jusqu'à $t = \ell$, en utilisant une analyse suivant la même logique.

Références

1. A. Westfield, *F5* steganographic algorithm, *Proc. of the Information Hiding* 4^{th} *International Workshop*, vol. 2137, pp. 289-302, 2001.
2. Y. Denneulin, J.-L. Roch, E. Tannier, *Théorie des codes*, January (2000), pp. 41-46.
3. J. Fridrich, *Steganography in Digital Media - Principles, Algorithms, and Applications*, Cambridge Univ. Press, 2009.
4. M.I. Garcia-Planas, D. Magret, M.E. Montoro, *Two parametric quasi-cyclic codes as hyperinvariant subspaces.* Cybernetics and physics Journal, **2**, (2), pp. 90-96, (2013).
5. M.I. Garcia-Planas, Rl M. Soudi, L.E. Um. *Convolutional codes under control theory point of view. Analysis of output-observability.* Recent Advances in Circuits, Communications & Signal Processing, pp. 131-137, (2013).
6. H. Jouhari, New Steganographic Schemes using Binary and Quaternary Codes, Ph.D thesis, Université Mohammed V-Agdal, Morocco ; (2013)
7. H. Jouhari and El M. Souidi, *Application of Cyclic Codes over* $\mathbb{Z}_4$ *in Steganography.* Journal of Applied Mathematical Sciences, vol.6, 2012, N139, pp 6911-6925.
8. H. Jouhari and El M. Souidi, *Steganographic Scheme Using The* Z_4*-Linear Goethals Codes.* Proceedings of the Third International Conference on Digital Inforation Processing and Communications, Dubai 2013, UAE, pp. 114-121.
9. M.O. Medeni and E.M. Souidi, *A Novel Steganographic Protocol from Error-correcting Codes.* Journal of Information Hiding and Multimedia Signal Processing, vol.1, 2010.

considérons e, la séquence représentant les bits modifiés de ce bloc de couverture, tout en ayant chacun des objets de ce protocole donné par sa représentation polynomiale, le stego object is donné par :

$$r(X) = v(X) + e(X) \tag{2}$$

A partir de ces deux équations 1 and 2, nous avons :

$$m - v \times H^t = e \times H^t \tag{3}$$

qui nous donne la formule d'extraction.

Example 1. Dans l'algorithme $F5$ [1], la technique utilisée avec un $[n, n-k, 1]$-code consiste encapsuler k bits sur une séquence de couverture de longueur n-length en changeant au maximum 1 bit.

Cette méthode est appelée le codage du syndrome, et d'un point de vue stéganographique, nous avons besoin de trouver un nombre minimal de flips of $e(X)$ afin de diminuer la distortion.

4 Stéganographie et codage convolutif

Ici, nous rappelons quelques notions faisant partie de la construction d'un protocole stéganographique basé sur le codage.

4.1 Les raisons et avantages

En fonction de ce qui est déjà connu du processus stéganographique traditionnel, notre objectif est de proposer une méthode stéganographique qui est implémentable en s'appuyant sur le codage/décodage convolutif. Il est déjà connu qu'il existe plusieurs protocoles stéganographiques définis sur les codes en blocs correcteurs d'erreurs, à l'aide des algorithmes de décodage, connus à l'origine, pour detecter et corriger des erreurs, afin d'introduire un minimum d'erreurs, le moins possible [10]. Nous utilisons une approche similaire, pour ce qui est de notre propre modèle stéganographique implémenté ici, avec la particularité d'être basé sur le codage convolutif, qui requiert essentiellement d'être implémenté de manière séquentielle, par exemple ecapsulation d'un message secret lors de l'envoi d'un autre, par exemple fichier ou image, durant une séquence temps indéterminée ou semi-infinie. En effet, l'idée générale est d'introduire le moins de déformation possible aux séquences initiales de couverture, dans le but d'encapsuler une autre séquence numérique (préferablement de longueur plus petite).

