Itérations chaotiques pour la sécurité de l'information dissimulée

Nicolas Friot

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UFC UNIVERSITÉ DE FRANCHE-COMTÉ

Itérations Chaotiques pour la Sécurisation de l'Information Dissimulée

NICOLAS FRIOT

femto-st SCIENCES & TECHNOLOGIES

SPIM

Thèse de Doctorat

UFC

école doctorale **sciences pour l'ingénieur et microtechniques**

UNIVERSITÉ DE FRANCHE-COMTÉ

N° 8 8 8

THÈSE présentée par

NICOLAS FRIOT

pour obtenir le

Grade de Docteur de
l'Université de Franche-Comté

Spécialité : **Informatique**

Itérations Chaotiques pour la Sécurisation de l'Information Dissimulée

Unité de Recherche :
Institut FEMTO-ST, UMR 6174, CNRS - Département d'Informatique des Systèmes Complexes (DISC)

Soutenue publiquement le 5 Juin 2014 devant le Jury composé de :

HAMAMACHE KHEDDOUCI	Rapporteur	Professeur à l'Université Claude Bernard Lyon 1
PIERRE SPITERI	Rapporteur	Professeur Émérite à l'INP Toulouse
JACQUES M. BAHI	Directeur de thèse	Professeur à l'Université de Franche-Comté
CHRISTOPHE GUYEUX	Examinateur	Maître de Conférences HDR à l'Université de Franche-Comté
GUILLAUME BONFANTE	Examinateur	Maître de Conférences HDR à l'Université de Lorraine - École des Mines de Nancy
LHASSANE IDOUMGHAR	Examinateur	Maître de Conférences HDR à l'Université de Haute Alsace
RAPHAËL COUTURIER	Examinateur	Professeur à l'Université de Franche-Comté

Remerciements

En tout premier lieu, je tiens à remercier mon directeur de thèse, le **Professeur Jacques BAHI** pour son soutien éclairé et les conseils qu'il m'a prodigués, tout au long de l'élaboration de cette thèse. Il a su, à chaque instant, me recommander les bons choix à opérer, particulièrement lorsqu'il fallait prendre des décisions sur des orientations majeures. C'est pour moi un réel honneur d'avoir travaillé avec lui et d'avoir pu bénéficier de son expérience. Les connaissances, les savoir-faire, les méthodologies et peut-être surtout « la vision » qu'il a su me transmettre, ont été une des clés de l'aboutissement de cette thèse.

Je tiens également à remercier **Christophe GUYEUX, Maître de Conférences**, avec qui j'ai mené la plupart de mes travaux de recherche. Il a toujours su prendre, quel que soit le moment, le temps nécessaire pour me guider dans la voie de cet art subtil et complexe qu'est la recherche fondamentale. Mes échanges avec Christophe ont toujours été très féconds et ses avis très pertinents. Sa remarquable maîtrise des mathématiques lui permet d'aller d'emblée à l'essentiel et de faire preuve d'une sagacité exceptionnelle dans la réalisation des travaux scientifiques. Christophe a su me transmettre ses bonnes pratiques de raisonnement et son souci constant de l"efficacité. Je l'en remercie vivement, car c'est aujourd'hui toujours un réel atout pour moi de continuer à tirer profit de son enseignement pour conduire mes travaux scientifiques et mes projets.

Je remercie aussi tous les autres **membres de l'équipe AND**[1] avec lesquels j'ai été amené à travailler durant l'élaboration de cette thèse, et tout particulièrement le directeur de l'équipe, le **Professeur Raphaël COUTURIER**, qui m'a accueilli avec bienveillance. Je remercie **Jean-Luc ANTHOINE**, **Jean-Claude CHARR**, **Jean-François COUCHOT**, **Karine DESCHINKEL**, **Stéphane DOMAS**, **Arnaud GIERSCH**, **Mourad HAKEM**, **David LAIYMANI**, **Abdallah MAKHOUL**, **Gilles PERROT** et **Michel SALOMON**.

Je remercie de même **Fabrice AMBERT** et **Pierre-Cyrille HEAM**, tous deux membres de l'équipe VESONTIO. Ils étaient régulièrement présents dans les locaux belfortains du département informatique (DISC) de l'institut FEMTO-ST et m'ont, eux aussi, fait un très bon accueil.

Je remercie, bien évidemment, les **membres de l'IUT de Belfort-Montbéliard** pour la

1. Équipe Algorithmique Numérique Distribuée du Département d'Informatique des Systèmes Complexes (DISC) de l'Institut FEMTO-ST, CNRS, UMR 6174

bienveillance et la compréhension dont ils ont fait preuve à mon égard pendant toute la durée de l'élaboration de ma thèse. Je remercie particulièrement le Directeur de l'IUT, **Olivier PREVOT**, et je pense bien sûr aussi à **Christelle MARC** et **Christelle REINA**, au service communication ; à **Sylvie ALTMAYER**, **Figen BASAN** et **Nathalie WILTZ** au service financier ; à **Chantal BRIGNON**, **Martine CANOVAS**, **Julie DENÊTRE** et **Isabelle SCHMITT** au service administratif ; sans oublier **Martine BAILLEUL**, qui a récemment changé de service et travaille aujourd'hui à la Maison de l'Université à Besançon. J'adresse également mes remerciements à la Commission de la Recherche de l'IUT qui a, notamment, financé l'édition de la brochure de la future société *Stégosécu ISIS*, entreprise innovante pour la valorisation économique des nos résultats de recherche.

Il m'importe de remercier également ma collègue, désormais Docteur en informatique à Bordeaux, **Lilia Ziane KHODJA** dont la collaboration m'a été précieuse sur bien des plans. Je souligne notamment sa présence à mes côtés et auprès des collègues doctorants de notre université de Franche-Comté pour nous accompagner dans des nos projets communs. Cette présence, alliée à son engagement et à son travail , nous a permis de créer une association « *Le Ruban de Moebius* », dont l'objectif est la valorisation des compétences des doctorants et docteurs et le développement de synergies entre secteur industriel, secteur universitaire et tissu associatif : *www.ruban-moebius.fr*[2].

C'est dans ce même contexte d'engagement universitaire que je tiens, pareillement, à remercier **l'ensemble des collègues administrateurs de l'université** avec lesquels j'ai pu avoir, à maintes reprises, de fructueux échanges au Conseil Scientifique, devenu aujourd'hui Conseil Académique.

Je remercie encore **l'ensemble des membres du Conseil d'Unité** de notre laboratoire et son directeur, **Nicolas CHAILLET**, ainsi que la directrice du département informatique, **Olga KOUCHNARENKO**, qui a soutenu mes travaux et projets avec un esprit critique avisé et toujours constructif. Ma gratitude s'adresse encore aux **collègues membres du Conseil de l'École Doctorale SPIM** et à son directeur, **Philippe LUTZ**. Ce fut pour moi un réel plaisir et une fierté de contribuer au rayonnement de notre école et à la valorisation des compétences des doctorants.

Je n'oublie pas de remercier toutes celles et ceux qui m'ont accompagné dans chaque démarche des projets ANR (Agence Nationale de la Recherche), notamment en ce qui concerne le projet APACHE que nous avons mis en place avec **Vincent LEFEBVRE** et **Gianni SANTINELLI** de la société Tages Solidshield SAS, en région PACA ; ainsi que **Jean-François MANCEAU** et **Thérèse LEBLOIS** du département MN2S de l'institut FEMTO-ST.

Je remercie aussi les collègues qui nous ont accompagnés dans la mise en place des projets de valorisation et de création d'entreprises innovantes, projets que j'ai portés durant cette thèse. La nature de ces projets et le sérieux des **membres du comité de pilotage** m'ont permis de décrocher le titre de « Lauréat 2012 » du concours du Ministère de l'Enseignement Supérieur et de la Recherche, en partenariat avec la BPI, pour la « Création

2. Le réseau professionnel, l'insertion et l'innovation participative sont les trois axes forts de ce projet qui s'est concrétisé après plus de douze mois de travail, à partir des deux éléments clés qui furent notre point de départ. Tout d'abord les Assises de l'Enseignement Supérieur et de la Recherche, ensuite la publication du Rapport Louis Gallois "Pacte pour la compétitivité des industries françaises". Cette réalisation a été possible grâce à la contribution et au soutien des 15 autres membres fondateurs de l'association avec lesquels nous nous sommes réunis chaque semaine, en dehors et en plus de nos temps de travail respectif, durant ces douze mois de la phase préparatoire. À ce titre, je tiens à remercier chacun des membres fondateurs et à signaler leur investissement conséquent.

d'entreprises de technologies innovantes. Grand fut pour moi l'honneur de représenter notre équipe-projet auprès des plus hautes instances universitaires et territoriales.

Dans le cadre de cet accompagnement, je pense encore à **Philippe PICART**, directeur du « Service valorisation de l'université », à **Blandine TATIN**, directrice de « L'incubateur d'entreprises innovantes de Franche-Comté », à **Pascale BRENET**, directrice de « L'Institut d'Administration des Entreprises (IAE) », pour ses précieux conseils lors de la formation à l'entreprenariat et l'innovation que j'ai suivie dans son institut. Je pense aussi à **Catherine RIBLET**, directrice du centre de formation KISEL, à Belfort, dont l'expertise en matière de gestion et de management d'une entreprise m'a été fort utile ; à sa « littéraire » associée, **Pierrine LABRIET**, ainsi qu'à **Monique RODRIGUEZ** enseignante en anglais. Catherine, Pierrine et Monique étaient toutes les trois toujours disponibles pour relire et corriger nos travaux, dans leur domaine respectif de compétences.

Je remercie également **Daniel SEIGNEUR**, pour le soutien qu'il a pu m'apporter dans le cadre des projets de valorisation que nous avons conduits avec le laboratoire, notamment le projet *Stégosécu ISIS*, et le montage vidéo qu'il a réalisé dans le cadre du concours des Entrepreneuriales 2010 auquel nous avons participé.

Je remercie enfin ma famille et mes amis qui m'ont permis, par leur affection, leur sollicitude et leurs encouragements de lever mes doutes et d'aller de l'avant. Je pense à **Michèle**, à **Mercèdes**, à **Yannick** et à tous les autres qui, de près ou de loin, m'ont apporté leur soutien sans failles.

En hommage à mon grand-père.

À ma famille.

Sommaire

I

INTRODUCTION GÉNÉRALE

CHAPITRE 1

Introduction

> Une théorie est vraie si elle est énonçable selon les règles de la logique formelle, et si ses conséquences sont vérifiables par tout observateur.
>
> JACQUES ATTALI, ÉCRIVAIN, ÉCONOMISTE, SCIENTIFIQUE, HOMME POLITIQUE (1943-)

1.1/ PRÉSENTATION GÉNÉRALE

Les systèmes dynamiques discrets, œuvrant en itérations chaotiques ou asynchrones, se sont révélés être des outils particulièrement intéressants à utiliser en sécurité informatique. En effet, ces itérations peuvent être programmées de manière efficace et il a été montré que, sous certaines conditions, ces dernières se comportaient d'une façon imprévisible. Pour être plus précis, elles satisfont les propriétés de chaos topologique, telles qu'elles ont été définies par Devaney, et vérifient, de plus, les propriétés de mélange, d'entropie topologique, d'expansivité et de transitivité forte.

Pour tirer profit de la qualité du désordre généré par ces itérations chaotiques, des applications ont été proposées par l'équipe AND [1] dans les domaines suivants : fonctions de hachage, génération de nombres pseudo-aléatoires (PRNG), stéganographie, tatouage numérique, et chiffrement de données. Or, ces domaines sensibles sont en évolution permanente [2]. Bien que cette sécurité soit usuellement établie par des arguments probabilistes, une approche originale, en termes d'imprévisibilité, a été récemment introduite par

1. Équipe Algorithmique Numérique Distribuée du Département d'Informatique des Systèmes Complexes (DISC) de l'Institut FEMTO-ST, CNRS, UMR 6174

2. Par exemple, le National Institute for Standards and Technologies (NIST, USA) a récemment lancé un appel d'offres à la communauté scientifique, pour trouver de nouvelles fonctions de hachage et de nouveaux PRNG plus sûrs que ceux existants, récemment révélés faillibles.

l'équipe AND. Cette nouvelle approche a permis l'étude d'algorithmes à base d'itérations chaotiques, dont la sécurité s'est avérée d'un niveau tout à fait satisfaisant.

L'objectif de cette thèse consiste à approfondir l'étude de la sécurité des algorithmes de dissimulation d'informations à base d'itérations chaotiques, en prouvant que tous les outils proposés sont effectivement sûrs, selon l'approche probabiliste. Il nous fallait aussi proposer de nouvelles méthodes pour résoudre les problèmes inhérents aux algorithmes de dissimulation d'informations, antérieurement proposés par l'équipe AND, et d'en étudier la sécurité d'un point de vue probabiliste et topologique. Une fois cette sécurité clairement établie, mon travail consistait à reprendre les prototypes existants (en langage Python) et à améliorer leurs performances afin de les rendre utilisables pour des applications réelles.

Il m'a ensuite été demandé d'étudier dans quelle mesure ces résultats de recherche pourraient être valorisés au travers de projets économiques, incluant la perspective d'une création d'entreprise de technologies innovantes.

C'est la conduite de cette double étude, à la fois scientifique et économique, qui a animé mes travaux pendant toute l'élaboration de ma thèse et qui lui confère sa spécificité.

1.2/ CONTRIBUTIONS

Avant les travaux exposés dans cette thèse, il n'était possible, s'agissant du tatouage numérique par itérations chaotiques, que d'extraire un seul bit à chaque itération (on parle de « One bit watermarking »). C'est pourquoi il nous a été demandé de proposer des algorithmes de dissimulation d'informations à la fois sûrs et capables de lever cette restriction.

Notre travail s'est déroulé en plusieurs phases.

Nous nous sommes d'abord intéressés à comprendre ce que signifiait la notion de sécurité dans le domaine de la dissimulation d'informations. À cette fin, nous avons commencé à étudier les travaux de François Cayre et Patrick Bas [24, 25], qui introduisaient une notion de sécurité liée à un contexte probabiliste : la *stégo-sécurité*.
Nous avons alors élargi notre champ de connaissances en étudiant les travaux de C. Guyeux [41], selon lesquels l'imprévisibilité d'un processus lui confère potentiellement une sécurité accrue. Ce qui nous a amenés à poursuivre nos recherches dans le cadre de l'étude proposée par ce dernier, c'est-à-dire la théorie mathématique et topologique du chaos de Devaney [30]. Cette étude a conduit l'équipe à définir une nouvelle notion de sécurité, dite *sécurité topologique*. À partir de cette notion, plusieurs niveaux de sécurité ont pu être dégagés en fonction des propriétés topologiques dont le processus à étudier relève. La conséquence étant que chaque propriété topologique supplémentaire confère au processus une meilleure capacité à faire face à un contexte d'attaques particulier.

Nous avons ensuite étudié le processus de tatouage numérique à un bit (« One bit watermarking ») basé sur les itérations chaotiques, noté $\mathcal{CIW}_1$, et présenté initialement dans [41].

Nous avons finalement cherché à améliorer ce processus pour qu'il soit plus performant. Ces travaux ont donné lieu à la production d'améliorations successives, résolvant au fur et à mesure les problèmes se présentant, jusqu'à atteindre une méthode de dissimulation d'informations complète. La stégo-sécurité de cette méthode a été prouvée et permet, de

plus, de dissimuler un message de taille quelconque. La sécurité topologique, même si elle a été abordée, est encore en cours d'approfondissement.

De façon plus détaillée, les contributions apportées dans cette thèse sont les suivantes. Après avoir analysé plus en profondeur les propriétés du schéma $\mathcal{CIW}_1$ [11], nous avons prouvé sa stégo-sécurité et étudié son expansivité et son mélange topologique. Ces résultats ont montré que, contrairement à l'étalement de spectre [18], le processus $\mathcal{CIW}_1$ peut contrer un adversaire dans certaines catégories d'attaques bien définies. Cette première étude a aussi été l'occasion de mesurer combien une image se dégradait à l'issue de l'embarquement d'un message secret, dans le cadre de la dissimulation par itérations chaotiques.

Cependant, comme nous l'avons déjà mentionné, le schéma $\mathcal{CIW}_1$ ne permet d'insérer qu'un bit par image. Une telle technique, si elle est robuste, pourrait éventuellement être utile en matière de tatouage numérique, mais il serait alors nécessaire d'améliorer la capacité d'embarquement du processus $\mathcal{CIW}_1$. C'est pourquoi nous avons cherché à atteindre ce but. Nos travaux nous ont alors amenés à introduire un nouvel algorithme de dissimulation d'informations, appelé $\mathcal{CIS}_2$. Nous avons prouvé que cette méthode était stégo-sûre et topologiquement-sûre.L'exposant de Lyapunov du schéma $\mathcal{CIS}_2$ a été mesuré, ce qui a permis de montrer que les conséquences d'une erreur sur la clé secrète d'embarquement ne pouvaient être prédites. D'autres propriétés topologiques ont été démontrées, telles que la compacité, la complétude et la perfection de l'espace des phases, ou encore la forte transitivité et le mélange topologique du processus. La constante de sensibilité a été évaluée. Nous avons finalement introduit une nouvelle modélisation formelle du procédé augurant de riches conséquences.

Nous avons également cherché à mettre en œuvre, en pratique, notre algorithme. Pour ce faire, nous avons identifié des contraintes applicatives. Le but de ces contraintes était de garantir l'extraction intègre du message original à partir du contenu stéganographié. Sur la base de ces contraintes et sur la base de l'algorithme $\mathcal{CIS}_2$, nous avons alors pu développer le programme $\mathcal{CIS}_5$ qui a fait l'objet dépôt logiciel auprès de l'Agence pour la Protection des Programmes [65]. Des propriétés ont également été mises en évidence, comme l'exactitude et l'exhaustivité de l'implémentation algorithmique du procédé $\mathcal{CIS}_2$.

Nous avons ensuite cherché à lever ces contraintes applicatives. Nous avons alors conçu un nouveau procédé, $\mathcal{DI}_3$, présentant certains avantages pratiques. L'algorithme mis en jeu a été étudié. Sa stégo-sécurité a été prouvée, sa robustesse établie, et nous l'avons stéganalysé en utilisant des outils d'intelligence artificielle capables de séparer les images naturelles (anodines) des images stéganographiées (altérées). Ces outils se sont comportés sensiblement de la même manière que les meilleurs stéganographieurs actuels. Cependant, la technique utilisée par $\mathcal{DI}_3$ n'étant pas basée sur des itérations chaotiques, nous n'avons pas réussi à en mesurer sa sécurité topologique, même si, dans l'absolu, une telle mesure est toujours possible.

Nous avons par ailleurs conçu le processus $\mathcal{CIS}_3$. Dès la formalisation mathématique formelle de ce nouveau processus, nous avons cherché à lever les contraintes applicatives inhérentes au schéma $\mathcal{CIS}_2$. La stégo-sécurité de $\mathcal{CIS}_3$ a également été établie. Avec ce processus, aucune contrainte applicative supplémentaire n'est nécessaire pour garantir l'extraction intègre du message original à partir du contenu stéganographié. Cependant, un problème persistait. L'espace des phases du système dynamique discret, modélisant le processus, était fini. La sécurité de l'algorithme, selon l'approche topologique exposée dans ce rapport, s'en trouvait donc dégradée. Cette constatation a nécessité de proposer

un quatrième processus dénommé $\mathcal{CIS}_4$.

Ce dernier processus a été pleinement défini et sa stégo-sécurité prouvée. Il a été modélisé sous la forme d'un système dynamique discret continu sur un espace métrique pleinement défini. La preuve de sécurité topologique du processus $\mathcal{CIS}_4$ reste à établir.

Précisons plus en détail, à présent, les contributions que nous avons apportées à la perspective de valorisation économique de nos travaux de recherche.

Cette perspective nous a conduits à proposer la création d'une entreprise de technologies innovantes que nous avons baptisée *Stégosécu ISIS* [44]. La future société proposera la commercialisation d'une *plateforme de protection des documents numériques*. Il est à noter que le projet *Stégosécu ISIS* a été primé en 2012 par le Ministère de l'Enseignement Supérieur et de la Recherche et par la Banque Publique d'Investissement. Il est donc important de préciser que, parallèlement aux travaux scientifiques qui ont été conduits durant cette thèse, nous avons également travaillé sur des aspects plus économiques. Nous avons cherché de quelle manière nous pouvions valoriser économiquement nos résultats de recherches via la future société *Stégosécu ISIS*, société à laquelle seront transférées nos technologies, dont notamment le programme $\mathcal{CIS}_5$.

Un projet ANR [5] a également été proposé dans ce contexte. Nous avons baptisé ce projet *APACHE* [3] (développement d'une puce électronique pour l'authentification du matériel informatique et des composants électroniques). Le projet reprend plusieurs résultats issus des travaux de cette thèse, notamment pour les fonctionnalités de la puce liées aux itérations chaotiques, à la stéganographie et au tatouage numérique. Le projet *APACHE* reprend également d'autres résultats du laboratoire sur les générateurs de nombres pseudo-aléatoires et les fonctions de hachage.

Ainsi, durant toute cette thèse, deux axes d'étude ont-ils été menés conjointement. Le premier axe (scientifique) est détaillé dans ce rapport. Les études menées dans le cadre du second axe (économique) feront l'objet d'un dossier annexe qui sera produit à l'issue de la soutenance de thèse, en vue de l'obtention du *Diplôme Universitaire (DU) Entrepreneuriat et Innovation* de l'Institut d'Administration des Entreprises (IAE) de l'Université de Franche-Comté. Ce *Diplôme Universitaire*, associé à une thèse, est prévu dans le dispositif *Docteur PEPITE (Pôles Étudiants pour l'Innovation, le Transfert et l'Entrepreneuriat)* [29].

Dans ce contexte, le dossier complémentaire décrira en détail le projet *Stégosécu ISIS*. Le dossier exposera précisément les concepts et les innovations mis en jeu dans le produit phare de la future société : une *plateforme de protection des documents numériques*. Le dossier dressera un rappel de l'état de l'art scientifique sur lequel s'appuie cette plateforme (rappels de certains résultats issus des contributions de cette thèse). Le dossier se focalisera ensuite sur les variables économiques de la future société. Il exposera la stratégie générale de l'entreprise ainsi que sa stratégie marketing et financière. Le programme de Recherche et Développement (R&D) ainsi que le lien entre les deux projets (*APACHE* et *Stégosécu ISIS*) seront également présentés dans ce dossier complémentaire.

Mais, comme ces contributions à vocation économique dépassent le cadre de ce rapport de thèse, nous nous limitons à ces quelques précisions indicatives apportées en introduction.

3. Acoustic PUFs to Authenticate Computer Hardware and Electronics : PUFs acoustiques pour l'authentification du matériel informatique et des composants électroniques

1.3/ PLAN DE CETTE THÈSE

Ce rapport de thèse est consacré à la recherche dans les domaines de la science de l'information dissimulée, de la sécurisation de l'information et de la protection des données numériques. Le rapport est découpé en cinq parties.

La première partie de ce rapport est constituée par cette introduction et par un court chapitre exposant les notations (*cf.* chapitre 2) qui seront communément utilisées dans tout le rapport.

La seconde partie, quant à elle, est vouée à l'état de l'art.

Le chapitre 3 présente les deux domaines les plus importants de l'information dissimulée : la stéganographie et le tatouage numérique. Le chapitre suivant (*cf.* chapitre 4), propose quelques rappels de topologie et se termine par la présentation de la théorie du chaos selon Devaney. Le chapitre 5 détaille ce que l'on entend par stéganalyse. Dans ce chapitre sont évoquées les catégories d'attaques repérées par la littérature, les notions de stégo-sécurité et de sécurité-topologique. Ce cinquième chapitre se clôt par un exemple concret d'étude de sécurité d'une technique classique de dissimulation d'informations : l'étalement de spectre. Le dernier chapitre de l'état de l'art (*cf.* chapitre 6), quant à lui, se focalise sur les itérations chaotiques. Ce chapitre rappelle les propriétés topologiques de ce type d'itérations. Au travers de l'algorithme $\mathcal{CIW}_1$, ce chapitre présente également un cas d'application de ces itérations au domaine de la dissimulation d'informations.

La troisième partie de cette thèse est consacrée à nos contributions.

Le chapitre 7 contient les preuves de stégo-sécurité, d'expansivité et de mélange topologique du procédé $\mathcal{CIW}_1$. Ces résultats sont comparés avec l'étalement de spectre, et la distorsion introduite par le procédé est mesurée. Le chapitre 8 présente le processus $\mathcal{CIS}_2$ et expose la preuve de stégo-sécurité du processus, ainsi que la construction de l'espace des phases et du système dynamique discret permettant d'établir le caractère chaotique, selon Devaney, du procédé. Cette étude topologique du schéma $\mathcal{CIS}_2$ est poursuivie au chapitre suivant (*cf.* chapitre 9), via la production d'une semi-conjugaison topologique rendant possible l'évaluation de son exposant de Lyapunov. Le schéma $\mathcal{DI}_3$ est, quant à lui, introduit au chapitre 10. Ce schéma est étudié théoriquement (stégo-sécurité) et pratiquement (évaluation de la robustesse et utilisation de stéganalyseurs), grâce à la production des algorithmes le définissant. Le chapitre 11 concerne la présentation et l'étude de stégo-sécurité du processus $\mathcal{CIS}_3$. Le dernier chapitre de contributions (*cf.* chapitre 12), quant à lui, introduit notre dernier algorithme, à savoir le schéma $\mathcal{CIS}_4$. Ce chapitre présente notre dernière preuve de stégo-sécurité et expose le modèle mathématique formel qui permettra la future évaluation de sa sécurité topologique.

Cette thèse se termine par une partie contenant un chapitre de conclusion (*cf.* chapitre 13) avec bilan et perspectives, ainsi que la bibliographie.

1.4/ VALORISATION SCIENTIFIQUE DES TRAVAUX DE RECHERCHE

Nous terminons ce chapitre introductif par la liste des différentes contributions issues de cette thèse. Ces contributions portent sur nos résultats scientifiques ou sur leur valorisation économique *(un astérisque signale que l'ordre des auteurs est alphabétique)*.

1.4.1/ SIX PUBLICATIONS DANS DES CONFÉRENCES INTERNATIONALES

1. **SECRYPT'2013 :** Jacques M. Bahi, Nicolas Friot et Christophe Guyeux*. *A new secure process for steganography : $\mathcal{CIS}_2$. Evaluation of the Lyapunov coefficient*. SECRYPT'2013, International Conference on Security and Cryptography, July 2013, Reykjavík en Iceland. Publication.
2. **IIH-MSP'2013 :** Jacques M. Bahi, Jean-François Couchot, Nicolas Friot et Christophe Guyeux*. *Quality Studies of an Invisible Chaos-Based Watermarking Scheme with Message Extraction*. IIH-MSP'2013, 9-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, Beijing, China, Octobre 2013. IEEE Computer Society. Publication.
3. **IIH-MSP'2012 :** Jacques M. Bahi, Nicolas Friot et Christophe Guyeux*. *Lyapunov exponent evaluation of a digital watermarking scheme proven to be secure*. IIH-MSP'2012, 8-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, Piraeus-Athens, Greece, July 2012. IEEE Computer Society. Publication.
4. **INTERNET'2012 :** Jacques M. Bahi, Jean-François Couchot, Nicolas Friot et Christophe Guyeux*. *A Robust Data Hiding Process Contributing to the Development of a Semantic Web*. INTERNET'2012, 4-th Int. Conf. on Evolving Internet, Venice, Italy, June 2012. Publication et exposé.
5. **SECRYPT'2011 :** Nicolas Friot, Christophe Guyeux et Jacques M. Bahi. *Chaotic iterations for steganography - Stego-security and topological-security*. SECRYPT'2011, International Conference on Security and Cryptography, Seville, Spain, July 2011. Publication et exposé.
6. **IIH-MSP'2010 :** Christophe Guyeux, Nicolas Friot et Jacques M. Bahi. *Chaotic iterations versus spread-spectrum : topological and stego security*. IIH-MSP'2010, 6-th Int. Conf. On Intelligent Information Hiding and Multimedia Signal Processing, pages 208-211, Darmstadt, Germany, October 2010. Publication et exposé.

1.4.2/ UN DÉPÔT LOGICIEL

1. $\mathcal{CIS}_5$ **:** Dépôt logiciel $\mathcal{CIS}_5$ auprès de l'Agence pour la Protection des Programmes [65] – APP – 2011. Nicolas Friot, Christophe Guyeux, and Jacques M. Bahi. *CIS5, Programme sécurisé de stéganographie basé sur les itérations chaotiques*, January 2012. Note : Produit logiciel. Numéro de dépôt APP : IDDN.FR.001.040023.00.S.P.2012.000.10800 (logiciel œuvre de l'Université de Franche-Comté).

1.4.3/ DEUX PUBLICATIONS EN SÉMINAIRES NATIONAUX ET EN WORKSHOPS

1. **IHTIAP'2012 :** Jacques M. Bahi, Jean-François Couchot, Nicolas Friot, and Christophe Guyeux*. *Application of Steganography for Anonymity through the Internet*. In IHTIAP'2012, 1-st Workshop on Information Hiding Techniques for Internet Anonymity and Privacy, Venice, Italy, June 2012. Publication.
2. **JCS'2011 :** Nicolas Friot, Christophe Guyeux, and Jacques M. Bahi. *A new secure process for steganography : CI2. Stego and topological security*. JCS'2011,

Journées Codes et Stéganographie - Janvier 2011, Écoles Militaires de Saint-Cyr Coëtquidan. Publication et exposé.

1.4.4/ TROIS PRIX OBTENUS

1. **Best Paper Award** au Workshop IHTIAP'2012.
2. **Lauréat 2012** du Concours *Création d'Entreprise de Technologie Innovante* du Ministère de l'Enseignement Supérieur et de la Recherche en partenariat avec la Banque Publique d'Investissement. Future société *Stégosécu ISIS*, www.stegosecu.fr.
3. Projet *APACHE*, **labélisé par les pôles de compétitivité** : Micro-techniques, Véhicule du Futur et Solutions Communicantes Sécurisées (SCS) à Sophia Antipolis en PACA.

1.4.5/ RAPPORTS ET PRÉSENTATIONS INTERNES

1. **Divers :** Autres présentations internes au laboratoire, Institut FEMTO-ST DISC, CNRS, UMR 6174. Journées des Doctorants, présentations des résultats à l'équipe, etc.

1.4.6/ TROIS SALONS PROFESSIONNELS

1. **Innovact'2013 :** Nicolas Friot et Lilia Ziane Khodja, Présentation de la future société *Stégosécu ISIS*, de ses produits innovants, de son équipe et de son savoir-faire. Innovact'2013, The European forum for innovative start-ups - 17ème édition, 26 et 27 mars 2013, Centre des Congrès, Reims, *www.innovact.com*
2. **Sofins'2013 :** Nicolas Friot. Présentation de la future société *Stégosécu ISIS* et des mesures engagées par pour travailler avec les forces spéciales - Sofins'2013, Le premier rendez-vous des forces spéciales avec l'univers industriel et de la recherche, 9, 10 et 11 avril 2013, Camp de Souge, *www.sofins.fr*
3. **Expert-TIC'2013 :** Nicolas Friot. Lors du Carrefour des possibles, présentation de la future société *Stégosécu ISIS* et des mesures engagées pour travailler avec la DGA, la DCRI et les forces spéciales. Expert-TIC'2013, Solutions numériques pour l'entreprise, 12 et 13 Septembre 2013 à Besançon, Parc Micropolis, *www.expertic.fr*

CHAPITRE 2

Notations

> Chaque progrès donne un nouvel espoir, suspendu à la solution d'une nouvelle difficulté. Le dossier n'est jamais clos.
>
> CLAUDE LEVI-STRAUSS, ANTHROPOLOGUE ET ETHNOLOGUE (1908-2009)

Dans tout le document, pour prévenir tout conflit et pour éviter des écritures illisibles, nous avons pris pour convention les notations suivantes, utilisées habituellement en mathématiques discrètes :

- Le $n-$ième terme de la suite s sera noté s^n.
- La $i-$ième composante du vecteur v sera notée v_i.
- La $k-$ième composée de la fonction f sera notée f^k. Ainsi, $f^k = f \circ f \circ \ldots \circ f$, k fois.
- La dérivée de la fonction f sera notée f'.

Par exemple, soit $u : \mathbb{N} \to \mathbb{R}^2$ une suite de $\mathbb{R}^2$. Alors u^0 désigne le premier terme de cette suite ; c'est un vecteur à deux composantes : u_1^0 et u_2^0.

D'autre part, $\mathbb{B}$ désignera l'ensemble $\{0; 1\}$ muni de ses lois usuelles d'algèbre de Boole (addition, multiplication et négation booléennes), et l'on notera de plus $\wedge$ le *et*, $\vee$ le *ou*, et $\oplus$ le *ou exclusif booléen*. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ sont les notations habituelles des ensembles respectifs suivants : entiers naturels, entiers relatifs, nombres rationnels et nombres réels. L'ensemble $\mathcal{X}^{\mathcal{Y}}$ est l'ensemble des applications de $\mathcal{Y}$ dans $\mathcal{X}$, et donc $\mathcal{X}^{\mathbb{N}}$ désigne l'ensemble des suites à valeurs dans $\mathcal{X}$. Enfin, $[\![a; b]\!] = \{a, a + 1, \ldots, b\}$ est l'ensemble des entiers compris entre a et b.

Nous utiliserons, de plus, la notation $\lfloor x \rfloor$ pour désigner la partie entière d'un réel x, c'est-à-dire le plus grand entier inférieur à x. $\lceil x \rceil$ désignera, pour sa part, le plus petit entier supérieur à x. La partie décimale de x est, quant à elle, notée $\{x\}$. Enfin, soit $(n, p) \in \mathbb{Z}^2$, alors le reste de la division euclidienne de n par p est noté $n \mod p$.

II

ÉTAT DE L'ART

CHAPITRE 3

La science de l'information dissimulée

> Un problème créé ne peut être résolu en réfléchissant de la même manière qu'il a été créé.
>
> ALBERT EINSTEIN, PHYSICIEN (1879-1955)

Ce premier chapitre vise à introduire quelques notions servant de base à tous les travaux à venir. Nous présenterons ici de manière générale et sans technique les tenants et les aboutissants de la science de l'information dissimulée. Cette introduction générale exposera, notamment, les deux sous-domaines de cette science que sont la stéganographie et le tatouage numérique. Ce chapitre se terminera par quelques rappels de topologie mathématique, rappels utiles à la définition de la sécurité topologique introduite dans le chapitre suivant.

3.1/ PRÉSENTATION GÉNÉRALE

Nous appellerons *information dissimulée* (« information hiding ») toute étude sur les techniques visant à insérer, de manière discrète, une quelconque information au sein d'un contenant donné. Nous ne nous intéresserons qu'aux contenus et contenants numériques. Cette information dissimulée regroupe des sous-domaines variés, tels que la stéganographie et la stéganalyse, le tatouage numérique, le fingerprinting, *etc.* Certains de ces sous-domaines sont définis ci-après. Pour les autres, nous renvoyons le lecteur aux travaux de Stefan Katzenbeisser [48].

L'information dissimulée consistant à cacher un message dans un autre, il faut avant toutes choses commencer par donner des noms à ces deux types de messages. Nous ne formulerons pas ici de définitions précises, nous souhaitons juste introduire qualitativement les objets de notre étude [41].

Définition 1 : Hôte, média de couverture, tatouage et filigrane

Supposons que l'on camoufle un message dans un support donné, par une quelconque technique d'information dissimulée.

- Le support recevant le message s'appelle l'*hôte*, ou encore la *couverture* ou *média de couverture*.
- Le message caché est aussi appelé la *marque*, le *filigrane*, ou le *tatouage*.

3.2/ STÉGANOGRAPHIE VERSUS TATOUAGE NUMÉRIQUE

Les deux domaines les plus étudiés de l'information dissimulée sont la stéganographie et le tatouage numérique (digital watermarking). Nous les présentons dans ce qui suit.

3.2.1/ LA STÉGANOGRAPHIE

3.2.1.1/ PRÉSENTATION

La stéganographie est très ancienne, il suffit pour s'en convaincre de se référer à l'encre sympathique et aux acrostiches qui en sont deux exemples très classiques. Cette technique a pour but de cacher au sein d'un message primaire une information secondaire. Ce message invisible doit être uniquement accessible à des personnes propriétaires d'une information secrète, c'est à dire à des initiés. L'enjeu est donc d'éviter que des adversaires détectent la présence d'une information camouflée dans un support d'apparence anodine.

Nous donnons ci-dessous la définition de la stéganographie proposée par Christian Cachin [23], qui est, en quelque sorte, le père de la stéganalyse.

Définition 2 : Stéganographie

La stéganographie est l'art et la science de communiquer de telle sorte que la présence d'un message ne peut pas être détectée.

En d'autres termes, la *stéganographie* est la technique consistant à cacher des messages de sorte que, en dehors de l'expéditeur et du destinataire, nul ne puisse en soupçonner l'existence. L'objet de la stéganographie est donc de faire passer inaperçu un message dans un autre, et non de rendre un message uniquement intelligible à qui-de-droit – ce qui est le rôle de la cryptographie [41]. Cette technologie, quoique très ancienne, connaît un regain d'intérêt à l'ère du numérique [33].

Pour qu'une technique de stéganographie soit viable, il faut bien sûr que les adversaires ne puissent découvrir le contenu caché. Mais cela ne suffit pas : pour bien faire, il faudrait que ces adversaires ne puissent même pas savoir qu'un contenu est caché, et qu'ils ne puissent pas empêcher sa transmission (volontairement ou non). Enfin, ils ne devraient pas être capables d'envoyer une fausse information, en tentant d'usurper l'identité de l'expéditeur via le canal de communication secret.

Nous reviendrons plus en détails concernant ces pré-requis spécifiques à la stéganographie. Signalons cependant, dès à présent, l'application type de cette stéganographie : la

création d'un canal secret d'échange [33].

3.2.1.2/ APPLICATIONS DE LA STÉGANOGRAPHIE

Canal secret et anonymat On pense souvent que les personnes qui recherchent l'anonymat sur Internet ont quelque chose de mal ou de honteux à cacher. Ainsi, les logiciels tels que les proxys ou Tor [66, 27], permettant de garantir la vie privée et l'anonymat, ne seraient-ils utilisés que par des terroristes, des pédophiles, des marchands d'armes... De tels outils devraient donc être interdits. Cependant, le terrorisme et la pédophilie existent indépendamment d'Internet. En outre, les actualités récentes nous rappellent que, dans de nombreux endroits à travers le monde, avoir une opinion différente de celle imposée par les leaders, notamment politiques et religieux, est quelque chose de dangereux. Le blogger saoudien Hamza Kashgari a, par exemple, été condamné à mort après ses tweets sur Mohammed [91]. De même, le printemps arabe et le conflit en Syrie ont mis en lumière les faits suivants. Tout d'abord, Internet est un média important, difficile à réduire au silence, utile à la démocratie, à sa transparence et aux efforts pour combattre la corruption. Ensuite, revendiquer ses opinions, faire du journalisme ou de la politique, s'avère être une activité dangereuse, encore à l'heure actuelle, dans différents États de la planète.

De ce fait, différents logiciels ont émergé ces dernières décennies dans le but de préserver la vie privée sur Internet. Leurs auteurs considèrent en effet que la liberté d'expression est un droit fondamental qui doit être protégé, que les journalistes doivent être capables d'informer la communauté sans danger, et qu'il n'est pas normal qu'on puisse risquer sa vie en défendant les droits de l'homme. L'outil le plus connu est sans doute Tor, qui route le trafic internet à travers un réseau de serveurs bénévoles dans le monde. Un autre exemple de ce type est donné par Perseus [31], un plugin firefox [32] qui protège les données personnelles, sans pour autant enfreindre les réglementations nationales en matière de cryptographie. Finalement, les serveurs proxys anonymes déployés dans le monde peuvent aussi préserver l'anonymat.

Ces trois solutions ne sont pas sans défauts. Par exemple, la requête n'est pas anonyme au niveau du serveur proxy d'anonymisation, ce qui déplace simplement le problème : ces serveurs proxy sont-ils dignes de confiance ? Perseus peut être cassé à l'aide d'ordinateurs suffisamment puissants. Enfin, en raison de sa renommée et de sa conception particulière, Tor est ciblé par de nombreuses attaques et présente diverses faiblesses (par exemple, il ne protège pas contre le contrôle du trafic à ses extrémités).

La stéganographie est une alternative à ces approches, permettant une forme d'anonymat sur Internet [42]. Elle peut être utilisée de plusieurs façons : création de canaux secrets dans des images d'arrière-plan de sites internets, au sein des galeries photographiques sur Facebook, dans les flux audios et vidéos, ou dans les caractères non interprétés dans les codes source HTML.

3.2.2/ LE TATOUAGE NUMÉRIQUE

3.2.2.1/ PRÉSENTATION

Le tatouage numérique, en anglais « digital watermarking », n'impose pas exactement les mêmes contraintes ; il doit son développement à la transition actuelle vers le tout numérique. Les données numériques ont pour avantages sur leurs homologues analogiques d'être d'une manipulation plus sûre, d'un stockage moins coûteux, d'une transmission plus rapide et d'une indexation plus facile. *A contrario*, la copie à l'identique devient aisée, ce qui conduit à divers problèmes liés à la propriété intellectuelle. Le tatouage numérique (digital watermarking) est apparu pour lutter contre cette fraude : il consiste à rajouter sur un média (image, son ou vidéo) une marque imperceptible et robuste. Imperceptible, pour ne pas détériorer le média concerné. Robuste, pour pouvoir être détectée même après traitement et modification du média – qu'il soit involontaire ou résultant d'une attaque du droit d'auteur. Ici, le but n'est pas de transmettre une information secrète, mais de marquer l'appartenance : la stéganographie recherche la dissimulation, quand le marquage désire avant tout la robustesse.

Comme rappelé dans [41], le tatouage numérique est une discipline récente, née au début des années 90. Le terme fut introduit en 1992 par Andrew Tirkel et Charles Osborne [80] et les premiers articles portant sur ce domaine sont dus à Tanaka *et ses coauteurs* [78], et Caronni, Tirkel *et ses coauteurs* [48]. Enfin, Ingemar Cox [28] popularisa l'étalement de spectre, une de ses plus célèbres techniques.

Ce tatouage numérique (ou marquage) peut se définir de la manière suivante, due à Teddy Furon [39] :

> **Définition 3 : Tatouage numérique**
>
> Le *tatouage numérique* est l'art de cacher des métadonnées dans du contenu numérique de manière robuste.

Cette définition en entraîne une autre, celle de robustesse. Comme rappelé dans [41], la définition de Kalker [47] fait autorité [39, 62] :

> **Définition 4 : Robustesse (tatouage)**
>
> Le *tatouage robuste* est le mécanisme consistant à créer un canal de communication multiplexé dans le contenu original, et dont la capacité se dégrade comme une fonction continue de la dégradation du contenu tatoué.

Suivant Teddy Furon, remarquons que le terme « cacher » a plusieurs sens, selon les auteurs. Il signifie tantôt que l'embarquement des métadonnées ne cause aucune distorsion perceptible, et tantôt sous-entend « d'une manière sûre » [39]. Cette « sécurité » dans le tatouage peut se formuler comme suit (Kalker [47]) :

> **Définition 5 : Sécurité (tatouage)**
>
> La sécurité se réfère à l'incapacité qu'ont les utilisateurs non autorisés d'avoir accès au canal de tatouage. Un tel accès fait référence à la tentative de supprimer, détecter et estimer, écrire et modifier les bits de tatouage.

Cette définition bien acceptée [39, 62] est cependant trop générale et pas assez tech-

nique pour être utile. Comme nous le verrons par la suite, des définitions plus mathématiques ont été proposées ces dix dernières années pour donner plus de consistance à cette notion. Avant cela, nous signalerons le fait que la robustesse est certes intrinsèquement reliée à la définition de tatouage, mais que dans certains cas cette robustesse est justement rejetée. Il s'agit de tatouages dits « fragiles », et nous les présentons ci-dessous.

3.2.2.2/ CATÉGORIES DE TATOUAGES NUMÉRIQUES

À la différence de la stéganographie, le tatouage peut être visible [41].

Tatouages robustes, visibles ou non Il est fréquent que les photographes ajoutent un tatouage visible en forme de copyright aux versions de prévisualisation (basse résolution) de leurs photographies, afin d'éviter que ces dernières ne soient utilisées à la place des versions haute résolution payantes. Les photos haute résolution vendues par l'agence possèdent elles aussi un tatouage, mais invisible. Ce dernier ne dégrade donc pas le contenu visuel, mais permet de détecter l'éventuelle source d'une utilisation ou d'une distribution illégale. Le tatouage peut, en effet, être l'identifiant de l'acheteur : en cas d'utilisation non-autorisée, l'agence peut alors se retourner contre ce dernier.

Tatouage fragile *A contrario*, il existe aussi des tatouages dits *fragiles* [41]. Ce sont des tatouages invisibles, qui sont utilisés pour détecter toute modification du support hôte, par exemple, pour vérifier que le contenu n'a pas été modifié par un tiers (comme dans les sceaux numériques ajoutés aux preuves apportées au dossier, lors d'un procès, quand ces dernières sont numériques).

3.2.2.3/ APPLICATIONS DU TATOUAGE NUMÉRIQUE

Le tatouage numérique a un champ d'applications assez vaste dont quelques exemples sont donnés ci-dessous.

Copyright Le tatouage numérique permet habituellement d'ajouter des informations de copyright ou d'autres messages de vérification à un fichier, à un signal audio ou vidéo, à une une image ou à tout autre document numérique. Le message caché est un ensemble de bits, dont le contenu dépend de l'application : nom ou identifiant du créateur, du propriétaire, de l'acheteur...

La marque est généralement de petite taille. Dans le détail, elle peut être constituée par :

1. Des informations reliées à la possibilité de dupliquer le document : *copie illimitée*, *copie interdite*, ou *une copie autorisée*. À charge pour le matériel adéquat de permettre ou non la copie, en fonction du marquage détecté.
2. Le nom du propriétaire du document, prévenant ainsi toute usurpation d'une œuvre.
3. L'ayant-droit du document. Le tatouage s'apparente alors à l'empreinte de l'utilisateur du support. Il s'agit là de *fingerprinting* évoqué précédemment dans ce rapport.

Ces marques doivent être telles qu'un nombre limité d'ayants-droit doivent ne pas pouvoir les retrouver (les supprimer) en comparant leurs copies.

Authentification de documents numériques La justice n'a aucun moyen, à l'exception des scellés, de déterminer si une preuve a été manipulée. Dans le cas où les preuves ont été numérisées, on peut insérer une marque fragile, qui sera détruite à la moindre tentative de manipulation de la preuve, ce qui en garantira son authenticité au cours d'un procès. Contrairement à la protection du droit d'auteur, on souhaite ici que le tatouage soit très fragile.

Enrichissement de données et gestion électronique documentaire (GED) L'enrichissement des données numériques est un autre exemple d'application des techniques de tatouage numérique. Par exemple, on peut ajouter diverses informations (composition du groupe, paroles, *etc.*) à une chanson numérisée. Comme ces données peuvent être altérées par l'utilisateur qui, sans penser à mal, réalise des opérations telles que la compression des données, un certain niveau de robustesse est requis.

Dans le même esprit, on pourrait vouloir cacher dans un scanner ou une radio anonyme des données sur le patient concerné, afin d'éviter tout risque d'erreur et permettre le développement de la télé-médecine tout en garantissant le secret médical. Cette situation relève à la fois de la stéganographie et du tatouage numérique, selon ce qui apparaît le plus important pour le manipulateur : garantir la confidentialité des données (stéganographie) ou éviter leur perte (tatouage).

Le Web sémantique Les moteurs de recherche sociaux sont souvent présentés comme une approche nouvelle qui sera utilisée dans la prochaine génération de moteurs de recherche. Dans cette conception, les contenus comme des photos ou des films sont étiquetés ("tagués") avec des descriptions mises en œuvre par les contributeurs. Les résultats fournis par les moteurs de recherches sont alors enrichis de ces descriptions. Ce marquage collaboratif, utilisé par exemple sur les sites Frick [2] et Delicious [1], peut participer au développement d'un Web sémantique dans lequel chaque page web contient des métadonnées lisibles par une machine, métadonnées qui décrivent son contenu. L'intérêt de cette approche réside dans la possibilité de réaliser des recherches sociales sans site web, ni bases de données : les descriptions sont directement dans les médias, quels que soient leurs formats. Pour atteindre cet objectif et intégrer ces métadonnées, les technologies fondées sur la dissimulation d'informations peuvent être utilisées.

3.2.2.4/ MÉTHODE D'ÉVALUATION

Le tatouage ne doit pas trop dégrader l'image hôte. Pour mesurer cette dégradation, on peut utiliser le PSNR, défini ci-dessous.

Définition 6 : PSNR

Le *PSNR* (*Peak Signal to Noise Ratio*), ou rapport signal sur bruit, est la mesure de distorsion suivante :

$$PSNR = 10 \cdot \log_{10}\left(\frac{d^2}{MSE}\right)$$

où

- d est la *dynamique du signal*. Dans le cas standard d'une image où les composantes d'un pixel sont codées sur 8 bits, d = 255.
- MSE (*Mean Squared Error*) est l'*erreur quadratique moyenne*. Elle est définie, pour deux images I_o et I_r de taille $m \times n$, comme étant égale à :

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I_o(i,j) - I_r(i,j))^2$$

Cette mesure de distorsion est aussi utilisée en stéganographie pour évaluer la discrétion d'un signal.

CHAPITRE 4

Rappels de topologie

> Toutes les idées sur lesquelles repose aujourd'hui la société ont été subversives avant d'être tutélaires.
>
> ANATOLE FRANCE, ÉCRIVAIN, PRIX NOBEL DE LITTÉRATURE (1844-1924)

Les études et les développements autour de la théorie du chaos n'ont cessé de s'opérer depuis la fin du XX^e siècle. Ce fut Robert L. Devaney qui le premier réussit à fournir une définition précise de la notion de chaos en mathématiques [30]. Même si plusieurs définitions parallèles, différentes et complémentaires furent fournies dès 1975, notamment dans les travaux suivants [71, 34, 35], c'est dans le cadre de la théorie de Devaney, proposée en 1989, que nous travaillerons.

Les études exposées dans la suite de ce rapport utiliseront largement la théorie de Devaney pour étudier la sécurité des algorithmes de dissimulation d'informations. Cette nouvelle approche de la sécurité en science de l'information dissimulée sera expliquée dans le chapitre 5. Le présent chapitre rappel donc les bases de la théorie de Devaney ainsi que les notions topologiques sur lesquelles cette théorie est fondée.

4.1/ ESPACES TOPOLOGIQUES

Nous rappelons, dans cette section, les types d'espaces qui nous intéressent, ainsi que les fonctions que l'on utilise sur ces espaces.

4.1.1/ ESPACES TOPOLOGIQUES, OUVERTS ET VOISINAGES

On introduit ici les termes de base de la topologie [72].

Définition 7 : Espace topologique

On appelle *espace topologique* la donnée d'un couple $(\mathcal{X}, \tau)$, où $\mathcal{X}$ est un ensemble et τ une famille de parties de $\mathcal{X}$, qu'on appelle des *ouverts*, vérifiant :

- $\varnothing, \mathcal{X} \in \tau$: l'ensemble vide et $\mathcal{X}$ sont des ouverts,
- une réunion quelconque d'ouverts est un ouvert,
- une intersection finie d'ouverts est un ouvert.

Définition 8 : Fermé

On appelle *fermé* le complémentaire d'un ouvert.

Notation 1 :

Soit A une partie d'un espace topologique $(\mathcal{X}, \tau)$. Le plus petit fermé contenant A existe toujours : c'est l'intersection de tous les fermés contenant A. On notera $\overline{A}$ cet ensemble, et l'on parlera de *fermeture topologique*, ou d'*adhérence* de A.

On peut munir un ensemble donné de plusieurs topologies différentes, ce qui conduit à la notion d'ordre suivante :

Définition 9 : Finesse des topologies

Une topologie $\tau \in \mathcal{P}(\mathcal{X})$ sur l'ensemble $\mathcal{X}$ est *plus fine* qu'une topologie $\tau' \in \mathcal{P}(\mathcal{X})$, si $\tau' \subset \tau$. On dira alors que τ' est *moins fine*, ou *plus grossière* que τ.

Remarque 1 :

Une topologie est donc plus fine qu'une autre si elle a plus d'ouverts. Cet ordre n'est pas total : il existe des topologies non comparables entre-elles.

Introduisons enfin la notion de voisinage :

Définition 10 : Voisinage

Soit $(\mathcal{X}, \tau)$ un espace topologique. On appelle *voisinage de* $x \in \mathcal{X}$ toute partie de $\mathcal{X}$ contenant un ouvert qui contient x.

4.1.2/ EXEMPLES DE TOPOLOGIES

Donnons dès à présent deux exemples de topologies sur un ensemble $\mathcal{X}$ donné [72] :

Définition 11 : Topologie discrète

La *topologie discrète* sur un ensemble $\mathcal{X}$ est la topologie $\tau = \mathcal{P}(\mathcal{X})$ de toutes les parties de $\mathcal{X}$.

La topologie discrète est la topologie de $\mathcal{X}$ possédant le plus d'ouverts. Elle est plus fine que toute autre topologie sur $\mathcal{X}$, et vérifie en particulier la propriété suivante : tout sous-ensemble de $\mathcal{X}$ est à la fois ouvert et fermé. Cette topologie tire son nom du fait que tous les points sont isolés (chaque point est à la fois ouvert et fermé). *A contrario* :

Définition 12 : Topologie grossière

La *topologie grossière* sur un ensemble $\mathcal{X}$ est la topologie $\tau = \{\varnothing, \mathcal{X}\}$.

C'est la topologie la moins fine qui soit sur $\mathcal{X}$, elle ne contient que deux ouverts.

4.1.3/ DISTANCES ET ESPACES MÉTRIQUES

Une manière commode de définir des topologies est d'utiliser des métriques.

Définition 13 : Distance

Sur un ensemble $\mathcal{X}$, une *distance*, aussi appelée *métrique*, est une application $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ possédant les propriétés suivantes :

Symétrie : $\forall x, y \in \mathcal{X}, d(x, y) = d(y, x)$.

Séparation : $\forall x, y \in \mathcal{X}, d(x, y) = 0 \Leftrightarrow x = y$.

Inégalité triangulaire : $\forall x, y, z \in \mathcal{X}, d(x, z) \leqslant d(x, y) + d(y, z)$.

Définition 14 : Espace métrique

On appelle *espace métrique* la donnée d'un couple $(\mathcal{X}, d)$, où $\mathcal{X}$ est un ensemble et d une distance sur $\mathcal{X}$.

Les espaces métriques, qui sont des espaces topologiques, possèdent des voisinages particuliers, appelés boules :

Définition 15 : Boule fermée, boule ouverte

Soit $(\mathcal{X}, d)$ un espace métrique. La *boule fermée* centrée en un point P, et de rayon réel r, est l'ensemble $\overline{\mathcal{B}}(P, r)$ des points dont la distance à P est inférieure ou égale à r : $\overline{\mathcal{B}}(P, r) = \{M \in \mathcal{X} \,/\, d(M, P) \leqslant r\}$.
La boule ouverte est l'ensemble $\mathcal{B}(P, r) = \{M \in \mathcal{X} \,/\, d(M, P) < r\}$.

4.2/ SYSTÈMES DYNAMIQUES DISCRETS ET ORBITES

Soit $f : \mathcal{X} \longrightarrow \mathcal{X}$ une application d'un espace topologique ou métrique $\mathcal{X}$ dans lui-même. On considère la suite des itérées définies par la relation de récurrence :

$$\begin{cases} x^0 \in \mathcal{X} \\ \forall n \in \mathbb{N}, x^{n+1} = f(x^n) \end{cases}$$

Le comportement de ces itérées dépend de la fonction f, et de l'espace sur lequel on itère. D'où la définition [34] :

Définition 16 : Système dynamique discret

Un *système dynamique discret* est un couple $(\mathcal{X}, f)$ formé par :

- un espace topologique non vide $(\mathcal{X}, \tau)$, appelé *espace des phases*,
- une fonction continue $f : \mathcal{X} \longrightarrow \mathcal{X}$, appelée *fonction successeur*.

La fonction f peut, dans certains cas, être inversée, ce qui permet alors de « remonter dans le temps ». On parle alors de *réversibilité* [34] :

Définition 17 : Système dynamique discret réversible

Un système dynamique discret $(\mathcal{X}, f)$ est dit *réversible* si f est un homéomorphisme (topologique), i.e. si f est une bijection bicontinue.

Enfin, lorsqu'on s'intéresse à la manière, dont un point x considéré, évolue au cours du temps, on parle d'orbite.

Définition 18 : Orbite

Pour $x \in \mathcal{X}$ donné, la suite $(f^{(n)}(x))_{n\in\mathbb{N}}$ est appelée *mouvement positif*, ou *orbite* de x. Elle est notée γ_x.

4.2.1/ PÉRIODICITÉ, ÉQUILIBRE ET RÉGULARITÉ

4.2.2/ PÉRIODICITÉ

La théorie du chaos cherche à savoir si le comportement d'un système dynamique discret $x^{n+1} = f(x^n)$ peut être prévu ou pas, c'est-à-dire si l'on peut deviner quelle va être l'orbite γ_x d'un point x donné. En ce sens, les points dont le comportement est le plus facile à appréhender sont les points périodiques et les points d'équilibre.

Définition 19 : Point périodique

Un point $p \in \mathcal{X}$ est dit *périodique* de *période* k si k est un entier naturel non nul tel que $f^k(p) = p$, et $\forall h \in [\![0; k-1]\!], f^h(p) \neq p$.

Notation 2 : Ensemble des points périodiques

On note $Per_k(f)$ l'ensemble des points k-périodiques de f, et $Per(f)$ l'ensemble des points périodiques de période quelconque.

La périodicité peut intervenir après une phase plus ou moins longue de transition, ce qui nous amène à introduire une variante à la précédente définition.

Définition 20 : Point ultimement périodique

Un point est *ultimement périodique* s'il existe deux entiers n et p tels que $f^{n+p}(x) = f^p(x)$. Soit n_0 le plus petit n vérifiant cela. L'ensemble $\{x, f(x), \ldots, f^{n_0}(x)\}$ est alors appelé *transitoire* de x, et n_0 est la *longueur du transitoire*.

4.2.3/ ÉQUILIBRE

Les points d'équilibre sont les points ayant la plus simple des orbites.

Définition 21 : Points d'équilibre

Les points périodiques de période 1 sont appelés *points fixes* de f, ou encore *points d'équilibre*. De même, les points ultimement périodiques de période 1 sont appelés *points ultimement fixes*.

4.2.4/ ESPACES DENSES ET SYSTÈMES RÉGULIERS

En topologie, le concept de densité d'un sous-ensemble A d'un espace topologique $\mathcal{X}$ permet de traduire l'idée que pour tout point x de $\mathcal{X}$ on peut trouver un point de A qui soit aussi proche de x que l'on souhaite [72].

Définition 22 : Espace Dense

Soit $\mathcal{X}$ un espace topologique et A un sous-ensemble de $\mathcal{X}$. On dit que A est *dense* dans $\mathcal{X}$ si pour tout élément x de $\mathcal{X}$, tout voisinage de x contient au moins un point de A.

Nous sommes maintenant en mesure de définir un premier aspect du chaos, à partir des points périodiques [34] :

Définition 23 : Systèmes Réguliers

Un système dynamique discret $(\mathcal{X}, f)$ est dit *régulier* si l'ensemble des points périodiques de f, $Per(f)$, est dense dans $\mathcal{X}$.

Remarque 2 : Caractérisation de la régularité

Dans un espace métrique $(\mathcal{X}, d)$, le système dynamique $(\mathcal{X}, f)$ est régulier si et seulement si $\forall x \in \mathcal{X}, \forall \varepsilon > 0, \exists p \in Per(f), d(x, p) < \varepsilon$.

Bien que le terme « régulier » semble s'opposer à l'idée de « chaos », on pourrait quand même considérer que, sous un certain angle, la définition ci-dessus serait susceptible d'engendrer un certain type particulier d'imprévisibilité : si l'on fait une simulation numérique de l'évolution du système, de petites erreurs dans les conditions initiales peuvent mener sur une orbite radicalement différente de celle qui fait l'objet de la simulation. Par exemple, passer d'une petite période vers une grande période, ou vers une absence de période.

Cette régularité ne permet évidemment pas de définir, à elle seule, une notion de chaos car il y a trop de cas problématiques. Avec cette seule propriété, des systèmes élémentaires seraient eux aussi chaotiques : l'identité, les permutations, etc.

4.3/ STABILITÉ

Après avoir introduit les notions relatives à la périodicité dans la section 4.2.2 et les notions relatives à l'équilibre dans la section 4.2.3, il nous reste à introduire quelques concepts relatifs à l'allure des orbites.

Commençons par définir ce qu'est une orbite stable.

Définition 24 : Orbite stable

Une orbite positive γ_x est dite *stable* si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall y \in \mathcal{X}, d(x,y) < \delta \Longrightarrow \forall n \in \mathbb{N}, d(\gamma_x^n, \gamma_y^n) < \varepsilon$$

x est alors appelé *point stable* de f.

En d'autres termes, si y est proche de x, alors l'orbite de y sera proche de celle de x. Au voisinage d'un point stable x, tous les points évoluent de la même manière : si l'on commet une petite erreur sur la condition initiale, on peut garantir que l'erreur entre le phénomène observé et l'évolution théorique reste minime.

Plus précisément, un système sera stable quand deux points proches conduiront à deux orbites similaires. Dans le cas contraire on aura à nouveau de l'imprévisibilité : instabilité, sensibilité ou expansivité. L'instabilité est donc simplement la négation de la stabilité.

Définition 25 : Instabilité

Un mouvement positif γ_x est *instable* si

$$\exists \varepsilon > 0, \forall \delta > 0, \exists y \in \mathcal{X}, \exists n \in \mathbb{N}, d(x,y) < \delta \text{ et } d(\gamma_x^n, \gamma_y^n) \geqslant \varepsilon$$

Un système ayant tous ses points stables est fortement prévisible : on le qualifiera de «stable». Dans la situation diamétralement opposée, on parlera de système instable :

Définition 26 : Système (in)stable

Un système dynamique discret est *stable* si toutes ses orbites positives sont stables. Il est dit *instable* si toutes ses orbites positives sont instables.

Remarque 3 : Orbite dense

On peut aussi parler d'orbite dense en lien avec la notion de densité d'espace introduite à la définition 22.

4.4/ L'INDÉCOMPOSABILITÉ D'UN SYSTÈME

La théorie du chaos cherche à définir ce qu'est un système dynamique ayant un comportement complexe, imprévisible. Un des prérequis est qu'il ne puisse pas se simplifier en sous-systèmes plus faciles à étudier. Diverses définitions topologiques permettent de donner corps à une telle idée. Nous détaillerons ici la transitivité, la transitivité forte, et le mélange topologique.

4.4.1/ TRANSITIVITÉ TOPOLOGIQUE

Définition 27 : Transitivité topologique

f est dite *topologiquement transitive* si, pour tout couple d'ouverts $U, V \subset \mathcal{X}$, il existe $k > 0$ tel que $f^k(U) \cap V \neq \varnothing$.

4.4.2/ TRANSITIVITÉ FORTE

Nous rappelons maintenant la définition de la transitivité forte.

Définition 28 : Transitivité forte

Un système itératif f est dit *fortement transitif* sur l'espace topologique $(\mathcal{X}, \tau)$ si et seulement si, pour tous points $A, B \in \mathcal{X}$ et tout voisinage V de B, $n_0 \in \mathbb{N}$ et $X \in V$ peuvent être trouvés tels que $f^{n_0}(X) = A$.

En d'autres mots, pour tout couple (x, y), il existe un point aussi proche que l'on veut de x pour lequel une de ses itérées est égale à y.

4.4.3/ MÉLANGE TOPOLOGIQUE

Le mélange topologique, quant à lui, se définit de la manière suivante.

Définition 29 : Mélange topologique

Un système itératif est dit topologiquement mélangeant si et seulement si pour tout couple d'ouverts disjoints $U, V \neq \varnothing$, $n_0 \in \mathbb{N}$ peut être trouvé tel que $\forall n \geqslant n_0, f^n(U) \cap V \neq \varnothing$.

4.5/ COMPACITÉ ET DE FORTE TRANSITIVITÉ

Dans cette section, nous introduisons une nouvelle notion topologique, appelée : compacité [72].

4.5.1/ RAPPELS CONCERNANT LA COMPACITÉ TOPOLOGIQUE

La notion de compacité, que nous rappelons dans cette partie, est notamment utile pour simplifier l'étude de certaines propriétés topologiques. Afin de donner la définition de la compacité topologique, nous devons tout d'abord introduire deux autres définitions préliminaires :

4.5.2/ ESPACE SÉPARÉ, ESPACE DE HAUSDORFF

Définition 30 : Espace séparé (espace de Hausdorff)

Un *espace de Hausdorff* (ou *espace séparé*) est un espace topologique dans lequel deux points distincts choisis de façon arbitraire admettent des voisinages disjoints. On dit alors que ces deux points sont *séparés*.

Remarque 4 :

Tous les espaces métriques sont des espaces de Hausdorff.

4.5.3/ Recouvrement d'un espace

Définition 31 : Recouvrement d'un espace

Soit $\mathcal{X}$ un espace topologique, alors un recouvrement $\mathcal{C}$ de $\mathcal{X}$ est une collection de sous-ensembles $(U_i)_{i\in I}$ de $\mathcal{X}$ dont la réunion est égale à l'espace $\mathcal{X}$ tout entier. Dans ce cas, nous dirons que $\mathcal{C}$ recouvre l'espace $\mathcal{X}$, ou que les ensembles $(U_i)_{i\in I}$ recouvrent $\mathcal{X}$.

$\mathcal{C}$ est appelé recouvrement ouvert si tous ses éléments sont des ensembles ouverts.

Un sous-recouvrement de $\mathcal{C}$ est un sous-ensemble de $\mathcal{C}$ qui recouvre encore l'espace $\mathcal{X}$.

4.5.4/ Compacité et caractérisation séquentielle

Définition 32 : Compacité

Un espace topologique de Hausdorff (ou espace séparé) est dit compact si tout recouvrement ouvert de l'espace admet un sous-recouvrement fini.

La caractérisation séquentielle de la compacité pour un espace métrique peut à présent être rappelée. Cette caractérisation pourra alors être utilisée dans le but de démontrer la forte transitivité du processus $\mathcal{CIS}_2$ (*cf.* chapitre 8 section 8.4.6) :

Proposition 1 : Caractérisation séquentielle de la compacité

Un espace métrique $(\mathcal{X}, d)$ est compact si et seulement si toute suite infinie de $\mathcal{X}$ admet une sous-suite convergente.

4.5.5/ Lien entre compacité et forte transitivité

La forte transitivité est introduite dans la définition 28 page précédente.

Nous pouvons à présent rappeler la proposition suivante, issue des travaux d'Enrico Formenti [34] :

Proposition 2 : Forte transitivité sur un espace métrique compact

Dans un espace métrique compact, transitivité et forte transitivité sont équivalentes.

4.6/ Suites de Cauchy et complétude

On introduit maintenant la notion d'espace complet. Ce sont des espaces métriques tels que certaines suites particulières, dites suites de Cauchy, convergent.

Définition 33 : Suite de Cauchy

Une suite $(x^n)_{n\in\mathbb{N}}$ d'un espace métrique $(\mathcal{X}, d)$ est dite *de Cauchy* si pour tout réel $\varepsilon > 0$, il existe un entier naturel N tel que pour tout couple d'entiers (p, q) vérifiant $p, q \geqslant N$, la distance $d(x^p, x^q)$ soit inférieure à ε : $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p, q > N, d(x^p, x^q) < \varepsilon$.

Une suite de Cauchy est donc une suite dont les termes se rapprochent à partir d'un certain rang. Ces suites sont les suites les plus susceptibles de converger. En fait, les suites convergentes sont des suites de Cauchy, mais la réciproque n'est pas toujours vraie. D'où la définition :

Définition 34 : Espace métrique complet

Un espace métrique $(\mathcal{X}, d)$ est dit *complet* si toute suite de Cauchy de $(\mathcal{X}, d)$ a une limite dans $(\mathcal{X}, d)$.

Intuitivement, un espace est complet s'il « n'a pas de trou », s'il « n'a aucun point manquant ». Par exemple, les nombres rationnels ne forment pas un espace complet, puisque $\sqrt{2}$ n'y figure pas alors qu'il existe une suite de Cauchy de nombres rationnels ayant cette limite. Il est toujours possible de « remplir les trous » amenant ainsi à la complétion d'un espace donné [83].

Dans le cadre de l'étude de la science de l'information dissimulée et de l'évaluation de la sécurité des processus, dont cette thèse fait l'objet, la propriété de complétude d'un espace s'avère, elle aussi, particulièrement utile. En effet, la complétude est une propriété qualitative supplémentaire dont peuvent faire preuve les schémas de dissimulation d'informations étudiés. Une telle propriété s'avère intéressante, par exemple, pour la comparaison avec d'autres processus topologiquement sûrs (*cf.* section 5.2.2) qui, eux, pourraient ne pas posséder cette propriété. La complétude permet donc d'apprécier davantage le niveau de sécurité des algorithmes étudiés.

Cette notion est reprise au chapitre 8 section 8.4.8 pour évaluer le niveau de sécurité du processus $\mathcal{CIS}_2$.

4.7/ SYSTÈMES DYNAMIQUES DISCRETS PARFAITS

4.7.1/ POINTS D'ACCUMULATION ET SYSTÈMES PARFAITS

Pour finir cette section, on introduit la définition d'espaces particuliers dits *parfaits*, dans lesquels la transitivité s'obtient dès que l'on a découvert un point d'orbite dense.

Définition 35 : Point d'accumulation

On dit qu'un point x d'un espace topologique $(\mathcal{X}, \tau)$ est un *point d'accumulation* d'une partie A de $\mathcal{X}$ si tout voisinage de x contient une infinité de points de A.

Définition 36 : Systèmes parfaits

Un système dynamique discret $(\mathcal{X}, f)$ est dit *parfait* si chaque point de $\mathcal{X}$ est un point d'accumulation dans $\mathcal{X}$.

Comme pour la complétude, la perfection des systèmes permet d'apprécier davantage le niveau de sécurité des algorithmes étudiés. Cette notion est notamment reprise au chapitre 8 section 8.4.8 pour évaluer le niveau de sécurité du processus $\mathcal{CIS}_2$.

En tenant compte de la définition d'orbite dense donnée dans la remarque 3 et de la définition de la transitivité donnée à la section 4.4, on peut montrer que :

4.7.2/ PERFECTION, TRANSITIVITÉ ET ORBITES DENSES

Proposition 3 : Perfection et transitivité

Si $(\mathcal{X}, f)$ est parfait et a une orbite dense, alors il est transitif.

Cette proposition est intéressante car elle nous donne un autre moyen de prouver l'une des propriétés nécessaires à la mise en évidence de systèmes chaotiques telle que l'a définie Devaney (*cf.* section 4.9).

La réciproque de cette propriété, démontrée par Enrico Formenti [34], est elle aussi intéressante, car, à partir d'un système transitif sur un espace compact, elle permet de prouver l'existence d'une orbite dense :

Proposition 4 : Existence d'une orbite dense

Supposons que l'espace métrique $(\mathcal{X}, d)$ est compact. Si $(\mathcal{X}, d)$ est transitif, alors il possède une orbite dense.

Démonstration. Reprise de la démonstration d'Enrico Formenti [34].

Si $(\mathcal{X}, d)$ est compact, alors il est séparable. Soit $(U_1, \ldots, U_n) \subset \mathcal{X}^n$ une base dénombrable d'ouverts. Construisons une orbite qui intersecte tous les U_n :

Par transitivité, $\exists\, n_0 \in \mathbb{N}, U_0 \cap f^{n_0}(U_1) \neq \varnothing$. Soient :

- V_0 ouvert tel que $\overline{V_0} \subset U_0 \cap f^{-n_0}(U_1)$,
- $\vdots$,
- V_i ouvert tel que $\overline{V_i} \subset U_0 \cap f^{-n_i}(U_i)$.

Par compacité, $V = \bigcup_{i \in \mathbb{N}} \overline{V_i}$ est non vide, et $\forall x \in V, \forall i \in \mathbb{N}, x \in f^{n_1}(x) \in U_i$. □

4.8/ PROPRIÉTÉS QUANTITATIVES

Les définitions précédentes sont purement qualitatives. Il existe cependant des moyens de mesurer quantitativement l'impact d'une erreur sur la condition initiale. A cette fin, la sensibilité et l'expansivité ont été introduites dans la littérature. L'expansivité se définit de la manière suivante.

4.8.1/ EXPANSIVITÉ ET CONSTANTE D'EXPANSIVITÉ

Définition 37 : Expansivité

Une fonction f possède la propriété d'*expansivité* si :

$$\exists \varepsilon > 0, \forall x \neq y, \exists n \in \mathbb{N}, d(f^n(x), f^n(y)) \geqslant \varepsilon.$$

ε est appelé la *constante d'expansivité* de f.

La constante d'expansivité vérifie la propriété suivante : une erreur arbitrairement petite sur toute condition initiale est *toujours* amplifiée jusqu'à atteindre, à un moment donné, ε. Dans le cas de la sensibilité définie ci-dessous, elle *peut* l'être.

4.8.2/ SENSIBILITÉ ET CONSTANTE DE SENSIBILITÉ

Définition 38 : Sensibilité aux conditions initiales

f a une *sensible dépendance aux conditions initiales* s'il existe $\delta > 0$ tel que, pour tout $x \in \mathcal{X}$ et pour tout voisinage V de x, il existe $y \in V$ et $n \geqslant 0$ tel que $d(f^n(x), f^n(y)) > \delta$. δ est appelé *constante de sensibilité* de f.

Remarque 5 :

Les autres propriétés des systèmes dynamiques discrets, telles que, par exemple, la transitivité, la forte transitivité, le mélange topologique, etc. (voir section 4.4) sont qualifiées de propriétés qualitatives.

4.9/ SYSTÈMES DYNAMIQUES DISCRETS, CHAOTIQUES POUR DEVANEY

4.9.1/ DÉFINITION

Nous sommes maintenant en mesure de donner la définition de Devaney :

Définition 39 : Fonction chaotique

Une fonction $f : \mathcal{X} \longrightarrow \mathcal{X}$ est dite *chaotique* sur $\mathcal{X}$ si :

1. $(\mathcal{X}, f)$ est régulière,
2. f est topologiquement transitive,
3. f a une sensible dépendance aux conditions initiales.

Lorsque f est chaotique, alors le système $(\mathcal{X}, f)$ est chaotique, et, selon Devaney, il est imprévisible en raison de sa sensible dépendance aux conditions initiales. Il ne peut pas être décomposé ou simplifié en deux sous-systèmes qui n'interagissent pas, du fait de la transitivité topologique. De plus, au milieu de ce comportement aléatoire, nous avons tout de même un élément de régularité : des comportements fondamentalement différents sont par conséquent possibles et surviennent de façon imprévisible.

Signalons pour finir le théorème de Banks [19], qui montre que seules les propriétés de régularité et de transitivité sont nécessaires pour satisfaire le chaos de Devaney.

Théorème 1 : Théorème de Banks

Si un système dynamique discret sur un espace métrique donné est à la fois régulier et transitif, alors il est sensible à sa condition initiale.

Notons que ce théorème ne donne aucune information sur la constante de sensibilité, il ne s'agit là que d'un résultat de nature qualitative.

4.9.2/ IMPACT DE LA TOPOLOGIE

La définition énoncée ci-dessus semble dépendre du choix de la topologie. Nous avons donc cherché à déterminer dans quelle mesure une topologie plus fine (contenant plus d'ouverts) modifiait ou non la propriété d'être chaotique selon Devaney. Le résultat, publié dans [9], se résume comme suit.

Théorème 2 : Finesse des topologies et chaos

Soit $\mathcal{X}$ un ensemble, et τ, τ' deux topologies sur $\mathcal{X}$ telles que τ' soit plus fine que τ. Soit $f : \mathcal{X} \to \mathcal{X}$, une fonction continue pour les deux topologies τ et τ'.
Si $(\mathcal{X}_{\tau'}, f)$ est chaotique, au sens de Devaney, alors $(\mathcal{X}_{\tau}, f)$ est aussi chaotique.

Démonstration. Soit ω_1, ω_2 deux ouverts de la topologie τ. Alors $\omega_1, \omega_2 \in \tau'$, car τ' est plus fine que τ. Comme f est τ'-transitive, alors $\exists n \in \mathbb{N}, \omega_1 \cap f^n(\omega_2) = \varnothing$. Par conséquent, f est aussi τ-transitive.

Nous allons à présent établir la régularité de $(\mathcal{X}_{\tau}, f)$, *i.e.*, pour tout $x \in \mathcal{X}$ et tout τ-voisinage V de x, il existe un point périodique f dans V. Soit $x \in \mathcal{X}$, et $V \in \mathcal{V}_{\tau}(x)$ un τ-voisinage de x. Par définition d'un voisinage, $\exists \omega \in \tau, x \in \omega \subset V$. Cependant $\tau \subset \tau'$, alors $\omega \in \tau'$, et donc $V \in \mathcal{V}_{\tau'}(x)$. Mais $(\mathcal{X}_{\tau'}, f)$ est régulière. Il existe alors un point périodique pour f dans V, et la régularité de $(\mathcal{X}_{\tau}, f)$ est prouvée.

Concernant la sensibilité aux conditions initiales, dans le cas des espaces métriques, le résultat est immédiat puisque toute boule ouverte au sens de la topologie τ est incluse dans une boule ouverte au sens de la topologie τ'. □

4.10/ CONTINUITÉ ET CARACTÉRISATION SÉQUENTIELLE

La notion de continuité se définit dans les espaces topologiques de la manière suivante [72] :

Définition 40 : Continuité (cas des espaces topologiques)

Soit f une application entre deux espaces topologiques. Elle est dite *continue en* x si pour tout voisinage V de $f(x)$, il existe un voisinage de x dont l'image par f est incluse dans V.

La continuité s'exprime plus simplement lorsque l'on considère des applications entre espaces métriques.

Définition 41 : Continuité (cas des espaces métriques)

Soient (X, d) et (X', d') deux espaces métriques. Soient $a \in X$ et $f : X \to X'$. On dit que l'application f est *continue en a* si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in X, d(x, a) \leq \eta \implies d'(f(x), f(a)) \leq \varepsilon.$$

Dans les espaces métriques, la continuité se montre facilement en utilisant la caractérisation séquentielle suivante :

Proposition 5 : Caractérisation séquentielle de la continuité

Soit $f : (X, d) \to (X', d')$ une application entre deux espaces métriques.
Alors f est continue en $a \in X$ si et seulement si pour toute suite $(x^n)_{n\in\mathbb{N}}$ convergeant vers a, la suite $(f(x^n))_{n\in\mathbb{N}}$ converge vers $f(a)$.

CHAPITRE 5

La stéganalyse

> Si la science évolue, c'est souvent parce qu'un aspect encore inconnu des choses se dévoile soudain.
>
> FRANÇOIS JACOB, CHERCHEUR EN BIOLOGIE, PRIX NOBEL (1920-2013)

Ce chapitre expose les différentes notions de sécurité proposées concernant la dissimulation d'informations.

5.1/ APPROCHE CLASSIQUE DE LA SÉCURITÉ POUR LA DISSIMULATION D'INFORMATIONS

5.1.1/ HISTORIQUE

Le premier travail fondamental, relatif à la sécurité en matière de dissimulation d'informations, a été réalisé par Christian Cachin au début des années 2000, dans le contexte de la stéganographie [23]. Dans cet article, Christian Cachin explique que, les tentatives d'un attaquant de faire la distinction entre une image innocente et un contenu stéganographique se résument finalement à un test d'hypothèses. Les propriétés basiques d'un système stéganographique sont définies par Christian Cachin en utilisant les notions d'entropie, d'information mutuelle et d'entropie relative. Au même moment, Thomas Mittelholzer a proposé le premier cadre théorique pour l'analyse de la sécurité de schémas de tatouage numérique [60].

Ces efforts pour apporter un cadre théorique à la sécurité pour la stéganographie et le tatouage numérique ont été poursuivis par Ton Kalker, qui a tenté de clarifier les concepts (robustesse, sécurité, etc.). Ton Kalker a également tenté de donner une classification des attaques portant sur le tatouage numérique [47]. Ce travail a été approfondi par Teddy Furon et ses co-auteurs, qui ont traduit le principe d'Auguste Kerckhoffs, communément

appelé principe de Kerckhoffs, (Alice et Bob ne doivent compter que sur un secret partagé précédemment en privé), utilisé dans le domaine de la cryptographie, au domaine de la dissimulation d'informations [38]. Ils ont utilisé la méthodologie de Diffie et Hellman, ainsi que le contexte cryptographique de Shannon [73] afin de donner une classification des attaques, selon le type d'information auquel Eve a accès [25, 63]. Ces catégories sont détaillées ci-dessous.

5.1.2/ LES CONFIGURATIONS D'ATTAQUES ET LEUR IMPACT

Le cadre habituel est celui du problème du prisonnier de Simmons [75] : Alice et Bob sont en prison et veulent concevoir un plan d'évasion en échangeant des messages cachés (filigranes) dans des documents hôtes d'allure innocente. Les documents sont transmis par une gardienne, Eve, qui les étudie, et peut choisir d'interrompre la communication quand elle le veut (*cf.* figure 5.1). La détermination du niveau de sécurité d'une technique de tatouage dépend alors du contexte dans lequel cette dernière est censée œuvrer et des catégories d'attaques auxquelles elle est censée résister.

FIGURE 5.1 – Le problème du prisonnier de Simmons [75]

Ces catégories peuvent être listées selon le type d'information auquel Eve a accès [25, 63] :

Attaque de l'objet tatoué seul. L'adversaire n'a accès qu'à des contenus tatoués pour réaliser ses attaques. Cette catégorie d'attaque est notée, dans la littérature, **WOA**, pour *Watermarked Only Attack*.

Attaque du message connu. L'adversaire a accès à des couples de contenus tatoués avec leurs messages cachés correspondants. Cette catégorie d'attaque est notée **KMA**, pour *Known Message Attack*.

Attaque de l'original connu. Eve a accès à des couples de contenus tatoués avec leurs messages originaux. Cette catégorie d'attaque est notée **KOA**, pour *Known Original Attack*.

Attaque du message constant. L'adversaire observe plusieurs contenus tatoués, et sait seulement que le message caché est le même dans tous les contenus. Cette catégorie d'attaque est notée **CMA**, pour *Constant Message Attack*.

Attaque de l'original estimé. L'attaquant n'a accès qu'à une estimation de l'original du signal hôte. Cette catégorie d'attaque est notée **EOA**, pour *Estimated Original Attacks*.

La classe d'attaque EOA a, quant à elle, été introduite dans [79].

Ces différentes attaques sont résumées dans le tableau 5.1.

Classe d'attaque	Contenu original	Contenu stéganographique	Message caché
WOA (Attaque de l'objet tatoué seul)		×	
KMA (Attaque du message connu)		×	×
KOA (Attaque de l'original connu)	×	×	
CMA (Attaque du message constant)			×
EOA (Attaque de l'original estimé)	× (Estimé)		

TABLE 5.1 – Classification des attaques dans le contexte de Kalker [47]

5.1.3/ APPROCHE PROBABILISTE DE CAYRE ET BAS

François Cayre et Patrick Bas ont ensuite proposé dans [24] une approche de la sécurité basée sur la théorie des probabilités, dans le contexte WOA, valable à la fois pour la stéganographie et le tatouage numérique. Le cadre est une fois de plus le problème du prisonnier de Simmons (*cf.* figure 5.1).

5.1.3.1/ NOTATIONS PRÉLIMINAIRES

Commençons par introduire quelques notations dans le cadre du problème du prisonnier [24] :

- N_c est la taille disponible (en bits) dans les vecteurs hôtes, pour y cacher des messages, N_v est la taille du vecteur hôte, et N_0 est le nombre de contenus observés.
- $\mathbf{X}$ est un ensemble de vecteurs représentant une collection de N_0 contenus originaux (hôtes), chaque élément x de $\mathbf{X}$ étant un vecteur tiré aléatoirement dans un espace E de dimension N_v.
- $\mathbf{Y}$ est un ensemble de vecteurs représentant une collection de N_0 contenus tatoués, chaque élément y de $\mathbf{Y}$ étant un vecteur aléatoire à N_v coordonnées.
- $\mathbb{K} \subset \mathbb{R}$ est un ensemble fini de taille N_K, dit « ensemble des clés ».
- Enfin, $e : \mathbb{K} \times E \to E$ représente une fonction de tatouage.

Le respect du principe de Kerckhoffs [49] appliqué à la fonction de tatouage (ou fonction d'embarquement) utilisée par Alice et Bob, permet de supposer qu'Alice et Eve peuvent toutes deux construire une estimation parfaite des différentes lois de probabilités entrant ici en jeu : celle des contenus originaux, celle des contenus tatoués, *etc.* Listons plus précisément ces lois de probabilités [24] :

- $p(X)$ est la loi de probabilité de N_0 contenus originaux (les hôtes). Bien que nous nous placions dans la configuration WOA, nous supposons quand même Eve capable de modéliser la loi de X : $p(X) = p(x_0, \ldots, x_{N_0-1})$. Cette hypothèse est le pire

des cas, du point de vue d'Alice et de Bob, pour lequel *Eve a pu estimer quelle pouvait être la loi des contenus originaux, rien qu'en observant attentivement les contenus tatoués qu'elle transporte*.

- $p(Y)$ est la loi de probabilité de N_0 contenus tatoués, chaque contenu l'ayant été avec une clé différente. Cette loi peut être obtenue par Eve en utilisant suffisamment la fonction de tatouage qu'elle a à sa disposition (principe de Kerckhoffs).
- $p(Y_K)$ est la loi de probabilité de N_0 contenus tatoués, chaque contenu l'ayant été avec une seule et unique clé secrète $K \in \mathbb{K}$. Ce modèle peut être construit par Eve, simplement en observant les contenus qu'elle transporte, vu que l'*on suppose qu'Alice et Bob ne changent pas leur clé secrète K au cours de leurs échanges*.
- $p(Y|K_i)$ est la loi de probabilité de N_0 contenus retatoués, chaque contenu tatoué (par Alice et Bob) ayant été retatoué par Eve en utilisant la même clé connue $K_i \in \mathbb{K}$. Ce modèle peut être construit par Eve, car l'on suppose le principe de Kerckhoffs respecté : l'adversaire a accès à la fonction de tatouage e, et peut retatouer, avec sa propre clé K_i, des messages tirés aléatoirement dans une collection Y de contenus déjà tatoués.

Comme les supports hôtes sont supposés indépendants les uns des autres, les lois précédentes sont en fait des lois marginales : $p(X) = p(x_0) \times \ldots \times p(x_{N_0-1})$. La même chose vaut pour $p(Y)$ et $p(Y|K)$.

5.1.3.2/ QUATRE CLASSES DE SÉCURITÉ

Le but d'Eve est de trouver la constante K_e qui maximise la probabilité $p(Y_K|K_e)$, c'est-à-dire la clé K_e qui a été la plus sûrement utilisée pour tatouer les Y_K.

Une des originalités de l'approche de Cayre et Bas consiste à considérer que la geôlière peut être passive ou *active* (Eve peut détecter la présence d'un message caché, estimer ce dernier, puis agir éventuellement sur la communication). Cette considération conduit les auteurs de [24] à considérer quatre niveaux de sécurité dans le WOA, le plus faible niveau étant l'insécurité :

Définition 42 : Insécurité

La fonction de marquage e est *non-sûre* (insecure) si et seulement si $\exists K_1 \in \mathbb{K}, p(Y|K_1) = p(Y_K)$, et $\forall K_2 \in \mathbb{K}, p(Y|K_2) \neq p(Y_K)$.

Dans ce cas de figure, il existe une unique clé K_1 dont le modèle associé $p(Y|K_1)$ des contenus tatoués avec cette clé, correspond exactement au modèle $p(Y_K)$ des observations. Ainsi, la méthode consistant à rechercher, par exemple de manière exhaustive, la clé aboutissant à la meilleure correspondance, a des chances sérieuses d'aboutir : l'estimation de K_e est possible [24].

La deuxième définition introduite par Cayre et Bas est la *key-security* :

Définition 43 : Sécurité de la clé (key-security)

La fonction de marquage e est *key-secure* si et seulement si $\exists \mathcal{S}_K \subset \mathbb{K}, card(\mathcal{S}_K) > 1, \forall K_1 \in \mathcal{S}_K, p(Y|K_1) = p(Y_K)$.

Dans la définition précédente, $K \in \mathcal{S}_K$, et $\mathcal{S}_K$ correspond à l'ensemble des clés qui ne modifient pas le modèle probabiliste des observations. Si le schéma est non-sûr, alors un tel S_K n'existe pas. On peut constater aussi que, dans ce cas, quand bien même il est impossible d'estimer la clé secrète, il est cependant possible de trouver $\mathcal{S}_K$, ce qui est un risque (bien qu'un attaquant ne puisse pas retrouver K à partir de $\mathcal{S}_K$ [24]).

Dans ce cas de figure, la sécurité d'un schéma utilisant la clé privé K repose sur la taille de $\mathcal{S}_K$. Cette propriété de sécurité est aussi reliée, dans une certaine mesure, à la robustesse, dans le sens où une telle sécurité autorise une faible distorsion du support. Cette classe est le niveau de sécurité le plus élémentaire, selon Cayre et Bas [24], au moins lorsque l'on ne désire pas permettre d'accès non autorisé en lecture/écriture dans le canal secret d'échange (Eve est passive).

Dans le cas particulier où $\mathcal{S}_K = \mathbb{K}$, on parle alors de « subspace-security ».

Définition 44 : Sécurité du sous-espace (subspace-security)

La fonction de marquage e est *subspace-secure* si et seulement si $\forall K_1 \in \mathbb{K}, p(Y|K_1) = p(Y_K)$.