Afin d'arriver, nous nous devons de couvrir quelques points essentiels :

1. les conditions pour que le schéma stéganographique soit établi, ce qui signifie que les fonctions *emb* et *rec* soient appropriées ;
2. les conditions pour les sous-sections modifiées, qui concernent la borne sur la quantité de bits á modifier, ceux qui ont étés altérés durant l'encapulation, dans le but d'en modifier le moins possible ;

3. la borne classique d'imperceptibilité, tout comme le rayon d'embedding. En effet, nous devons donner des conditions sur le code convolutif afin de conserver l'imperceptibilité du changement sur la couverture, indépendamment de la séquence de couverture ou du message à cacher ;
4. l'interêt et les avantages, comparé à la stéganographie basé sur les codes en blocs.

10. C. Munuera, *Steganography From A Coding Theory Point Of View.* Department of Applied Mathematics, University of Valladolid.

11. G. J. Simmons. The prisoners' problem and the subliminal channel. In D. Chaum, editor, Advances in Cryptology, Proceedings of CRYPTO' 83, Santa Barbara, CA, August 22-24, pp. 51-67. Plenum Press, New York, 1984.

Revendications

1. Méthode de dissimulation de l'information en sténographie caractérisée en ce que le programme algorithmique utilise des familles de code convulutif pour l'insertion et l'extraction de l'information sans distorsion apparente à l'œil humain et œil de la machine ;

2. Méthode de dissimulation de l'information selon la revendication 1 caractérisée en ce que la code s'applique à tout type de support

3. Méthode de dissimulation de l'information selon la revendication 1 et 2 caractérisé en ce que le logiciel fonctionne sur tout type de terminal.

MA 38026A1

ROYAUME DU MAROC

OFFICE MAROCAIN DE LA PROPRIETE INDUSTRIELLE ET COMMERCIALE

المملكة المغربية

المكتب المغربي للملكية الصناعية والتجارية

RAPPORT DE RECHERCHE PRELIMINAIRE AVEC OPINION SUR LA BREVETABILITE

Renseignements relatifs à la demande	
N° de la demande : 38026	Date de dépôt : 21/04/2015
Déposant : Université Mohammed V Rabat	Date de Priorité :
Intitulé de l'invention : Méthode de dissimulation de l'information par codes convolutifs en stéganographie.	

Le présent document est le rapport de recherche préliminaire avec opinion sur la brevetabilité établi par l'OMPIC conformément à l'article 43 et notifié au déposant conformément à l'article 43.1 de la loi 17/97 relative à la protection de la propriété industrielle.

- Le présent rapport est constitué de 3 pages (la présente page incluse)
- Les documents cités par l'examinateur dans la partie Rapport de recherche sont joints au présent document

Le présent rapport contient des indications relatives aux éléments suivants :

Partie 1 : Considérations générales
- ☒ Cadre 1 : Base du présent rapport
- ☐ Cadre 2 : Priorité
- ☐ Cadre 3 : Titre et/ou Abrégé tel qu'ils sont définitivement arrêtés

Partie 2 : Rapport de recherche

Partie 3 : Opinion sur la brevetabilité
- ☒ Cadre 4 : Remarques de clarté
- ☒ Cadre 5 : Déclaration motivée quand à la Nouveauté, l'Activité Inventive et l'Application Industrielle
- ☐ Cadre 6 : Observations à propos de certaines revendications dont aucune recherche significative n'a pu être effectuée
- ☐ Cadre 7 : Défaut d'unité d'invention

Examinateur: BAMI MOHAMMED

Téléphone: 05 22 58 64 14

Email : bami@ompic.ma

Date d'établissement du rapport : 13/10/2015

ROYAUME DU MAROC Unité Brevets, Dessins et Modèles OFFICE MAROCAIN DE LA PROPRIETE INDUSTRIELLE ET COMMERCIALE