En cas de « subspace-security », Eve ne sera pas capable de distinguer la bonne clé de toute autre clé, même en cas d'étude exhaustive. Il lui est donc impossible d'estimer le sous-espace $\mathcal{S}_K$ associé à la clé secrète K, car Y et K sont indépendants. La subspace-security entraîne la key-security, et conduit au fait qu'il n'y a aucune fuite d'information entre les contenus tatoués et la clé secrète. En effet, dans ce cas de figure, la clé secrète est équivalente à toute autre clé, et donc Eve ne peut rien déduire de ses observations.

D'un autre côté, si le schéma considéré est key-secure, mais pas subspace-secure, alors Eve pourra estimer, à partir d'un nombre suffisant d'observations, le sous-espace $\mathcal{S}_K$, mais pas K. Elle pourra alors concentrer tous ses efforts sur $\mathcal{S}_K$.

Enfin, la *stégo-sécurité* est ce qui peut se produire de pire pour Eve :

Définition 45 : Stégo-sécurité

La fonction de tatouage e est *stégo-sûre* (stego-secure) si et seulement si $\forall K_i \in \mathbb{K}, p(Y|K_i) = p(X)$.

La stégo-sécurité signifie notamment que la connaissance de K n'aidera pas à faire la différence entre $p(X)$ et $p(Y)$. Cette définition entraîne la propriété suivante :

$$p(Y|K_1) = \cdots = p(Y|K_{N_k}) = p(Y) = p(X).$$

On peut montrer que cette propriété est équivalente à une divergence de Kullback-Leibler nulle, ce qui est la définition du *secret parfait*, définition proposée par Cachin [23].

Notons D_{KL} la divergence de Kullback-Leibler, alors :

Proposition 6 : Interprétation de la stégo-sécurité

Stégo-sécurité $\Longrightarrow$ Secret parfait $(D_{KL}(p(Y)|p(X)) = 0)$.

Dans ce cas de figure, il est impossible pour Eve de savoir si le contenu qu'elle étudie est passé par la fonction de tatouage, ou pas : elle ne peut tirer aucune information à partir des contenus transmis.

Signalons enfin que, si le schéma considéré est subspace-secure, mais pas stégo-sûr, alors Eve, qui ne pourra estimer ni K, ni $\mathcal{S}_K$, pourra quand même distinguer les contenus innocents des contenus marqués. En d'autres termes, d'un point de vue stéganalytique, le schéma ne serait pas sûr, car la fonction d'embarquement ne respecterait pas le secret parfait de Cachin. En revanche, le schéma sera sûr du point de vue du tatouage numérique.

5.2/ APPROCHE TOPOLOGIQUE DE LA STÉGANALYSE

5.2.1/ APPORTS DE L'APPROCHE TOPOLOGIQUE POUR LA SÉCURITÉ

Les notions de sécurité présentées dans la section précédente n'ont de sens que dans la configuration WOA, et ont forcément pour cadre le problème du prisonnier de Simmons [75]. De plus, il s'agit dans ces études d'un cadre particulier de ce problème : les messages véhiculés par Eve ont été tatoués avec une seule et unique clé secrète pour tous les messages. Ce cadre est relativement restrictif, comme l'ont indiqué Cayre et Bas eux-mêmes [24] : « Comme dans les autres travaux de ce genre, nous considérons qu'Alice et Bob n'utilisent qu'une seule clé. Certes, dans les applications réelles, en particulier dans la stéganographie, il est hautement souhaitable de modifier la clé à chaque communication entre Alice et Bob. »

En outre, l'existence d'une loi de probabilité pour la couverture est nécessaire et, comme l'a déclaré Cachin [23], « Supposer l'existence d'une loi de probabilité pour la couverture semble rendre notre modèle peu réaliste en pratique » : il n'y a pas de modèle canonique, pour la couverture, qui puisse être utilisé dans tous les scénarios. Cette existence même peut faire défaut, les couvertures n'ayant aucune nécessité à se plier à pareille exigence. Enfin, Alice et Bob peuvent chercher soit à induire en erreur Eve, en l'incitant par un quelconque moyen à supposer une mauvaise loi de probabilité, soit s'intéresser à des ensembles de couvertures tels qu'Eve ne puisse jamais déterminer $p(\mathbf{X})$.

Pour répondre à ces problèmes et afin d'être en mesure d'étudier les configurations d'attaques de type KMA, KOA, et CMA, nous avons proposé un nouveau cadre théorique pour l'étude de la sécurité des systèmes de dissimulation d'informations.

Cette nouvelle approche est fondée sur la topologie et la théorie du chaos. Les concepts utilisés dans cette approche ont été rappelés au chapitre 4.

Dans ce contexte, un algorithme de dissimulation d'informations est considéré comme une machine dont le mécanisme est public. Cette machine reçoit le message à cacher, la clé secrète et le support hôte à utiliser, puis renvoie le contenu tatoué. Avec cette approche, la sécurité du système dépend du comportement imprévisible (désordonné) de la machine : il y a une faille de sécurité si un adversaire est capable de prédire les lieux où le filigrane peut se trouver, c'est-à-dire prédire l'image de la machine pour une entrée quelconque.

Pour donner plus de consistance à cette nouvelle approche et à la notion d'imprévisibilité (de désordre), il a été prouvé dans [41] que :possible.

Théorème 3 : Modélisation d'un algorithme par un système dynamique discret

Toute méthode de dissimulation d'informations peut être modélisée sous la forme d'un système dynamique discret de la forme :

$$x^0 \in \mathcal{X}, x^{n+1} = f(x^n).$$

Par conséquent, on peut donc relier l'imprévisibilité du schéma de tatouage à certains aspects topologiques de sa fonction associée f, aspects issus de la théorie mathématique du chaos rappelée précédemment. Ce nouveau cadre théorique définit une certaine ap-

proche de la sécurité pour la dissimulation d'informations. Il respecte le principe de Kerckhoffs. Il est basé sur une description topologique, alors que la plupart des études dans ce domaine ont généralement utilisé la théorie des probabilités [62, 39]. Le but de cette recherche est avant tout de combler l'absence de notion de sécurité dans les configurations CMA, KOA et KMA. Accessoirement, on obtiendra de ce fait un outil supplémentaire permettant d'évaluer la sécurité dans le cadre WOA ; ce qui ne nous semble pas sans intérêt : ne peut-on penser qu'en matière de sécurité, plus grands sont le nombre, la variété et la différence de points de vue, mieux c'est ?

Ainsi, contrairement aux modèles proposés dans la section précédente, la sécurité topologique, qui sera définie au chapitre suivant, peut être utilisée dans les configurations KOA, KMA et CMA. De plus, dans le cas particulier des algorithmes de tatouage basés sur le chaos, on sera en mesure de vérifier si l'affirmation d'un comportement chaotique tient ou non : une telle vérification semble naturelle, de même qu'évaluer la force de ce comportement chaotique semble intéressant quand on commence une étude de sécurité d'un algorithme supposé topologiquement-sûr.

5.2.2/ SÉCURITÉ TOPOLOGIQUE

Pour vérifier si un schéma de dissimulation d'informations S est topologiquement sécurisé ou pas, S doit être écrit comme un processus itératif $x^{n+1} = f(x^n)$ sur un espace métrique $(\mathcal{X}, d)$. Comme vu précédemment, cette formulation est toujours possible. Et alors,

Définition 46 : Sécurité topologique

Un système de dissimulation d'informations S est dit topologiquement-sécurisé sur $(\mathcal{X}, d)$ si, son processus itératif associé, a un comportement chaotique au sens de Devaney.

5.2.3/ CARACTÉRISATION DE LA SÉCURITÉ-TOPOLOGIQUE, NIVEAUX DE SÉCURITÉ

Nous avons vu à la section précédente comment définir une nouvelle approche de sécurité en science de l'information dissimulée : *la sécurité topologique*. Ainsi, un schéma de dissimulation d'informations est-il sécurisé s'il est imprévisible [12]. Son processus itératif associé doit donc satisfaire la propriété du chaos de Devaney, et son niveau de sécurité topologique augmente avec le nombre de propriétés topologiques qu'il satisfait.

Précisons quelque peu ce que l'on entend exactement par « niveau de sécurité topologique ».

Dans le chapitre 4 voué aux rappels de certains concepts issus de la topologie mathématiques, nous avons exposé plusieurs propriétés topologiques utiles à l'évaluation qualitative et quantitative du niveau de sécurité des schémas de dissimulation d'informations topologiquement sûrs.

Dans leurs travaux, portant sur une approche probabiliste de l'évaluation de la sécurité [24] (*cf.* section 5.1.3), François Cayre et Patrick Bas ont mis en évidence plusieurs niveaux de sécurité (insécurité, key-security, subspace-security et Stégo-sécurité) pour cette approche.

De la même manière que François Cayre et Patrick Bas, nous avons proposé plusieurs niveaux de sécurités pour notre nouvelle approche topologique. Le niveau de sécurité premier est la sécurité topologique. Les niveaux de sécurité supérieurs sont caractérisés par les propriétés topologiques vérifiées par l'algorithme étudié (*cf.* chapitre 4).

Nous donnons ici une synthèse des différentes propriétés dont nous disposons pour caractériser le niveau de sécurité topologique des processus de dissimulation d'informations.

S'agissant des *propriétés qualitatives*, nous disposons à présent des propriétés suivantes : la régularité, la transitivité, la transitivité forte, le mélange topologique, la compacité, la complétude et la perfection.

S'agissant des *propriétés quantitatives*, nous disposons à présent des propriétés suivantes : la sensibilité aux conditions initiales, l'expansivité et l'exposant de Lyapunov (comme nous le verrons au chapitre 9).

5.3/ L'ÉTALEMENT DE SPECTRE

5.3.1/ PRÉSENTATION DE LA TECHNIQUE

Nous rappelons maintenant une méthode bien connue de la littérature, à savoir la stéganographie par étalement de spectre. En guise d'exemple de ce qui précède, et à des fins de comparaison avec nos méthodes à venir, nous en étudierons la sécurité dans ce qui suit.

Introduisons quelques notations pour commencer. Soit $x \in \mathbb{R}^{N_v}$ un vecteur hôte, dans lequel on souhaite cacher un message $m \in \{0;1\}^{N_c}$. Un générateur de nombres pseudo-aléatoires (PRNG) est utilisé afin d'obtenir des vecteurs dits « porteurs de secret » : $\left\{u^i \in \mathbb{R}^{N_v}, i \in [\![0; N_c - 1]\!]\right\}$.

L'*étalement de spectre* est une famille de techniques dans laquelle le message m est dissimulé dans l'hôte x, pour obtenir le contenu tatoué y, défini par :

$$y = x + w, \tag{5.1}$$

où $+$ est la somme de vecteurs de $\mathbb{R}^{N_v}$, et où le filigrane w a été construit à partir de m, de l'une des manières suivantes, selon la technique retenue.

Étalement de spectre dit « classique » : $w = \sum_{i=0}^{N_c-1} \gamma(-1)^{m^i} u^i$, où $\gamma \in \mathbb{R}$ est un niveau de distorsion fixé. On parle dans ce cas de « modulation BPSK » [24].

Étalement de spectre amélioré : $w = \sum_{i=0}^{N_c-1} \left((-1)^{m^i}\alpha - \lambda \frac{< x, u^i >}{||u^i||^2}\right) u^i$, où $\alpha \in \mathbb{R}$ et $\lambda \in \mathbb{R}$ sont calculés pour réaliser une distorsion moyenne acceptable, et minimiser la probabilité d'erreur lors de l'extraction du filigrane [24]. Cette technique se note aussi ISS (*Improved Spread Spectrum*, voir [54]).

Étalement de spectre dit « naturel » : $w = \sum_{i=0}^{N_c-1} -\left(1 + \eta\,(-1)^{m^i} \frac{< x, u^i >}{| < x, u^i > |}\right) \frac{< x, u^i >}{||u^i||^2}\, u^i$, où $\eta \in \mathbb{R}$ sert à nouveau à fixer un certain niveau de distorsion donné.

L'étalement de spectre naturel, est souvent appelé « Natural Watermarking ». Nous le noterons $\mathcal{NW}$ dans la suite de ce rapport.

Nous allons dans ce qui suit reformuler ces techniques de dissimulation, et ce afin de pouvoir étudier leur sécurité.

5.3.2/ MODÉLISATION DES TECHNIQUES D'ÉTALEMENT DE SPECTRE

Supposons que la taille des filigranes soit bornée par une valeur finie N_b donnée :

$$\max\left(\left\{\|w\|_\infty \,/\, w \in \mathbb{R}^{N_v}\right\}\right) \leqslant \mathsf{N}_b.$$

Cette borne peut être aussi grande que l'on veut ; cependant une très grande valeur de N_b est en contradiction avec les objectifs de dissimulation des données.

Soient $\overline{\mathcal{X}} = \left(\llbracket -\mathsf{N}_b, \mathsf{N}_b \rrbracket^{N_v}\right)^{\mathbb{N}} \times \mathbb{R}^{N_v}$ et $\overline{G}((S,E)) = (\overline{\sigma}(S); \bar{\iota}(S) + E)$, où :

- $\overline{\sigma}$ est la fonction *décalage* (shift) définie par :

$$\begin{array}{rccc} \overline{\sigma}: & \left(\llbracket -\mathsf{N}_b; \mathsf{N}_b \rrbracket^{N_v}\right)^{\mathbb{N}} & \longrightarrow & \left(\llbracket -\mathsf{N}_b; \mathsf{N}_b \rrbracket^{N_v}\right)^{\mathbb{N}} \\ & (S^n)_{n\in\mathbb{N}} & \longmapsto & (S^{n+1})_{n\in\mathbb{N}}, \end{array}$$

- et la fonction *initiale* $\bar{\iota}$ est l'application qui transforme une suite en son premier terme :

$$\begin{array}{rccc} \bar{\iota}: & \left(\llbracket -\mathsf{N}_b; \mathsf{N}_b \rrbracket^{N_v}\right)^{\mathbb{N}} & \longrightarrow & \llbracket -\mathsf{N}_b; \mathsf{N}_b \rrbracket^{N_v} \\ & (S^n)_{n\in\mathbb{N}} & \longmapsto & S^0. \end{array}$$

Nous sommes dorénavant en mesure de modéliser l'étalement de spectre sous la forme d'un système itératif, permettant d'en faire son étude de sécurité topologique :

Proposition 7 : Modélisation de l'étalement de spectre sous forme d'itérations chaotiques

Les techniques d'étalement de spectre sont les résultats de N_c itérations du système dynamique :

$$\begin{cases} X^0 \in \overline{\mathcal{X}}, \\ X^{n+1} = \overline{G}(X^n), \end{cases}$$

où X^0 dépend des données initiales et de la technique choisie (cette dépendance est précisée à la section suivante), et le média tatoué est la deuxième coordonnée de X^{N_c}.

Remarquons que la deuxième coordonnée de X^k correspond à l'image hôte après k modifications, tandis que sa première coordonnée explique comment modifier l'hôte à la prochaine itération.

5.3.3/ CONDITIONS INITIALES ET VARIANTES DES TECHNIQUES D'ÉTALEMENT DE SPECTRE

On constate immédiatement que choisir la technique d'étalement de spectre revient à choisir la condition initiale du système de la proposition 7, ce que l'on résume dans les résultats suivants :

Proposition 8 : Modélisation de l'étalement de spectre classique

L'étalement de spectre classique correspond aux itérations de la proposition 7, où la condition initiale $X^0 = (S^0, E^0)$ est définie ainsi : E^0 est le vecteur hôte x, et S^0 est la suite $\left((-1)^{m^0}\gamma\, u^0, (-1)^{m^1}\gamma\, u^1, \ldots, (-1)^{m^{N_c-1}}\gamma\, u^{N_c-1}\right)$ complétée indéfiniment par des vecteurs nuls de $\mathbb{R}^{N_v}$.

Proposition 9 : Modélisation de l'étalement de spectre amélioré

La technique d'étalement de spectre amélioré correspond au système de la proposition 7, où la condition initiale $X^0 = (S^0, E^0)$ est définie par : E^0 est le vecteur hôte x, et S^0 est la suite $\left(\left((-1)^{m^i}\alpha - \lambda\frac{< x, u^i >}{\|u^i\|^2}\right) u^i\right)_{i=0,\ldots,N_c-1}$, complétée avec des vecteurs nuls.

Proposition 10 : Modélisation de l'étalement de spectre naturel

La technique d'étalement de spectre dit naturel correspond au système de la proposition 7, où la condition initiale $X^0 = (S^0, E^0)$ est définie par : E^0 est le vecteur hôte x, et S^0 est la suite $\left(-\left(1 + \eta(-1)^{m^i}\frac{< x, u^i >}{| < x, u^i > |}\right)\frac{< x, u^i >}{\|u^i\|^2}\, u^i\right)_{i=0,\ldots,N_c-1}$, complétée, comme ci-dessus, avec des vecteurs nuls.

5.3.4/ SÉCURITÉ DE L'ÉTALEMENT DE SPECTRE

Nous rappelons le résultat suivant [24] :

Théorème 4 : Stégo-sécurité de l'étalement de spectre naturel ($\mathcal{NW}$)

L'étalement de spectre naturel $\mathcal{NW}$, avec un paramètre de distorsion $\eta = 1$, est une technique de dissimulation d'informations stégo-sûre.

Les théorèmes suivants relatifs à la sécurité topologique, quant à eux, ont été prouvés dans [41]. Le lecteur pourra se référer à ces travaux pour obtenir le détail des démonstrations permettant d'obtenir ces résultats.

Théorème 5 : Sécurité topologique de l'étalement de spectre

Les techniques d'étalement de spectre sont topologiquement-sécurisées. Cet étalement est de plus fortement transitif et topologiquement mélangeant, mais n'est pas expansif sur $(\overline{X}, \overline{d})$. Sa constante de sensibilité est supérieure ou égale à $\frac{N_b}{2}$, où N_b est un majorant de la taille des filigranes (messages) à embarquer dans le contenu hôte.

CHAPITRE 6

Les itérations chaotiques appliquées à la dissimulation d'informations

> Il est hélas devenu évident aujourd'hui que notre technologie a dépassé notre humanité.
>
> ALBERT EINSTEIN, PHYSICIEN (1879-1955)

6.1/ LES ITÉRATIONS CHAOTIQUES

Les itérations chaotiques définissent le cadre théorique à partir duquel nos algorithmes de dissimulation ont été développés. Nous rappelons, dans cette section, les principales définitions et propriétés de ces dernières, telles qu'exposées dans [10].

6.1.1/ DÉFINITIONS

Nous considérons ici un *système* avec un nombre fini d'éléments (ou de *cellules*). Nous noterons N ce nombre d'éléments. La caractéristique principale de ce système réside dans le fait que chaque cellule a un *état booléen*.

Définition 47 : Stratégie chaotique

. Une suite d'éléments appartenant à $[\![1;\mathrm{N}]\!]$ est appelée une *stratégie chaotique*. L'ensemble de toutes les stratégies chaotiques est noté $\mathbb{S}_\mathrm{N}$ (ou simplement $\mathbb{S}$, si la borne N se déduit du contexte).

Remarque 6 :

Dans la suite de ce rapport, les stratégies chaotiques pourront être désignées plus simplement par le terme de *stratégie*.

Donnons maintenant la définition des itérations chaotiques :

Définition 48 : Itérations chaotiques

Soit $f : \mathbb{B}^{\mathbb{N}} \longrightarrow \mathbb{B}^{\mathbb{N}}$ une fonction et soit $S \in \mathbb{S}$ une stratégie chaotique. Les *itérations chaotiques* sont définies par

$$x_i^n = \begin{cases} x_i^{n-1} & \text{si } S^n \neq i \\ \left(f(x^{n-1})\right)_{S^n} & \text{si } S^n = i. \end{cases}$$

En d'autres termes, à la n-ième itération, seule la S^n-ième cellule est "itérée".

Notons que, dans une formulation plus générale, S^n peut être un sous-ensemble de composants et $f(x^{n-1})_{S^n}$ peut être remplacé par $f(x^k)_{S^n}$, où $k < n$. Cette formulation peut décrire, par exemple, la transmission des retards (se référer notamment à [6] ou [7]). Dans ce contexte, nous proposons d'ailleurs une telle formalisation pour l'un des schémas de dissimulation d'informations que nous avons mis au point. Il s'agit du processus $\mathcal{CIS}_2$. Cette formulation est donnée au chapitre 8 à la section 8.5.

Les itérations chaotiques sont un cas particulier des systèmes itératifs qui ont fortement à voir avec les « systèmes dynamiques discrets ». Le terme « systèmes itératifs » désigne un certain nombre d'objets différents, suivant la branche des sciences que l'on considère. Ainsi, en mathématiques appliquées, ce terme possède plusieurs sens. Les itérations chaotiques ont été introduites par D. Chazan et W. Miranker [26], puis ensuite popularisées notamment par Bertsekas [21], Bahi [13, 14], etc. Ces itérations dites chaotiques ont été étudiées particulièrement pour leur convergence sur $\mathbb{R}^{\mathbb{N}}$ par J.-C. Miellou [58, 59] et F. Robert [70]. En ce qui concerne le cadre discret, on pourra se référer à [22]. On pourra donc se reporter à ces différents travaux si l'on souhaite obtenir une définition plus générale des itérations chaotiques.

6.1.2/ NOMBRE DE STRATÉGIES CHAOTIQUES

Il est possible d'évaluer la taille de $\mathbb{S}$. Pour ce faire, il faut d'abord rappeler quelques notions de dénombrabilité.

6.1.2.1/ RAPPELS

Définition 49 : Dénombrabilité

Un ensemble E est dit *dénombrable* s'il existe une bijection de E sur une partie de l'ensemble des entiers naturels $\mathbb{N}$. Dans le cas contraire, cet ensemble est dit *indénombrable*.

Pour les ensembles infinis, c'est-à-dire qui peuvent être mis en bijection avec une partie stricte d'eux-mêmes, on préfère parler de puissance plutôt que de cardinalité. Rappelons ci-dessous ce qu'est la *puissance du continu*.

Définition 50 : Puissance du continu

Un ensemble a la *puissance du continu* c s'il peut être mis en bijection avec $\mathbb{R}$.

Exemple 1 : Deux ensembles ayant la puissance du continu

L'ensemble $\mathcal{P}(\mathbb{N})$ des parties de $\mathbb{N}$, comme l'ensemble $\mathbb{N}^{\mathbb{N}}$ des suites d'entiers, ont tous deux la puissance du continu.

6.1.2.2/ PUISSANCE DE $\mathbb{S}$

Rappelons le résultat prouvé dans [41] :

Proposition 11 : Taille de l'ensemble des stratégies chaotiques

L'ensemble $\mathbb{S}$ des stratégies chaotiques est infini indénombrable, il a la puissance du continu $c = card(\mathbb{R})$.

6.1.3/ APPROCHE PRATIQUE

En pratique, c'est-à-dire sur machine (ordinateur), on n'itère qu'un nombre fini de fois. Cela revient à considérer les stratégies chaotiques ayant un nombre fini de termes. Il n'est pas certain qu'il faille distinguer l'approche théorique de l'approche pratique. Cela dit, cette distinction est aisée et permet d'éviter certaines discussions sur le passage non trivial de la théorie aux nombres machines.

6.1.3.1/ DÉFINITION ET NOTATION

Définition 51 : Stratégie chaotique finie

On appelle stratégie chaotique finie toute suite finie de $[\![1;N]\!]$.

Notation 3 : Ensemble des stratégies chaotiques finies

On note $\tilde{\mathbb{S}}_N$ l'ensemble des stratégies chaotiques finies des systèmes à N cellules. On préférera la notation $\tilde{\mathbb{S}}$ lorsqu'il n'y a pas ambiguïté sur le nombre de cellules.

6.1.3.2/ NOMBRE DE STRATÉGIES CHAOTIQUES FINIES

Commençons par un petit rappel.

Définition 52 : Ensemble des décimaux

On appelle *nombre décimal* un nombre réel n'ayant qu'un nombre fini de décimales.

Notation 4 : Ensemble des nombres décimaux

L'ensemble des nombres décimaux est noté $\mathbb{D}$. Il est infini dénombrable.

Comme $\tilde{\mathbb{S}}$ est l'ensemble des suites entières dont les termes sont bornés par N, *et dont le nombre de termes est fini non borné*, c'est un ensemble infini dénombrable [41] :

Proposition 12 : Taille de l'ensemble des stratégies chaotiques finies

L'ensemble $\tilde{\mathbb{S}}$ des stratégies chaotiques finies est infini dénombrable.

6.2/ TOPOLOGIE DES ITÉRATIONS CHAOTIQUES

Dans cette section, nous donnons quelques éléments de preuve établissant les propriétés topologiques des itérations chaotiques. Comme nos systèmes sont inspirés par ces travaux [43, 11, 10], les preuves, détaillées dans ce rapport, suivront le même schéma.

Nous allons tout d'abord introduire certaines notions et terminologies.

Définition 53 : Adaptateur de stratégies chaotiques

Soit $k \in \mathbb{N}^*$. Un *adaptateur de stratégies chaotiques* est une suite numérique dont tous les éléments appartiennent à l'intervalle entier $[\![0, k-1]\!]$. L'ensemble de toutes les stratégies chaotiques dont les termes appartiennent à $[\![0, k-1]\!]$ est noté $\mathbb{S}_k$.

Nous pouvons généraliser la définition précédente au cas d'un ensemble quelconque.

Définition 54 : Adaptateur de stratégies sur un ensemble $\mathscr{E}$

Soit $k \in \mathbb{N}^*$. Un *adaptateur de stratégies* sur un ensemble $\mathscr{E}$ est une suite numérique dont tous les éléments appartiennent à $\mathscr{E}$. L'ensemble de toutes les stratégies d'un ensemble $\mathscr{E}$ est noté $\mathbb{S}_{\mathscr{E}}$.

Définition 55 : Mesure booléenne discrète

La *mesure booléenne discrète* est l'application $\delta : \mathbb{B} \longrightarrow \mathbb{B}$ définie par $\delta(x, y) = 0 \Leftrightarrow x = y$.

Définition 56 : Fonction initiale

Soit $k \in \mathbb{N}^*$. La *fonction initiale* est l'application i_k définie par :

$$\begin{array}{rccc} i_k : & \mathbb{S}_k & \longrightarrow & [\![0, k-1]\!] \\ & (S^n)_{n\in\mathbb{N}} & \longmapsto & S^0 \end{array}$$

Définition 57 : Fonction décalage

Soit $k \in \mathbb{N}^*$. La *fonction décalage* est l'application σ_k définie par :

$$\begin{array}{rccc} \sigma_k : & \mathbb{S}_k & \longrightarrow & \mathbb{S}_k \\ & (S^n)_{n\in\mathbb{N}} & \longmapsto & (S^{n+1})_{n\in\mathbb{N}} \end{array}$$

Définition 58 : Fonction F_f

Étant donnée une fonction $f : \mathbb{B}^N \to \mathbb{B}^N$, la fonction F_f est définie par :

$$\begin{array}{rl} F_f : & [\![0; N-1]\!] \times \mathbb{B}^N \longrightarrow \mathbb{B}^N \\ & (k, E) \longmapsto \left(E_j.\delta(k,j) + f(E)_k.\overline{\delta(k,j)}\right)_{j \in [\![0;N-1]\!]} \end{array}$$

Définition 59 : Espace des phases $\mathcal{X}_1$

L'espace des phases utilisé pour les itérations chaotiques est par définition : $\mathcal{X}_1 = \mathbb{S}_N \times \mathbb{B}^N$.

Définition 60 : Application G_f

Étant donné une fonction $f : \mathbb{B}^N \to \mathbb{B}^N$, l'application G_f est définie par :

$$\begin{array}{rccc} G_f : & \mathcal{X}_1 & \longrightarrow & \mathcal{X}_1 \\ & (S, E) & \longmapsto & \left(\sigma_N(S), F_f(i_N(S), E)\right) \end{array}$$

À partir de ces définitions, les itérations chaotiques peuvent être décrites par les itérations du système dynamique discret suivant :

Définition 61 : Système dynamique discret pour les itérations chaotiques

$$\begin{cases} X^0 \in \mathcal{X}_1 \\ \forall k \in \mathbb{N}^*, X^{k+1} = G_f(X^k) \end{cases}$$

Enfin, une nouvelle distance d_1 entre deux points a été définie par :

Définition 62 : Distance d_1 sur $\mathcal{X}_1$

$\forall (S, E), (\check{S}, \check{E}) \in \mathcal{X}_1$, $d_1((S, E); (\check{S}, \check{E})) = d_{\mathbb{B}^N}(E, \check{E}) + d_{\mathbb{S}_N}(S, \check{S})$, où :

- $d_{\mathbb{B}^N}(E, \check{E}) = \sum_{k=0}^{N-1} \delta(E_k, \check{E}_k) \in [\![0; N]\!]$
- $d_{\mathbb{S}_N}(S, \check{S}) = \frac{9}{N} \sum_{k=1}^{\infty} \frac{|S^k - \check{S}^k|}{10^k} \in [0; 1].$

sont respectivement deux distances sur $\mathbb{B}^N$ et $\mathbb{S}_N$ ($\forall N \in \mathbb{N}^*$).

Remarque 7 : Interprétation de la distance d_1

Cette nouvelle distance a été introduite dans [10] afin de respecter les exigences suivantes :

1. Lorsque le nombre de cellules différentes, entre deux systèmes, augmente, leur distance devrait augmenter elle aussi.
2. Si deux systèmes présentent les mêmes cellules et que leurs stratégies chaotiques respectives commencent avec le même terme, la distance entre ces deux points doit être faible, car l'évolution de ces deux systèmes sera identique pendant un certain temps.

La distance présentée ci-dessus suit ces recommandations.
En effet, si la valeur entière $\lfloor d(X,Y)\rfloor$ est égale à n, alors les systèmes E et $\check{E}$ diffèrent en n cellules. En outre, $d(X,Y)-\lfloor d(X,Y)\rfloor$ est une mesure de la différence entre les stratégies chaotiques S et $\check{S}$. Plus précisément, cette partie flottante est inférieure à 10^{-k} si et seulement si les k premiers termes des deux stratégies chaotiques sont égaux. En outre, si les k premiers chiffres sont non nuls, alors les k premiers termes des deux stratégies chaotiques sont différents.

Il a finalement été démontré que :

Proposition 13 : Continuité de G_f

G_f est une fonction continue sur $(\mathcal{X}_1, d_1)$, pour tout $f : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$.

Nous pouvons à présent rappeler la définition de la fonction booléenne principalement utilisée dans nos travaux [11] :

Définition 63 : Négation vectorielle booléenne

La *négation vectorielle booléenne* est la fonction définie par :

$$\begin{array}{rccc} f_0 : & \mathbb{B}^{\mathsf{N}} & \longrightarrow & \mathbb{B}^{\mathsf{N}} \\ & (b_0, \cdots, b_{\mathsf{N}-1}) & \longmapsto & (\overline{b_0}, \cdots, \overline{b_{\mathsf{N}-1}}) \end{array}$$

Sur l'espace métrique $(\mathcal{X}_1, d_1)$, G_{f_0} satisfait les trois conditions du chaos de Devaney : régularité, transitivité, et sensibilité aux conditions initiales. Nous pouvons donc rappeler le théorème suivant, démontré dans [11] :

Théorème 6 : G_{f_0} est une fonction chaotique

G_{f_0} est une bien fonction chaotique sur $(\mathcal{X}_1, d_1)$ au sens de Devaney (*cf.* section 4.9).

Jusqu'à présent, lorsque nous parlions d'« itérations chaotiques », le terme « chaotiques » n'avait rien à voir avec la théorique du chaos de Devaney. Ce terme provenais simplement de la littérature dans laquelle on rencontre habituellement ce qualificatif pour désigner ce type particulier d'itérations. Grâce au théorème précédent, nous sommes à présent assuré que le terme est finalement bien choisi, puisque ces itérations chaotiques sont bien chaotiques au sens de Devaney.

Remarquons enfin que [10] :

Proposition 14 : Cardinalité de l'espace des phases $\mathcal{X}_1$

L'espace des phases $\mathcal{X}_1$ a, au moins, la puissance du continu.

Cette dernière proposition nous montre qu'à partir d'un objet naturellement fini l'« ordinateur », grâce aux itérations chaotiques, il est possible de passer à un espace infini indénombrable.

6.3/ ITÉRATIONS CHAOTIQUES POUR LA DISSIMULATION D'INFORMATIONS

Pour expliquer comment utiliser les itérations chaotiques à des fins de dissimulation d'informations, nous devons préalablement définir la notion d'*importance d'un coefficient donné*.

6.3.1/ COEFFICIENTS LES PLUS ET LES MOINS SIGNIFICATIFS

Nous pouvons tout d'abord remarquer que certains termes du contenu hôte original x ont moins de signification que d'autres. Ils peuvent ainsi être modifiés sans que cette modification ne soit remarquée. C'est pourquoi nous introduisons la *fonction de signification*, qui attache un poids à chaque terme définissant le support numérique, en fonction de sa position t.

Définition 64 : Fonction de signification

Une *fonction de signification* est une suite numérique réelle $(u^k)^{k\in\mathbb{N}}$.

Exemple 2 : Exemple de fonction de signification

Considérons un ensemble d'images stockées en niveaux de gris au format "graymap portable" (P3-PGM) : chaque pixel a 256 niveaux de gris possibles, *i.e.* est mémorisé avec huit bits. Dans ce contexte, nous considérons $u^k = 8 - (k \mod 8)$ comme le k-ième terme d'une fonction de signification $(u^k)^{k\in\mathbb{N}}$.
Intuitivement, dans chaque groupe de huit bits (*i.e.*, pour chaque pixel) le premier bit a une importance égale à 8, alors que le dernier bit a une importance égale à 1. Cet exemple est conforme à l'idée que changer le premier bit induit un changement plus fort sur l'image que de changer le dernier bit.

Pour un contenu hôte donné x, les MSC sont alors définis, dans l'écriture binaire de x, par les rangs de x qui décrivent la partie pertinente de l'image, alors que les LSC traduisent ses parties les moins importantes. Ces deux notions sont illustrées sur la figure 6.1, et définies ci-après.

Définition 65 : Importance des coefficients

Soit $(u^k)^{k\in\mathbb{N}}$ une fonction de signification, m et M deux réels tels que $m < M$.

- Les *coefficients les plus significatifs (most significant coefficients, ou MSCs)* de x sont définis par le vecteur fini :

$$u_M = \left(k \;\middle|\; k \in \mathbb{N} \text{ et } u^k \geqslant M \text{ et } k \leq |\, x \,|\right).$$

- Les *coefficients les moins significatifs (least significant coefficients, ou LSCs)* de x sont définis par le vecteur fini :

$$u_m = \left(k \;\middle|\; k \in \mathbb{N} \text{ et } u^k \leq m \text{ et } k \leq |\, x \,|\right).$$

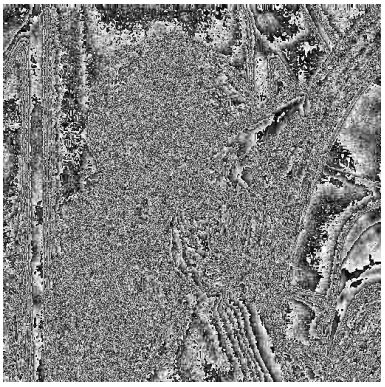
- Les *coefficients passifs (passive coefficients)* de x sont le vecteur fini :

$$u_p = \left(k \;\middle|\; k \in \mathbb{N} \text{ et } u^k \in]m; M[\text{ et } k \leq |\, x \,|\right).$$

Quand les MSC et les LSC représentent une suite de bits, ils sont aussi appelés "Bits les plus significatifs" (MSB) et "Bits les moins significatifs" (LSB). Cependant, les coefficients des MSC et LSC peuvent aussi être, par exemple, ceux de la DCT (discrete cosine transform) utilisée dans le format d'image JPEG.

(a) Lena originale

(b) MSC de Lena



(c) LSC $\times 17$ de Lena

FIGURE 6.1 – Coefficients les plus et les moins significatifs de Lena.

6.3.2/ PRÉSENTATION DU PROCÉDÉ $\mathcal{CIW}_1$

Un premier algorithme de tatouage numérique avait été proposé par notre équipe avant notre début de thèse. Cet algorithme, présenté dans [11], était basé sur les itérations chaotiques, mais ne permettait d'insérer qu'un bit par image. Nous le rappelons dans ce qui suit.

Introduisons les notations suivantes :

- Soit $N \in \mathbb{N}$ la taille du média de couverture.
- Soit $K \in [0;1]$ une clé d'embarquement.
- Soit $X \in \mathbb{B}^{\mathsf{N}}$ les N *Least Significant Coefficients (LSC)* d'un média de couverture C donné.
- Soit X_0 l'état initial de X.
- Soit $(S^n)_{n\in\mathbb{N}} \in [\![1,\mathsf{N}]\!]^{\mathbb{N}}$ une stratégie chaotique, qui dépend, d'une part, du message à dissimuler $M \in [0;1]$, d'autre part, de la clé K.
- Soit $f_0 : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$ la négation vectorielle booléenne.

Alors le média tatoué est C, pour lequel les LSC ont été remplacés par $Y_K = X^N$, où :

$$\begin{cases} X^0 = X \\ \forall n < N, X^{n+1} = G_{f_0}(X^n). \end{cases}$$

(a) Lena originale.

(b) Lena tatouée.

FIGURE 6.2 – Dissimulation d'informations par itérations chaotiques

6.3.3/ DIFFÉRENTS TYPES DE STRATÉGIES CHAOTIQUES

La méthode présentée précédemment dépend d'une stratégie chaotique fournie en entrée de l'algorithme. Dans [43], nous avons proposé deux façons de générer $(S^n)_{n\in\mathbb{N}}$. La première façon consiste à choisir une stratégie chaotique indépendante du média de couverture C, ce que nous appelons *"itérations chaotiques avec stratégie indépendante" ou* $\mathcal{CIIS}$.

La seconde manière, quant à elle, consiste à choisir une stratégie chaotique dépendante du média de couverture, ce que l'on appelle *"itérations chaotiques avec stratégie dépendante" ou* $\mathcal{CIDS}$.

6.3.3.1/ STRATÉGIE CHAOTIQUE DE TYPE $\mathcal{CIIS}$

Dans cette section nous considérons un premier type de stratégies chaotiques. Ce type est défini par l'indépendance de la stratégie vis à vis du média d'embarquement. Il est appelé $\mathcal{CIIS}$ (Chaotic Iterations with Independant Strategy) ou itérations chaotiques avec stratégie indépendante.

Nous allons tout d'abord donner la définition de la fonction chaotique linéaire par morceaux (*Piecewise Linear Chaotic Map* ou *PLCM*. Se référer à [74] pour plus de détails sur cette fonction.

Définition 66 : Fonction chaotique linéaire par morceaux (PLCM)

La *fonction chaotique linéaire par morceaux (PLCM)* est définie par

$$F(x,p) = \begin{cases} x/p & \text{si} \quad x \in [0;p], \\ (x-p)/(\frac{1}{2}-p) & \text{si} \quad x \in \left[p;\frac{1}{2}\right], \\ F(1-x,p) & \text{sinon,} \end{cases}$$

où $p \in \left]0;\frac{1}{2}\right[$ est un « *paramètre de contrôle* ». Alors le terme général de la stratégie chaotique $(S^n)_n$ dans un contexte de type $\mathcal{CIIS}$ est défini par l'expression suivante :

$$S^n = \lfloor \mathsf{N} \times K^n \rfloor + 1,$$

où :

$$\begin{cases} p \in \left[0;\frac{1}{2}\right] \\ K^0 = M \otimes K \\ K^{n+1} = F(K^n,p), \forall n \leq N_0 \end{cases}$$

dans laquelle, rappelons-le, $\otimes$ désigne le « *ou exclusif bit à bit* » (XOR) entre deux parties flottantes de deux nombres (*i.e.*, entre leur représentation en chiffres binaires).

6.3.3.2/ STRATÉGIE CHAOTIQUE DE TYPE $\mathcal{CIDS}$

Dans cette section nous considérons un second type de stratégies chaotiques. Ce type est défini, contrairement au type précédent, par la dépendance de la stratégie chaotique vis à vis du média d'embarquement. Il est appelé $\mathcal{CIDS}$ (Chaotic Iterations with Dependant Strategy) ou itérations chaotiques avec stratégie dépendante.

Avec les mêmes notations que précédemment, on définit la stratégie chaotique dans la configuration $\mathcal{CIDS}$ de la manière suivante : $\forall k \leqslant N$,

- si $k \leqslant \mathsf{N}$ et $X^k = 1$, alors $S^k = k$,
- sinon $S^k = 1$.

Dans cette situation, si $N \geqslant \mathsf{N}$, alors seuls deux contenus tatoués sont possibles, respectivement : $Y_K = (0,0,\cdots,0)$ et $Y_K = (1,0,\cdots,0)$. Ces deux stratégies chaotiques, que nous avons définies, nous ont conduits à nos premières contributions, rappelées dans le chapitre suivant.

III

CONTRIBUTIONS À LA SCIENCE DE L'INFORMATION DISSIMULÉE

CHAPITRE 7

Comparaison entre le $\mathcal{CIW}_1$ et l'étalement de spectre $\mathcal{NW}$: la stégo-sécurité

> Le savant n'est pas l'homme qui fournit les vraies réponses, c'est celui qui pose les vraies questions.
>
> CLAUDE LEVI-STRAUSS, ANTHROPOLOGUE ET ETHNOLOGUE (1908-2009)

7.1/ INTRODUCTION

Nous avons rappelé précédemment qu'en plus d'être stégo-sécurisée, la technique de tatouage numérique naturel ($\mathcal{NW}$) est topologiquement-sûre (*cf.* théorème 5). Cette technique possède, en outre, des propriétés supplémentaires (*c. f.* chapitre 5 théorème 5), issues de son niveau d'imprévisibilité. Ces propriétés sont la forte transitivité, le mélange topologique qu'elle engendre et sa constante de sensibilité qui est supérieure ou égale à $\frac{N_b}{2}$, où N_b est un majorant de la taille des filigranes (messages) à embarquer dans le contenu hôte. Toutefois, les processus d'étalement de spectre de type $\mathcal{NW}$ ne sont pas expansifs, ce qui est problématique dans le cadre de l'attaque du message constant (*Constant Message Attack* ou *CMA*) ainsi que dans la configuration de l'attaque du message connu (*Known Message Attack* ou *KMA*) [18].

Lors de précédents travaux, rappelés au chapitre précédent, l'équipe a montré que le $\mathcal{CIW}_1$ était plus sûr que $\mathcal{NW}$, au sens topologique. En effet, contrairement à ce dernier, $\mathcal{CIW}_1$ possède toutes les propriétés de désordre topologique, dont notamment l'expansivité (*cf.* chapitre 3 section 4.8.1). Cette technique a, de ce fait, la capacité de résister aux attaques *KMA* et *CMA*. Cependant, le processus $\mathcal{NW}$ était connu pour être stégo-sûr, aussi fallait-il s'assurer qu'il en était de même de $\mathcal{CIW}_1$. Cette vérification par la preuve a été l'objet de mon premier travail. Ce travail a été publié dans [43]. Nous en avons aussi profité pour montrer le mélange topologique du $\mathcal{CIW}_1$ et en mesurer son expansivité.

Cette dernière propriété de nature topologique restait à établir pour cet algorithme de dissimulation d'informations.

7.2/ LA STÉGO-SÉCURITÉ DU SCHÉMA $\mathcal{CIW}_1$

7.2.1/ PREUVE DE STÉGO-SÉCURITÉ

Dans [43], nous avons montré que :

Proposition 15 : Stégo-sécurisé du $\mathcal{CIW}_1$

Le schéma $\mathcal{CIW}_1$ est stégo-sûr pour toutes les stratégies chaotiques de type $\mathcal{CIIS}$.

Démonstration. Supposons que $X \sim \mathbf{U}\left(\mathbb{B}^N\right)$ dans une configuration CIIS. Nous allons montrer par récurrence que $\forall n \in \mathbb{N}, X^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$.

L'initialisation est immédiate, puisque $X^0 = X \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Supposons maintenant la propriété $X^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$ vraie au rang n. Soit $e \in \mathbb{B}^N$ et $\mathbf{B}_k = (0, \cdots, 0, 1, 0, \cdots, 0) \in \mathbb{B}^N$ (le chiffre 1 est en position k). Alors $P\left(X^{n+1} = e\right) = \sum_{k=1}^{N} P\left(X^n = e + \mathbf{B}_k, S^n = k\right)$. Ces deux événements sont indépendants dans une configuration $\mathcal{CIIS}$, donc : $P\left(X^{n+1} = e\right) = \sum_{k=1}^{N} P\left(X^n = e + \mathbf{B}_k\right) \times P\left(S^n = k\right)$. Du fait de l'hypothèse de récurrence : $P\left(X^{n+1} = e\right) = \frac{1}{2^N} \sum_{k=1}^{N} P\left(S^n = k\right)$. L'ensemble des événements $\{S^n = k\}$ pour $k \in [\![1; N]\!]$ est une partition de l'univers des possibles, donc $\sum_{k=1}^{N} P\left(S^n = k\right) = 1$.

Finalement, $P\left(X^{n+1} = e\right) = \frac{1}{2^N}$, ce qui conduit à $X^{n+1} \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Ce résultat est vrai $\forall n \in \mathbb{N}$, nous avons donc prouvé que,

$$\forall K \in [0; 1], Y_K = X^{N_0} \sim \mathbf{U}\left(\mathbb{B}^N\right) \text{ quand } X \sim \mathbf{U}\left(\mathbb{B}^N\right)$$

□

Nous allons maintenant prouver que :

Proposition 16 : Cas d'absence de stégo-sécurité

Le schéma $\mathcal{CIW}_1$ n'est pas stégo-sûr pour les stratégies chaotiques de type $\mathcal{CIDS}$.

Démonstration. Du fait de la définition des stratégies chaotiques de type $\mathcal{CIDS}$ (*cf.* chapitre 6 section 6.3.3), nous avons $P(Y_K = (1, 1, \cdots, 1)) = 0$. Donc il n'y a pas uniforme répartition des stégo-contenus Y_K. □

7.2.2/ DISCUSSION

7.2.2.1/ DISTRIBUTION DES LSC

Nous avons supposé que $x^0 \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$ pour démontrer la stégo-sécurité du processus de dissimulation d'informations. Cette hypothèse est assez restrictive, mais elle peut être obtenue, au moins partiellement, de deux manières différentes. Soit en utilisant un canal qui semble aléatoire (par exemple, lorsque l'on applique un test du χ^2) et qui peut être trouvé dans le média. Soit encore en procédant comme suit : avant d'embarquer le message secret, tous les LSC orignaux sont remplacés par des LSC générés aléatoirement, en espérant qu'une telle modification soit considérée par l'attaquant comme du bruit.

Nous pouvons remarquer que, considérant l'anonymisation des données pour le respect de la vie privée sur Internet, nous sommes dans le cadre de l'attaque par filigrane seul (*WOA*). Dans ce contexte, l'attaquant a uniquement accès aux contenus stéganographiques, sans connaissance du média original avant l'embarquement du message dans le canal aléatoire (LSC).

7.2.2.2/ DISTRIBUTION DES STRATÉGIES CHAOTIQUES S

Nous avons aussi supposé que la stratégie chaotique suivait une distribution uniforme. Cette hypothèse n'est pas aussi restrictive, dans la mesure où n'importe quel générateur de nombres pseudo-aléatoires (*PRNG*), cryptographiquement sûr, satisfait cette propriété. Avec de tels PRNG, il est impossible, dans un temps polynomial, de faire la distinction entre des nombres réellement aléatoires et des nombres fournis par ce type de générateur. Par exemple, le *Blum Blum Shub (BBS)* [46] ou *ISAAC* [45] conviennent.

7.3/ ÉVALUATIONS DE PROPRIÉTÉS TOPOLOGIQUES

Dans cette section, nous rappelons quelques-uns des résultats établis dans [41].

7.3.1/ SENSIBILITÉ AUX CONDITIONS INITIALES DE $\mathcal{CIW}_1$

Nous savons que cette propriété est une conséquence de la régularité et de la transitivité (*cf.* théorème de Banks au chapitre 3 théorème 1). Cependant, cela ne nous dit pas quelle est la constante de sensibilité. C'est pourquoi la sensibilité a été redémontré complètement.

Proposition 17 : Sensibilité de $\mathcal{CIW}_1$

G_{f_0} est sensible aux conditions initiales sur $(\mathcal{X}_1, d_1)$ et sa constante de sensibilité est supérieure ou égale à $\mathsf{N} - 1$.

Démonstration. Issue de [41].
Pour plus de lisibilité, définissons $\mathcal{E} : \mathcal{X}_1 \to \mathbb{B}^{\mathsf{N}}$ par $\mathcal{E}(S, E) = E$.

Dans cette démonstration, nous noterons indifféremment $\mathbb{S}$ ou $\mathcal{S}$ l'ensemble des stratégies chaotiques introduit dans la définition 47 du chapitre 6.

Soit $\check{X} = (\check{S}, \check{E}) \in \mathcal{X}_1$. On recherche $\widetilde{X} = (\widetilde{S}, \widetilde{E}) \in \mathcal{X}_1$ tel que $d\left((\check{X}, \widetilde{X})\right) \leqslant \delta$ et $\exists n_0 \in \mathbb{N}$, $d_1\left(G_{f_0}^{(n_0)}(\check{X}); G_{(f_0)}^{n_0}(\widetilde{X})\right) \geqslant \mathsf{N} - 1$. Posons $k_0 = \lfloor -\log_{10}(\delta) \rfloor + 1$, tel que si une stratégie S coïncide avec $\check{S}$ sur les k_0 premiers termes $\left(S \in \left\{S \in \mathcal{S} \,/\, \forall k \leqslant k_0, S^k = \check{S}^k\right\}\right)$, alors $(S, \check{E})$ et $(\check{S}, \check{E})$ sont proches à δ près $(d_1\left((S, \check{E}), (\check{S}, \check{E})\right) \leqslant \delta)$.

Soit $\mathcal{J} = \left\{i \in [\![1, \mathsf{N}]\!] \,/\, \mathcal{E}\left(G_{f_0}^{(k_0)}(\check{S}, \check{E})\right)_i \neq \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{S}, \check{E})\right)_i\right\}$ l'ensemble des indices des cellules différentes entre les états de $G_{f_0}^{(k_0)}(\check{S}, \check{E})$ et de $G_{f_0}^{(k_0+\mathsf{N})}(\check{S}, \check{E})$, et soit p le nombre de différences : $p = card\,(\mathcal{J})$. Si $p = \mathsf{N}$, alors le point $(\widetilde{S}, \widetilde{E}) \in \mathcal{X}_1$ défini par :

1. $\widetilde{E} = \check{E}$, pour avoir les mêmes cellules que $\check{X}$, et donc être à une distance inférieure à 1 de $\check{X}$,
2. $\forall k \leqslant k_0, \widetilde{S}^k = \check{S}^k$, pour être proche de $\check{X}$ à δ près,
3. $\forall k \in [\![1, \mathsf{N}]\!], \widetilde{S}^{k_0+k} = k$, pour avoir ensuite un état dont les cellules sont toutes différentes de $\check{X}$,
4. $\forall k > k_0 + \mathsf{N}, \widetilde{S}^k = 1$, pour finir de définir pleinement $(\widetilde{S}, \widetilde{E})$,

satisfait $d_1\left((\widetilde{S}, \widetilde{E}); (\check{S}, \check{E})\right) < \delta$ (c'est-à-dire que $\widetilde{X}$ est proche de $\check{X}$), et $\forall i \in [\![1, \mathsf{N}]\!]$, $\mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{S}; \widetilde{E})\right)_i \neq \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{S}; \check{E})\right)_i$, (c'est-à-dire que l'itérée $k_0 + \mathsf{N}$ de $\widetilde{X}$ est éloignée d'une distance supérieure à N de l'itérée $k_0 + \mathsf{N}$ de $\check{X}$), donc le résultat est obtenu.

Sinon, soit $j_1 < j_2 < ... < j_p$ les éléments de $\mathcal{J}$ et $j_0 \notin \mathcal{J}$. Alors $\widetilde{X} = (\widetilde{E}, \widetilde{S}) \in \mathcal{X}_1$ défini, pour les mêmes raisons que précédemment, par :

1. $\widetilde{E} = \check{E}$,
2. $\forall k \leqslant k_0, \widetilde{S}^k = \check{S}^k$,
3. $\forall k \in [\![1, p]\!], \widetilde{S}^{k_0+k} = j_k$,
4. $\forall k \in \mathbb{N}^*, \widetilde{S}^{k_0+p+k} = j_0$.

est tel que $d_1(\check{X}, \widetilde{X}) < \delta$. De plus, $\forall i \in [\![1, p]\!], \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{X})\right)_{j_i} \neq \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{X})\right)_{j_i}$, car :

- $\forall i \in [\![1, \mathsf{N}]\!], \mathcal{E}\left(G_{f_0}^{(k_0)}(\check{X})\right)_i = \mathcal{E}\left(G_{f_0}^{(k_0)}(\widetilde{X})\right)_i$, du fait de la définition de k_0.
- $\forall i \in [\![1, p]\!], j_i \in \mathcal{J} \Rightarrow \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{X})\right)_{j_i} = \mathcal{E}\left(G_{f_0}^{(k_0)}(\check{X})\right)_{j_i}$, d'après la définition de $\mathcal{J}$.
- $\forall i \in [\![1, p]\!], j_i$ apparaît exactement une fois dans $\widetilde{S}^{k_0}, \widetilde{S}^{k_0+1}, ..., \widetilde{S}^{k_0+\mathsf{N}}$, donc

$$\mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{X})\right)_{j_i} \neq \mathcal{E}\left(G_{f_0}^{(k_0)}(\widetilde{X})\right)_{j_i}.$$

Enfin, $\forall i \in [\![1, \mathsf{N}]\!] \setminus \{j_0, j_1, ..., j_p\}, \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{X})\right)_i \neq \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{X})\right)_i$, car :

- $\forall i \in [\![1, \mathsf{N}]\!], \mathcal{E}\left(G_{f_0}^{(k_0)}(\check{X})\right)_i = \mathcal{E}\left(G_{f_0}^{(k_0)}(\widetilde{X})\right)_i$,
- $i \notin \mathcal{J} \Rightarrow \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{X})\right)_i \neq \mathcal{E}\left(G_{f_0}^{(k_0)}(\check{X})\right)_i$,
- $i \notin \{\widetilde{S}^{k_0}, \widetilde{S}^{k_0+1}, ..., \widetilde{S}^{k_0+\mathsf{N}}\} \Rightarrow \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{X})\right)_i = \mathcal{E}\left(G_{f_0}^{(k_0)}(\widetilde{X})\right)_i$.

Donc, dans ce cas, $\forall i \in [\![1, \mathsf{N}]\!] \setminus \{j_0\}, \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\widetilde{S}; \widetilde{E})\right)_i \neq \mathcal{E}\left(G_{f_0}^{(k_0+\mathsf{N})}(\check{S}; \check{E})\right)_i$ et le résultat est obtenu. □

7.3.2/ EXPANSIVITÉ DE $\mathcal{CIW}_1$

Proposition 18 : Expansivité de $\mathcal{CIW}_1$

G_{f_0} est expansif sur $\mathcal{X}$, sa constante d'expansivité étant égale à 1.

Démonstration. Issue de [41].
Si $(S,E) \neq (\check{S};\check{E})$, alors soit $E \neq \check{E}$, donc au moins une cellule n'est pas dans le même état dans E et $\check{E}$. En conséquence de quoi, la distance entre (S,E) et $(\check{S};\check{E})$ est supérieure ou égale à 1. Ou $E = \check{E}$. Alors les stratégies chaotiques S et $\check{S}$ ne sont pas égales.

Soit n_0 le premier indice à partir duquel les termes des suites S et $\check{S}$ diffèrent. Alors $\forall k < n_0, \tilde{G}^k_{f_0}(S,E) = \tilde{G}^k_{f_0}(\check{S},\check{E})$, et $\tilde{G}^{n_0}_{f_0}(S,E) \neq \tilde{G}^{n_0}_{f_0}(\check{S},\check{E})$. Comme $E = \check{E}$, la cellule qui a changé dans E à l'itérée numéro n_0 n'est pas la même que celle qui a changé dans $\check{E}$. Et donc la distance entre $\tilde{G}^{n_0}_{f_0}(S,E)$ et $\tilde{G}^{n_0}_{f_0}(\check{S},\check{E})$ est supérieure ou égale à 2. □

7.3.3/ CAS DU MÉLANGE TOPOLOGIQUE

Proposition 19 : Mélange topologique du $\mathcal{CIW}_1$

$\tilde{G}_{f_0}$ est topologiquement mélangeant sur $(\mathcal{X}_1, d_1)$.

Ce résultat est une conséquence immédiate du lemme suivant :

Lemme 1 :

Pour toute boule ouverte B de $\mathcal{X}_1$, il existe un indice n tel que $G^{(n)}_{f_0}(B) = \mathcal{X}_1$.

Démonstration. En effet, soit $B = \mathcal{B}((S,E),\varepsilon)$ une telle boule ouverte, dont on peut supposer le rayon inférieur à 1 (qui peut le plus, peut le moins). Les éléments de B ont donc tous le même état E, et sont tels qu'il existe un indice $k\,(= -\log_{10}(\varepsilon))$ vérifiant que :

- Toutes les stratégies chaotiques de B ont les k mêmes premiers termes.
- Au-delà du rang k, tous les termes sont possibles.

Donc, au bout de k itérées, l'état du système est maintenant $G^{(k)}_{f_0}(S,E)_2$, et toutes les stratégies chaotiques sont possibles (tout point de la forme $(G^{(k)}_{f_0}(S,E)_2, \hat{\mathrm{S}})$, avec $\hat{\mathrm{S}} \in \mathcal{S}$ quelconque, est atteignable à partir de B.

Soit maintenant un point quelconque (S',E') de $\mathcal{X}$. On va prouver qu'on peut l'atteindre à partir d'un élément de B. En effet, soit s la liste des cellules différentes entre $G^{(k)}_{f_0}(S,E)_2$ et E', et $|s|$ sa taille. Le point de B dont l'état est E et dont la stratégie chaotique :

- coïncide avec S sur les k premiers termes,
- se poursuit avec les éléments de s,
- se termine avec S'

est tel que l'image par $G^{(k+|s|)}_{f_0}$ est exactement (S',E'). □

7.4/ ÉTALEMENT DE SPECTRE ET ITÉRATIONS CHAOTIQUES

Contrairement aux techniques d'étalement de spectre, les itérations chaotiques sont donc expansives, et cette propriété est aussi héritée par le $\mathcal{CIW}_1$.

Donnons maintenant quelques conséquences de l'évaluation de ces propriétés qualitatives et quantitatives. Tout d'abord, la propriété d'expansivité renforce considérablement la sensibilité du processus de dissimulation d'informations. Une telle sensibilité est primordiale afin de réduire drastiquement les avantages que Eve peut tirer de l'analyse des échanges dans une attaque de type "message connu" *(KMA)* ou de type "original connu" *(KOA)* [25].

Par exemple, dans une situation d'expansivité, il est impossible d'avoir une estimation du filigrane en déplaçant le message (ou le média de couverture) un peu à la manière d'un curseur : le curseur sera trop sensible et les changements seront trop importants pour être utiles. *A contrario*, une très grande constante d'expansivité ε est inadaptée : les médias de couverture seront fortement modifiés tandis que le filigrane restera indétectable. En effet, considérons deux exemplaires du même média de couverture tatoué deux fois avec deux filigranes différents. Ainsi $d(X,Y) < 1$ pour la distance définie précédemment. Toutefois, en raison de l'expansivité, $\exists n \in \mathbb{N}, d\left(G^n(X);G^n(Y)\right) \geqslant \varepsilon$. Ainsi, $d_\infty\left(G^n(X)_1;G^n(Y)_1\right) \geqslant \varepsilon - 1$, donc soit $d_\infty\left(X_1;G^n(X)_1\right) \geqslant \dfrac{\varepsilon-1}{2}$, soit $d_\infty\left(Y_1;G^n(Y)_1\right) \geqslant \dfrac{\varepsilon-1}{2}$. Si ε est grand, alors au moins un des deux médias tatoués sera très différent de son média de couverture d'origine. Par conséquent, nous pouvons conclure de la précédente étude que les itérations chaotiques ont un niveau d'expansivité approprié qui garantit, dans une certaine mesure, sa sécurité.

Pour résumer, les processus de dissimulation d'informations basés sur des itérations chaotiques sont topologiquement sûrs, et possèdent la propriété supplémentaire de forte transitivité, propriété dont fait preuve aussi l'étalement de spectre. Toutefois, les processus de dissimulation d'informations basés sur les itérations chaotiques ont une plus grande constante de sensibilité que l'étalement de spectre, et, contrairement à ce dernier, ils sont expansifs (avec une constante de sensibilité égale à 1). En outre, les itérations chaotiques sont topologiquement mélangeantes, ce qui prouve qu'elles semblent plus appropriées que l'étalement de spectre pour résister aux attaques dans les configurations de type KOA, KMA et Constant-Message Attaque (CMA).

7.5/ ÉVALUATION DE LA DISTORSION

Nous avons finalement comparé l'évolution du PSNR lorsque l'on applique le $\mathcal{CIW}_1$, sur près de 1000 images en niveaux de gris, de taille 512x512, issues de la base de données BOSS [77, 20]. Les tests ont été réalisés, d'une part, avec une clé d'embarquement constante, d'autre part, avec une clé d'embarquement générée aléatoirement. Nous avons obtenu dans [15] les résultats décrits dans le tableau 7.5 page suivante,

Stratégie chaotique	Moyenne des PSNR	Écart-type des PSNR
Constante	21.6653	0.03604
Générée aléatoirement	13.3909	10.5690

TABLE 7.1 – Évaluation du PSNR du processus de dissimulation d'informations $\mathcal{CIW}_1$ (tests réalisés sur 1000 images issues de la base de données BOSS [77]).

CHAPITRE 8

Étude du processus amélioré $\mathcal{CIS}_2$

L'art de la science est de prévoir et non, comme on l'a dit souvent, de comprendre.

PIERRE LECOMTE DU NOUY, ÉCRIVAIN, MATHÉMATICIEN ET BIOPHYSICIEN (1883-1947)

Le chapitre précédent présentait nos premières contributions, relatives à l'établissement de nouvelles propriétés de nature topologique pour l'algorithme $\mathcal{CIW}_1$. Ces contributions ont été réalisées en collaboration avec C. Guyeux. Dans ce chapitre, nous étudions un deuxième algorithme de dissimulation d'informations, que nous avons nommé $\mathcal{CIS}_2$ pour « Chaotic Iterations based Steganographier ». Cet algorithme permet d'insérer plus d'un bit. Il généralise le schéma $\mathcal{CIW}_1$ présenté dans [11]. Le schéma $\mathcal{CIS}_2$ a été publié dans [37] et son implémentation pratique, $\mathcal{CIS}_5$, a fait l'objet d'un dépôt logiciel [61] auprès de l'Agence pour la Protection des Programmes [65].

8.1/ PRÉSENTATION DU PROCESSUS $\mathcal{CIS}_2$

Notre processus de dissimulation d'informations est fondé sur des itérations chaotiques « améliorées » avec substitution et mélange du message. Cette substitution et ce mélange nous poussent à introduire d'abord quelques notations.

Notation 5 : Pour le processus $\mathcal{CIS}_2$

Nous désignerons par :

- $x^0 \in \mathbb{B}^{\mathsf{N}}$ le vecteur des N coefficients les moins significatifs (LSC) d'un média de couverture C donné.
- $m^0 \in \mathbb{B}^{\mathsf{P}}$ la marque à embarquer dans x^0.
- $S_p \in \mathbb{S}_N$ une stratégie chaotique appelée **stratégie de placement**.
- $S_c \in \mathbb{S}_P$ une stratégie chaotique appelée **stratégie de choix**.
- $S_m \in \mathbb{S}_P$ une stratégie chaotique appelée **stratégie de mélange**.

Définissons maintenant le nouveau processus de dissimulation d'informations.

Définition 67 : $\mathcal{CIS}_2$

Avec les notations précédentes, le processus $\mathcal{CIS}_2$ est le résultat des itérations décrites ci-dessous. $\forall (n,i,j) \in \mathbb{N}^* \times [\![0;\mathsf{N}-1]\!] \times [\![0;\mathsf{P}-1]\!]$:

$$\begin{cases} x_i^n = \begin{cases} x_i^{n-1} & \text{si } (S_p)^n \neq i \\ m_{(S_c)^n} & \text{si } (S_p)^n = i. \end{cases} \\ \\ m_j^n = \begin{cases} m_j^{n-1} & \text{si } (S_m)^n \neq j \\ \\ \overline{m_j^{n-1}} & \text{si } (S_m)^n = j. \end{cases} \end{cases}$$

où $\overline{m_j^{n-1}}$ correspond à la négation booléenne de m_j^{n-1}.
Le contenu stéganographique est C, pour lequel les LSC ont été remplacés par le vecteur booléen $y = x^{\mathsf{P}} \in \mathbb{B}^{\mathsf{N}}$.

8.2/ ÉTUDE DE STÉGO-SÉCURITÉ DU $\mathcal{CIS}_2$

Dans [37], nous avons montré que :

Proposition 20 : Stégo-sécurité de $\mathcal{CIS}_2$

Le processus $\mathcal{CIS}_2$ est stégo-sûr.

Stégo-sécurité de $\mathcal{CIS}_2$. Dans le cadre du contexte de $\mathcal{CIS}_2$, supposons que $x^0 \sim \mathbf{U}\left(\mathbb{B}^N\right)$ et $m^0 \sim \mathbf{U}\left(\mathbb{B}^P\right)$. Nous allons alors démontrer par récurrence sur l'entier n que $\forall n \in \mathbb{N}, x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$.

L'initialisation est évidente au regard des hypothèses d'uniforme répartition de x^0 et de m^0.

Supposons à présent que pour un certain rang n : $x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Pour un entier $k \in \mathbb{B}^N$ donné, on note $\tilde{k}_i \in \mathbb{B}^N$ le vecteur défini par : $\forall i \in [\![0;\mathsf{N}-1]\!]$, si $k = (k_0, k_1, \ldots, k_i, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1})$, alors $\tilde{k}_i = \left(k_0, k_1, \ldots, \overline{k_i}, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1}\right)$.

Soit $E_{i,j}$ l'événement probabiliste suivant :

$$\forall (i,j) \in [\![0;\mathsf{N}-1]\!] \times [\![0;\mathsf{P}-1]\!], E_{i,j} = S_p^{n+1} = i \wedge S_c^{n+1} = j \wedge m_j^{n+1} = k_i \wedge \left(x^n = k \vee x^n = \tilde{k}_i\right),$$

et $p = P\left(x^{n+1} = k\right)$. Donc :

$$p = P\left(\bigvee_{i\in[\![0;\mathsf{N}-1]\!], j\in[\![0;\mathsf{P}-1]\!]} E_{i,j}\right).$$

Nous pouvons introduire à présent la notation suivante : $P_1(i) = P\left(S_p^{n+1} = i\right)$, $P_2(j) = P\left(S_c^{n+1} = j\right)$, $P_3(i,j) = P\left(m_j^{n+1} = k_i\right)$, et $P_4(i) = P\left(x^n = k \vee x^n = \tilde{k}_i\right)$.
Ces quatre événements sont indépendants dans le contexte $\mathcal{CIS}_2$, donc

$$p = \sum_{i\in[\![0;\mathsf{N}-1]\!], j\in[\![0;\mathsf{P}-1]\!]} P_1(i)P_2(i)P_3(i,j)P_4(i).$$

D'après la proposition 15, $P\left(m_j^{n+1} = k_i\right) = \frac{1}{2}$. Comme les deux événements sont incompatibles :

$$P\left(x^n = k \vee x^n = \tilde{k}_i\right) = P\left(x^n = k\right) + P\left(x^n = \tilde{k}_i\right).$$

Alors, en utilisant l'hypothèse de récurrence : $P\left(x^n = k\right) = \frac{1}{2^N}$, et $P\left(x^n = \tilde{k}_i\right) = \frac{1}{2^N}$.
Soit S la somme définie par :

$$S = \sum_{i\in[\![0;\mathsf{N}-1]\!], j\in[\![0;\mathsf{P}-1]\!]} P_1(i)P_2(j).$$

Alors $p = 2 \times \frac{1}{2} \times \frac{1}{2^N} \times S = \frac{1}{2^N} \times S$.
Nous pouvons à présent évaluer S :

$$\begin{array}{rcl} S & = & \sum_{i\in[\![0;\mathsf{N}-1]\!], j\in[\![0;\mathsf{P}-1]\!]} P_1(i)P_2(j) \\ & = & \sum_{i\in[\![0;\mathsf{N}-1]\!]} P_1(i) \times \sum_{j\in[\![0;\mathsf{P}-1]\!]} P_2(j). \end{array}$$

L'ensemble des événements $\left\{S_p^{n+1} = i\right\}$ pour $i \in [\![0;N-1]\!]$ et l'ensemble des événements $\left\{S_c^{n+1} = j\right\}$ pour $j \in [\![0;P-1]\!]$ forment une partition de l'univers des possibles, d'où $S = 1$.
Finalement, $P\left(x^{n+1} = k\right) = \frac{1}{2^N}$, ce qui nous a conduit à $x^{n+1} \sim \mathbf{U}\left(\mathbb{B}^N\right)$.
Ce résultat étant vrai $\forall n \in \mathbb{N}$, nous avons prouvé que le contenu stéganographié y suit une loi de répartition uniforme dans l'ensemble des contenus stéganographiques possibles, d'où $y \sim \mathbf{U}\left(\mathbb{B}^N\right)$ quand $x \sim \mathbf{U}\left(\mathbb{B}^N\right)$. □

Remarque 8 : Distribution du message m

Afin de démontrer la stégo-sécurité du processus de dissimulation d'informations $\mathcal{CIS}_2$, nous avons supposé que $m \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{P}}\right)$. Cette hypothèse n'est pas réellement restrictive. En effet, crypter le message avant son embarquement dans les LSC du média de couverture est suffisant pour atteindre ce but.

8.3/ MODÈLE TOPOLOGIQUE

Dans cette section, nous démontrons que le $\mathcal{CIS}_2$ peut être modélisé sous la forme d'un système dynamique discret sur un espace topologique, ou plus précisément sur un espace métrique. Nous démontrerons dans la section suivante que ce $\mathcal{CIS}_2$ est un cas de chaos topologique, tel que le définit Devaney. Ces travaux ont été publiés dans [37].

8.3.1/ FONCTION D'ITÉRATIONS ET ESPACE DES PHASES

Suivant le canevas de la preuve du $\mathcal{CIW}_1$, nous commençons par introduire quelques définitions et notations utiles par la suite.

Définition 68 : Fonction F

Soit

$$\begin{array}{rl} F: & [\![0;\mathsf{N}-1]\!]\times\mathbb{B}^{\mathsf{N}}\times[\![0;\mathsf{P}-1]\!]\times\mathbb{B}^{\mathsf{P}}\longrightarrow\mathbb{B}^{\mathsf{N}} \\ & (k,x,\lambda,m)\longmapsto\left(\delta(k,j).x_j+\overline{\delta(k,j)}.m_\lambda\right)_{j\in[\![0;\mathsf{N}-1]\!]} \end{array}$$

où + et . sont, comme d'habitude, les opérations de somme et produits booléens.

Considérons l'espace des phases $\mathcal{X}_2$ défini de la manière suivante.

Définition 69 : Espace des phases $\mathcal{X}_2$

On pose :

$$\mathcal{X}_2=\mathbb{S}_{\mathsf{N}}\times\mathbb{B}^{\mathsf{N}}\times\mathbb{S}_{\mathsf{P}}\times\mathbb{B}^{\mathsf{P}}\times\mathbb{S}_{\mathsf{P}},$$

où $\mathbb{S}_N$ et $\mathbb{S}_P$ sont les ensembles introduits dans la section 8.1.

On définit alors la fonction $\mathcal{G}_{f_0}$ comme suit :

Définition 70 : Fonction $\mathcal{G}_{f_0}$

$\mathcal{G}_{f_0}:\mathcal{X}_2\longrightarrow\mathcal{X}_2$ est définie par

$$\mathcal{G}_{f_0}\left(S_p,x,S_c,m,S_m\right)=$$

$$\left(\sigma_N(S_p),F(i_N(S_p),x,i_P(S_c),m),\sigma_P(S_c),G_{f_0}(m,S_m),\sigma_P(S_m)\right)$$

On peut alors directement en déduire la formalisation suivante pour le processus $\mathcal{CIS}_2$.

Proposition 21 : Système dynamique discret pour $\mathcal{CIS}_2$

Le processus de dissimulation d'informations $\mathcal{CIS}_2$ peut être décrit par les itérations du système dynamique discret suivant :

$$\begin{cases} X^0\in\mathcal{X}_2 \\ X^{k+1}=\mathcal{G}_{f_0}(X^k). \end{cases}$$

Donnons maintenant un peu plus d'informations sur l'espace des phases.

8.3.2/ ÉTUDE DE L'ESPACE DES PHASES

8.3.2.1/ CARDINALITÉ DE $\mathcal{X}_2$

En comparant $\mathcal{X}_2$ et $\mathcal{X}_1$, nous pouvons en tirer le résultat suivant :

Proposition 22 : Cardinalité de $\mathcal{X}_2$

L'espace des phases $\mathcal{X}_2$ a, au moins, la puissance du continu.

Démonstration. Soit φ la fonction définie de la manière suivante :

$$\begin{array}{rccc} \varphi : & \mathcal{X}_1 & \longrightarrow & \mathcal{X}_2 \\ & (S,x) & \longmapsto & (S,x,0,0,0). \end{array}$$

φ étant clairement injective, on peut en déduire que la puissance de $\mathcal{X}_2$ est supérieure ou égale à la celle de $\mathcal{X}_1$. Et finalement, $\mathcal{X}_2$ a au moins la puissance du continu. □

Remarque 9 :

Ce résultat est indépendant du nombre de cellules du système.

8.3.2.2/ UNE NOUVELLE DISTANCE SUR $\mathcal{X}_2$

Montrons maintenant que l'espace des phases peut être vu comme un espace métrique. On définit pour cela une nouvelle distance sur $\mathcal{X}_2$, de la manière suivante.

Définition 71 : Distance d_2 sur $\mathcal{X}_2$

$\forall X, \check{X} \in \mathcal{X}_2$, si $X = (S_p, x, S_c, m, S_m)$ et $\check{X} = (\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})$, alors :

$$\begin{array}{rcl} d_2(X,\check{X}) & = & d_{\mathbb{B}^N}(x,\check{x}) + d_{\mathbb{B}^P}(m,\check{m}) \\ & + & d_{\mathbb{S}_N}(S_p,\check{S_p}) + d_{\mathbb{S}_P}(S_c,\check{S_c}) + d_{\mathbb{S}_P}(S_m,\check{S_m}), \end{array}$$

où $d_{\mathbb{B}^N}$, $d_{\mathbb{B}^P}$, $d_{\mathbb{S}_N}$, et $d_{\mathbb{S}_P}$ sont les mêmes distances que dans la définition 62.

$\mathcal{CIS}_2$ est donc un système dynamique discret sur un espace métrique. Avant de pouvoir être en mesure d'étudier son chaos, il nous faut vérifier qu'on itère bien une fonction continue (pour la métrique de $\mathcal{X}_2$). Il s'agit là de l'objet de la section suivante.

8.3.3/ CONTINUITÉ DE $\mathcal{CIS}_2$

Afin de démontrer que le processus de dissimulation d'informations $\mathcal{CIS}_2$ est un exemple de chaos topologique au sens de Devaney, la fonction $\mathcal{G}_{f_0}$ doit être continue sur l'espace métrique $(\mathcal{X}_2, d_2)$. Nous allons le démontrer.

Proposition 23 : Continuité de $\mathcal{G}_{f_0}$

$\mathcal{G}_{f_0}$ est une fonction continue sur $(\mathcal{X}_2, d_2)$.

Continuité de $\mathcal{G}_{f_0}$. Nous allons utiliser ici le principe de la continuité séquentielle.

Soit $((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)_{n\in\mathbb{N}}$ une suite de points de l'espace des phases $\mathcal{X}_2$, qui converge vers (S_p, x, S_c, m, S_m). Nous allons prouver que $\left(\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)\right)_{n\in\mathbb{N}}$ converge vers $\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)$. Rappelons que pour tout n, $(S_p)^n$, $(S_c)^n$ et $(S_m)^n$ sont des stratégies chaotiques, nous considérons donc une suite de stratégies chaotiques (*i.e.*, une suite de suites).

Comme $d_2(((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n), (S_p, x, S_c, m, S_m))$ converge vers 0, chaque distance $d_{\mathbb{B}^\mathbb{N}}(x^n, x)$, $d_{\mathbb{B}^\mathbb{P}}(m^n, m)$, $d_{\mathbb{S}_N}((S_p)^n, S_p)$, $d_{\mathbb{S}_P}((S_c)^n, S_c)$, et $d_{\mathbb{S}_P}((S_m)^n, S_m)$ converge vers 0. Mais $d_{\mathbb{B}^\mathbb{N}}(x^n, x)$ et $d_{\mathbb{B}^\mathbb{P}}(m^n, m)$ sont des entiers naturels, alors $\exists n_0 \in \mathbb{N}, \forall n \geqslant n_0, d_{\mathbb{B}^\mathbb{N}}(x^n, x) = 0$ et $\exists n_1 \in \mathbb{N}, \forall n \geqslant n_1, d_{\mathbb{B}^\mathbb{P}}(m^n, m) = 0$.

En d'autres termes, il existe un rang $n_3 = Max(n_0, n_1) \in \mathbb{N}$ à partir duquel les cellules ne changent plus d'état : $\exists n_3 \in \mathbb{N}, n \geqslant n_3 \Longrightarrow (x^n = x) \wedge (m^n = m)$.

De plus, $d_{\mathbb{S}_\mathbb{N}}((S_p)^n, S_p) \longrightarrow 0$, $d_{\mathbb{S}_\mathbb{P}}((S_c)^n, S_c) \longrightarrow 0$, et $d_{\mathbb{S}_\mathbb{P}}((S_m)^n, S_m) \longrightarrow 0$, donc il existe $n_4, n_5, n_6 \in \mathbb{N}$ tels que

- $\forall n \geqslant n_4, d_{\mathbb{S}_\mathbb{N}}((S_p)^n, S_p) < 10^{-1}$,
- $\forall n \geqslant n_5, d_{\mathbb{S}_\mathbb{P}}((S_c)^n, S_c) < 10^{-1}$,
- $\forall n \geqslant n_6, d_{\mathbb{S}_\mathbb{P}}((S_m)^n, S_m) < 10^{-1}$.

Soit $n_7 = Max(n_4, n_5, n_6)$. Pour $n \geqslant n_7$, toutes les stratégies $(S_p)^n$, $(S_c)^n$, et $(S_m)^n$ ont le même premier terme, qui est respectivement $(S_p)_0$,$(S_c)_0$ et $(S_m)_0$:$\forall n \geqslant n_7$,

$$((S_p)^n_0 = (S_p)_0) \wedge ((S_c)^n_0 = (S_c)_0) \wedge ((S_m)^n_0 = (S_m)_0).$$

Soit $n_8 = Max(n_3, n_7)$. À partir du rang n_8, les états de x^n et x d'une part, et m^n et m d'autre part, sont identiques. De plus, les stratégies $(S_p)^n$ et S_p, $(S_c)^n$ et S_c, et $(S_m)^n$ et S_m commencent avec le même premier terme.

De ce fait, les états de $\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)$ et $\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)$ sont égaux, et donc, après le terme de rang n_8, la distance d_2 entre ces deux points est strictement inférieure à 3.10^{-1}, donc strictement inférieure à 1.

Nous allons à présent démontrer que la distance entre $\left(\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)\right)$ et $\left(\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)\right)$ converge vers 0. Soit $\varepsilon > 0$.

- Si $\varepsilon \geqslant 1$, nous avons vu que la distance entre $\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)$ et $\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)$ est strictement inférieure à 1 à partir du terme de rang n_8 (même état).

- Si $\varepsilon < 1$, alors $\exists k \in \mathbb{N}, 10^{-k} \geqslant \frac{\varepsilon}{3} \geqslant 10^{-(k+1)}$. Comme $d_{\mathbb{S}_\mathbb{N}}((S_p)^n, S_p)$, $d_{\mathbb{S}_\mathbb{P}}((S_c)^n, S_c)$ et $d_{\mathbb{S}_\mathbb{P}}((S_m)^n, S_m)$ converge vers 0, nous avons :
  - $\exists n_9 \in \mathbb{N}, \forall n \geqslant n_9, d_{\mathbb{S}_\mathbb{N}}((S_p)^n, S_p) < 10^{-(k+2)}$,
  - $\exists n_{10} \in \mathbb{N}, \forall n \geqslant n_{10}, d_{\mathbb{S}_\mathbb{P}}((S_c)^n, S_c) < 10^{-(k+2)}$,
  - $\exists n_{11} \in \mathbb{N}, \forall n \geqslant n_{11}, d_{\mathbb{S}_\mathbb{P}}((S_m)^n, S_m) < 10^{-(k+2)}$.

  Soit $n_{12} = Max(n_9, n_{10}, n_{11})$. Alors, après le terme de rang n_{12}, les $k+2$ premiers termes de $(S_p)^n$ et S_p, de $(S_c)^n$ et S_c, et de $(S_m)^n$ et S_m, sont égaux.

Par conséquent, les $k+1$ premiers termes des stratégies de $\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)$ et $\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)$ sont les mêmes (en raison du décalage de stratégie) et d'après les définitions de $d_{\mathbb{S}_N}$ et $d_{\mathbb{S}_P}$:

$$d_2\left(\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n); \mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)\right)$$

est égal à :

$$d_{\mathbb{S}_N}((S_p)^n, S_p) + d_{\mathbb{S}_P}((S_c)^n, S_c) + d_{\mathbb{S}_P}((S_m)^n, S_m),$$

qui est inférieur à $3.10^{-(k+1)} \leqslant 3.\frac{\varepsilon}{3} = \varepsilon$. Soit $N_0 = max(n_8, n_{12})$. Nous pouvons affirmer que :

$$\forall \varepsilon > 0, \exists N_0 \in \mathbb{N}, \forall n \geqslant N_0,$$

$$d_2\left(\mathcal{G}_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n); \mathcal{G}_{f_0}(S_p, x, S_c, m, S_m)\right) \leqslant \varepsilon.$$

$\mathcal{G}_{f_0}$ est par conséquent continue sur $(\mathcal{X}_2, d_2)$. □

8.4/ $\mathcal{CIS}_2$ EST CHAOTIQUE

Afin de démontrer que nous nous trouvons dans un cas de figure du chaos topologique de Devaney, nous devons vérifier que le système dynamique discret considéré présente les propriétés de régularité, transitivité et sensibilité aux conditions initiales.

8.4.1/ RÉGULARITÉ

Proposition 24 : Régularité de $\mathcal{G}_{f_0}$

Les points périodiques de $\mathcal{G}_{f_0}$ sont denses dans $\mathcal{X}_2$.

Démonstration. Soit $(\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m}) \in \mathcal{X}_2$ et $\varepsilon > 0$. Nous allons chercher à déterminer un point périodique $(\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m})$ satisfaisant la condition $d_2((\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m}); (\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m})) < \varepsilon$.

Comme ε peut être strictement inférieure à 1, nous devons choisir $\widetilde{x} = \check{x}$ et $\widetilde{m} = \check{m}$. Nous définissons $k_0(\varepsilon) = \lfloor -log_{10}(\frac{\varepsilon}{3}) \rfloor + 1$ et nous considérons l'ensemble : $\mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)} = \left\{S = (S_p, S_c, S_m) \in \mathbb{S}_N \times \mathbb{S}_P \times \mathbb{S}_P / ((S_p)^k = \check{S_p}^k) \wedge ((S_c)^k = \check{S_c}^k)) \wedge ((S_m)^k = \check{S_m}^k)), \forall k \leqslant k_0(\varepsilon)\right\}$.

Alors, $\forall (S_p, S_c, S_m) \in \mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)}$, $d_2((S_p, \check{x}, S_c, \check{m}, S_m); (\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})) < 3.\frac{\varepsilon}{3} = \varepsilon$. Il reste à choisir $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)}$ tel que $(\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m}) = (\widetilde{S_p}, \check{x}, \widetilde{S_c}, \check{m}, \widetilde{S_m})$ soit un point périodique pour $\mathcal{G}_{f_0}$.

Soit $\mathcal{J} = \{i \in [\![0; N-1]\!] / x_i \neq \check{x}_i$, où $(S_p, x, S_c, m, S_m) = \mathcal{G}_{f_0}^{k_0}(\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})\}$, $\lambda = card(\mathcal{J})$, et $j_0 < j_1 < ... < j_{\lambda-1}$ les éléments de $\mathcal{J}$.

1. Nous allons tout d'abord construire trois stratégies : S_p^*, S_c^*, et S_m^*, comme suit,
   (a) $(S_p^*)^k = \check{S_p}^k$, $(S_c^*)^k = \check{S_c}^k$, et $(S_m^*)^k = \check{S_m}^k$, si $k \leqslant k_0(\varepsilon)$.
   (b) Nous expliquons ensuite comment remplacer $\check{x}_{j_q}$, $\forall q \in [\![0; \lambda - 1]\!]$:
   Avant toute chose on doit remplacer $\check{x}_{j_0}$:

i. Si $\exists \lambda_0 \in [\![0; \mathrm{P}-1]\!] / \check{x}_{j_0} = m_{\lambda_0}$, alors on peut choisir $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = \lambda_0$, $(S_m^*)^{k_0+1} = \lambda_0$, et donc I_{j_0} sera égale à 1.

ii. Si un tel λ_0 n'existe pas, on choisit : $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = 0$, $(S_m^*)^{k_0+1} = 0$, $(S_p^*)^{k_0+2} = j_0$, $(S_c^*)^{k_0+2} = 0$, $(S_m^*)^{k_0+2} = 0$, et $I_{j_0} = 2$.
Tous les $\check{x}_{j_q}$ sont remplacés de la même manière. Les autres termes de S_p^*, S_c^*, et S_m^* sont construits de manière identique, et les valeurs de I_{j_q} sont définies suivant le même principe. Enfin, posons $\gamma = \sum_{q=0}^{\lambda-1} I_{j_q}$.

(c) Finalement, soit $(S_p^*)^k = (S_p^*)^j$, $(S_c^*)^k = (S_c^*)^j$, et $(S_m^*)^k = (S_m^*)^j$, où $j \leqslant k_0(\varepsilon) + \gamma$ satisfait la condition suivante $j \equiv k$ [mod $(k_0(\varepsilon) + \gamma)$], si $k > k_0(\varepsilon) + \gamma$.

Alors, $\mathcal{G}_{f_0}^{k_0(\varepsilon)+\gamma}(S_p^*, \check{x}, S_c^*, \check{m}, S_m^*) = (S_p^*, \check{x}, S_c^*, m, S_m^*)$. Soit

$$\mathcal{K} = \left\{ i \in [\![0; \mathrm{P}-1]\!] / m_i \neq \check{m}_i, \text{ où } \mathcal{G}_{f_0}^{k_0(\varepsilon)+\gamma}(S_p^*, \check{x}, S_c^*, \check{m}, S_m^*) = (S_p^*, \check{x}, S_c^*, m, S_m^*) \right\},$$

$\mu = card(\mathcal{K})$, et $r_0 < r_1 < ... < r_{\mu-1}$ les éléments de $\mathcal{K}$.

2. Nous allons maintenant construire les stratégies $\widetilde{S_p}$, $\widetilde{S_c}$, $\widetilde{S_m}$.

(a) Premièrement, soit $\widetilde{S_p}^k = (S_p^*)^k$, $\widetilde{S_c}^k = (S_c^*)^k$, et $\widetilde{S_m}^k = (S_m^*)^k$, si $k \leqslant k_0(\varepsilon) + \gamma$.

(b) Expliquons comment remplacer $\check{m}_{r_q}, \forall q \in [\![0; \mu-1]\!]$. Commençons déjà par remplacer $\check{m}_{r_0}$:

i. Si $\exists \mu_0 \in [\![0; \mathrm{N}-1]\!] / \check{x}_{\mu_0} = m_{r_0}$, alors on peut choisir $\widetilde{S_p}^{k_0+\gamma+1} = \mu_0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$. Dans cette situation, on définit $J_{r_0} = 1$.

ii. Si un tel μ_0 n'existe pas, alors on peut choisir : $\widetilde{S_p}^{k_0+\gamma+1} = 0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$, $\widetilde{S_p}^{k_0+\gamma+2} = 0$, $\widetilde{S_c}^{k_0+\gamma+2} = r_0$, $\widetilde{S_m}^{k_0+\gamma+2} = 0$, $\widetilde{S_p}^{k_0+\gamma+3} = 0$, $\widetilde{S_c}^{k_0+\gamma+3} = r_0$, et finalement $\widetilde{S_m}^{k_0+\gamma+3} = 0$. Soit alors $J_{r_0} = 3$.
Les autres $\check{m}_{r_q}$ sont remplacés comme précédemment, les autres termes de $\widetilde{S_p}$, $\widetilde{S_c}$, et $\widetilde{S_m}$ sont construits suivant un principe similaire, et les valeurs de J_{r_q} sont définies de la même façon. Notons alors $\alpha = \sum_{q=0}^{\mu-1} J_{r_q}$.

(c) Finalement, soit $\widetilde{S_p}^k = \widetilde{S_p}^j$, $\widetilde{S_c}^k = \widetilde{S_c}^j$, et $\widetilde{S_m}^k = \widetilde{S_m}^j$ où $j \leqslant k_0(\varepsilon) + \gamma + \alpha$ satisfait $j \equiv k$ [mod $(k_0(\varepsilon) + \gamma + \alpha)$], si $k > k_0(\varepsilon) + \gamma + \alpha$.

Alors, $\mathcal{G}_{f_0}^{k_0(\varepsilon)+\gamma+\alpha}(\widetilde{S_p}, \check{x}, \widetilde{S_c}, \check{m}, \widetilde{S_m}) = (\widetilde{S_p}, \check{x}, \widetilde{S_c}, \check{m}, \widetilde{S_m})$. Donc, $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)}$ défini comme précédemment est tel que $(\widetilde{S_m}, \check{x}, \widetilde{S_m}, \check{m}, \widetilde{S_m})$ est un point périodique, de période $k_0(\varepsilon) + \gamma + \alpha$, qui est ε–proche de $(\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})$.