Partie 1 : Considérations générales

Cadre 1 : base du présent rapport

Les pièces suivantes de la demande servent de base à l'établissement du présent rapport :

- Description
  19 pages
- Revendications
  1-3

Partie 2 : Rapport de recherche

Classement de l'objet de la demande :

CIB : G06T1/00
CPC : G06T1/00

Bases de données électroniques consultées au cours de la recherche :

EPOQUE, Espacenet, Orbit

Catégorie*	Documents cités avec, le cas échéant, l'indication des passages pertinents	N° des revendications visées
X	US6859545 B1 22/02/2005 Chung Shan Institute Of Science And Technology	1-3
X	US8761391 B2 24/06/2014 Trent J. Brundage, Hugh L. Brunk	1-3
X	Decoding alogorithm for convolutional codes under linear systems point of view 2012 Souid, Planas ET AL	1-3

***Catégories spéciales de documents cités :**

-« **X** » document particulièrement pertinent ; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
-« **Y** » document particulièrement pertinent ; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
-« **A** » document définissant l'état général de la technique, non considéré comme particulièrement pertinent
-« P » documents intercalaires ; Les documents dont la date de publication est située entre la date de dépôt de la demande examinée et la date de priorité revendiquée ou la priorité la plus ancienne s'il y en a plusieurs
-« E » Éventuelles demandes de brevet interférentes. Tout document de brevet ayant une date de dépôt ou de priorité antérieure à la date de dépôt de la demande faisant l'objet de la recherche (et non à la date de priorité, cf. B-VI, 3 et B-XI, 4), mais publié postérieurement à cette date et dont le contenu constituerait un état de la technique pertinent pour la nouveauté

Partie 3 : Opinion sur la brevetabilité

Cadre 4 : Clarté

L'objet des revendications 1-3 portant sur une méthode est tel qu'il est impossible de ressortir les étapes de la méthode pour lesquelles une protection est revendiquée. Il est indispensable de définir les étapes essentielles pour la dissimulation de l'information dans le corps de la revendication.
Les revendications 1-3 tentent de définir l'objet par le résultat recherché. En tout état de cause, cette formulation n'est pas acceptable en l'espèce, puisqu'il semble possible de définir l'objet en des termes plus concrets, c'est-à-dire en exposant comment l'effet (Rev1 : l'extraction de l'information sans distorsion apparente ; Rev : 2 le code s'applique à tout support ; Rev3 : le logiciel fonctionne sur tout type de terminal).

Cadre 5 : Déclaration motivée quand à la Nouveauté, l'Activité Inventive et l'Application Industrielle

Nouveauté (N)	Revendications aucune Revendications 1-3	Oui Non
Activité inventive (AI)	Revendications aucune Revendications 1-3	Oui Non
Possibilité d'application Industrielle (PAI)	Revendications 1-3 Revendications aucune	Oui Non

Il est fait référence aux documents suivants. Les numéros d'ordre qui leur sont attribués ci après seront utilisés dans toute la suite de la procédure :

D1 : US6859545 B1

1. **Nouveauté (N) & Activité inventive (AI) :**

Le document D1 (voir figures 1-7) divulgue une méthode de dissimulation de l'information en sténographie caractérisée en ce que le programme algorithmique utilise des familles de code convolutif pour l'insertion et l'extraction de l'information sans distorsion apparente à l'œil humain et de la machine. Le code s'applique à tout support et fonctionne sur tout type de terminal.

Le document D2 invalide aussi le critère de la nouveauté des revendications 1-3.

L'objet des revendications 1-3 ne satisfait pas les exigences de nouveauté et d'activité inventive selon les dispositions des arts. 26 et 28 de la loi 17/97 telle que modifiée et complétée par la loi 23/13.

2. **Possibilité d'application industrielle (PAI) :**

L'objet de la présente invention présente une utilité déterminée, probante et crédible au sens de l'art.29 de la loi 17/97 telle que modifiée et complétée par la loi 23/13.