En conclusion, $(\mathcal{X}_2, \mathcal{G}_{f_0})$ est régulier. □

8.4.2/ TRANSITIVITÉ

Montrons maintenant que :

Proposition 25 : Transitivité de $\mathcal{G}_{f_0}$

$(\mathcal{X}_2, \mathcal{G}_{f_0})$ est topologiquement transitive.

Transitivité de $\mathcal{G}_{f_0}$. On définit tout d'abord $\mathcal{X} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{N}}$, tel que $\mathcal{X}(S_p, x, S_c, m, S_m) = x$ et $\mathcal{M} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{P}}$, tel que $\mathcal{M}(S_p, x, S_c, m, S_m) = m$. Soient $\mathcal{B}_A = \mathcal{B}(X_A, r_A)$ et $\mathcal{B}_B = \mathcal{B}(X_B, r_B)$ deux boules ouvertes de l'espace métrique $\mathcal{X}_2$, avec $X_A = ((S_p)_A, x_A, (S_c)_A, m_A, (S_m)_A)$ et $X_B = ((S_p)_B, x_B, (S_c)_B, m_B, (S_m)_B)$. Nous allons alors chercher à déterminer $\widetilde{X} = (\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m})$ dans $\mathcal{B}_A$ tel que $\exists n_0 \in \mathbb{N}, \mathcal{G}_{f_0}^{n_0}(\widetilde{X}) \in \mathcal{B}_B$.
$\widetilde{X}$ doit être dans $\mathcal{B}_A$ et r_A peut être strictement inférieur à 1, donc $\widetilde{x} = x_A$ et $\widetilde{m} = m_A$. Soit $k_0 = \lfloor -\log_{10}(\frac{r_A}{3}) + 1 \rfloor$. On peut noter que $\mathcal{S}_{X_A,k_0} = \Big\{(S_p, S_c, S_m) \in \mathbb{S}_{\mathsf{N}} \times (\mathbb{S}_{\mathsf{P}})^2 / \forall k \leqslant k_0, (S_p^k = (S_p)_A^k) \wedge (S_c^k = (S_c)_A^k) \wedge (S_m^k = (S_m)_A^k))\Big\}$. Alors $\forall (S_p, S_c, S_m) \in \mathcal{S}_{X_A,k_0}, (S_p, \widetilde{x}, S_c, \widetilde{m}, S_m) \in \mathcal{B}_A$. Soit :

$$\mathcal{J} = \left\{ i \in [\![0, \mathsf{N}-1]\!] / \check{x}_i \neq \mathcal{X}(X_B)_i, \text{ où } (\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m}) = \mathcal{G}_{f_0}^{k_0}(X_A) \right\},$$

$\lambda = card(\mathcal{J})$, et $j_0 < j_1 < ... < j_{\lambda-1}$ les éléments de $\mathcal{J}$.

1. Nous allons tout d'abord construire trois stratégies : S_p^*, S_c^*, et S_m^* de la manière suivante.
   (a) $(S_p^*)^k = (S_p)_A^k$, $(S_c^*)^k = (S_c)_A^k$, et $(S_m^*)^k = (S_m)_A^k$, si $k \leqslant k_0$.
   (b) Expliquons comment remplacer $\mathcal{X}(X_B)_{j_q}$, $\forall q \in [\![0; \lambda - 1]\!]$. Avant toute chose, nous devons remplacer $\mathcal{X}(X_B)_{j_0}$:
      i. Si $\exists \lambda_0 \in [\![0; \mathsf{P}-1]\!] / \mathcal{X}(X_B)_{j_0} = \check{m}_{\lambda_0}$, alors on peut choisir $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = \lambda_0$, $(S_m^*)^{k_0+1} = \lambda_0$, et donc I_{j_0} sera égale à 1.
      ii. Si un tel λ_0 n'existe pas, on choisit : $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = 0$, $(S_m^*)^{k_0+1} = 0$, $(S_p^*)^{k_0+2} = j_0$, $(S_c^*)^{k_0+2} = 0$, $(S_m^*)^{k_0+2} = 0$ et finalement nous pouvons noter $I_{j_0} = 2$.
      Tous les $\mathcal{X}(X_B)_{j_q}$ sont remplacés de manière identique. Les autres termes de S_p^*, S_c^*, et S_m^* sont construits de façon identique, et les valeurs de I_{j_q} sont définies de la même manière. Soit $\gamma = \sum_{q=0}^{\lambda-1} I_{j_q}$.
   (c) $(S_p^*)^k = (S_p^*)^j$, $(S_c^*)^k = (S_c^*)^j$ et $(S_m^*)^k = (S_m^*)^j$ où $j \leqslant k_0 + \gamma$ satisfait la condition suivante : $j \equiv k$ [mod $(k_0 + \gamma)$], si $k > k_0 + \gamma$.

   Alors,$\mathcal{G}_{f_0}^{k_0+\gamma}((S_p^*, x_A, S_c^*, m_A, S_m^*)) = (S_p^*, x_B, S_c^*, m, S_m^*)$. Soit :

   $$\mathcal{K} = \left\{ i \in [\![0; \mathsf{P}-1]\!] / m_i \neq \mathcal{M}(X_B)_i, \text{ où } (S_p^*, x_B, S_c^*, m, S_m^*) = \mathcal{G}_{f_0}^{k_0+\gamma}((S_p^*, x_A, S_c^*, m_A, S_m^*)) \right\},$$

   $\mu = card(\mathcal{K})$, et $r_0 < r_1 < ... < r_{\mu-1}$ les éléments de $\mathcal{K}$.
2. Nous allons dans un second temps construire trois autres stratégies, à savoir $\widetilde{S_p}$, $\widetilde{S_c}$, $\widetilde{S_m}$, et de la manière suivante :
   (a) $\widetilde{S_p}^k = (S_p^*)^k$, $\widetilde{S_c}^k = (S_c^*)^k$, et $\widetilde{S_m}^k = (S_m^*)^k$, si $k \leqslant k_0 + \gamma$.
   (b) Nous allons à présent expliquer comment remplacer $\mathcal{M}(X_B)_{r_q}, \forall q \in [\![0; \mu - 1]\!]$. Premièrement, nous devons remplacer $\mathcal{M}(X_B)_{r_0}$:
      i. Si $\exists \mu_0 \in [\![0; \mathsf{N}-1]\!] / \mathcal{M}(X_B)_{r_0} = (x_B)_{\mu_0}$, alors on peut choisir $\widetilde{S_p}^{k_0+\gamma+1} = \mu_0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$, et J_{r_0} sera égale à 1.
      ii. Si un tel μ_0 n'existe pas, on choisit : $\widetilde{S_p}^{k_0+\gamma+1} = 0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$, $\widetilde{S_p}^{k_0+\gamma+2} = 0$, $\widetilde{S_c}^{k_0+\gamma+2} = r_0$, $\widetilde{S_m}^{k_0+\gamma+2} = 0$, $\widetilde{S_p}^{k_0+\gamma+3} = 0$, $\widetilde{S_c}^{k_0+\gamma+3} = r_0$, $\widetilde{S_m}^{k_0+\gamma+3} = 0$, et nous pouvons alors noter $J_{r_0} = 3$.

Tous les $\mathcal{M}(X_B)_{r_q}$ sont remplacés de manière similaire. Les autres termes de $\widetilde{S_p}$, $\widetilde{S_c}$, et $\widetilde{S_m}$ sont construits de manière identique, et les valeurs de J_{r_q} sont définies de la même manière. Soit alors $\alpha = \sum_{q=0}^{\mu-1} J_{r_q}$.

(c) $\forall k \in \mathbb{N}^*, \widetilde{S_p}^{k_0+\gamma+\alpha+k} = (S_p)_B^k$, $\widetilde{S_c}^{k_0+\gamma+\alpha+k} = (S_c)_B^k$, et $\widetilde{S_m}^{k_0+\gamma+\alpha+k} = (S_m)_B^k$.

Alors, $\mathcal{G}_{f_0}^{k_0+\gamma+\alpha}(\widetilde{S_p}, x_A, \widetilde{S_c}, m_A, \widetilde{S_m}) = X_B$, avec $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{X_A,k_0}$, donc $\widetilde{X} = (\widetilde{S_p}, x_A, \widetilde{S_c}, m_A, \widetilde{S_m}) \in \mathcal{X}_2$ est tel que $\widetilde{X} \in \mathcal{B}_A$ et $\mathcal{G}_{f_0}^{k_0+\gamma+\alpha}(\widetilde{X}) \in \mathcal{B}_B$, ce qui montre la transitivité cherchée. □

8.4.3/ SENSIBILITÉ AUX CONDITIONS INITIALES

Montrons enfin que :

Proposition 26 : Sensibilité de $\mathcal{G}_{f_0}$

$(\mathcal{X}_2, \mathcal{G}_{f_0})$ a une sensible dépendance aux conditions initiales.

Démonstration. $\mathcal{G}_{f_0}$ est régulière et transitive. Donc, d'après le théorème 1 (théorème de Banks), $\mathcal{G}_{f_0}$ a une sensible dépendance aux conditions initiales. □

8.4.4/ CHAOS TOPOLOGIQUE DE DEVANEY

En conclusion, $(\mathcal{X}_2, \mathcal{G}_{f_0})$ est topologiquement régulier, transitif et a une sensible dépendance aux conditions initiales. Nous avons donc démontré le résultat suivant :

Théorème 7 : $\mathcal{G}_{f_0}$ est chaotique

$\mathcal{G}_{f_0}$ est une fonction chaotique sur l'espace métrique $(\mathcal{X}_2, d_2)$ au sens de Devaney.

Nous pouvons alors affirmer que :

Théorème 8 : Sécurité topologique de $\mathcal{CIS}_2$

$\mathcal{CIS}_2$ est topologiquement-sûr.

8.4.5/ ÉVALUATION DE LA CONSTANTE DE SENSIBILITÉ

On rappelle qu'un système est sensible aux conditions initiales si pour chaque x, il existe des points arbitrairement proches de x dont les orbites respectives sont séparées au moins de ε pendant l'évolution du système (définition 38 page 33). Tous les points voisins de x ne sont pas forcément séparés de ε pendant l'évolution du système : il suffit qu'il en existe au moins un dans chaque boule ouverte de centre x. Comme nous l'avons vu à la section précédente, nous savons que cette propriété est une conséquence de la régularité et de la transitivité (théorème de Banks 1). Cependant, cela ne nous dit pas quelle est la constante de sensibilité. C'est pourquoi l'on va redémontrer la sensibilité « à la main » :

Proposition 27 : Constante de sensibilité de CIS_2

$\mathcal{G}_{f_0}$ est sensible aux conditions initiales sur $(\mathcal{X}_2, d_2)$ et sa constante de sensibilité est supérieure ou égale à N + P − 1.

Démonstration. On définit tout d'abord $\mathcal{X} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{N}}$, tel que $\mathcal{X}(S_p, x, S_c, m, S_m) = x$ et $\mathcal{M} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{P}}$, tel que $\mathcal{M}(S_p, x, S_c, m, S_m) = m$.
Soit $\varepsilon > 0$. Considérons $\varepsilon < 1$. Soit $\check{X} = (\check{S_p}, \check{x}, \check{S_c}, m, \check{S_m}) \in \mathcal{X}_2$. On recherche $\widetilde{X} = (\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, m, \widetilde{S_m}) \in \mathcal{X}_2$ tel que $d_2\left((\check{X}, \widetilde{X})\right) \leqslant \varepsilon$ et $\exists n_0 \in \mathbb{N}$, $d_2\left(\mathcal{G}_{f_0}^{n_0}(\check{X}); \mathcal{G}_{f_0}^{n_0}(\widetilde{X})\right) \geqslant \mathsf{N} + \mathsf{P} - 1$.

Nous définissons $k_0(\varepsilon) = \lfloor -log_{10}(\frac{\varepsilon}{3}) \rfloor + 1$ et nous considérons l'ensemble : $\mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)} = \left\{ S = (S_p, S_c, S_m) \in \mathbb{S}_{\mathsf{N}} \times \mathbb{S}_{\mathsf{P}} \times \mathbb{S}_{\mathsf{P}} / ((S_p)^k = \check{S_p}^k) \wedge ((S_c)^k = \check{S_c}^k)) \ \wedge ((S_m)^k = \check{S_m}^k)), \forall k \leqslant k_0(\varepsilon) \right\}$.

Alors, $\forall (S_p, S_c, S_m) \in \mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)}$, $d_2((S_p, \check{x}, S_c, \check{m}, S_m), (\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})) < 3.\frac{\varepsilon}{3} = \varepsilon$. Les points $(S_p, \check{x}, S_c, \check{m}, S_m)$ et $(\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m})$ sont donc proches à ε près. Il reste à choisir $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{\check{S_p}, \check{S_c}, \check{S_m}, k_0(\varepsilon)}$ tel que $\exists n_0 \in \mathbb{N}$, $d_2\left(\mathcal{G}_{f_0}^{n_0}(\check{X}); \mathcal{G}_{f_0}^{n_0}(\widetilde{X})\right) \geqslant \mathsf{N} + \mathsf{P} - 1$.

Soit $\mathcal{J} = \left\{ i \in [\![0, \mathsf{N} - 1]\!] \ / \ \mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_i = \mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_i \right\}$ l'ensemble des indices des cellules égales entre les états de $\mathcal{G}_{f_0}^{k_0}(\check{X})$ et de $\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})$, et soit λ le nombre de valeurs égales : $\lambda = card\,(\mathcal{J})$. Soient $j_0 < j_1 < ... < j_{\lambda-1}$ les éléments de $\mathcal{J}$.

Nous allons tout d'abord expliquer comment obtenir des cellules toutes différentes pour les deux états de $\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})$ et $\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(X^*)$ où $X^* = (S_p^*, \check{x}, S_c^*, \check{m}, S_m^*)$.

1. Nous allons tout d'abord construire trois stratégies : S_p^*, S_c^*, et S_m^*, comme suit,
   (a) $(S_p^*)^k = \check{S_p}^k$, $(S_c^*)^k = \check{S_c}^k$, et $(S_m^*)^k = \check{S_m}^k$, si $k \leqslant k_0(\varepsilon)$. (Pour que X^* soit proche de $\check{X}$ à ε près).
   (b) Nous expliquons ensuite comment remplacer $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2q)}(\check{X})\right)_{j_q}$, $\forall q \in [\![0; \lambda - 1]\!]$:
   Avant toute chose (pour $q = 0$) on doit remplacer $\mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{j_0}$:
      i. Si $\exists \lambda_0 \in [\![0; \mathsf{P} - 1]\!] / \mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{j_0} = \mathcal{M}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{\lambda_0}$, alors on peut choisir : $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = \lambda_0$, $(S_m^*)^{k_0+1} = \lambda_0$ (pour obtenir la négation $\overline{\mathcal{M}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{\lambda_0}}$ à la place de $\mathcal{M}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{\lambda_0}$), $(S_p^*)^{k_0+2} = j_0$, $(S_c^*)^{k_0+2} = \lambda_0$, $(S_m^*)^{k_0+2} = \lambda_0$ (pour obtenir la négation de $\mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{j_0}$), et donc I_{j_0} sera égale à 2.
      ii. Si un tel λ_0 n'existe pas (alors $\mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{j_0}$ est différent de tous les $\mathcal{M}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_i$, en particulier $\mathcal{M}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_0$), on choisit : $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = 0$, $(S_m^*)^{k_0+1} = 1$ (pour obtenir la négation de $\mathcal{X}\left(\mathcal{G}_{f_0}^{k_0}(\check{X})\right)_{j_0}$), $(S_p^*)^{k_0+2} = j_0$, $(S_c^*)^{k_0+2} = 0$, $(S_m^*)^{k_0+2} = 0$ (pour avoir une seconde opération "qui ne change rien" à ce niveau), et donc, dans ce cas aussi, $I_{j_0} = 2$.

   Tous les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2q)}(\check{X})\right)_{j_q}$ sont remplacés de la même manière. Les autres termes de S_p^*, S_c^*, et S_m^* sont construits de manière identique, et les valeurs de I_{j_q} sont définies suivant le même principe, toutes égales à 2. Enfin, posons $\gamma = \sum_{q=0}^{\lambda-1} I_{j_q} = 2\lambda$.

(c) Il nous faut donc à présent nous occuper des $(2\mathsf{N} - 2\lambda)$ opérations restantes (pour obtenir un nombre d'itérations identique de part et d'autre). Chaque opération doit laisser les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+\gamma)}(\check{X})\right)_i$ inchangés. Alors : $\forall k \in [\![\lambda+1;\mathsf{N}]\!]$:

i. Si $\exists \lambda_k \in [\![0;\mathsf{P}-1]\!] / \mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_0 = \mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_{\lambda_k}$, alors soit $\lambda'_k \neq \lambda_k$, on peut choisir :
$(S_p^*)^{k_0+2k-1} = 0$, $(S_c^*)^{k_0+2k-1} = \lambda_k$, $(S_m^*)^{k_0+2k-1} = \lambda'_k$, $(S_p^*)^{k_0+2k} = 0$, $(S_c^*)^{k_0+2k} = \lambda_k$, $(S_m^*)^{k_0+2k} = \lambda'_k$.

ii. Si un tel λ_k n'existe pas (alors $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_0$ est différent de tous les $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_i$, en particulier $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_0$), on choisit : $(S_p^*)^{k_0+2k-1} = 0$, $(S_c^*)^{k_0+2k-1} = 0$, $(S_m^*)^{k_0+2k-1} = 0$ (pour obtenir la négation de $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_0$), $(S_p^*)^{k_0+2k} = 0$, $(S_c^*)^{k_0+2k} = 0$, $(S_m^*)^{k_0+2k} = 0$ (pour revenir, finalement, à la valeur initiale de $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2k-2)}(\check{X})\right)_0$).

(d) $\forall k > k_0 + 2\mathsf{N}$, $(S_p^*)^k = 1$, $(S_c^*)^k = 1$ et $(S_m^*)^k = 1$ pour finir de définir pleinement X^*.

Alors, $d_2\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X}), \mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(X^*)\right) \geq \mathsf{N}$ avec $d_2(\check{X}, X^*) < \varepsilon$ puisque, par construction, $(S_p^*, S_c^*, S_m^*) \in \mathcal{S}_{\check{S_p},\check{S_c},\check{S_m},k_0(\varepsilon)}$.

Soit $\mathcal{K} = \left\{ i \in [\![0, \mathsf{P}-1]\!] \ / \ \mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_i = \mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N}+3\mathsf{P})}(\check{X})\right)_i \right\}$ l'ensemble des indices des cellules égales entre les messages de $\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})$ et de $\mathcal{G}_{f_0}^{(k_0+2\mathsf{N}+3\mathsf{P})}(\check{X})$, et soit μ le nombre de valeurs égales : $\mu = card(\mathcal{K})$. Soient $r_0 < r_1 < ... < r_{\mu-1}$ les éléments de $\mathcal{K}$.

2. Nous allons maintenant construire les stratégies $\widetilde{S_p}$, $\widetilde{S_c}$, $\widetilde{S_m}$.

(a) Premièrement, soit $\widetilde{S_p}^k = (S_p^*)^k$, $\widetilde{S_c}^k = (S_c^*)^k$, et $\widetilde{S_m}^k = (S_m^*)^k$, si $k \leqslant k_0(\varepsilon) + 2\mathsf{N}$ (pour que les propriétés du point 1 restent valable).

(b) Expliquons comment remplacer $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N}+2q)}(\check{X})\right)_{r_q}, \forall q \in [\![0;\mu-1]\!]$. Commençons déjà $(q=0)$ par remplacer $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}$:

i. Si $\exists \mu_0 \in [\![0;\mathsf{N}-1]\!] / \mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0} = \mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{\mu_0}$, alors on peut choisir $\widetilde{S_p}^{(k_0+2\mathsf{N}+1)} = \mu_0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+1)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+1)} = r_0$. On choisit ensuite $\widetilde{S_p}^{(k_0+2\mathsf{N}+2)} = \mu_0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+2)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+2)} = r_0$ et, finalement, on choisit $\widetilde{S_p}^{(k_0+2\mathsf{N}+3)} = \mu_0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+3)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+3)} = r_0$. Après ces trois opérations, $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{\mu_0}$ reste inchangé et on obtient la négation $\overline{\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}}$ à la place de $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}$. Dans cette situation, on définit $J_{r_0} = 3$.

ii. Si un tel μ_0 n'existe pas (alors $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}$ est différent de tous les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_i$, en particulier $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_0$), alors on peut choisir $\widetilde{S_p}^{(k_0+2\mathsf{N}+1)} = 0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+1)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+1)} = r_0$. On peut choisir ensuite $\widetilde{S_p}^{(k_0+2\mathsf{N}+2)} = 0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+2)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+2)} = 0$, et, finalement $\widetilde{S_p}^{(k_0+2\mathsf{N}+3)} = 0$, $\widetilde{S_c}^{(k_0+2\mathsf{N}+3)} = r_0$, $\widetilde{S_m}^{(k_0+2\mathsf{N}+3)} = 0$. Après ces trois opérations, $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_0$ reste inchangé et on obtient la négation $\overline{\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}}$ à la place de $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathsf{N})}(\check{X})\right)_{r_0}$. Soit alors $J_{r_0} = 3$.

Les autres $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N})}(\check{X})\right)_{r_q}$ sont remplacés comme précédemment, les autres termes de $\widetilde{S_p}$, $\widetilde{S_c}$, et $\widetilde{S_m}$ sont construits suivant un principe similaire, et les valeurs de J_{r_q} sont définies de la même façon, toutes égales à 3. Notons alors $\alpha = \sum_{q=0}^{\mu-1} J_{r_q} = 3\mu$.

(c) Il nous faut donc à présent nous occuper des $(3\mathrm{P} - 3\mu)$ opérations restantes (pour obtenir un nombre d'itérations identique de part et d'autre). Chaque opération doit, dans la mesure du possible, laisser les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_i$ ainsi que les $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_j$ inchangés. Il y a deux cas possibles, soit $(3\mathrm{P} - 3\mu)$ est paire, soit $(3\mathrm{P} - 3\mu)$ est impaire.

i. Si $(3\mathrm{P} - 3\mu)$ est paire, nous arrivons à laisser inchangé tous les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_i$ ainsi que les $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_j$. Pour ce faire, il suffit d'opérer, successivement, les deux opérations suivantes (on note k le rang considéré) :

A. Si $\exists \mu_k \in [\![0; \mathrm{N}-1]\!] / \mathcal{M}\left(\mathcal{G}_{f_0}^{k}(\check{X})\right)_0 = \mathcal{X}\left(\mathcal{G}_{f_0}^{k}(\check{X})\right)_{\mu_k}$, on peut choisir $(S_p^*)^{k+1} = \mu_k$, $(S_c^*)^{k+1} = 0$, $(S_m^*)^{k+1} = 1$, on peut choisir ensuite $(S_p^*)^{k+2} = \mu_k$, $(S_c^*)^{k+2} = 0$, $(S_m^*)^{k+2} = 1$.

B. Si un tel μ_k n'existe pas (alors $\mathcal{M}\left(\mathcal{G}_{f_0}^{k}(\check{X})\right)_0$ est différent de tous les $\mathcal{X}\left(\mathcal{G}_{f_0}^{k}(\check{X})\right)_i$, en particulier $\mathcal{X}\left(\mathcal{G}_{f_0}^{k}(\check{X})\right)_0$), on choisit : $(S_p^*)^{k+1} = 0$, $(S_c^*)^{k+1} = 0$, $(S_m^*)^{k+1} = 0$, et, ensuite, on choisit $(S_p^*)^{k+2} = 0$, $(S_c^*)^{k+2} = 0$, $(S_m^*)^{k+2} = 0$.

ii. Si $(3\mathrm{P} - 3\mu)$ est impaire, nous arrivons, de la même manière qu'au point précédent, à laissé inchangé tous les $\mathcal{X}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_i$ ainsi que les $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_j$, sauf peut le dernier des $\mathcal{M}\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+\alpha)}(\check{X})\right)_j$.

(d) $\forall k > k_0 + 2\mathrm{N} + 3\mathrm{P}$, $\widetilde{S_p}^k = 1$, $\widetilde{S_c}^k = 1$ et $\widetilde{S_m}^k = 1$ pour finir de définir pleinement $\widetilde{X}$.

Alors, $d_2\left(\mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+3\mathrm{P})}(\check{X}), \mathcal{G}_{f_0}^{(k_0+2\mathrm{N}+3\mathrm{P})}(\widetilde{X})\right) \geq \mathrm{N} + \mathrm{P} - 1$ (-1 pour prendre en compte le cas où l'on a pas réussi à laisser identique la dernière composante du message tel que décrit au point 2(c)ii) avec $d_2(\check{X}, \widetilde{X}) < \varepsilon$ puisque, par construction, $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{\check{S_p},\check{S_c},\check{S_m},k_0(\varepsilon)}$.

Donc, la constante de sensibilité de $\mathcal{G}_{f_0}$ est égale à $\mathrm{N} + \mathrm{P} - 1$. □

8.4.6/ COMPACITÉ ET DE FORTE TRANSITIVITÉ

Dans cette section nous nous intéressons à l'étude de la compacité de l'espace métrique $\mathcal{X}_2$. Ce concept sera ensuite utilisé pour établir la forte transitivité, une nouvelle propriété qualitative de la sécurité topologique du processus de dissimulation d'informations $\mathcal{CIS}_2$. Les concepts de compacité et de forte transitivité sont définis à la section 4.5.

8.4.6.1/ COMPACITÉ DE $\mathcal{X}_2$

Donc, en utilisant la caractérisation séquentielle de la compacité donnée dans la proposition 1, il est alors possible de démontrer que :

Théorème 9 : Compacité de $\mathcal{X}_2$

L'espace métrique $(\mathcal{X}_2, d_2)$ est compact.

Démonstration. La preuve suivante est une preuve constructive [84, 88]. Il sera expliqué, dans cette partie, comment il est toujours possible de construire une sous-suite convergente à partir de n'importe quelle suite.

Soit $(Y^n)_n = ((S_p)^n, x^n, (S_c), m^n, (S_m)^n))_n \in (\mathcal{X}_2)^{\mathbb{N}}$ une suite de points de $\mathcal{X}_2$.

1. Construction du premier terme de la limite séquentielle :

   (a) Construction de I_1 et n_x :
   Comme $\mathbb{B}^{\mathsf{N}}$ est un ensemble fini et comme $\forall n \in \mathbb{N}, x^n \in \mathbb{B}^{\mathsf{N}}$, $\exists \lambda_x \in \mathbb{B}^{\mathsf{N}}$ tel que λ_x apparaît un nombre infini de fois dans la seconde projection de la suite $(Y^n)_n$. Soit n_x le premier indice où la valeur λ_x apparaît. Soit I_1 l'ensemble des termes Y^n de la suite $(Y^n)_n$ tels que $x^n = x^{n_x}$. On note $I_1 = \{Y^n \ / \ n \in \mathbb{N} \wedge x^n = x^{n_x}\}$.

   (b) Construction de I_2 et n_m :
   Comme I_1 est un ensemble infini, $\mathbb{B}^{\mathsf{P}}$ est un ensemble fini, et $\forall n \in \mathbb{N}, m^n \in \mathbb{B}^{\mathsf{P}}$, $\exists \lambda_m \in \mathbb{B}^{\mathsf{P}}$ tel que λ_m apparaît un nombre infini de fois dans la seconde projection de la suite $(Y^n)_n$. Soit n_m le premier indice où cette valeur λ_m apparaît. Nous notons $I_2 = \{Y^n \in I_1 \ / \ n \in \mathbb{N} \wedge x^n = x^{n_m} \wedge m^n = m^{n_m}\}$.

   (c) Construction de I_3 et p_0 :
   Les premiers termes $\left((S_p)^n\right)^0$ des stratégies $(S_p)^n$ de tuples de I_2 appartiennent à $[\![1, \mathsf{N}]\!]$, et I_2 est un ensemble infini. Donc, $\exists k_p \in [\![1, \mathsf{N}]\!]$ pour lequel il y a un nombre infini de stratégies de I_2 dont le premier terme est égal à k_p. Soit p_0 le plus petit entier n tel que $Y^n \in I_2$ et $\left((S_p)^n\right)^0 = k_p$. Soit $I_3 = \{Y^n \in I_2 \ / \ n \in \mathbb{N} \wedge x^n = x^{p_0} \wedge m^n = m^{p_0} \wedge \left((S_p)^n\right)^0 = \left((S_p)^{p_0}\right)^0\}$

   (d) Construction de I_4 et c_0, puis I_5 et m_0 :
   De la même manière que pour les étapes précédentes, et pour des raisons analogues, nous pouvons construire les deux ensembles I_4 et $I_5 = \{Y^n \in I_4 \ / \ n \in \mathbb{N} \wedge x^n = x^{m_0} \wedge m^n = m^{m_0} \wedge \left((S_p)^n\right)^0 = \left((S_p)^{m_0}\right)^0 \wedge ((S_c)^n)^0 = ((S_c)^{m_0})^0 \wedge ((S_m)^n)^0 = ((S_m)^{m_0})^0\}$

2. La construction de toutes les autres valeurs de la suite limite et de toutes les autres valeurs m_k utilisées dans la sous-suite obéit exactement au même modèle que précédemment.

Soit $l = \left(\left(\left((S_p)^{m_k}\right)^k, x^{p_0}, ((S_c)^{m_k})^k, m^{p_0}, ((S_m)^{m_k})^k\right)_{k \in \mathbb{N}}\right)$,
alors la sous-suite $\left((S_p)^{m_k}, x^{m_k}, (S_c)^{m_k}, m^{m_k}, (S_m)^{m_k}\right)$ converge vers l. □

8.4.6.2/ FORTE TRANSITIVITÉ DE $\mathcal{CIS}_2$

La forte transitivité est introduite dans la définition 28 page 29.

Donc, d'après la proposition donnée à la section 4.5.5, et d'après les deux théorèmes 9 et 7, nous pouvons en déduire le résultat suivant :

Théorème 10 : Forte transitivité de $\mathcal{G}_{f_0}$

$\mathcal{G}_{f_0}$ est fortement transitif sur $(\mathcal{X}_2, d_2)$.

Par conséquent, $\mathcal{CIS}_2$ est aussi fortement transitif.

Cette nouvelle propriété qualitative de $\mathcal{CIS}_2$ permet donc à ce processus de dissimulation d'informations de mieux résister aux attaques dans les configurations de type KOA, KMA, CMA et WOA telles que définies dans [24] et rappelées à la section 5.1.2 dans le contexte du tatouage numérique.

8.4.7/ MÉLANGE TOPOLOGIQUE DE $\mathcal{CIS}_2$

La définition du mélange topologique est donnée dans la définition 29 page 29.

Proposition 28 : Mélange topologique de $\mathcal{G}_{f_0}$

$\mathcal{G}_{f_0}$ est topologiquement mélangeant sur $(\mathcal{X}_2, d_2)$.

Ce résultat est une conséquence immédiate du lemme suivant.

Lemme 2 :

Pour toute boule ouverte B de $\mathcal{X}_2$, un indice n peut être trouvé tel que $\mathcal{G}^n_{f_0}(B) = \mathcal{X}_2$.

Démonstration. On définit tout d'abord $\mathcal{X} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{N}}$, tel que $\mathcal{X}(S_p, x, S_c, m, S_m) = x$ et $\mathcal{M} : \mathcal{X}_2 \to \mathbb{B}^{\mathsf{P}}$, tel que $\mathcal{M}(S_p, x, S_c, m, S_m) = m$.

Soit $X' = (S'_p, x', S'_c, m', S'_m) \in \mathcal{X}_2$ et $\varepsilon > 0$. Soit $\mathcal{B} = B(X', \varepsilon)$ une boule ouverte, pour laquelle le rayon peut être considéré comme strictement inférieur à 1. Tous les éléments de $\mathcal{B}$ ont le même état x', le même message m' (d'après la définition de la distance d_2) et sont tels qu'un entier $k(\varepsilon) = \lfloor -log_{10}(\frac{\varepsilon}{3}) \rfloor + 1$ satisfait :

- toutes les stratégies chaotiques S_p, S_c et S_m de B ont les mêmes $k(\varepsilon)$ premiers termes,
- après l'indice $k(\varepsilon)$, toutes les valeurs sont possibles.

Soit $X'' = (S''_p, x'', S''_c, m'', S''_m) \in \mathcal{X}_2$.

Nous allons alors chercher à déterminer $\widetilde{X} = (\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m})$ dans $\mathcal{B}$ tel que $\exists n_0 \in \mathbb{N}, \mathcal{G}^{n_0}_{f_0}(\widetilde{X}) = X''$.
$\widetilde{X}$ doit être dans $\mathcal{B}$ avec $\varepsilon < 1$, donc $\widetilde{x} = x'$ et $\widetilde{m} = m'$. Soit $k_0(\varepsilon) = \lfloor -\log_{10}(\frac{\varepsilon}{3}) + 1 \rfloor$.

Nous définissons l'ensemble $\mathcal{S}_{X',k_0}$ par :

$$\mathcal{S}_{X',k_0} = \left\{ (S_p, S_c, S_m) \in \mathbb{S}_{\mathsf{N}} \times (\mathbb{S}_{\mathsf{P}})^2 / \forall k \leqslant k_0, (S^k_p = (S'_p)^k) \wedge (S^k_c = (S'_c)^k) \wedge (S^k_m = (S'_m)^k)) \right\}$$

Alors $\forall (S_p, S_c, S_m) \in \mathcal{S}_{X',k_0}, (S_p, x', S_c, m', S_m) \in \mathcal{B}$.

Soit

$$\mathcal{J} = \left\{ i \in [\![0, \mathsf{N}-1]\!] / \check{x}_i \neq \mathcal{X}(X'')_i, \text{ où } (\check{S_p}, \check{x}, \check{S_c}, \check{m}, \check{S_m}) = \mathcal{G}^{k_0}_{f_0}(S_p, x', S_c, m', S_m) \right\},$$

$\lambda = card(\mathcal{J})$, et $j_0 < j_1 < ... < j_{\lambda-1}$ les éléments de $\mathcal{J}$.

1. Nous allons tout d'abord construire trois stratégies : S_p^*, S_c^*, et S_m^* de la manière suivante.
   (a) $(S_p^*)^k = (S_p')^k$, $(S_c^*)^k = (S_c')^k$, et $(S_m^*)^k = (S_m')^k$, si $k \leqslant k_0$.
   (b) Expliquons comment remplacer $\mathcal{X}(X'')_{j_q}$, $\forall q \in [\![0; \lambda - 1]\!]$. Avant toute chose, nous devons remplacer $\mathcal{X}(X'')_{j_0}$:
      i. Si $\exists \lambda_0 \in [\![0; \mathrm{P} - 1]\!] / \mathcal{X}(X'')_{j_0} = \breve{m}_{\lambda_0}$, alors on peut choisir $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = \lambda_0$, $(S_m^*)^{k_0+1} = \lambda_0$, et donc I_{j_0} sera égale à 1.
      ii. Si un tel λ_0 n'existe pas, on choisit : $(S_p^*)^{k_0+1} = j_0$, $(S_c^*)^{k_0+1} = 0$, $(S_m^*)^{k_0+1} = 0$, $(S_p^*)^{k_0+2} = j_0$, $(S_c^*)^{k_0+2} = 0$, $(S_m^*)^{k_0+2} = 0$ et finalement nous pouvons noter $I_{j_0} = 2$.
      Tous les $\mathcal{X}(X'')_{j_q}$ sont remplacés de manière identique. Les autres termes de S_p^*, S_c^*, et S_m^* sont construits de façon identique, et les valeurs de I_{j_q} sont définies de la même manière. Soit $\gamma = \sum_{q=0}^{\lambda-1} I_{j_q}$.
   (c) $(S_p^*)^k = (S_p^*)^j$, $(S_c^*)^k = (S_c^*)^j$ et $(S_m^*)^k = (S_m^*)^j$ où $j \leqslant k_0 + \gamma$ satisfait la condition suivante : $j \equiv k$ [mod $(k_0 + \gamma)$], si $k > k_0 + \gamma$.

   Alors,$\mathcal{G}_{f_0}^{k_0+\gamma}((S_p^*, x', S_c^*, m', S_m^*)) = (S_p^*, x'', S_c^*, m, S_m^*)$. Soit :

   $$\mathcal{K} = \left\{ i \in [\![0; \mathrm{P} - 1]\!] / m_i \neq \mathcal{M}(X'')_i, \text{ où } (S_p^*, x'', S_c^*, m, S_m^*) = \mathcal{G}_{f_0}^{k_0+\gamma}((S_p^*, x', S_c^*, m', S_m^*)) \right\},$$

   $\mu = card(\mathcal{K})$, et $r_0 < r_1 < ... < r_{\mu-1}$ les éléments de $\mathcal{K}$.
2. Nous allons dans un second temps construire trois autres stratégies, à savoir $\widetilde{S_p}$, $\widetilde{S_c}$, $\widetilde{S_m}$, et de la manière suivante :
   (a) $\widetilde{S_p}^k = (S_p^*)^k$, $\widetilde{S_c}^k = (S_c^*)^k$, et $\widetilde{S_m}^k = (S_m^*)^k$, si $k \leqslant k_0 + \gamma$.
   (b) Nous allons à présent expliquer comment remplacer $\mathcal{M}(X'')_{r_q}, \forall q \in [\![0; \mu - 1]\!]$. Premièrement, nous devons remplacer $\mathcal{M}(X_B)_{r_0}$:
      i. Si $\exists \mu_0 \in [\![0; \mathrm{N} - 1]\!] / \mathcal{M}(X_B)_{r_0} = (x_B)_{\mu_0}$, alors on peut choisir $\widetilde{S_p}^{k_0+\gamma+1} = \mu_0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$, et J_{r_0} sera égal à 1.
      ii. Si un tel μ_0 n'existe pas, on choisit : $\widetilde{S_p}^{k_0+\gamma+1} = 0$, $\widetilde{S_c}^{k_0+\gamma+1} = r_0$, $\widetilde{S_m}^{k_0+\gamma+1} = r_0$, $\widetilde{S_p}^{k_0+\gamma+2} = 0$, $\widetilde{S_c}^{k_0+\gamma+2} = r_0$, $\widetilde{S_m}^{k_0+\gamma+2} = 0$, $\widetilde{S_p}^{k_0+\gamma+3} = 0$, $\widetilde{S_c}^{k_0+\gamma+3} = r_0$, $\widetilde{S_m}^{k_0+\gamma+3} = 0$, et nous pouvons alors noter $J_{r_0} = 3$.
      Tous les $\mathcal{M}(X'')_{r_q}$ sont remplacés de manière similaire. Les autres termes de $\widetilde{S_p}$, $\widetilde{S_c}$, et $\widetilde{S_m}$ sont construits de manière identique, et les valeurs de J_{r_q} sont définies de la même manière. Soit alors $\alpha = \sum_{q=0}^{\mu-1} J_{r_q}$.
   (c) $\forall k \in \mathbb{N}^*, \widetilde{S_p}^{k_0+\gamma+\alpha+k} = (S_p'')^k$, $\widetilde{S_c}^{k_0+\gamma+\alpha+k} = (S_c'')^k$, et $\widetilde{S_m}^{k_0+\gamma+\alpha+k} = (S_m'')^k$.

Alors, $\mathcal{G}_{f_0}^{k_0+\gamma+\alpha}(\widetilde{S_p}, x', \widetilde{S_c}, m', \widetilde{S_m}) = X''$, avec $(\widetilde{S_p}, \widetilde{S_c}, \widetilde{S_m}) \in \mathcal{S}_{X',k_0}$, donc $\widetilde{X} = (\widetilde{S_p}, x', \widetilde{S_c}, m', \widetilde{S_m}) \in \mathcal{X}_2$ est tel que $\widetilde{X} \in \mathcal{B}$ et $\mathcal{G}_{f_0}^{k_0+\gamma+\alpha}(\widetilde{X}) = X''$, ce qui montre le lemme attendu. □

8.4.8/ COMPLÉTUDE TOPOLOGIQUE ET PERFECTION

Dans cette section, nous étudions la complétude topologique et la perfection de l'espace métrique $\mathcal{X}_2$.

8.4.8.1/ ÉTUDE DE LA COMPLÉTUDE

La définition d'un espace métrique complet est donnée à la section 4.6.

Proposition 29 : Complétude de $\mathcal{X}_2$

$(\mathcal{X}_2, d_2)$ est un espace métrique complet.

Démonstration. Soit $(X^n)_{n\in\mathbb{N}}$ une suite de Cauchy de l'espace $\mathcal{X}_2$:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n, p \in [\![N; +\infty[\![, d_2(X^n, X^p) < \varepsilon.$$

Nous utiliserons la notation suivante : $\forall n \in \mathbb{N}, X^n = ((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n)$.

Alors, pour $\varepsilon = 1$, il existe un rang N_1 tel que

$$\forall n, p \in [\![N_1, +\infty[\![, d_2(X^n, X^p) < 1$$

En particulier, l'état du système et le message ne changent plus après le rang N_1, notons $\tilde{x}$ et $\tilde{m}$ leurs valeurs respectives. Ils valent respectivement $\tilde{x}$ et $\tilde{m}$ (d'après la définition 71 introduisant la distance d_2).

De même, pour $\varepsilon = 10^{-k}$, il existe un rang N_k tel que

$$\forall n, p \in [\![N_k, +\infty[\![, d_2(X^n, X^p) < 10^{-k}$$

Donc, par définition de la distance d_2, puisque $\varepsilon < 1$, alors, à partir du rang N_k, $d_{\mathbb{B}^{\mathbb{N}}}(x^n, x^p) = 0$ et $d_{\mathbb{B}^{\mathbb{P}}}(m^n, m^p) = 0$. À partir de ce même rang :

$$d_{\mathbb{S}_{\mathbb{N}}}((S_p)^n, (S_p)^p) + d_{\mathbb{S}_{\mathbb{P}}}((S_c)^n, (S_c)^p) + d_{\mathbb{S}_{\mathbb{P}}}((S_m)^n, (S_m)^p) < 10^{-k}$$

Donc :

$$\begin{cases} d_{\mathbb{S}_{\mathbb{N}}}((S_p)^n, (S_p)^p) < 10^{-k} \\ d_{\mathbb{S}_{\mathbb{P}}}((S_c)^n, (S_c)^p) < 10^{-k} \\ d_{\mathbb{S}_{\mathbb{P}}}((S_m)^n, (S_m)^p) < 10^{-k} \end{cases}$$

Par conséquent, à partir du rang N_k, par définition des distances $d_{\mathbb{S}_{\mathbb{N}}}$ et $d_{\mathbb{S}_{\mathbb{P}}}$, les stratégies de $(S_p)^n$ et $(S_p)^p$ ont les k mêmes premiers termes. Il en va de même pour les stratégies $(S_c)^n$ et $(S_c)^p$, et pour les stratégies $(S_m)^n$ et $(S_m)^p$.

On construit alors, par un procédé diagonal, les trois stratégies $\widetilde{S_p}$, $\widetilde{S_c}$ et $\widetilde{S_m}$ de la manière suivante :

$$\begin{cases} \left(\widetilde{S_p}\right)^k = \left((S_p)^{N_k}\right)^k \\ \left(\widetilde{S_c}\right)^k = \left((S_c)^{N_k}\right)^k \\ \left(\widetilde{S_m}\right)^k = \left((S_c)^{N_k}\right)^k \end{cases}$$

Alors X^n converge vers $(\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m})$.

En effet, soit $\varepsilon > 0$, alors il existe $N = -\log_{10}(\varepsilon), n \geqslant N \Rightarrow d(X, (\widetilde{S_p}, \widetilde{x}, \widetilde{S_c}, \widetilde{m}, \widetilde{S_m}) < \varepsilon.$ □

8.4.8.2/ ÉTUDE DE LA PERFECTION

La définition d'un espace métrique parfait est donnée à la section 4.7.

Nous pouvons alors montrer que :

Proposition 30 : Perfection de $\mathcal{G}_f$

$(\mathcal{X}_2, \mathcal{G}_f)$ est un système dynamique discret parfait.

Démonstration. Soit $X = (S_p, x, S_c, m, S_m) \in \mathcal{X}_2$, et $r > 0$ Alors l'ensemble

$$\left\{(\widetilde{S_p}, x, \widetilde{S_c}, m, \widetilde{S_m}) \in \mathcal{X}_2 \ / \ \forall k \leqslant -\log_{10}(r) + 1, \widetilde{S_p}^k = (S_p)^k \text{ et } \widetilde{S_c}^k = (S_c)^k \text{ et } \widetilde{S_m}^k = (S_m)^k\right\}$$

est infini et inclus dans $B(X, r)$ (boule ouverte de l'espace métrique $\mathcal{X}_2$ de centre X et de rayon r qui peut). Donc X est un point d'accumulation. D'où le résultat. □

8.5/ AUTRE MODÈLE MATHÉMATIQUE POUR LE $\mathcal{CIS}_2$

Dans la présente section, nous introduisons une seconde formalisation du procédé de dissimulation d'informations $\mathcal{CIS}_2$. Nous n'avons pas eu le temps d'exploiter plus avant ce nouveau cadre formel, mais il nous semble riche de conséquences. Cette formalisation fait intervenir des suites de parties d'ensembles entiers, plutôt que des suites d'entiers.

L'idée consiste à considérer un vecteur contenant à la fois les LSCs du support hôte et le message à cacher, comme suit :

$$\mathscr{X} = \begin{pmatrix} x_0 \\ \vdots \\ x_{\mathsf{N}-1} \\ m_0 \\ \vdots \\ m_{\mathsf{P}-1} \end{pmatrix}.$$

Le processus modifie alors plusieurs coordonnées de ce vecteur, ce qui nous a conduit à introduire la notion de stratégie généralisée, rappelée ci-dessous :

Définition 72 : Stratégie chaotique généralisée

Soit $\mathsf{N} \in \mathbb{N}$. On appelle stratégie chaotique généralisée sur l'ensemble $[\![0; \mathsf{N}-1]\!]$ toute suite de parties de $[\![0; \mathsf{N}-1]\!]$.

Définition 73 : Ensemble des stratégie chaotique généralisées

Soit $\mathsf{N} \in \mathbb{N}$. L'ensemble des stratégies chaotiques généralisées sur $[\![0; \mathsf{N}-1]\!]$ est noté $\mathbb{S}^G_{[\![0;\mathsf{N}-1]\!]}$, et donc :

$$\mathbb{S}^G_{[\![0;\mathsf{N}-1]\!]} = (\mathscr{P}([\![0; \mathsf{N}-1]\!]))^{\mathbb{N}}$$

Les stratégies généralisées pour la deuxième modélisation du processus $\mathcal{CIS}_2$ sont alors :

$$(S'_n)_{n\in\mathbb{N}} = \{0; 1; 2; \ldots; \mathsf{N}-1; (S_m)^n + \mathsf{N}\}$$

où $(S_m)_{n\in\mathbb{N}}$ est la stratégie de mélange introduite dans les notations 5. Enfin, la fonction d'itération pour la deuxième modélisation du processus $\mathcal{CIS}_2$ se définit de la manière suivante :

$$\mathscr{F}_{(S_p,S_c)}(X,k) = \left(\begin{pmatrix} \overline{\delta_{(S_p^k)0}} x_0 + \delta_{(S_p^k)0} \sum_{i=0}^{\mathsf{P}} \left(\delta_{(S_c^k)i} m_i\right) \\ \vdots \\ \overline{\delta_{(S_p^k)j}} x_j + \delta_{(S_p^k)j} \sum_{i=0}^{\mathsf{P}} \left(\delta_{(S_c^k)i} m_i\right) \\ \vdots \\ \overline{\delta_{(S_p^k)(\mathsf{N}-1)}} x_{(\mathsf{N}-1)} + \delta_{(S_p^k)j} \sum_{i=0}^{\mathsf{P}} \left(\delta_{(S_c^k)i} m_i\right) \\ \overline{m_0} \\ \vdots \\ \overline{m_j} \\ \vdots \\ \overline{m_{\mathsf{P}-1}} \end{pmatrix} ; k+1 \right)$$

où $(S_p)_{n\in\mathbb{N}}$ et $(S_c)_{n\in\mathbb{N}}$ sont respectivement les stratégies de placement et de choix introduites dans les notations 5 pour le processus $\mathcal{CIS}_2$.

8.6/ MISE EN ŒUVRE PRATIQUE DE $\mathcal{CIS}_2$

Dans ce chapitre de contributions, nous avons présenté, de manière formelle, notre premier système de dissimulation d'informations. Il est topologiquement-sûr et stégo-sécurisé. Il est donc en mesure de résister à des attaques de types KMA, CMA et KOA (*cf.* chapitre 5 section 5.1.2). Ces résultats ont été obtenus après avoir étudié le comportement topologique de ce schéma de dissimulation d'informations. Cet algorithme $\mathcal{CIS}_2$ se trouve être le troisième système de dissimulation d'informations prouvé sûr.

À présent, si on analyse précisément l'algorithme $\mathcal{CIS}_2$ dans une perspective d'application concrète, on constate que celui-ci présente certaines limitations qui se traduisent de la manière suivante :

- Du fait du nombre de stratégies considérées, il est nécessaire de réaliser trois opérations (*cf.* notation 5) à chaque itération du système.
- Il existe certaines clés d'embarquement qui ne permettent pas d'extraire le message original de façon intègre à partir du contenu stéganographique.

Apportons tout d'abord quelques précisions sur le premier point qui pourrait, à première vue, apparaître comme une limitation de l'algorithme. Cela n'est finalement pas vraiment le cas. En effet, compte-tenu de la puissance actuelle des ordinateurs, le fait que l'algorithme implique 3 opérations pour chaque itération n'est finalement pas très gênant, ni

limitatif en terme de rapidité d'exécution du processus. Les quelques cas qui requièrent une rapidité d'exécution extrême sont finalement assez rares. Le cas le plus représentatif serait sans doute celui de l'informatique embarquée et des micro-contrôleurs [89]. Cependant, aujourd'hui, même les micro-contrôleurs ont eux aussi une vitesse d'exécution suffisamment rapide pour exécuter un tel algorithme. Une étude précise de la complexité algorithmique de $\mathcal{CIS}_2$ permettrait, très certainement, de quantifier précisément les choses à ce niveau.

Intéressons-nous, à présent, au second point relatif à l'existence de certaines clés d'embarquement qui ne permettent pas d'extraire le message original de façon intègre à partir du contenu stéganographique. Notons que cette extraction n'est pas toujours nécessaire dans le contexte du tatouage numérique, il l'est, en revanche, dans le contexte de la stéganographie. Les sections qui suivent exposent la solution que nous avons proposée pour lever cette limitation afin que l'algorithme $\mathcal{CIS}_2$ puisse être utilisé en pratique pour des applications concrètes.

Pour y parvenir, nous avons identifié des contraintes applicatives supplémentaires pour que l'algorithme $\mathcal{CIS}_2$ puisse être utilisé en pratique. Ces nouvelles contraintes permettent de retrouver le message original de façon intègre suite à la phase d'extraction du message à partir du contenu stéganographique.

Dans les sections qui suivent, nous étudions précisément ces contraintes supplémentaires.

8.6.1/ CONTRAINTES APPLICATIVES POUR LE PROCESSUS $\mathcal{CIS}_2$

En reprenant le processus, tel qu'il a été présenté dans la section 8.1, le contenu stéganographique est le vecteur booléen $y = x^l \in \mathbb{B}^{\mathsf{N}}$. Ce contenu stéganographique permettra d'extraire le message original de façon intègre à condition que les contraintes suivantes soient appliquées :

1. Le nombre d'itérations l est suffisamment grand (voir détails ci-dessous).

2. Soit $\mathfrak{I}(S_p)$ l'ensemble $\{S_p^1, S_p^2, \ldots, S_p^l\}$ de cardinalité k, $k \leq l$ (les répétitions sont enlevées dans un ensemble). Cet ensemble contient tous les éléments de x qui ont été modifiés au cours du processus d'itération. Considérons $\mathfrak{I}(S_c)_{|D}$ défini par $\{S_c^{d_1}, S_c^{d_2}, \ldots, S_c^{d_k}\}$ où d_i est la dernière itération qui a modifié l'élément $i \in \mathfrak{I}(S_p)$. Nous avons besoin que cet ensemble soit égal à $[\![0; \mathsf{P}-1]\!]$.

Nous allons faire quelques remarques sur les contraintes indiquées ci-dessus. La première implique que le nombre d'itérations soit supérieur à un seuil donné. Cette exigence a des raisons pratiques et théoriques. En théorie, la possibilité de passer avec succès les tests statistiques est directement liée à ce nombre d'itérations. Mais, dans la pratique, cette valeur est limitée par la taille du contenu de l'hôte.

La deuxième contrainte, quant à elle, aborde l'exhaustivité et l'exactitude de la méthode, comme détaillé dans la section suivante.

8.6.2/ ÉTUDE D'EXACTITUDE ET EXHAUSTIVITÉ

Sans altération (attaque) du contenu stéganographique, le schéma de dissimulation d'informations doit veiller à ce que l'utilisateur puisse, d'une part, toujours extraire un message à partir du contenu stéganographique dès lors que l'utilisateur possède la bonne clé d'extraction. L'algorithme doit également, d'autre part, veiller à ce que le message extrait (filigrane) soit intègre par rapport à sa version originale embarquée initialement dans le contenu hôte. Ces deux exigences correspondent respectivement à l'étude de l'exhaustivité (complétude algorithmique) et de l'exactitude de l'approche proposée. Pour réaliser cette étude, nous montrons tout d'abord l'évaluation suivante :

Proposition 31 : Condition nécessaire et suffisante pour l'extraction

Dans la section 8.6.1, le point 2 est une condition nécessaire et suffisante pour permettre l'extraction du message à partir du média d'embarquement.

Démonstration. Pour ce qui est de la condition suffisante, soit d_i la dernière itération (date) pour laquelle l'élément $i \in \Im(S_p)$ de x a été modifié :

$$d_i = \max\{j | S_p^j = i\}.$$

Soit $D = \{d_i | i \in \Im(S_p)\}$. L'ensemble $\Im(S_c)_{|D}$ est donc la restriction de l'image de S_c à D.

L'hôte, qui résulte de ce schéma d'itération, est donc $(x_0^l, \ldots, x_{N-1}^l)$ où x_i^l est soit $x_i^{d_i}$ si i appartient à $\Im(S_p)$, soit x_i^0 dans le cas contraire. De plus, pour chaque $i \in \Im(S_p)$, l'élément $x_i^{d_i}$ est égal à $m_{S_c^{d_i}}^{d_i-1}$. Grâce à la contrainte du point 2, tous les indices $j \in [\![0; P-1]\!]$ appartiennent à $\Im(S_c)_{|D}$. Soit alors $j \in [\![0; P-1]\!]$ tel que $S_c^{d_i} = j$. Nous avons donc tous les éléments m_j du vecteur m. Considérons maintenant certains $m_j^{d_i-1}$. Alors les valeurs de m_j^0 peuvent être immédiatement déduites par calcul, dans S_c, en comptabilisant le nombre de fois où la composante j a été commutée avant $d_i - 1$.

Il importe d'étudier également la condition nécessaire. Si $\Im(S_c)_{|D} \subsetneq [\![0; P-1]\!]$, alors il existe $j \in [\![0; P-1]\!]$ qui n'appartient pas à $\Im(S_c)_{|\Im(S_p)}$. Donc m_j n'est pas présent dans x^l et le message ne peut pas être extrait. □

Lorsque la contrainte, décrite au point 2, est satisfaite, nous obtenons un système qui trouve toujours le message d'origine, à condition que le média tatoué n'ait pas été modifié (altéré ou attaqué).

Dans ce contexte, l'exactitude et l'exhaustivité sont établies.

Grâce à la contrainte, décrite au point 2, la cardinalité k de $\Im(S_p)$ est plus grande que P. Par ailleurs, la cardinalité de D serait plus petite que P et serait similaire à la cardinalité de $\Im(S_c)_{|D}$, ce qui est contradictoire.

Un bit d'indice j du message original m^0 est donc embarqué au moins deux fois dans x^l. En comptant le nombre de fois où ce bit a été échangé dans S_m, la valeur de m_j peut être déduite à de nombreux endroits. Sans altération du média stéganographique (attaque), toutes ces valeurs sont égales et le message est immédiatement obtenu. Après une attaque, la valeur de m_j est obtenue comme valeur moyenne de toutes ses occurrences. Le schéma de dissimulation d'informations est donc complet. Remarquons que si le média de couverture n'est pas attaqué, le message retourné est toujours égal au message

original compte-tenu de la définition de la fonction de signification (*cf.* section 6.3.1 du chapitre 6).

8.6.3/ IMPLÉMENTATION CONCRÈTE DE $\mathcal{CIS}_2$

Grâce aux études menées dans cette section, nous avons pu proposer une implémentation concrète de l'algorithme $\mathcal{CIS}_2$, sur la base d'une adaptation de la méthode, tenant compte des contraintes applicatives exposées à la section 8.6.1. Cette implémentation a fait l'objet d'un dépôt logiciel [61] auprès de l'Agence pour la Protection des Programmes [65]. Nous avons baptisé ce programme $\mathcal{CIS}_5$.

Précisons que ce programme sera utilisé, avec d'autres, dans le cadre du développement de la *plateforme de protection des documents numériques* prévue dans le projet de création de la future société *Stégosécu-ISIS* [44].

CHAPITRE 9

Exposant de Lyapunov du $\mathcal{CIS}_2$

> Ne me dites pas que ce problème est difficile. S'il n'était pas difficile, ce ne serait pas un problème.
>
> MARÉCHAL FOCH, OFFICIER ET ACADÉMICIEN (1851-1929)

Le schéma de stéganographie $\mathcal{CIS}_2$ appartient à la catégorie restreinte des algorithmes étant à la fois stégo-sûrs et topologiquement sécurisés. Du fait de sa stégo-sécurité, ce schéma est capable de faire face à un adversaire dans la configuration de l'attaque du filigrane seul *WOA*. Sa sécurité topologique renforce sa capacité à résister aux autres catégories d'attaques, issues du problème du prisonnier de Simmons, telles que le *KMA* ou le *KOA*. Dans ce chapitre, nous approfondissons l'étude des propriétés topologiques du $\mathcal{CIS}_2$, via une description de ce schéma comme itérations sur la droite des réels, et par l'investigation de son exposant de Lyapunov. Les résultats, publiés dans [8], laissent à penser que ce processus est capable de décourager un attaquant malicieux dans la configuration "Attaque de l'Original Estimé" (EOA) [79] (*cf.* chapitre 5 section 5.1.2).

9.1/ NOUVEAUX RAPPELS DE TOPOLOGIE

9.1.1/ L'EXPOSANT DE LYAPUNOV

Certains systèmes dynamiques sont très sensibles à de petites variations de leurs conditions initiales. Cette caractéristique est illustrée par des propriétés de nature topologique, telles que la constante de sensibilité aux conditions initiales ou l'expansivité, précédemment et respectivement introduites dans les définitions 38 et 37. Cependant, ces variations peuvent rapidement prendre des proportions gigantesques, car elles peuvent grandir de façon exponentielle, et aucune constante introduite jusqu'ici ne peut quantifier ce fait. Alexander Lyapunov a examiné ce phénomène et a introduit un expo-

sant [57, 55, 8, 40] qui mesure la vitesse à laquelle ces petites variations peuvent croître.

Définition 74 : Exposant de Lyapunov

Considérons une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$. *L'exposant de Lyapunov* du système composé par :

$$\begin{cases} x^0 \in \mathbb{R} \\ x^{n+1} = f(x^n) \end{cases}$$

est défini par :

$$\lambda(x_0) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left| f' \left(x^{i-1} \right) \right|.$$

Considérons un système dynamique avec une erreur infinitésimale portant sur la condition initiale x_0. Quand l'exposant de Lyapunov est positif, cette erreur augmentera exponentiellement (situation de chaos), tandis qu'elle diminuera si $\lambda(x_0) \leqslant 0$.

Remarque 10 :

L'exposant de Lyapunov vient compléter la liste des propriétés quantitatives définies en état de l'art à la section 4.8.

9.1.2/ EXPOSANT DE LYAPUNOV ET DISSIMULATION D'INFORMATIONS

Notre équipe a précédemment établi que, dans le but d'élargir la connaissance que l'on peut avoir du niveau de sécurité d'un schéma de dissimulation d'informations, la quantité de désordre générée par le schéma peut aussi être mesurée en évaluant l'exposant de Lyapunov [57, 55, 8, 40]. Cette évaluation permet en effet de caractériser la capacité du schéma à faire face à un attaquant dans le contexte EOA [79] (*cf.* chapitre 5 section 5.1.2).

En effet, dans une configuration de type *EOA*, l'attaquant a uniquement accès a une estimation des contenus originaux. Avec cette seule connaissance, l'attaquant ne devrait pas être en mesure de retrouver quelque information que ce soit à propos du message caché ou de la clé secrète. Les deux autres notions de sensibilité et d'expansivité issues de la sécurité topologique du processus sont pertinentes pour faire face à des attaques dans cette configuration. Cependant, ces deux propriétés topologiques ne donnent pas de quantification précise de la sensibilité du schéma, ce qui justifie la considération de l'exposant de Lyapunov dans ce qui suit.

Avant de pouvoir évaluer cet exposant, il nous faut introduire au préalable la notion de semi-conjugaison topologique.

9.1.3/ LA SEMI-CONJUGUAISON TOPOLOGIQUE

Parfois, au lieu d'essayer de prouver des propriétés directement sur le système lui même, il est préférable de réduire le problème initial à un autre problème dont les caractéristiques sont connues ou apparaîssent plus accessibles. Un tel principe de réduction, dans la théorique mathématique du chaos, est une semi-conjugaison.

Définition 75 : Semi-conjugaison topologique

Le système dynamique discret $(\mathcal{X}, f)$ est *topologiquement semi-conjugué* au système $(\mathcal{Y}, g)$ s'il existe une fonction $\varphi : \mathcal{X} \longrightarrow \mathcal{Y}$, à la fois continue et surjective, telle que :

$$\varphi \circ f = g \circ \varphi,$$

ce qui rend commutatif le diagramme suivant [34].

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{f} & \mathcal{X} \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ \mathcal{Y} & \xrightarrow[g]{} & \mathcal{Y} \end{array}$$

Dans ce cas, le système $(\mathcal{Y}, g)$ est appelé un *facteur* du système $(\mathcal{X}, f)$.

Comme l'a montré Enrico Formenti, différents comportements dynamiques sont hérités par les facteurs de systèmes [34]. Ils sont résumés dans la proposition suivante :

Proposition 32 : Héritage des propriétés pour les facteurs des semi-conjugaisons

Soit $(\mathcal{Y}, g)$ un facteur du système $(\mathcal{X}, f)$. Alors :

1. pour tout $j \leqslant k$, $p \in Per_k(f) \Longrightarrow \varphi(p) \in Per_j(g)$, où $Per_n(h)$ représente l'ensemble des points périodiques de période n pour la fonction d'itération h.
2. $(\mathcal{X}, f)$ régulière $\Longrightarrow (\mathcal{Y}, g)$ régulière,
3. $(\mathcal{X}, f)$ transitive $\Longrightarrow (\mathcal{Y}, g)$ transitive.

Donc si $(\mathcal{X}, f)$ est chaotique telle que l'a défini Devaney, alors le système $(\mathcal{Y}, g)$ est lui aussi chaotique.

9.2/ UNE SEMI-CONJUGAISON TOPOLOGIQUE ENTRE $\mathcal{X}_2$ ET $\mathbb{R}$

9.2.1/ PRÉLIMINAIRE

On présente dans cette section une semi-conjugaison topologique entre notre processus $\mathcal{CIS}_2$ œuvrant sur $\mathcal{X}_2$ et son équivalent sur $\mathbb{R}$. Les raisons d'être de cette semi-conjugaison sont multiples :

- En avoir une vision plus simple.
- Permettre de mieux comprendre leur dynamique, notamment en les étudiant avec les outils de l'analyse classique : continuité, dérivabilité, tracés de courbes, etc.
- Rendre possible le calcul de l'exposant de Lyapunov, réalisé à la section 9.3, car il nous faut une fonction dérivable pour cela.
- Faciliter la comparaison théorique de notre processus $\mathcal{CIS}_2$ avec d'autres systèmes dynamiques, usuellement définis sur $\mathbb{R}$. Cette comparaison se fera principalement sous l'angle du désordre topologique, en tirant les conclusions du chapitre intitulé "De la relativité du chaos " dans la thèse de C.Guyeux [41].

– Faciliter la comparaison pratique de la sécurité de nos algorithmes avec ceux déjà existants.

En utilisant la semi-conjugaison topologique, nous démontrons que $\mathcal{CIS}_2$ modélisé par $\mathcal{G}_{f_0}$ sur $\mathcal{X}_2$ peut être décrit par des itérations sur un intervalle réel. Comme nos recherches sont inspirées par le travail de [8], les preuves détaillées dans ce chapitre suivent le même canevas.

Pour ce faire, nous devons introduire plusieurs notations et terminologies.

9.2.2/ L'ESPACE DES PHASES EST UN INTERVALLE RÉEL

Soit $\mathcal{X}_{(\mathsf{N};\mathsf{P})} = \mathbb{S}_N \times \mathbb{B}^{\mathsf{N}} \times \mathbb{S}_P \times \mathbb{B}^{\mathsf{P}} \times \mathbb{S}_P$.

(Pour des raisons de meilleure compréhension, nous allons supposer que $\mathsf{N} = 3$ et $\mathsf{P} = 2$. Ainsi $\mathsf{N} + \mathsf{P} = 5$ et $\mathsf{NP}^2 = 12$. Cependant, une formulation équivalente de ce qui suit peut aisément être obtenue en remplaçant les bases de numération 5 et 12 par n'importe quelle autre base $(\mathsf{N} + \mathsf{P})$ et (NP^2). N doit simplement être plus grand que P.)

Définition 76 : Fonction ψ pour la semi-conjugaison topologique

La fonction $\psi : [\![1, \mathsf{N}]\!] \times [\![1, \mathsf{P}]\!] \times [\![1, \mathsf{P}]\!] \to [\![0, \mathsf{NP}^2 - 1]\!]$ est définie de la manière suivante : $\psi\left(S_p^i, S_c^i, S_m^i\right) = (S_p^i - 1)\mathsf{P}^2 + (S_c^i - 1)\mathsf{P} + (S_m^i - 1)$.

Cette fonction a pour but de convertir un triplet de stratégies en une simple stratégie d'entiers, exprimée dans une base de numération différente. Évidemment, ψ est une fonction bijective, l'opération inverse étant notée ψ^{-1}. Les trois projections de ψ^{-1} sont désignées par :

– $\psi_1^{-1}\left(\psi\left(S_p^i, S_c^i, S_m^i\right)\right) = S_p^i$,

– $\psi_2^{-1}\left(\psi\left(S_p^i, S_c^i, S_m^i\right)\right) = S_c^i$,

– $\psi_3^{-1}\left(\psi\left(S_p^i, S_c^i, S_m^i\right)\right) = S_m^i$.

Le tableau 9.1 précise les valeurs de la fonction ψ dans le cadre de l'exemple retenu.

Nous sommes dorénavant en mesure d'introduire notre semi-conjugaison.

Base N = 3	Base P = 2	Base P = 2	Base $NP^2 = 12$
S_p^i	S_c^i	S_m^i	$\psi\left(S_p^i, S_c^i, S_m^i\right)$
1	1	1	0
1	1	2	1
1	2	1	2
1	2	2	3
2	1	1	4
2	1	2	5
2	2	1	6
2	2	2	7
3	1	1	8
3	1	2	9
3	2	1	10
3	2	2	11

TABLE 9.1 – Illustration de la fonction ψ (voir la définition 76).

Définition 77 : Fonction φ pour la semi-conjugaison topologique

Nous définissons

$$\varphi : \quad \mathcal{X}_{(3;2)} = \mathbb{S}_3 \times \mathbb{B}^3 \times \mathbb{S}_2 \times \mathbb{B}^2 \times \mathbb{S}_2 \quad \longrightarrow \quad [0, 2^5[$$

$$\left(S_p, E, S_c, M, S_m\right) \quad \longmapsto \quad \varphi\left(S_p, E, S_c, M, S_m\right),$$

de la manière suivante : Si $\left(S_p, E, S_c, M, S_m\right) =$
$\left((S_p^0, S_p^1, \ldots); (E_0, E_1, E_2, E_3); (S_c^0, S_c^1, \ldots); (M_0, M_1); (S_m^0, S_m^1, \ldots)\right)$,
alors $\varphi\left(S_p, E, S_c, M, S_m\right)$ est le nombre réel :

- dont la partie entière est $\sum_{k=0}^{2} 2^{4-k} E_k + \sum_{k=3}^{4} 2^{4-k} M_{k-3}$, à savoir, les chiffres binaires de e sont $E_0\ E_1\ E_2\ M_0\ M_1$.
- dont la partie décimale est égale à :
  $0, \psi\left(S_p^0, S_c^0, S_m^0\right)\ \psi\left(S_p^1, S_c^1, S_m^1\right)\ \psi\left(S_p^2, S_c^2, S_m^2\right)\ \ldots = \sum_{k=1}^{+\infty} 12^{-k} S^{k-1}$.
  (Cette partie décimale est donc exprimée en base 12.)

On remarque que φ réalise une association entre un point de $\mathcal{X}_{(3;2)}$ et un nombre réel de l'intervalle $[0, 2^5[$. Nous devons à présent traduire, sur cet intervalle réel, le processus de dissimulation d'informations $C\mathcal{IS}_2$, qui est représenté par la fonction $\mathcal{G}_{f_0}$. Pour ce faire, nous introduisons deux fonctions intermédiaires sur $[0, 2^5[$. Elles seront notées e et s.

Définition 78 : Fonctions e et s pour la semi-conjugaison topologique

Soit $x \in [0, 2^5[$ et :

- $e_0, \ldots, e_4$ les chiffres binaires de la partie entière de x : $\lfloor x \rfloor = \sum_{k=0}^{4} 2^{4-k} e_k$.
- $(s^k)_{k\in\mathbb{N}}$ les chiffres de x, exprimés en base 12, où la décomposition décimale choisie pour x est celle qui n'a pas un nombre infini de chiffres 11 : $x = \lfloor x \rfloor + \sum_{k=0}^{+\infty} s^k 12^{-k-1}$.

e et s sont alors définies de la manière suivante :

$$\begin{array}{rccc} e : & [0,2^5[& \longrightarrow & \mathbb{B}^3 \times \mathbb{B}^2 \\ & x & \longmapsto & ((e_0,e_1,e_2);(e_3,e_4)) \end{array}$$

et

$$\begin{array}{rccc} s : & [0,2^5[& \longrightarrow & [\![0,11]\!]^{\mathbb{N}}. \\ & x & \longmapsto & (s^k)_{k\in\mathbb{N}} \end{array}$$

Nous sommes maintenant en mesure de pouvoir définir la fonction g, dont le but est de traduire le processus de dissimulation d'informations $\mathcal{CIS}_2$ représenté par $\mathcal{G}_{f_0}$ sur un intervalle de $\mathbb{R}$.

Définition 79 : Fonction g pour la semi-conjugaison topologique

$g : [0, 2^5[\longrightarrow [0, 2^5[$ est telle que $g(x)$ est un nombre réel de $[0, 2^5[$ défini de la manière suivante :

- sa partie entière a une décomposition binaire égale à $e'_0, \ldots, e'_4$, avec $\forall i \in [\![0, 2]\!]$:

$$e'_i = \begin{cases} e(x)_i & \text{if } i \neq \psi_1^{-1}\left(s^0\right) \\ e(x)_{2+\psi_2^{-1}(s^0)} & \text{if } i = \psi_1^{-1}\left(s^0\right) \end{cases}$$

  et $\forall i \in [\![3, 4]\!]$:

$$e'_i = \begin{cases} e(x)_i & \text{if } i \neq \psi_3^{-1}\left(s^0\right) \\ e(x)_i + 1 \pmod 2 & \text{if } i = \psi_3^{-1}\left(s^0\right), \end{cases}$$

- sa partie décimale est quant à elle égale à $s(x)^1, s(x)^2, \ldots$

En d'autres mots, si $x = \sum_{k=0}^{4} 2^{4-k} e_k + \sum_{k=0}^{+\infty} s^k \, 12^{-k-1}$, alors :

$$g(x) = \sum_{k=0}^{2} 2^{4-k} \left[e_k \left(\delta(k, \psi_1^{-1}(s^0)) + 1 \pmod 2 \right) + e_{2+\psi_2^{-1}(s^0)} \left(\delta(k, \psi_1^{-1}(s^0)) \right) \right]$$

$$+ \sum_{k=3}^{4} 2^{4-k} (e_k + \delta(k, \psi_3^{-1}(s^0) \pmod 2)) + \sum_{k=0}^{+\infty} s^{k+1} 12^{-k-1},$$

où δ est la mesure discrète booléenne.

De nombreuses distances peuvent être définies sur l'ensemble $[0, 2^5[$, la plus utilisée étant la distance euclidienne $\Delta(x, y) = |y - x|^2$. Cette distance euclidienne ne reproduit

pas exactement la notion de proximité introduite par notre première distance d_2 sur $\mathcal{X}_2$. En effet, d_2 est plus riche que Δ. C'est la raison pour laquelle nous devons introduire la distance suivante.

Définition 80 : Distance sur $\mathbb{R}$ pour la semi-conjugaison topologique

Soient $x, y \in [0, 2^5[$, D représente la fonction de $[0, 2^5[^2$ dans $\mathbb{R}^+$ définie par : $D(x, y) = D_e\left(e(x), e(y)\right) + D_s\left(s(x), s(y)\right)$, où :

$$D_e(e, \check{e}) = \sum_{k=0}^{4} \delta(e_k, \check{e}_k), \quad \text{et} \quad D_s(s, \check{s}) = \sum_{k=1}^{\infty} \frac{|s^k - \check{s}^k|}{12^k}.$$

Proposition 33 : D est une distance

D est une distance sur $[0, 2^5[$.

Démonstration. Les trois axiomes définissant une distance doivent être vérifiés.

- $D \geqslant 0$, car tous les éléments sont positifs dans sa définition. Si $D(x, y) = 0$, alors $D_e(x, y) = 0$, et ainsi la partie entière de x et y sont égales (elles ont la même décomposition binaire). De plus, $D_s(x, y) = 0$, donc $\forall k \in \mathbb{N}^*, s(x)^k = s(y)^k$. En d'autres termes, x et y ont le même k-ième chiffre décimal, $\forall k \in \mathbb{N}^*$. Et donc $x = y$.
- Évidemment, $\forall x, y, D(x, y) = D(y, x)$.
- Finalement, l'inégalité triangulaire est obtenue par le fait que les deux distances δ et $|x - y|$ la satisfont.

□

La convergence des suites, au sens de la distance D, n'est pas la même que la convergence usuelle relative à la distance euclidienne. Par exemple, si $x^n \to x$ au sens de la distance D, alors nécessairement la partie entière de chaque x^n est égale à la partie entière de x (au moins à partir d'un certain rang) et la partie décimale de x^n correspond à celle de x "aussi loin que nécessaire". Pour illustrer ce fait, une comparaison entre D et la distance euclidienne est donnée figure 9.1.

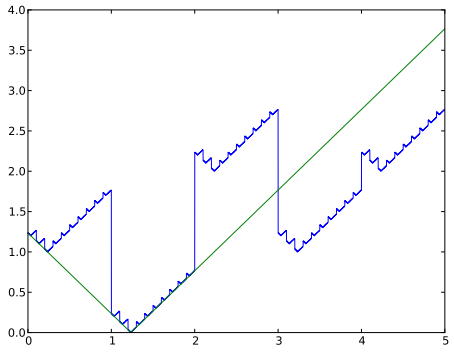
Il est maintenant possible de définir une semi-conjugaison topologique entre $\mathcal{X}_2$ et un intervalle de $\mathbb{R}$. En effet, φ a été construite afin d'être continue et surjective, et donc nous obtenons le théorème suivant.

Théorème 11 : Semi-conjugaison topologique entre $\mathcal{X}_2$ et $[0, 2^5[$

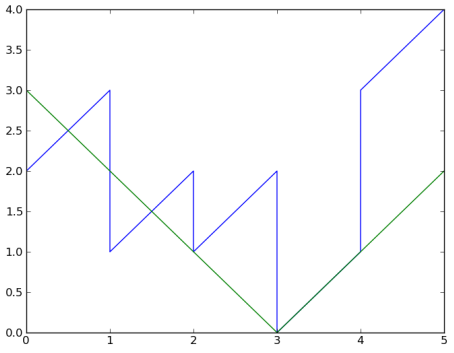
Le processus de dissimulation d'informations $\mathcal{CIS}_2$ représenté par $\left(\mathcal{G}_{f_0}, \mathcal{X}_2\right)$ peut être considéré comme de simples itérations sur $\mathbb{R}$, ce qui est illustré par la semi-conjugaison donnée ci-après :

$$\begin{array}{ccc} \left(\mathcal{X}_{(3;2)}, d_2\right) & \xrightarrow{\mathcal{G}_{f_0}} & \left(\mathcal{X}_{(3;2)}, d_2\right) \\ \varphi \downarrow & & \downarrow \varphi \\ \left([0, 2^5[, D\right) & \xrightarrow[g]{} & \left([0, 2^5[, D\right) \end{array}$$

En d'autres mots, $\mathcal{X}_2$ est, à peu de choses près, égal à $[0, 2^{N+P}[$.



(a) Fonction $x \to dist(x; 1,234)$ sur l'intervalle $(0; 5)$.



(b) Fonction $x \to dist(x; 3)$ sur l'intervalle $(0; 5)$.

FIGURE 9.1 – Comparaison entre D (en bleu) et la distance euclidienne (en vert).

9.2.3/ ÉTUDE DE g

Nous allons à présent étudier le processus $\mathcal{CIS}_2$ décrit par la fonction réelle g.

Commençons par rappeler une illustration du comportement des itérations chaotiques (figures 9.2, 9.3 et 9.4, ces dernières étant tirées de [41].) Elle permet de donner quelques indications sur la manière dont le nouveau schéma $\mathcal{CIS}_2$ peut se comporter, sachant que ce schéma est lui même fondé sur le principe des itérations chaotiques.

Nous pouvons en déduire que la fonction g est une fonction linéaire par morceaux, ce que prouve le résultat suivant.

Proposition 34 : Dérivabilité et pente de g

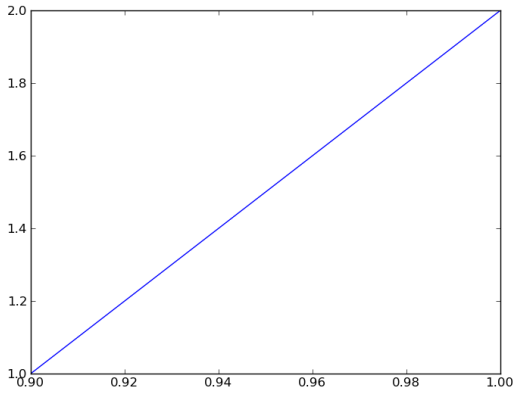
Le processus $\mathcal{CIS}_2$ représenté par la fonction g définie sur $\mathbb{R}$ possède des fonctions dérivées à tous les ordres sur l'intervalle $[0, 2^5[$, excepté en 385 points de l'ensemble I defini par :

$$I = \left\{ \frac{n}{12} \ / \ n \in [\![0; 2^5 \times 12]\!] \right\}.$$

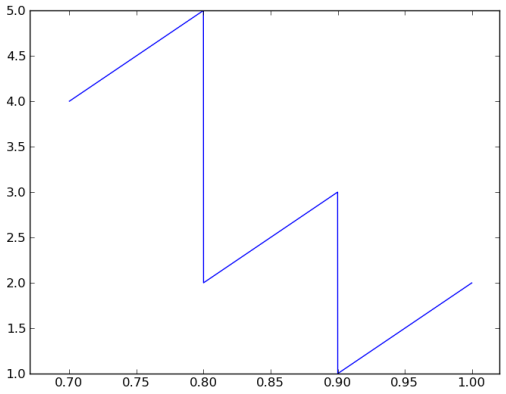
De plus, sur chaque intervalle de la forme $\left[\frac{n}{12}, \frac{n+1}{12}\right[$, avec $n \in [\![0; 2^5 \times 12[\![$, g est une fonction linéaire ayant une pente égale à 12 : $\forall x \notin I, g'(x) = 12$.

Démonstration. Soit $I_n = \left[\frac{n}{12}, \frac{n+1}{12}\right[$, avec $n \in [\![0; 2^5 \times 12]\!]$. Tous les points de I_n ont la même partie entière e et la même partie décimale s^0 : sur l'ensemble I_n, les fonctions $e(x)$ et $x \mapsto s(x)^0$ de la définition 78 ne dépendent que de n. Ainsi toutes les images $g(x)$ de ces points x :

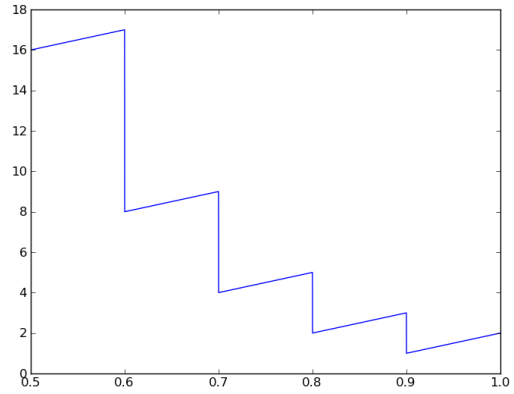
– ont la même partie entière, qui est e, excepté probablement le bit numéro s^0. En d'autres termes, cet entier a approximativement la même décomposition binaire que e, la seule exception étant le chiffre s^0 (ce chiffre est donc soit $e + 2^{12-s^0}$, soit $e - 2^{12-s^0}$, selon la parité de s^0, *i.e.* il vaut $e + (-1)^{s^0} \times 2^{12-s^0}$).



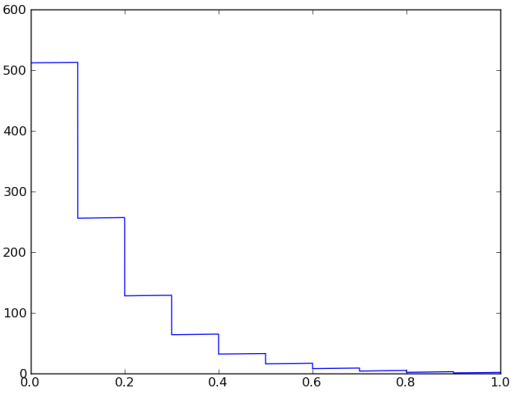
Itérations chaotiques sur l'intervalle $(0,9;1)$.



Itérations chaotiques sur l'intervalle $(0,7;1)$.

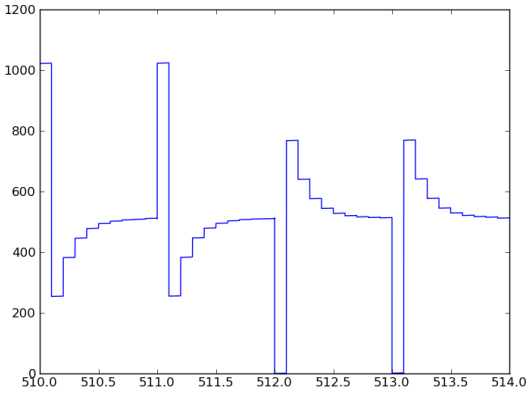


Itérations chaotiques sur l'intervalle $(0,5;1)$.

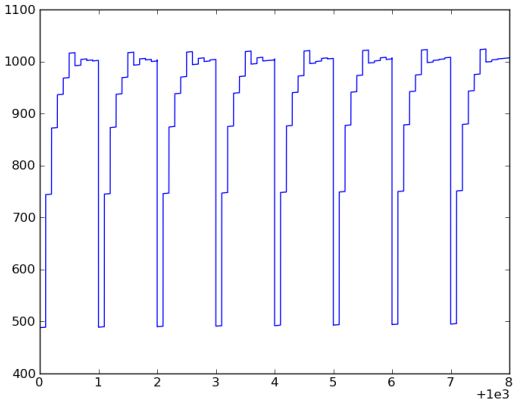


Itérations chaotiques sur l'intervalle $(0;1)$.

FIGURE 9.2 – Représentation des itérations chaotiques, fondement du processus $\mathcal{CIS}_2$.



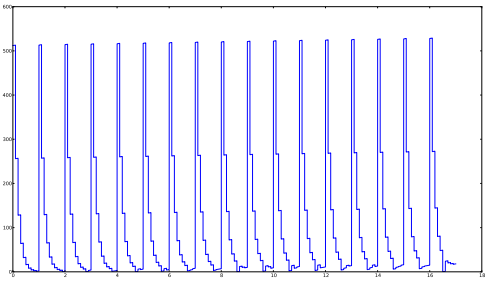
Itérations sur l'intervalle $(510;514)$.



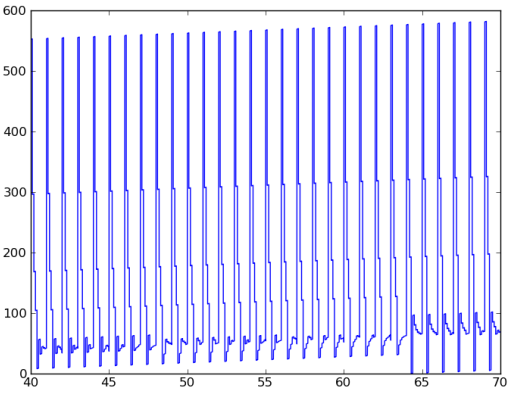
Itérations sur l'intervalle $(1000;1008)$.

FIGURE 9.3 – Itérations chaotiques sur un petit intervalle.

– Un décalage à gauche a été appliqué sur la partie décimale y, faisant disparaître ainsi le premier chiffre commun s^0. En d'autres mots, y a été changé en $12 \times y - s^0$.



Itérations chaotiques sur l'intervalle $(0; 16)$.



Itérations chaotiques sur l'intervalle $(40; 70)$.

FIGURE 9.4 – Aspect général des itérations chaotiques.

En résumé, l'action de g sur les points de I se manifeste de la manière suivante : d'abord, par une multiplication par 12, et ensuite par l'addition de la même constante à chaque terme, ce qui nous donne : $\frac{1}{12}\left(e + (-1)^{s^0} \times 2^{12-s^0}\right) - s^0$. □

Remarque 11 : g est une fonction linéaire par morceaux

Finalement, la fonction g représentant le processus de dissimulation $\mathcal{CIS}_2$ est un élément d'une large famille de fonctions qui sont à la fois chaotiques et linéaires par morceaux (à l'instar de la fonction tente [82]).

9.2.4/ COMPARAISON DES DEUX DISTANCES SUR $[0, 2^5[$

Les deux propositions suivantes permettent de comparer nos deux distances sur $[0, 2^5[$:

Proposition 35 : Continuité pour D vis à vis de Δ

Id : $\left([0, 2^5[, \Delta \right) \to \left([0, 2^5[, D \right)$ n'est pas une fonction continue.

Démonstration. La suite $x^n = 1,999\ldots999$ constituée par n 9 comme partie décimale est telle que :

- $\Delta(x^n, 2) \to 0$.
- Mais $D(x^n, 2) \geqslant 1$, et donc $D(x^n, 2)$ ne converge pas vers 0.

La caractérisation séquentielle de la continuité permet de conclure la démonstration. □

À l'inverse :

Proposition 36 : Continuité pour Δ vis à vis de D

Id : $\left([0, 2^5[, D \right) \to \left([0, 2^5[, \Delta \right)$ est une fonction continue.

Démonstration. Si $D(x^n, x) \to 0$, alors $D_e(x^n, x) = 0$ au moins pour n supérieur à un certain rang donné, comme D_e ne retourne que des entiers. Ainsi, après ce rang donné, les parties entières de tous les x^n sont égales à la partie entière de x.

De plus, $D_s(x^n, x) \to 0$, alors $\forall k \in \mathbb{N}^*, \exists N_k \in \mathbb{N}, n \geqslant N_k \Rightarrow D_s(x^n, x) \leqslant 10^{-k}$. Cela signifie que pour tout k, un indice N_k peut être trouvé de telle sorte que $\forall n \geqslant N_k$, tous les x^n aient les mêmes k premiers chiffres, qui sont les chiffres de x. Nous pouvons en déduire la convergence suivante : $\Delta(x^n, x) \to 0$, et finalement le résultat. □

La conclusion de ces propositions est que la distance proposée est plus précise que la distance euclidienne. En d'autres termes :

Corollaire 1 : Finesse de D vis à vis de Δ

D est plus fine que la distance euclidienne Δ.

Ce corollaire peut être reformulé de la manière suivante.

- La topologie produite par Δ est un sous-ensemble de la topologie produite par D.
- D a plus d'espaces ouverts que Δ.
- Il est plus difficile de converger pour la topologie τ_D issue de D, que de converger pour celle issue de la distance Δ, qui est alors notée τ_Δ.

Évaluons maintenant plus en profondeur le désordre topologique de notre schéma de dissimulation d'informations, basé sur des itérations chaotiques, qui est dorénavant décrit par des itérations sur $\mathbb{R}$ de la fonction g introduite dans la définition 79.

9.3/ CHAOS DU $\mathcal{CIS}_2$ SUR $\mathbb{R}$

Nous avons précédemment rappelé que le processus $\mathcal{CIS}_2$ représenté par $\left(\mathcal{G}_{f_0}, \mathcal{X}_2\right)$ est chaotique au sens de la formulation de Devaney. Nous pouvons en déduire que le processus $\mathcal{CIS}_2$ est aussi chaotique sur $\mathbb{R}$, lorsque l'on considère la topologie de l'ordre, car :

- $\left(\mathcal{G}_{f_0}, \mathcal{X}_2\right)$ et $\left(g, [0, 2^5[_D\right)$ sont semi-conjugués par φ,
- Ensuite $\left(g, [0, 2^5[_D\right)$ est un système chaotique, au sens de Devaney, car la semi-conjugaison préserve cette caractéristique.
- Mais la topologie générée par D est plus fine que la topologie générée par la distance euclidienne Δ – qui est la topologie de l'ordre.
- D'après le théorème 2, nous pouvons donc en déduire que le processus $\mathcal{CIS}_2$ modélisé par g reste chaotique, au sens de Devaney, pour la topologie de l'ordre sur $\mathbb{R}$.

Ce résultat peut être formulé de la manière suivante.

Théorème 12 : $\mathcal{CIS}_2$ est chaotique sur $\mathbb{R}$ pour la topologie de l'ordre

Le processus $\mathcal{CIS}_2$ modélisé par g sur $\mathbb{R}$ est chaotique au sens de la formulation de Devaney, lorsque $\mathbb{R}$ est muni de sa topologie usuelle, qui est la topologie de l'ordre.

En fait, ce résultat est plus faible que le théorème établissant le chaos pour la topologie plus fine induite par la distance d_2. Cependant, le théorème 12 est quand même, dans une certaine mesure, important.

En effet, avant les travaux de notre équipe, les études de chaos des algorithmes se faisaient sur la droite réelle. En suivant une démarche nouvelle, notre équipe a contourné le problème des erreurs d'arrondis, donc des pertes éventuelles de chaos, en étudiant le schéma $\mathcal{CIS}_2$ sur un ensemble spécifique ($\mathcal{X}_2$ au lieu de $\mathbb{R}$). Cette démarche autre avait pour objectif de rester le plus proche possible du mode de fonctionnement des ordinateurs : manipulation exclusive d'entiers bornés. Les propriétés de désordre, prouvées théoriquement, seront ainsi préservées lors de l'exécution de calculs sur ordinateur. Les erreurs d'arrondis ne poseront alors plus problème.

Toutefois, on peut se demander si cette nouvelle démarche ne conduit pas à un désordre de moins bonne qualité, car la topologie utilisée lors des preuves de chaos n'était pas la topologie usuelle. En d'autres termes, nous avions peut-être remplacé une situation de bon désordre, perdue lors des calculs sur ordinateur, par une autre situation de désordre préservée mais de mauvaise qualité. Le théorème 12 prouve exactement le contraire.

9.4/ ÉVALUATION DE L'EXPOSANT DE LYAPUNOV

Nous terminons ce chapitre et l'étude théorique du processus $\mathcal{CIS}_2$ en évaluant son exposant de Lyapunov λ, rappelé dans la définition 74.

Soit $\mathcal{L} = \left\{x^0 \in [0, 2^5[\ / \ \forall n \in \mathbb{N}, x^n \notin I\right\}$, où I est l'ensemble des points de l'intervalle réel où g n'est pas dérivable (comme expliqué dans la proposition 34). Alors,

Théorème 13 : Exposant de Lyapunov du processus $\mathcal{CIS}_2$

$\forall x^0 \in \mathcal{L}$, l'exposant de Lyapunov du $\mathcal{CIS}_2$ ayant x^0 pour condition initiale est égal à $\lambda(x^0) = \ln(12) > 0$.

Démonstration. g est une fonction linéaire par morceaux, avec une pente égale à 12 ($g'(x) = 12$ sur tout intervalle où la fonction g est dérivable). Alors $\forall x \in \mathcal{L}$,

$$\begin{aligned}
\lambda(x) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left| g'\left(x^{i-1}\right)\right| \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} \ln |12| \\
&= \lim_{n\to+\infty} \frac{1}{n} n \ln |12| = \ln 12.
\end{aligned}$$

□

Remarque 12 : Calculabilité de l'exposant de Lyapunov de $\mathcal{CIS}_2$

L'ensemble des conditions initiales pour lesquelles cet exposant n'est pas calculable est un ensemble dénombrable. Il s'agit des conditions initiales ayant une valeur d'itération égale à un nombre de la forme $\frac{n}{12}$, avec $n \in \mathbb{N}$.
De plus, pour un système ayant N + P cellules (un nombre de LSC égal à N et un message secret à embarquer ayant une taille de P), nous trouverons, *mutatis mutandis*, un ensemble infini indénombrable de conditions initiales $x^0 \in \left[0; 2^{N+P}\right[$ telles que $\lambda(x^0) = \ln(NP^2)$.

Il est ainsi envisageable de rendre l'exposant de Lyapunov du schéma $\mathcal{CIS}_2$ aussi grand que possible, selon le nombre de coefficients les moins significatifs (LSC), que l'on décide de considérer dans le média de couverture et selon la taille du message à embarquer.

Cet exposant permet de quantifier l'amplification de l'erreur sur la condition initiale exacte (le média sans le filigrane est *a priori* inconnu de l'adversaire) après plusieurs itérations du processus de dissimulation d'informations. Rappelons que, grâce à ses propriétés topologiques, ce schéma est déjà dans une certaine mesure capable de faire face à un adversaire dans des configurations de type *KMA*, *KOA*, et *CMA*.

Le résultat ci-dessus implique, de plus, une certaine résistance face aux attaques de type *EOA* [79] (*cf.* section 5.1.2).

Par conséquent, nous avons donc rendu disponible une nouvelle évaluation quantitative du schéma de dissimulation d'informations $\mathcal{CIS}_2$. Cette nouvelle propriété quantitative de l'algorithme vient compléter la liste des propriétés topologiques dont il faisait déjà preuve (*cf.* section 8.4).

En utilisant la semi-conjugaison topologique, ainsi que la nouvelle topologie décrite dans ce travail, il sera alors possible, dans des travaux futurs, de comparer le comportement chaotique de $\mathcal{CIS}_2$ sur $\mathbb{R}$ avec le comportement d'autres schémas de dissimulation d'informations, décrits et étudiés sur ce même espace ($\mathbb{R}$).

9.5/ NÉCESSITÉ D'UN NOUVEAU MODÈLE FORMEL

Afin de permettre une utilisation pratique de l'algorithme $\mathcal{CIS}_2$, nous avons vu au chapitre 8 (*cf.* section 8.6) qu'il était nécessaire de proposer des contraintes applicatives supplémentaires.

Toutefois, ces contraintes applicatives peuvent potentiellement remettre en cause les études de sécurité qui ont été conduites au chapitre 8 et dans le présent chapitre.

C'est pourquoi, dans les chapitres qui suivent (*cf.* chapitres 11 et 12), nous avons jugé utile de proposer de nouveaux algorithmes de dissimulation d'informations. Nous avons alors cherché comment lever les contraintes applicatives, dès la mise en place des modèles formels inhérents aux nouveaux algorithmes.

CHAPITRE 10

$\mathcal{DI}_3$, version optimisée du processus $\mathcal{CIS}_2$

L'Internet représente une menace pour ceux qui savent et qui décident. Parce qu'il donne accès au savoir autrement que par le cursus hiérarchique.

JACQUES ATTALI, ÉCRIVAIN, ÉCONOMISTE, SCIENTIFIQUE, HOMME POLITIQUE (1943-)

Dans ce chapitre nous présentons un nouvel algorithme qui est inspiré du schéma $\mathcal{CIS}_2$ exposé au chapitre 8. Ce nouvel algorithme sera désigné sous le nom de $\mathcal{DI}_3$. Dans ce chapitre, nous verrons que les clés d'embarquement prévues pour l'algorithme $\mathcal{DI}_3$ requièrent une seule stratégie alors que trois sont nécessaires pour $\mathcal{CIS}_2$. Nous verrons également que, dans le schéma $\mathcal{DI}_3$, il n'y a pas d'opération de mélange du message. De ce fait, ce nouveau schéma semble plus rapide que l'algorithme $\mathcal{CIS}_2$, ce qui peut être intéressant pour des applications particulières où le temps d'exécution est un paramètre critique. Toutefois, ce gain de rapidité se fait au détriment de la sécurité, car contrairement à $\mathcal{CIS}_2$, l'algorithme $\mathcal{DI}_3$ n'est pas topologiquement-sûr. Son étude théorique a été menée dans [15], tandis que son évaluation pratique a été publiée dans [16].

10.1/ UN NOUVEAU PROCESSUS : LE $\mathcal{DI}_3$

La présentation du $\mathcal{DI}_3$ nécessite l'introduction de deux nouvelles définitions sur les suites finies.

Définition 81 : Support d'une suite finie

Le *support d'une suite finie* S de n termes est l'ensemble fini $\mathscr{S}(S) = \left\{S^k, k < n\right\}$ contenant toutes les valeurs distinctes de S. Sa cardinalité est telle que $\#\mathscr{S}(S) \leqslant n$.

Définition 82 : Suite injective, surjective et bijective

Une suite finie $S \in \mathbb{S}_{\mathsf{N}}$ de n termes est *injective* si $n = \#\mathscr{S}(S)$. Elle est *surjective* si $N = \#\mathscr{S}(S)$. Finalement, elle est bijective si et seulement si elle est à la fois injective et surjective, et alors : $n = N = \#\mathscr{S}(S)$.

L'injectivité de S reflète donc le fait que tous les n termes de la suite sont distincts. La surjectivité, quant à elle, signifie que toutes les valeurs de l'ensemble $[\![1;\mathsf{N}]\!]$ sont atteintes au moins une fois. Nous allons maintenant introduire les notations utiles à la définition du nouveau procédé.

Notation 6 : Notations pour le processus $\mathcal{DI}_3$

Nous désignerons par :

- $\mathsf{P} \in \mathbb{N}^*$ la taille, en nombre de bits, du message à embarquer dans le média de couverture C.
- $m \in \mathbb{B}^{\mathsf{P}}$ est le message à dissimuler dans x^0.
- $x^0 \in \mathbb{B}^{\mathsf{N}}$ le vecteur des N coefficients les moins significatifs (LSC) d'un média de couverture C. Les LSC sont supposés être uniformément distribués, et de taille supérieure à P.
- $\lambda \in \mathbb{N}^*$ est le nombre d'itérations à réaliser, il est tel que $\lambda > \mathsf{P}$.
- Finalement, $S \in \mathbb{S}_{\mathsf{P}}$ est une stratégie telle que la suite finie $\left\{S^k, k \in [\![\lambda - \mathsf{P} + 1; \lambda]\!]\right\}$ soit injective.

La nouvelle procédure est définie par un processus itératif appliqué sur les LSC du média de couverture de la manière suivante :

Définition 83 : Schéma de dissimulation d'informations $\mathcal{DI}_3$

Le processus $\mathcal{DI}_3$ est défini par les itérations suivantes : $\forall (n,i) \in \mathbb{N}^* \times [\![0;\mathsf{N}-1]\!]$,

$$x_i^n = \begin{cases} x_i^{n-1} & \text{if } S^n \neq i \\ m_{S^n} & \text{if } S^n = i. \end{cases}$$

Le contenu stéganographique est le vecteur booléen $y = x^\lambda \in \mathbb{B}^{\mathsf{N}}$, qui remplacera les anciens LSC. En d'autres termes, les LSC du média de couverture sont remplacés par le vecteur y.

10.2/ L'ÉTUDE DE STÉGO-SÉCURITÉ

Nous allons prouver que,

Proposition 37 : Sécurité probabiliste du $\mathcal{DI}_3$

Le processus de dissimulation d'informations $\mathcal{DI}_3$ est stégo-sûr.

Démonstration. Supposons que $x^0 \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$, $m \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{P}}\right)$, et $S \sim \mathbf{U}\left(\mathbb{S}_{\mathsf{P}}\right)$ dans une configuration $\mathcal{DI}_3$, où $\mathbf{U}(X)$ décrit l'uniforme répartition de X. Nous allons démontrer par un

raisonnement par récurrence que $\forall n \in \mathbb{N}, x^n \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$. L'initialisation est évidente d'après l'hypothèse d'uniforme répartition.

Supposons à présent que la propriété suivante $x^n \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$ est vraie pour un certain rang n : $P\left(x^n = k\right) = \frac{1}{2^N}$.

Pour un entier donné $k \in \mathbb{B}^N$, on note $\tilde{k}_i \in \mathbb{B}^N$ le vecteur défini par :
$\forall i \in [\![0; \mathsf{N}-1]\!]$, si $k = (k_0, k_1, \ldots, k_i, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1})$, alors $\tilde{k}_i = \left(k_0, k_1, \ldots, \overline{k_i}, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1}\right)$, où $\overline{x}$ est la négation du bit x.

Soit p défini par : $p = P\left(x^{n+1} = k\right)$. Soit E_j et E deux événements définis par : $\forall j \in [\![0; \mathsf{P}-1]\!], E_j = (x^n = \tilde{k}_j) \wedge (S^n = j) \wedge (m_{S^n} = k_j)$,$E = (x^n = k) \wedge (m_{S^n} = x_{S^n})$. Donc, $p = P\left(E \vee \bigvee_{j=0}^{\mathsf{N}-1} E_j\right)$.

D'une part, $\forall j \in [\![0; \mathsf{P}-1]\!]$, l'événement E_j est une conjonction des sous-événements $(S^n = j)$ et des autres sous-événements. $\forall j \in [\![0; \mathsf{P}-1]\!]$, tous les sous-événements $(S^n = j)$ sont clairement deux à deux disjoints, donc tous les événements E_j sont deux à deux disjoints.

D'autre part, $\forall j \in [\![0; \mathsf{P}-1]\!]$, les événements E_j et E sont disjoints car, dans E_j, il apparaît une conjonction de sous-événements $(x^n = \tilde{k}_j)$ avec les autres sous-événements, alors que dans E il apparaît une conjonction de sous-événements $(x^n = k)$ avec les autres sous-événements. De plus, les deux sous-événements $(x^n = \tilde{k}_j)$ et $(x^n = k)$ sont clairement disjoints. Par conséquent, en utilisant la loi de probabilité, portant sur la réunion d'événements disjoints, nous pouvons affirmer que : $p = P(E) + \sum_{j=0}^{N} P(E_j)$.

Évaluons à présent les probabilités $P(E)$ et $P(E_j)$.

1. Comme les deux événements $(x^n = k)$ et $(m_{S^n} = x_{S^n})$ concernent deux suites différentes, ils sont clairement indépendants. En utilisant l'hypothèse de récurrence, $P(x^n = k) = \frac{1}{2^N}$. Donc,

$$\begin{aligned} p(E) &= P(x^n = k) \times P(m_{S^n} = x_{S^n}) \\ &= \tfrac{1}{2^{\mathsf{N}}} \times [P(m_{S^n} = 0)P(x_{S^n} = 0) \\ &\quad + P(m_{S^n} = 1)P(x_{S^n} = 1)] \\ &= \tfrac{1}{2^{\mathsf{N}}} \times [P(m_{S^n} = 0)P(x_{S^n} = 0) \\ &\quad + P(m_{S^n} = 1)(1 - P(x_{S^n} = 0))] \\ &= \tfrac{1}{2^{\mathsf{N}}} \times \left[\tfrac{1}{2}P(x_{S^n} = 0) + \tfrac{1}{2}(1 - P(x_{S^n} = 0))\right] \\ &= \tfrac{1}{2^{\mathsf{N}+1}}. \end{aligned}$$

2. Comme les trois événements $(x^n = \tilde{k}_j)$, $(S^n = j)$, et $(m_n = k_j)$ sont fondés sur trois stratégies clairement indépendantes, on a

$$\begin{aligned} P(E_j) &= P(x^n = \tilde{k}_j) \times P(S^n = j) \times P(m_{S^n} = k_j) \\ &= \tfrac{1}{2^{\mathsf{N}}} \times \tfrac{1}{\mathsf{P}} \times \tfrac{1}{2} \\ &= \tfrac{1}{\mathsf{P}} \times \tfrac{1}{2^{\mathsf{N}+1}}, \end{aligned}$$

grâce à l'hypothèse d'uniforme répartition de S et m. En conséquence :

$$\begin{aligned} p &= P(E) + \sum_{j=0}^{\mathsf{P}} P(E_j) \\ &= \tfrac{1}{2^{\mathsf{N}+1}} + \sum_{j=0}^{\mathsf{P}} \left(\tfrac{1}{\mathsf{P}} \times \tfrac{1}{2^{\mathsf{N}+1}}\right) \\ &= \tfrac{1}{2^{\mathsf{N}}}. \end{aligned}$$

Finalement, $P\left(x^{n+1} = k\right) = \frac{1}{2^N}$, ce qui nous conduit à $x^{n+1} \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$. Ce résultat est vrai $\forall n \in \mathbb{N}$, nous avons donc bien démontré que le contenu stéganographique y est

uniformément distribué dans l'ensemble des contenus stéganographique possibles : $y \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$ lorsque $x \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$. □

10.3/ IMPLÉMENTATION

Ce nouveau schéma est décrit ici par trois algorithmes principaux :

1. Le premier, détaillé dans l'algorithme 1 page suivante, permet de générer la stratégie d'embarquement du système. Cette stratégie est une partie de la clé d'embarquement qui contient en outre le choix des LSCs et le nombre d'itérations à réaliser.
2. Le second, détaillé dans l'algorithme 2 page ci-contre, permet d'embarquer le message dans les LSCs du média de couverture en utilisant une stratégie particulière. Cette stratégie a été, d'une part, générée par le premier algorithme, d'autre part, le même nombre d'itérations a été utilisé.
3. Le dernier programme, détaillé dans l'algorithme 3 page 110, permet d'extraire le message secret à partir des LSCs du média (le contenu stéganographique) en utilisant la stratégie qui est, avec la taille du message, une partie de la clé d'extraction.

En plus de ces trois algorithmes, deux autres fonctions complémentaires doivent être utilisées (*cf.* remarque 13 de la présente page) :

1. La première, détaillée dans l'algorithme 4 page 110, permet d'extraire les MSC, les LSC, et les coefficients passifs à partir du contenu hôte. Son implémentation est basée sur le concept de fonction de signification.
2. La dernière, détaillée dans l'algorithme 5 page 111, permet de reconstruire le nouveau contenu hôte (le contenu stéganographique) à partir des coefficients MSC, LSC, et des coefficients passifs correspondants. Son implémentation est donc aussi basée sur le concept de fonction de signification : elle réalise l'opération inverse de la précédente.

Remarque 13 :

Les deux algorithmes précédents doivent, en conséquence, être ajustés en fonction de chaque contexte d'application se référant : soit à une description spatiale des images, soit à une description fréquentielle, etc. En effet, la notion de fonction de signification n'a pas de sens intrinsèque.
Illustrons notre propos en considérant un domaine de description spatial pour les images en niveaux de gris.
Dans ce contexte, la fonction de signification (*cf.* algorithme 4 page 110) attribuera un poids plus fort aux premiers bits de chaque pixel. Dans ce cas, la fonction de signification pourra, par exemple, fournir les LSC à partir des 3 derniers bits de chaque pixel, fournir les MSC à partir des 3 premiers bits et, finalement, fournir les coefficients passifs à partir des bits restants, c'est à dire les 2 bits centraux.

Par conséquent, si l'on considère une description fréquentielle des images à l'aide de la transformée en cosinus discrets [87] ou de la transformée en ondelettes discrètes [90], l'implémentation de la fonction de signification sera bien évidemment complètement différente.

Algorithme 1: *strategie*(N, P, λ)

```
/* Cette fonction a pour but de générer la stratégie S du système,
   stratégie qui est une suite d'entiers appartenant à [[0, P − 1]], telle que
   (S_{n0}, ..., S_{n0+P−1}) soit injective sur [[0, P − 1]].                  */
```

Data : N : Le nombre de LSC.
Data : P : La taille (en nombre de bits) du message à embarquer dans les LSC.
Data : λ : Le nombre d'itérations à réaliser.
Result : S : La stratégie, suite d'entiers $(S_0, S_1, \ldots)$.
begin
  $n_0 \longleftarrow L - P + 1$;
  `// Indice d'itération caractéristique.`
  `/* Test préliminaire */`
  if $P > N$ *OR* $n_0 < 0$ **then**
    return *ERROR*
  $S \longleftarrow$ Tableau de taille λ, toutes les valeurs sont initialisées à 0;
  $cpt \longleftarrow 0$;
  while $cpt < n_0$ **do**
    $S_{cpt} \longleftarrow$ Entier aléatoire appartenant à $[\![0, P-1]\!]$.;
    $cpt \longleftarrow cpt + 1$;
  $A \longleftarrow$ On génère un arrangement de $[\![0, P-1]\!]$;
  for $k \in [\![0, P-1]\!]$ **do**
    $S_{n_0+k} \longleftarrow A_k$;
  return S

Algorithme 2: *embarquement*(LSC, M, S, λ)

```
/* Cette fonction a pour but d'embarquer le message M dans les LSC du
   média de couverture en utilisant la stratégie S comme clé
   d'embarquement.                                                            */
```

Data : LSC : La liste des LSC du média de couverture.
Data : M : Le message à dissimuler dans les LSC.
Data : S : La stratégie.
Data : λ : Le nombre d'itérations à réaliser.
Result : Les nouveaux LSCs contenant le message embarqué.
begin
  $N \longleftarrow$ Nombre de LSC
  $P \longleftarrow$ Taille du message M
  for $k \in [\![0, \lambda]\!]$ **do**
    $i \longleftarrow S_k$
    $LSC_i \longleftarrow M_i$
  return LSC

10.4/ STÉGANALYSE

Notre schéma de dissimulation d'informations a été comparé à l'état de l'art en matière de stéganographieurs, à savoir YASS [76], HUGO [64] et nsF5 [36].

Algorithme 3: $extraction(LSC, S, \lambda, P)$

```
/* Cette fonction a pour but d'extraire le message à partir des LSC du
   média de couverture, en utilisant la stratégie S comme clé
   d'embarquement et la taille du message P.                          */
```
Data : LSC : La liste des LSC du média de couverture.
Data : S : La stratégie.
Data : λ : Le nombre d'itérations à réaliser.
Data : P : La taille (en nombre de bits) du message à extraire.
Result : Le message extrait à partir des LSC.
begin
  $RS \longleftarrow$ La stratégie S écrite dans l'ordre inverse.
  $M \longleftarrow$ Tableau de taille P avec toutes les valeurs initialisées à 0.
  for $k \in [\![0, \lambda]\!]$ **do**
    $i \longleftarrow RS_k$
    $M_i \longleftarrow LSC_i$
  return M

Algorithme 4: $fonctionSignification(H)$

```
/* Ce programme doit implémenter la fonction de signification. Son but est
   de fournir les MSC, les LSC et les coefficients passifs correspondants,
   selon les choix de l'utilisateur.                                  */
```
Data : H : Le contenu hôte original.
Result : MSC : Les MSC du contenu hôte H.
Result : PC : Les coefficients passifs du contenu hôte H.
Result : LSC : Les LSC du contenu hôte H.
begin
  `/* Implémentation réalisée par l'utilisateur en fonction du contexte applicatif et des contraintes associées. */`
  return (MSC, PC, LSC)

La stéganalyse est basée sur la base de données d'images BOSS [77, 20]. Cette base est constituée d'un ensemble de 10000 images en niveaux de gris de taille 512x512. Nous avons sélectionné aléatoirement 50 images extraites de la base BOSS afin de constituer un ensemble de médias de couverture pour cette stéganalyse. Puisque YASS et nsF5 sont dédiés aux supports de format JPEG, toutes ces images ont, dans un premier temps, été converties à ce format grâce à la ligne de commande `mogrify` . Pour permettre la comparaison entre les schémas de stéganographie, la charge utile relative est toujours fixée à 0,1 bit par pixel. Sous cette contrainte, le message embarqué m est une séquence de 26214 bits générés aléatoirement. Cette étape a conduit à distinguer quatre ensembles de contenus stéganographiques, un pour chaque approche stéganographique.

Nous utilisons ensuite les outils de stéganalyse développés par l'équipe *HugoBreakers* [50, 51]. Ces outils sont basés sur des classifieurs par intelligence artificielle. Le classifieur considéré a remporté la compétition BOSS [20].

Le tableau 10.4 page ci-contre résume les résultats de cette stéganalyse. Les résultats sont exprimés sous la forme d'une erreur de probabilités du stéganalyseur. Les erreurs

Algorithme 5: $fonctionReconstruction(MSC, PC, LSC)$)

```
/* Cette fonction implémente l'opération inverse de la fonction de
   signification. Son but est de reconstruire le contenu hôte H à partir
   des MSC, des LSC et des coefficients passifs correspondants.
   L'implémentation de cette fonction doit elle aussi être faite sur
   mesure, selon les choix de l'utilisateur.                          */
```

Data : MSC : Les MSC du contenu hôte H.
Data : PC : Les coefficients passifs du contenu hôte H.
Data : LSC : Les LSC du contenu hôte H.
Result : H : Le nouveau contenu hôte reconstruit à partir de ces 3 types de coefficients.
begin

```
    /* Implémentation réalisée par l'utilisateur en fonction du contexte
       applicatif et de ses contraintes.                                */
```

    return (MSC, PC, LSC)

Stéganographieur	$\mathcal{DI}_3$	YASS	HUGO	NsF5
Probabilité d'erreur	0.4133	0.0067	0.495	0.47

TABLE 10.1 – Résultats de la stéganalyse par le stéganalyseur *HugoBreakers* [50, 51] pour 50 images extraites de la base BOSS [77, 20].

sont la moyenne des fausses alarmes et des détections manquées. Ainsi, une erreur proche de 0,5 signifie que le fait de décider si une image contient un contenu stéganographique est un choix aléatoire pour le stéganalyseur. Inversement, une petite erreur indique que le stéganalyseur peut facilement faire la distinction entre les contenus stéganographiques et les contenus qui ne le sont pas.

Ainsi, les meilleurs résultats sont-ils obtenus lorsque les valeurs sont proches de 0,5.

Le meilleur résultat est obtenu par HUGO, ce dernier étant proche de la perfection en termes stéganographiques pour le stéganalyseur considéré, dans la mesure où l'erreur est proche de 0,5. Cependant, notre stéganographieur semble prometteur, d'autant plus que nous n'avons effectué, pour l'instant, aucune optimisation de notre approche. Notons finalement que le stéganalyseur *HugoBreakers* devrait fournir de bien meilleurs résultats sur une base de données d'images plus large, par exemple lorsqu'on l'applique sur la base de données BOSS toute entière.

10.5/ ÉTUDE DE ROBUSTESSE

10.5.1/ DÉTERMINER SI UN MÉDIA EST ALTÉRÉ, MESURE DE SIMILARITÉ

Considérons un média y qui est tatoué avec un message m. Considérons y' une version altérée de y, c'est-à-dire, où plusieurs bits ont été modifiés. Soit m' le message qui a été extrait de y'.

Nous allons à présent vérifier à quel point le message extrait m' est éloigné du message d'origine m.

Pour ce faire, considérons les éléments suivants :

Notation 7 : Mesures de similarité

On définit :

- L'ensemble $M = \{i | m_i = 1\}$ associé au vecteur booléen du message m.
- L'ensemble M' défini de façon analogue à partir du message m'.

La plupart des *mesures de similarité*, comme par exemple celles décrites [92, 4, 68], dépendent des fonctions a, b, c, et d, toutes définies sur $\mathbb{B}^P \times \mathbb{B}^P$, à valeur dans $\mathbb{N}$ de la manière suivante :

- $a(m, m') = |M \cap M'|$.
- $b(m, m') = |M \setminus M'|$.
- $c(m, m') = |M' \setminus M|$.
- $d(m, m') = |\overline{M} \cap \overline{M'}|$.

où $|S|$ représente la cardinalité de l'ensemble S et $\overline{S}$ représente la complémentarité de l'ensemble S.

Dans ce qui suit, a, b, c, et d représentent respectivement $a(m, m')$, $b(m, m')$, $c(m, m')$, et $d(m, m')$.

D'après [69, 53] la mesure de Fermi-Dirac S_{FD} est celle qui a la plus grande capacité de discrimination, c'est-à-dire, celle qui permet une séparation claire entre les vecteurs corrélés et non corrélés. La mesure est rappelée ci-après dans le respect des scalaires a, b et c définis précédemment.

Définition 84 : Mesures de Fermi-Dirac S_{FD}

$$\begin{aligned} S_{FD}(\varphi) &= \frac{F_{FD}(\varphi) - F_{FD}(\frac{\pi}{2})}{F_{FD}(0) - F_{FD}(\frac{\pi}{2})}, \\ F_{FD}(\varphi) &= \frac{1}{1 + \exp(\frac{\varphi - \varphi_0}{\gamma})}, \\ \varphi &= \arctan(\frac{b + c}{a}) \end{aligned}$$

où φ_0 est égal à $\pi/4$ et γ est égal à 0.1.

La distance entre m et m' est alors calculée par $1 - S_{FD}(m, m')$ et est par conséquent un nombre réel de l'intervalle $[0; 1]$. Si une telle distance est inférieure à un seuil donné, y' sera alors déclaré tatoué, et non tatoué dans le cas contraire.

Dans la section suivante, appliquons cette approche pour la déterminition de la robustesse du processus de dissimulation d'informations $\mathcal{DI}_3$, face à différentes attaques ou transformations.

10.5.2/ PRINCIPE DE L'ÉTUDE

Nous étudions maintenant la robustesse de notre approche, travail qui a été publié dans [17]. Chaque expérimentation est construite sur un ensemble de 50 images, qui sont sélectionnées aléatoirement au sein de la base de données issue du challenge BOSS [20]. À nouveau, chaque média de couverture est une image numérique en niveaux de gris de taille 512×512 pixels. La charge utile relative est toujours fixée à 0,1 bit par pixel. Avec cette contrainte, le message embarqué m est une suite de 26214 bits générés aléatoirement.

En suivant ce qui se fait habituellement dans ce genre d'études, nous avons choisi quelques attaques classiques telles que le découpage, la compression, et la rotation de la stégo-image.

Le test de robustesse s'effectue comme suit :

- Nous appliquons des attaques successives sur les contenus stéganographiques.
- Grâce à la mesure de Fermi-Dirac S_{FD} définie à la section 10.5.1, nous effectuons alors le calcul de la différence entre le message extrait de l'image attaquée et le message original.
- Cette différence est exprimée en pourcentages.

10.5.3/ PRÉSENTATION DES RÉSULTATS OBTENUS

Concernant l'attaque par découpage, nous avons appliqué différents pourcentages de modification à l'image stéganographique (de 1% à 81%). La figure 10.5.3 (c) page suivante présente les effets d'une telle attaque. Pour la robustesse face aux compressions JPEG et JPEG 2000, les résultats sont présentés respectivement en figures 10.5.3 (a) et 10.5.3 (b) page suivante.

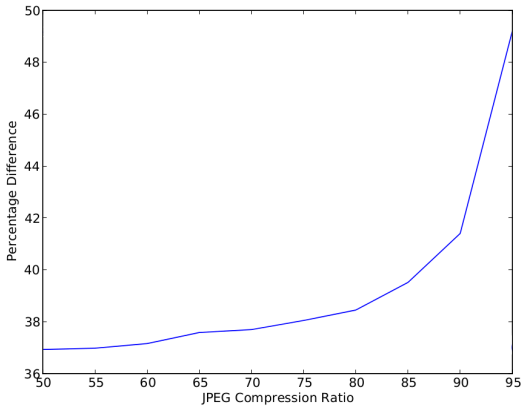
Les attaques de type transformations géométriques ont été également explorées, au travers de l'attaque par rotation : deux rotations opposées d'un angle θ sont appliquées successivement autour du centre de l'image. Dans cette transformation géométrique, les angles varient de 2 à 20 degrés. Les effets d'une telle attaque sont représentés figure 10.5.3 (d) page suivante.

À partir de ces expérimentations, on peut tirer la conclusion suivante : le processus de dissimulation d'informations $\mathcal{DI}_3$ ne présente pas d'inconvénients évidents, et résiste à ces attaques élémentaires car tous les pourcentages de différences sont effectivement inférieurs à 50%.

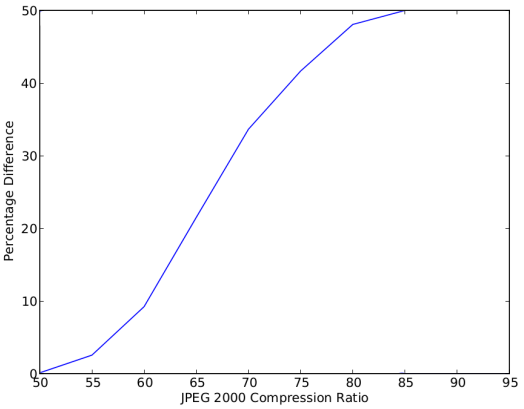
10.5.4/ ÉCHELLES VISUELLES DE DÉGRADATION DES IMAGES

Dans la section précédente (section 10.5), nous avons conduits une étude de robustesse de l'algorithme $\mathcal{DI}_3$ en respectant les principes reconnus par la communauté scientifique pour mener de telles études. Le principe, rappelons le, est d'évaluer les différences entre deux images à l'aide d'une mesure de similarité (*cf.* section 10.5.1).

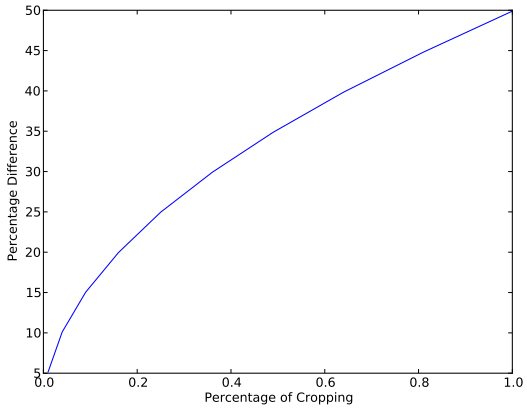
Toutefois, au regard de la définition de la robustesse d'un tatouage numérique (*cf.* définition 4 page 18), nous pouvons remarquer que la notion de robustesse est une notion subjective qui dépend du contexte et du type d'applications visés par le tatouage.



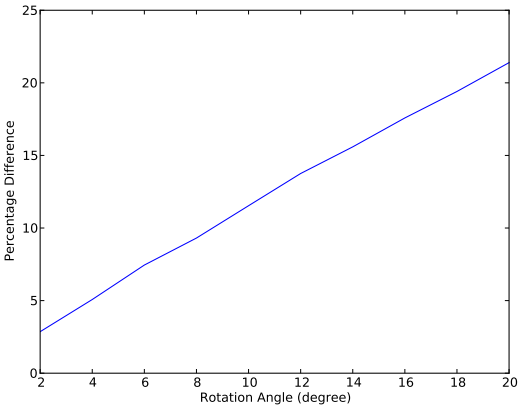
(a) Effet JPEG.



(b) Effet JPEG 2000.



(c) Attaque par découpage.



(d) Attaque par rotation.

FIGURE 10.1 – Robustesse du processus $\mathcal{DI}_3$ face à plusieurs attaques (50 images issues du répertoire BOSS [77, 20]).

Même si les mesures de similarité sont usuellement utilisées dans ce domaine, il peut être intéressant d'apprécier visuellement, de façon subjective, ce que l'on pourrait qualifier de « seuil de robustesse visuel ».

En effet, si l'on considère une image comme filigrane, un tel seuil peut être défini de façon subjective en utilisant une échelle visuelle des taux de différences, en termes de différences d'octets. Trois échelles ont été établies pour trois types d'images : les images en noir et blanc, les images en niveaux de gris et les images en couleur.

Nous pouvons tout d'abord remarquer que, dans le domaine du tatouage numérique, avec une image jouant le rôle de filigrane, un filigrane peut être considéré comme visuellement (subjectivement) identique à un autre filigrane « attaqué » (ayant été extrait une après attaque géométrique sur l'image stéganographique de couverture), si la dégradation du filigrane « attaqué » que l'on observe visuellement laisse apparaître une correspondance subjective (pour l'observateur) avec le premier filigrane. Il est important de noter que cette correspondance peut être faite également avec la version « en négatif » du filigrane original.

Les échelles visuelles présentées dans cette section permettent d'apprécier ce sentiment

subjectif, et fournissent aussi une illustration de ce que l'on entend par *version « en négatif » du filigrane original*.

10.5.4.1/ ÉCHELLE DANS LE DOMAINE DES IMAGES EN NOIR ET BLANC

L'échelle, dans le domaine des images en noir et blanc, est présentée dans la figure 10.2 page suivante.

Cette échelle permet de constater que des correspondances visuelles (subjectives) apparaîssent à l'observateur pour des taux de différences compris entre 0% et 20%, ou entre 80% et 100% dans le cas de l'image « en négatif ».

Donc, dans ce contexte, ce que nous pourrions qualifier de « seuil de robustesse visuel » peut être fixé à 20%.

10.5.4.2/ ÉCHELLE DANS LE DOMAINE DES IMAGES EN NIVEAUX DE GRIS

L'échelle, dans le domaine des images en niveaux de gris, est présentée dans la figure 10.3 page 117.

Cette échelle permet de constater que des correspondances visuelles (subjectives) apparaîssent à l'observateur pour des taux de différences compris entre 0% et 30%, ou entre 70% et 100% dans le cas de l'image « en négatif ».

Donc, dans ce contexte, ce que nous pourrions qualifier de « seuil de robustesse visuel » peut être fixé à 30%.

10.5.4.3/ ÉCHELLE DANS LE DOMAINE DES IMAGES EN COULEUR

L'échelle, dans le domaine des images en couleur, est présentée dans la figure 10.4 page 118.

Cette échelle permet de constater que des correspondances visuelles (subjectives) apparaîssent à l'observateur pour des taux de différences compris entre 0% et 40%, ou entre 60% et 100% dans le cas de l'image « en négatif ».

Donc, dans ce contexte, ce que nous pourrions qualifier de « seuil de robustesse visuel » peut être fixé à 40%.

10.5.4.4/ INTERPRÉTATION INTUITIVE

Ainsi, à partir des échelles visuelles fournies dans cette section, on constate que ce que nous avons appelé ici « seuil de robustesse visuel », dépend de la nature du filigrane. Plus le filigrane est « riche » (plus la charge utile pouvant être embarquée dans le média hôte est élevée), plus le « seuil de robustesse visuel » augmente, ce qui semble assez naturel par rapport à l'intuition que l'on peut avoir sur le sujet.

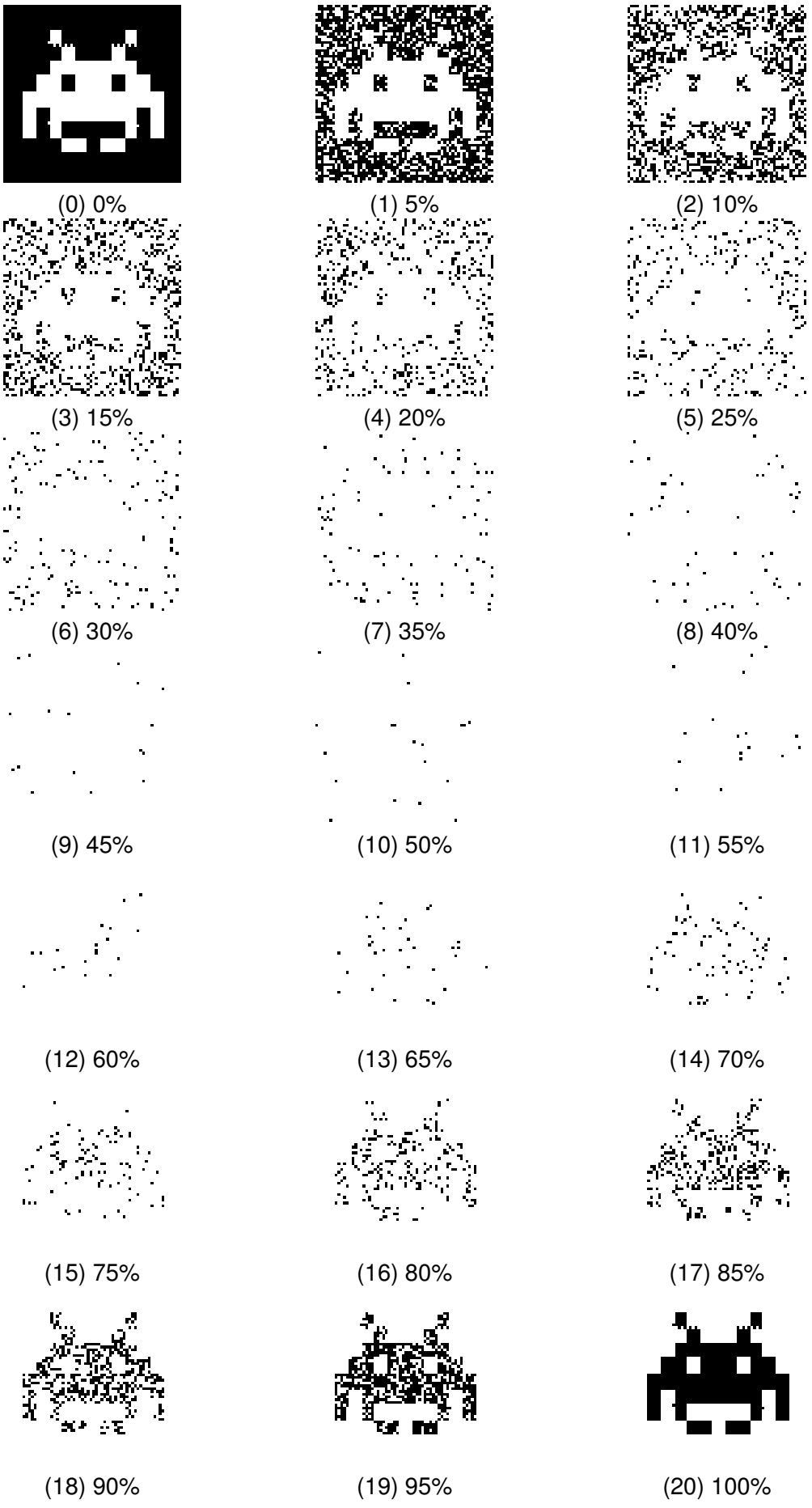


FIGURE 10.2 – Échelle visuelle des images en noir et blanc pour l'évaluation du seuil de robustesse. Le pourcentage correspond au taux de différences.

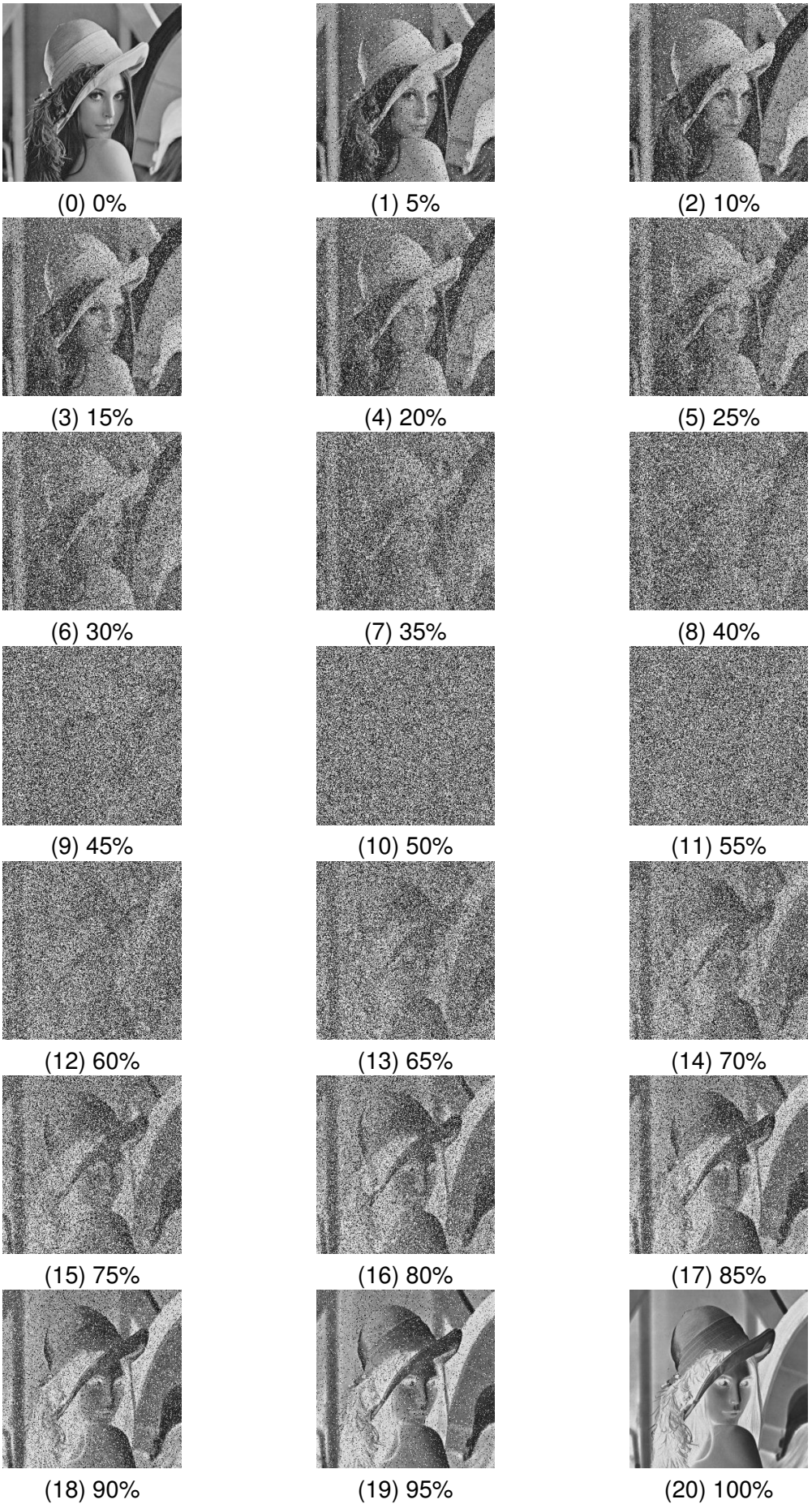


FIGURE 10.3 – Échelle visuelle des images en niveaux de gris pour l'évaluation du seuil de robustesse. Le pourcentage correspond au taux de différences.

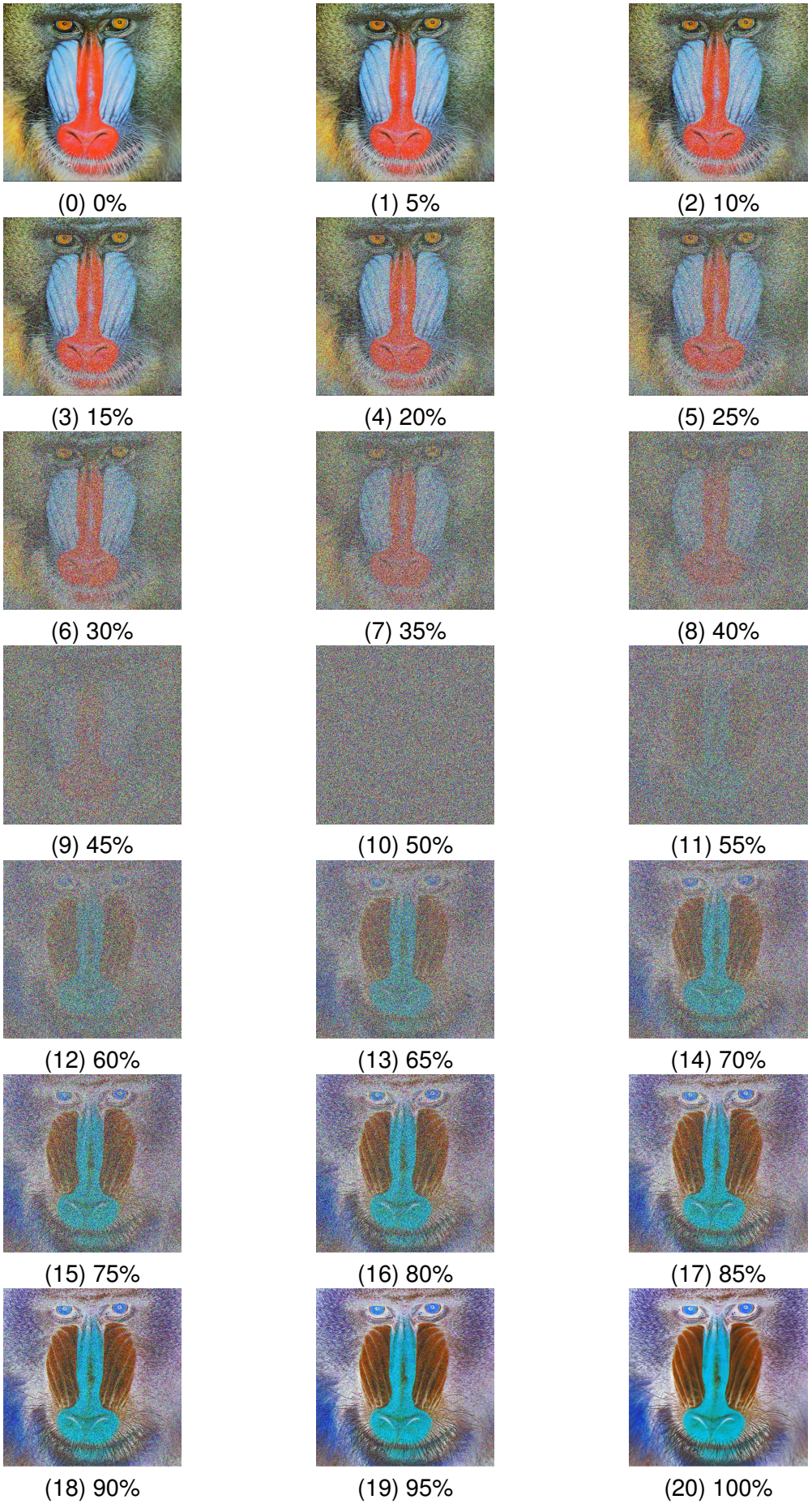


FIGURE 10.4 – Échelle visuelle des images en couleur pour l'évaluation du seuil de robustesse. Le pourcentage correspond au taux de différences.

CHAPITRE 11

Présentation du processus $\mathcal{CIS}_3$

> La science procède par révolutions et non par addition pure et simple. Cela tient aux théories qui sont toujours successives.
>
> CLAUDE BERNARD, MÉDECIN ET PHYSIOLOGISTE (1813-1878)

Dans le chapitre 8 nous avons vu que l'algorithme $\mathcal{CIS}_2$ nécessite plusieurs contraintes applicatives pour une utilisation pratique dans des cas concrets. Nous introduisons ici une première amélioration de $\mathcal{CIS}_2$. Le but de cette première amélioration est de proposer, par un cadre formel nouveau, une solution permettant de lever les contraintes applicatives inhérentes à l'algorithme $\mathcal{CIS}_2$. Pour ce faire, nous présentons dans ce chapitre un nouvel algorithme de dissimulation d'informations qui fera l'objet de futures publications. Ce processus alternatif sera noté $\mathcal{CIS}_3$.

11.1/ NOTATIONS ET TERMINOLOGIES

Commençons par introduire de nouvelles notations, utiles à la définition du processus $\mathcal{CIS}_3$.

11.1.1/ ARRANGEMENTS

Notation 8 : Ensemble des arrangements $\mathbb{A}_N^P$

Soit $P, N \in \mathbb{N}$. L'ensemble de tous les arrangements de P éléments de l'intervalle entier $[\![0; N-1]\!]$ est noté $\mathbb{A}_N^P$.

Remarque 14 : Convention pour l'ensemble des arrangements $\mathbb{A}_N^P$

Si P > N, alors $\mathbb{A}_N^P = \emptyset$

11.1.2/ PERMUTATIONS

Notation 9 : Groupe symétrique de degré N : $\mathfrak{S}_N$

Le groupe symétrique de degré N est noté $\mathfrak{S}_N$. Il contient l'ensemble des permutations de l'intervalle entier $[\![0; N-1]\!]$.

Définition 85 : Permutation circulaire sur $\mathscr{E}^k$: $\pi_{\mathscr{E}}$

Soit $\mathscr{E}$ un ensemble quelconque et $k \in \mathbb{N}$. La permutation circulaire sur l'ensemble $\mathscr{E}^k$ est la fonction $\pi_{\mathscr{E}^k}$ définie par :

$$\begin{array}{rccc} \pi_{\mathscr{E}^k} & : & \mathscr{E}^k & \longrightarrow & \mathscr{E}^k \\ & & (e^n)_{n\in[\![0;k-1]\!]} & \longmapsto & (C^n)_{n\in[\![0;k-1]\!]} \end{array}$$

$$\text{avec} \quad : \quad \left\{ \begin{array}{lll} C^n & = & e^{n+1} \quad si \quad n \in [\![0; k-2]\!] \\ C^{k-1} & = e^0 & \end{array} \right.$$

Définition 86 : Permutation circulaire sur $\mathbb{N}^k$:π_k

Soit $k \in \mathbb{N}$. La permutation circulaire sur $\mathbb{N}^k$ est la fonction π_k définie par :

$$\pi_k = \pi_{\mathbb{N}^k}$$

où $\pi_{\mathbb{N}^k}$ est la fonction introduite dans la définition 85.

11.2/ INTRODUCTION DE QUELQUES NOUVELLES DISTANCES

Nous introduisons ou rappelons maintenant plusieurs distances que nous utiliserons un peu plus loin dans le document.

Définition 87 : Distance sur $\mathbb{B}^N$: $d_{\mathbb{B}^N}$

La distance $d_{\mathbb{B}^N}$ sur l'ensemble $\mathbb{B}^N$ est l'application définie $\forall E, \check{E} \in \mathbb{B}^N$ par :

$$\begin{array}{rccc} d_{\mathbb{B}^N} & : & \mathbb{B}^N \times \mathbb{B}^N & \longrightarrow & [\![0; N]\!] \\ & & (E, \check{E}) & \longmapsto & \displaystyle\sum_{k=0}^{N-1} \delta(E_k, \check{E}_k) \end{array}$$

où δ est la distance discrète booléenne introduite précédemment.

Définition 88 : Distance sur $\mathbb{S}_N$: $d_{\mathbb{S}_N}$

La distance $d_{\mathbb{S}_N}$ sur l'ensemble $\mathbb{S}_N$ est l'application définie $\forall S, \check{S} \in \mathbb{S}_N$ par :

$$\begin{array}{rccc} d_{\mathbb{S}_N} & : & \mathbb{S}_N \times \mathbb{S}_N & \longrightarrow & [0;1] \\ & & (S, \check{S}) & \longmapsto & \frac{9}{N} \sum_{k=0}^{\infty} \frac{|S_k - \check{S}_k|}{10^k} \end{array}$$

Définition 89 : Distance sur $[\![0; N-1]\!]^P$: $d_{[\![0;N-1]\!]^P}$

La distance $d_{[\![0;N-1]\!]^P}$ sur l'ensemble $[\![0; N-1]\!]^P$ est l'application définie $\forall S, \check{S} \in [\![0; N-1]\!]^P$ par :

$$\begin{array}{rccc} d_{[\![0;N-1]\!]^P} & : & [\![0; N-1]\!]^P \times [\![0; N-1]\!]^P & \longrightarrow & [0;1] \\ & & (S, \check{S}) & \longmapsto & \frac{9}{N} \sum_{k=0}^{P-1} \frac{|S_k - \check{S}_k|}{10^k} \end{array}$$

Définition 90 : Distance sur $\mathbb{S}^P_{[\![0;N-1]\!]}$: $d_{\mathbb{S}^P_{[\![0;N-1]\!]}}$

La distance $d_{\mathbb{S}^P_{[\![0;N-1]\!]}}$ sur l'ensemble $\mathbb{S}^P_{[\![0;N-1]\!]}$ est l'application définie $\forall \Sigma, \check{\Sigma} \in \mathbb{S}^P_{[\![0;N-1]\!]}$ par :

$$\begin{array}{rccc} d_{\mathbb{S}^P_{[\![0;N-1]\!]}} & : & \mathbb{S}^P_{[\![0;N-1]\!]} \times \mathbb{S}^P_{[\![0;N-1]\!]} & \longrightarrow & [0;1] \\ & & (\Sigma, \check{\Sigma}) & \longmapsto & \frac{9}{N} \sum_{k=0}^{\infty} \frac{d_{[\![0;N-1]\!]^P}(\Sigma_k, \check{\Sigma}_k)}{10^k} \end{array}$$

où $d_{[\![0;N-1]\!]^P}$ est la distance introduite dans la définition 89.

11.3/ PRÉSENTATION DU PROCESSUS ALTERNATIF $\mathcal{CIS}_3$

11.3.1/ DIFFÉRENCES ENTRE LE $\mathcal{CIS}_2$ ET LE $\mathcal{CIS}_3$

Notre processus de dissimulation d'informations $\mathcal{CIS}_3$ est fondé sur son prédécesseur. La principale différence réside dans la définition de l'espace des phases, qui a été modifié afin de permettre l'extraction du message à l'issue de l'embarquement.

- Dans la définition de $\mathcal{CIS}_2$, la stratégie de placement $S_p \in \mathbb{S}_N$, alors que dans la définition de $\mathcal{CIS}_3$, $S'_p \in \mathbb{A}^P_N$.
- Dans la définition de $\mathcal{CIS}_2$, la stratégie de choix $S_c \in \mathbb{S}_P$, alors que dans la définition de $\mathcal{CIS}_3$, $S'_c \in \mathfrak{S}_P$,

où $\mathbb{A}^P_N$ et $\mathfrak{S}_P$ sont respectivement les ensembles introduits dans les notations 8 et 9.

Ces précisions étant fournies, venons-en à la définition du nouveau processus.

11.3.2/ DÉFINITION DU $\mathcal{CIS}_3$

Notation 10 : Pour le processus $\mathcal{CIS}_3$

Nous désignerons par :

- $x^0 \in \mathbb{B}^{\mathsf{N}}$ le vecteur des N coefficients les moins significatifs d'un média de couverture C donné.
- $m^0 \in \mathbb{B}^{\mathsf{P}}$ le message secret à embarquer dans x^0.
- $S'_p \in \mathbb{A}_{\mathsf{N}}^{\mathsf{P}}$ une stratégie appelée **stratégie de placement**.
- $S'_c \in \mathfrak{S}_{\mathsf{P}}$ une stratégie appelée **stratégie de choix**.
- $S'_m \in \mathbb{S}_{\mathsf{P}}$ une stratégie appelée **stratégie de mélange**.

Définition 91 : Processus $\mathcal{CIS}_3$

Le processus de dissimulation $\mathcal{CIS}_3$ se définit par les itérations suivantes : $\forall (n, i, j) \in \mathbb{N}^* \times [\![0; \mathsf{N}-1]\!] \times [\![0; \mathsf{P}-1]\!]$,

$$\begin{cases} x_i^n = \begin{cases} x_i^{n-1} & \text{si } (S'_p)^{(n \mod \mathsf{P})} \neq i \\ m_{(S'_c)^{(n \mod \mathsf{P})}} & \text{si } (S'_p)^{(n \mod \mathsf{P})} = i. \end{cases} \\ m_j^n = \begin{cases} m_j^{n-1} & \text{si } (S'_m)^n \neq j \\ \overline{m_j^{n-1}} & \text{si } (S'_m)^n = j. \end{cases} \end{cases}$$

où $\overline{m_j^{n-1}}$ correspond à la négation booléenne de m_j^{n-1}.
Le contenu stéganographique généré à partir du processus $\mathcal{CIS}_3$ est obtenu en remplaçant les N coefficients les moins significatifs (LSC) du contenu hôte initial par le vecteur $y \in \mathbb{B}^{\mathsf{N}}$ défini par :

$$\forall k \in \mathbb{Z}, y = x^{k\mathsf{P}}$$

On peut remarquer que, comme pour le $\mathcal{CIS}_2$, le processus de tatouage numérique à un bit $\mathcal{CIW}_1$ défini au chapitre 6 est appliqué sur le message secret $m^0 \in \mathbb{B}^{\mathsf{P}}$. Ainsi, toutes les propriétés établies pour le $\mathcal{CIW}_1$ s'appliquent aussi à m^0 et seront donc utilisables dans la suite de cette étude.

Étudions dorénavant la sécurité du nouveau procédé.

11.4/ ÉTUDE DE STÉGO-SÉCURITÉ

Nous allons démontrer que :

Proposition 38 : Stégo-sécurité de $\mathcal{CIS}_3$

Le processus $\mathcal{CIS}_3$ est stégo-sûr.

Stégo-sécurité de $\mathcal{CIS}_3$. Dans le cadre du contexte de $\mathcal{CIS}_3$, supposons que $x^0 \sim \mathbf{U}\left(\mathbb{B}^N\right)$

et $m^0 \sim \mathbf{U}\left(\mathbb{B}^P\right)$. Nous allons alors démontrer par récurrence sur l'entier n que $\forall n \in \mathbb{N}, x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$.

L'initialisation est évidente au regard des hypothèses d'uniforme répartition.

Supposons à présent que pour un certain rang n : $x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Pour un entier $k \in \mathbb{B}^N$ donné, on note $\tilde{k}_i \in \mathbb{B}^N$ le vecteur défini par : $\forall i \in [\![0; \mathsf{N}-1]\!]$, si $k = (k_0, k_1, \ldots, k_i, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1})$, alors $\tilde{k}_i = \left(k_0, k_1, \ldots, \overline{k_i}, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1}\right)$.

Soit $E_{i,j}$ l'événement probabiliste suivant : $\forall (i,j) \in [\![0; \mathsf{N}-1]\!] \times [\![0; \mathsf{P}-1]\!]$,
$E_{i,j} = (S'_p)^{(n+1 \mod \mathsf{P})} = i \wedge (S'_c)^{(n+1 \mod \mathsf{P})} = j \wedge m_j^{n+1} = k_i \wedge \left(x^n = k \vee x^n = \tilde{k}_i\right)$,
et soit $p = P\left(x^{n+1} = k\right)$. Donc :

$$p = P\left(\bigvee_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} E_{i,j}\right).$$

Nous pouvons introduire à présent la notation suivante : $P_1(i) = P\left((S'_p)^{(n+1 \mod \mathsf{P})} = i\right)$, $P_2(j) = P\left((S'_c)^{(n+1 \mod \mathsf{P})} = j\right)$, $P_3(i,j) = P\left(m_j^{n+1} = k_i\right)$, et $P_4(i) = P\left(x^n = k \vee x^n = \tilde{k}_i\right)$. Ces quatre événements sont indépendants dans le contexte de $\mathcal{CIS}_2$:

$$p = \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(i) P_3(i,j) P_4(i).$$

Or, d'après la proposition 15, $P\left(m_j^{n+1} = k_i\right) = \frac{1}{2}$. Comme les deux événements sont incompatibles, on a :

$$P\left(x^n = k \vee x^n = \tilde{k}_i\right) = P\left(x^n = k\right) + P\left(x^n = \tilde{k}_i\right).$$

En utilisant l'hypothèse de récurrence, nous trouvons : $P\left(x^n = k\right) = \frac{1}{2^N}$, et $P\left(x^n = \tilde{k}_i\right) = \frac{1}{2^N}$. Soit S la somme définie par :

$$S = \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(j).$$

Alors $p = 2 \times \frac{1}{2} \times \frac{1}{2^N} \times S = \frac{1}{2^N} \times S$. Nous pouvons à présent évaluer S :

$$\begin{array}{rcl} S & = & \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(j) \\ & = & \sum_{i \in [\![0;\mathsf{N}-1]\!]} P_1(i) \times \sum_{j \in [\![0;\mathsf{P}-1]\!]} P_2(j). \end{array}$$

L'ensemble des événements $\left\{(S'_p)^{(n+1 \mod \mathsf{P})} = i\right\}$ pour $i \in [\![0; \mathsf{N}-1]\!]$ et l'ensemble des événements $\left\{(S'_c)^{(n+1 \mod \mathsf{P})} = j\right\}$ pour $j \in [\![0; \mathsf{P}-1]\!]$ forment une partition de l'univers des possibles, d'où $S = 1$. Finalement, $P\left(x^{n+1} = k\right) = \frac{1}{2^{\mathsf{N}}}$, ce qui nous a conduits à $x^{n+1} \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$.

Ce résultat étant vrai $\forall n \in \mathbb{N}$, nous avons donc prouvé que le contenu stéganographié y suit une loi de répartition uniforme dans l'ensemble des contenus stéganographiques possibles, d'où $y \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$ quand $x \sim \mathbf{U}\left(\mathbb{B}^{\mathsf{N}}\right)$. □

11.5/ LIMITATION DU PROCESSUS $\mathcal{CIS}_3$

Nous avons certes solutionné notre problème d'extraction de message dissimulé, obtenant ainsi un algorithme complet et sûr, pour l'approche « stégo-sécurité ».

Cependant la comparaison avec les aux autres méthodes présentées dans cette thèse n'est pas pertinente car un problème persiste : nous ne pouvons raisonnablement pas étudier les propriétés topologiques du nouveau procédé. Expliquons en alors la raison.

Les nouveaux ensembles dans lesquels les stratégies de placement et les stratégies de choix prennent leurs valeurs sont des ensembles finis (*cf.* notation 10), alors que pour le processus $\mathcal{CIS}_2$, ces ensembles sont infinis. Nous aurions donc à étudier le système dynamique associé sur un ensemble fini, et sur de tels ensembles, les comportements chaotiques sont, sinon inexistants, à tout le moins dégénérés.

Nous voilà donc face à un nouveau problème. Grâce à l'introduction d'ensembles de permutations, toutes les clés d'embarquement prévues pour l'algorithme $\mathcal{CIS}_3$ permettent, sans contrainte applicative supplémentaire, l'extraction du message original intègre à partir du contenu stéganographique. Mais ce faisant, nous avons perdu le niveau de sécurité topologique dont relevait le processus précédent. Par conséquent, ce nouvel algorithme est encore améliorable.

Nous allons détailler au chapitre suivant, nos premières idées pour tirer parti des bénéfices des processus $\mathcal{CIS}_2$ et $\mathcal{CIS}_3$. Ce travail en cours conduirait à un quatrième schéma sécurisé au niveau topologique, permettant l'extraction du message, sans contrainte applicative supplémentaire, tout en conservant la propriété de stégo-sécurité des deux précédents processus.

CHAPITRE 12

Étude du processus $\mathcal{CIS}_4$

> Si vous outillez les gens et qu'ils utilisent leurs capacités naturelles et leurs curiosités, ils développeront les choses d'une manière qui vous surprendra bien au-delà de vos espérances.
>
> BILL GATES, INFORMATICIEN ET ENTREPRENEUR AMÉRICAIN (1955-)

Nous présentons dans ce chapitre notre dernière recherche pour perfectionner encore le processus de dissimulation d'informations à base d'itérations chaotiques « améliorées ». Le nouvel algorithme proposé devrait résoudre les problèmes des algorithmes $\mathcal{CIS}_2$ et $\mathcal{CIS}_3$. Il sera nommé $\mathcal{CIS}_4$ et fera, lui aussi, l'objet de futures publications.

12.1/ PRÉSENTATION DU PROCESSUS $\mathcal{CIS}_4$

Nous allons tout d'abord introduire les notations suivantes :

Notation 11 : Notations pour le processus $\mathcal{CIS}_4$

Soit $x^0 \in \mathbb{B}^{\mathsf{N}}$ le vecteur des N coefficients les moins significatifs d'un média de couverture C donné.

- Soit $m^0 \in \mathbb{B}^{\mathsf{P}}$ le message secret à embarquer dans x^0.
- Soit $S''_p \in \mathbb{S}_{\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}}$ une stratégie appelée **stratégie de placement**.
- Soit $S''_c \in \mathbb{S}_{\mathfrak{S}_{\mathsf{P}}}$ une stratégie appelée **stratégie de choix**.
- Soit finalement $S''_m \in \mathbb{S}_{\mathsf{P}}$ une stratégie appelée **stratégie de mélange**.

Le nouveau processus de dissimulation $\mathcal{CIS}_4$ est fondé sur ses deux prédécesseurs, la principale différence résidant dans la définition de l'espace. $\mathcal{CIS}_4$ a été revu afin de

permettre, à la fois, l'extraction du message, à l'issue de l'embarquement, et l'étude de sa sécurité topologique.

- Dans la définition du $\mathcal{CIS}_2$, la stratégie de placement $S_p \in \mathbb{S}_N$. De même, dans la définition du $\mathcal{CIS}_3$, $S'_p \in \mathbb{A}_{\mathsf{N}}^{\mathsf{P}}$, alors que pour $\mathcal{CIS}_4$, $S''_p \in \mathbb{S}_{\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}}$.
- Dans la définition du $\mathcal{CIS}_2$, la stratégie de choix $S_c \in \mathbb{S}_P$. De même, dans la définition du $\mathcal{CIS}_3$, $S'_c \in \mathfrak{S}_{\mathsf{P}}$, alors que dans la définition du $\mathcal{CIS}_4$, nous aurons $S''_c \in \mathbb{S}_{\mathfrak{S}_{\mathsf{P}}}$,

où $\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}$ et $\mathfrak{S}_{\mathsf{P}}$ sont respectivement les ensembles introduits dans les notations 8 et 9. Les ensembles $\mathbb{S}_{\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}}$ et $\mathbb{S}_{\mathfrak{S}_{\mathsf{P}}}$ correspondent quant à eux aux ensembles des suites d'arrangements et de permutations (*cf. chapitre 6 définition 54* sur les adaptateurs de stratégies).

Cette nouvelle définition de l'espace nous permet de passer d'espaces finis à des espaces infinis, rendant ainsi envisageable l'évaluation de la sécurité topologique du processus. Définissons maintenant notre dernier procédé.

Définition 92 : Processus $\mathcal{CIS}_4$

Le processus de dissimulation $\mathcal{CIS}_4$ est défini par les itérations suivantes : $\forall (n, i, j) \in \mathbb{N}^* \times [\![0; \mathsf{N}-1]\!] \times [\![0; \mathsf{P}-1]\!]$,

$$\begin{cases} x_i^n = \begin{cases} x_i^{n-1} & \text{si } \left((S''_p)^{\lfloor \frac{n}{\mathsf{P}} \rfloor}\right)^{(n \mod \mathsf{P})} \neq i \\ m_{\left((S''_c)^{\lfloor \frac{n}{\mathsf{P}} \rfloor}\right)^{(n \mod \mathsf{P})}-1} & \text{si } \left((S''_p)^{\lfloor \frac{n}{\mathsf{P}} \rfloor}\right)^{(n \mod \mathsf{P})} = i. \end{cases} \\ m_j^n = \begin{cases} m_j^{n-1} & \text{si } (S''_m)^n \neq j \\ \overline{m_j^{n-1}} & \text{si } (S''_m)^n = j. \end{cases} \end{cases}$$

où $\overline{m_j^{n-1}}$ correspond à la négation booléenne de m_j^{n-1}.

Le contenu stéganographique généré à partir du processus $\mathcal{CIS}_4$ est obtenu en remplaçant les N coefficients les moins significatifs (LSC) du contenu hôte initial par le vecteur $y \in \mathbb{B}^{\mathsf{N}}$ défini par :

$$\forall k \in \mathbb{Z}, y = x^{k\mathsf{P}}$$

Une fois encore, le processus de tatouage numérique à un bit $\mathcal{CIW}_1$, défini au chapitre 6, est appliqué sur le message secret $m^0 \in \mathbb{B}^{\mathsf{P}}$. Ainsi, toutes les propriétés établies pour $\mathcal{CIW}_1$ s'appliquent à m^0, et seront donc utilisables dans la suite de cette étude.

12.2/ ÉTUDE DE STÉGO-SÉCURITÉ

Nous allons démontrer que :

Proposition 39 : Stégo-sécurité de $\mathcal{CIS}_4$

Le processus $\mathcal{CIS}_4$ est stégo-sûr

Stégo-sécurité de $\mathcal{CIS}_4$. Dans le cadre du contexte de $\mathcal{CIS}_4$, supposons que $x^0 \sim \mathbf{U}\left(\mathbb{B}^N\right)$

et $m^0 \sim \mathbf{U}\left(\mathbb{B}^P\right)$. Nous allons alors démontrer par récurrence sur l'entier n que $\forall n \in \mathbb{N}, x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$.

L'initialisation est évidente au regard des hypothèses d'uniforme répartition.

Supposons à présent que pour un certain rang n : $x^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Pour un entier $k \in \mathbb{B}^N$ donné, on note $\tilde{k}_i \in \mathbb{B}^N$ le vecteur défini par : $\forall i \in [\![0; \mathsf{N}-1]\!]$,
si $k = (k_0, k_1, \ldots, k_i, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1})$, alors $\tilde{k}_i = \left(k_0, k_1, \ldots, \overline{k_i}, \ldots, k_{\mathsf{N}-2}, k_{\mathsf{N}-1}\right)$.

Soit $E_{i,j}$ l'événement probabiliste suivant : $\forall (i,j) \in [\![0; \mathsf{N}-1]\!] \times [\![0; \mathsf{P}-1]\!]$,
$E_{i,j} = (S'_p)^{(n+1 \mod \mathsf{P})} = i \wedge (S'_c)^{(n+1 \mod \mathsf{P})} = j \wedge m_j^{n+1} = k_i \wedge \left(x^n = k \vee x^n = \tilde{k}_i\right)$, et soit $p = P\left(x^{n+1} = k\right)$. Alors,

$$p = P\left(\bigvee_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} E_{i,j}\right).$$

Nous pouvons introduire à présent les notations suivantes : $P_1(i) = P\left(\left((S''_p)^{\lfloor \frac{n+1}{\mathsf{P}} \rfloor}\right)^{(n+1 \mod \mathsf{P})} = i\right)$, $P_2(j) = P\left(\left((S''_c)^{\lfloor \frac{n+1}{\mathsf{P}} \rfloor}\right)^{(n+1 \mod \mathsf{P})} = j\right)$, $P_3(i,j) = P\left(m_j^{n+1} = k_i\right)$, et $P_4(i) = P\left(x^n = k \vee x^n = \tilde{k}_i\right)$.

Ces quatre événements sont indépendants dans le contexte de $\mathcal{CIS}_4$, donc

$$p = \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(i) P_3(i,j) P_4(i).$$

D'après la proposition 15, $P\left(m_j^{n+1} = k_i\right) = \frac{1}{2}$. Comme les deux événements sont incompatibles :

$$P\left(x^n = k \vee x^n = \tilde{k}_i\right) = P\left(x^n = k\right) + P\left(x^n = \tilde{k}_i\right).$$

D'après l'hypothèse de récurrence : $P\left(x^n = k\right) = \frac{1}{2^N}$, et $P\left(x^n = \tilde{k}_i\right) = \frac{1}{2^N}$. Soit S la somme définie par :

$$S = \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(j).$$

Alors $p = 2 \times \frac{1}{2} \times \frac{1}{2^N} \times S = \frac{1}{2^N} \times S$.

Nous pouvons à présent évaluer S :

$$\begin{array}{rcl} S & = & \sum_{i \in [\![0;\mathsf{N}-1]\!], j \in [\![0;\mathsf{P}-1]\!]} P_1(i) P_2(j) \\ & = & \sum_{i \in [\![0;\mathsf{N}-1]\!]} P_1(i) \times \sum_{j \in [\![0;\mathsf{P}-1]\!]} P_2(j). \end{array}$$

L'ensemble des événements $\left\{\left((S''_p)^{\lfloor \frac{n+1}{\mathsf{P}} \rfloor}\right)^{(n+1 \mod \mathsf{P})} = i\right\}$ pour $i \in [\![0; N-1]\!]$ et l'ensemble des événements $\left\{\left((S''_c)^{\lfloor \frac{n+1}{\mathsf{P}} \rfloor}\right)^{(n+1 \mod \mathsf{P})} = j\right\}$ pour $j \in [\![0; P-1]\!]$ forment une partition de l'univers des possibles, d'où $S = 1$.

Finalement, $P\left(x^{n+1} = k\right) = \frac{1}{2^N}$, ce qui nous conduit à $x^{n+1} \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Ce résultat est vrai $\forall n \in \mathbb{N}$, nous avons donc prouvé que le contenu stéganographié y suit une loi de répartition uniforme dans l'ensemble des contenus stéganographiques possibles, d'où $y \sim \mathbf{U}\left(\mathbb{B}^N\right)$ quand $x \sim \mathbf{U}\left(\mathbb{B}^N\right)$. □

12.3/ UN MODÈLE TOPOLOGIQUE POUR $\mathcal{CIS}_4$

Dans cette section, et comme pour les précédents schémas, nous démontrons que le $\mathcal{CIS}_4$ peut être modélisé sous la forme d'un système dynamique discret sur un espace métrique.

12.3.1/ FONCTION D'ITÉRATIONS ET ESPACE DES PHASES

Définissons pour commencer la fonction suivante.

Définition 93 : Fonction F''

Soit

$$\begin{array}{rccc} F'' : & [\![0;\mathsf{N}-1]\!] \times \mathbb{B}^{\mathsf{N}} \times [\![0;\mathsf{P}-1]\!] \times \mathbb{B}^{\mathsf{P}} & \longrightarrow & \mathbb{B}^{\mathsf{N}} \\ & (k,x,\lambda,m) & \longmapsto & \left(\delta(k,j).x_j + \overline{\delta(k,j)}.m_\lambda\right)_{j\in[\![0;\mathsf{N}-1]\!]} \end{array}$$

L'espace des phases $\mathcal{X}_4$ est défini de la manière suivante :

Définition 94 : Espace des phases $\mathcal{X}_4$

$$\mathcal{X}_4 = \mathbb{S}_{\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}} \times \mathbb{B}^{\mathsf{N}} \times \mathbb{S}_{\mathfrak{S}_{\mathsf{P}}} \times \mathbb{B}^{\mathsf{P}} \times \mathbb{S}_{\mathsf{P}}$$

Nous avons de plus besoin des fonctions α et β, définies ci-après.

Définition 95 : Fonction $\alpha_{(\mathsf{N},\mathsf{P})}$ pour le processus $\mathcal{CIS}_4$

Soit $\mathsf{N},\mathsf{P} \in \mathbb{N}$. On définit une fonction auxiliaire $\alpha_{(\mathsf{N},\mathsf{P})}$ permettant de passer d'une suite d'arrangements à une suite classique dont les termes sont issus de cette suite d'arrangements, comme suit :

$$\begin{array}{rccc} \alpha_{(\mathsf{N},\mathsf{P})} : & \mathbb{S}_{\mathbb{A}_{\mathsf{N}}^{\mathsf{P}}} & \longrightarrow & \mathbb{S}_{\mathsf{N}} \\ & (e^n)_{n\in\mathbb{N}} & \longmapsto & \left(\left(e^{\lfloor \frac{n}{\mathsf{P}} \rfloor}\right)^{(n \mod \mathsf{P})}\right)_{n\in\mathbb{N}} \end{array}$$

La fonction $\alpha_{(\mathsf{N},\mathsf{P})}$ est injective par construction, mais n'est aucunement surjective : il existe des suites qui ne sont pas constituées de suites d'arrangements. Il suffit de prendre comme exemple la suite des décimales de π.

Définition 96 : Fonction β_{P} pour le processus $\mathcal{CIS}_4$

Soit $\mathsf{P} \in \mathbb{N}$. On définit une fonction auxiliaire β_{P} permettant de passer d'une suite de permutations à une suite classique dont les termes sont issus de cette suite de permutations.

$$\begin{array}{rccc} \beta_{\mathsf{N}} : & \mathbb{S}_{\mathfrak{S}_{\mathsf{P}}} & \longrightarrow & \mathbb{S}_{\mathsf{P}} \\ & (e^n)_{n\in\mathbb{N}} & \longmapsto & \left(\left(e^{\lfloor \frac{n}{\mathsf{P}} \rfloor}\right)^{(n \mod \mathsf{P})}\right)_{n\in\mathbb{N}} \end{array}$$

De même, la fonction β_{N} est injective, mais pas surjective. On définit finalement la fonction ${\mathcal{G}''}_{f_0}$ par :

Définition 97 : Fonction $\mathcal{G}''_{f_0}$

$$\mathcal{G}''_{f_0} : \mathcal{X}_4 \longrightarrow \mathcal{X}_4, \mathcal{G}''_{f_0}\left(S''_p, x, S''_c, m, S''_m\right) =$$
$$\left(\sigma_N\left(\alpha_{(\mathsf{N},\mathsf{P})}(S''_p)\right), F''\left(i_N\left(\alpha_{(\mathsf{N},\mathsf{P})}(S''_p)\right), x, i_P\left(\beta_{\mathsf{P}}(S''_c)\right), m\right),$$
$$\sigma_P\left(\beta_{\mathsf{P}}(S''_c)\right), G_{f_0}(m, S''_m), \sigma_P(S''_m)\Big)$$

Alors le processus de dissimulation d'informations $\mathcal{CIS}_4$ peut être décrit par les itérations du système dynamique discret suivant :

Définition 98 : Système dynamique discret pour $\mathcal{CIS}_4$

$$\begin{cases} X^0 \in \mathcal{X}_4 \\ X^{k+1} = \mathcal{G}''_{f_0}(X^k). \end{cases}$$

12.3.2/ CARDINALITÉ DE $\mathcal{X}_4$

En comparant $\mathcal{X}_4$ et $\mathcal{X}_1$, nous pouvons en tirer le résultat suivant :

Proposition 40 : Cardinalité de $\mathcal{X}_4$

L'espace des phases $\mathcal{X}_4$ a, au moins, la puissance du continu.

Démonstration. Soit $s''_A \in \mathbb{S}_{\mathbb{A}_\mathsf{N}^\mathsf{P}}$ une suite d'arrangements et soit $s''_P \in \mathbb{S}_{\mathfrak{S}_\mathsf{N}}$ une suite de permutations. Nous définissons alors la fonction $\varphi_{(s''_A, s''_P)}$ de la manière suivante :

$$\begin{array}{rccc} \varphi_{(s''_A,s''_P)} : & \mathcal{X}_1 & \longrightarrow & \mathcal{X}_4 \\ & (S, x) & \longmapsto & (s''_A, x, s''_P, 0, S) \end{array}$$

$\varphi_{(s''_A,s''_P)}$ est injective. Par conséquent la puissance de $\mathcal{X}_4$ est supérieure ou égale à la puissance de $\mathcal{X}_1$. Finalement, $\mathcal{X}_4$ a au moins la puissance du continu. □

Ce résultat est indépendant du nombre de cellules du système.

12.3.3/ UNE NOUVELLE DISTANCE SUR $\mathcal{X}_4$

On définit une nouvelle distance sur $\mathcal{X}_4$ de la manière suivante :

Définition 99 : Distance d_4 sur $\mathcal{X}_4$

$\forall X, \check{X} \in \mathcal{X}_4$, si $X = (S''_p, x, S''_c, m, S''_m)$ et $\check{X} = (\check{S''_p}, \check{x}, \check{S''_c}, \check{m}, \check{S''_m})$, alors :

$$\begin{array}{rcl} d_4(X, \check{X}) & = & d_{\mathbb{B}^\mathsf{N}}(x, \check{x}) + d_{\mathbb{B}^\mathsf{P}}(m, \check{m}) \\ & + & d_{\mathbb{S}_\mathsf{N}}(\alpha_{(\mathsf{N},\mathsf{P})}(S''_p), \alpha_{(\mathsf{N},\mathsf{P})}(\check{S''_p})) + d_{\mathbb{S}_\mathsf{P}}(\beta_{\mathsf{P}}(S''_c), \beta_{\mathsf{P}}(\check{S''_c})) + d_{\mathbb{S}_\mathsf{P}}(S''_m, \check{S''_m}), \end{array}$$

où $d_{\mathbb{B}^\mathsf{N}}$, $d_{\mathbb{B}^\mathsf{P}}$, $d_{\mathbb{S}_N}$, et $d_{\mathbb{S}_P}$ sont les mêmes distances que dans la définition 62, et les fonctions $\alpha_{(\mathsf{N},\mathsf{P})}$ et β_P sont respectivement introduites dans les définitions 95 et 96.

12.3.4/ CONTINUITÉ DE CIS_4

Afin de démontrer que le processus de dissimulation d'informations CIS_4 est un exemple de chaos topologique au sens de Devaney, la fonction $\mathcal{G}_{f_0}$ doit être continue sur l'espace métrique $(\mathcal{X}_2, d_2)$. Nous allons le démontrer.

Proposition 41 : Continuité de $\mathcal{G}''_{f_0}$

$\mathcal{G}''_{f_0}$ est une fonction continue sur $(\mathcal{X}_4, d_4)$.

Continuité de $\mathcal{G}''_{f_0}$. Nous allons une fois de plus utiliser le principe de la continuité séquentielle.

Soit $((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)_{n\in\mathbb{N}}$ une suite de points de l'espace des phases $\mathcal{X}_4$, qui converge vers $(S''_p, x, S''_c, m, S''_m)$. Nous allons prouver que $\left(\mathcal{G}''_{f_0}((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)\right)_{n\in\mathbb{N}}$ converge vers $\mathcal{G}''_{f_0}(S_p, x, S_c, m, S_m)$. Rappelons que pour tout n, $(S''_p)^n$, $(S''_c)^n$ et $(S''_m)^n$ sont des stratégies, respectivement issues des ensembles $\mathbb{S}_{\mathbb{A}^{\mathsf{P}}_{\mathsf{N}}}$, $\mathbb{B}^{\mathsf{P}}$ et $\mathbb{S}_{\mathsf{P}}$: nous considérons donc une suite de stratégies.

Comme $d_4(((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n), (S''_p, x, S''_c, m, S''_m))$ tend vers 0, chaque distance $d_{\mathbb{B}^{\mathsf{N}}}(x^n, x)$, $d_{\mathbb{B}^{\mathsf{P}}}(m^n, m)$, $d_{\mathbb{S}_{\mathsf{N}}}(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S''_p))$, $d_{\mathbb{S}_{\mathsf{P}}}(\beta_{\mathsf{P}}((S''_c)^n), \beta_{\mathsf{P}}(S''_c))$, et $d_{\mathbb{S}_{\mathsf{P}}}((S''_m)^n, S''_m)$ tend vers 0. Mais $d_{\mathbb{B}^{\mathsf{N}}}(x^n, x)$ et $d_{\mathbb{B}^{\mathsf{P}}}(m^n, m)$ sont des entiers naturels, donc $\exists n_0 \in \mathbb{N}, \forall n \geqslant n_0, d_{\mathbb{B}^{\mathsf{N}}}(x^n, x) = 0$ et $\exists n_1 \in \mathbb{N}, \forall n \geqslant n_1, d_{\mathbb{B}^{\mathsf{P}}}(m^n, m) = 0$.

En d'autres mots, il existe un rang $n_3 = Max(n_0, n_1) \in \mathbb{N}$ à partir duquel les cellules ne changent plus d'état : $\exists n_3 \in \mathbb{N}, n \geqslant n_3 \Longrightarrow (x^n = x) \wedge (m^n = m)$.

De plus, $d_{\mathbb{S}_{\mathsf{N}}}(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S''_p)) \longrightarrow 0$, $d_{\mathbb{S}_P}(\beta_{\mathsf{P}}((S''_c)^n), \beta_{\mathsf{P}}(S''_c)) \longrightarrow 0$, et $d_{\mathbb{S}_{\mathsf{P}}}((S''_m)^n, S''_m) \longrightarrow 0$, alors $\exists n_4, n_5, n_6 \in \mathbb{N}$,

- $\forall n \geqslant n_4, d_{\mathbb{S}_{\mathsf{N}}}(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S''_p)) < 10^{-1}$,
- $\forall n \geqslant n_5, d_{\mathbb{S}_P}(\beta_{\mathsf{P}}((S''_c)^n), \beta_{\mathsf{P}}(S''_c)) < 10^{-1}$,
- $\forall n \geqslant n_6, d_{\mathbb{S}_P}((S''_m)^n, S''_m) < 10^{-1}$.

Soit $n_7 = Max(n_4, n_5, n_6)$. Pour $n \geqslant n_7$, toutes les stratégies $\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n)$, $\beta_{\mathsf{P}}((S''_c)^n)$, et $(S''_m)^n$ ont le même premier terme, qui est respectivement $\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)_0)$,$\beta_{\mathsf{P}}((S''_c)_0)$ et $(S''_m)_0$: $\forall n \geqslant n_7$,

$$(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n_0) = \alpha_{(\mathsf{N},\mathsf{P})}((S''_p)_0)) \wedge (\beta_{\mathsf{P}}((S''c)^n_0) = \beta_{\mathsf{P}}((S''_c)_0)) \wedge ((S''_m)^n_0 = (S''_m)_0).$$

Soit $n_8 = Max(n_3, n_7)$. À partir du rang n_8, les états de x^n et x d'une part, et m^n et m d'autre part, sont identiques. De plus, les stratégies $\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n)$ et $\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)$, $\beta_{\mathsf{P}}((S''_c)^n)$ et $\beta_{\mathsf{P}}(S''_c)$, et $(S''_m)^n$ et S''_m commencent avec le même premier terme.

Par conséquent, les états de $\mathcal{G}''_{f_0}((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)$ et $\mathcal{G}''_{f_0}(S''_p, x, S''_c, m, S''_m)$ sont égaux, donc, après le terme de rang n_8, la distance d_4 entre ces deux points est strictement inférieure à 3.10^{-1}, donc strictement inférieure à 1.

Nous allons à présent démontrer que la distance entre $\left(\mathcal{G}''_{f_0}((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)\right)$ et $\left(\mathcal{G}''_{f_0}(S''_p, x, S''_c, m, S''_m)\right)$ converge vers 0. Soit $\varepsilon > 0$.

- Si $\varepsilon \geqslant 1$, nous avons vu que la distance entre $\mathcal{G}''_{f_0}((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)$ et $\mathcal{G}''_{f_0}(S''_p, x, S''_c, m, S''_m)$ est strictement inférieure à 1 à partir du terme de rang n_8 (même état).

– Si $\varepsilon < 1$, alors $\exists k \in \mathbb{N}, 10^{-k} \geqslant \frac{\varepsilon}{3} \geqslant 10^{-(k+1)}$. Comme $d_{\mathbb{S}_\mathsf{N}}(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S''_p))$, $d_{\mathbb{S}_\mathsf{P}}(\beta_\mathsf{P}((S''_c)^n), \beta_\mathsf{P}(S''_c))$ et $d_{\mathbb{S}_\mathsf{P}}((S''_m)^n, S''_m)$ converge vers 0, nous avons :
  – $\exists n_9 \in \mathbb{N}, \forall n \geqslant n_9, d_{\mathbb{S}_\mathsf{N}}(\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S''_p)) < 10^{-(k+2)}$,
  – $\exists n_{10} \in \mathbb{N}, \forall n \geqslant n_{10}, d_{\mathbb{S}_\mathsf{P}}(\beta_\mathsf{P}((S''_c)^n), \beta_\mathsf{P}(S''_c)) < 10^{-(k+2)}$,
  – $\exists n_{11} \in \mathbb{N}, \forall n \geqslant n_{11}, d_{\mathbb{S}_\mathsf{P}}((S''_m)^n, S''_m) < 10^{-(k+2)}$.

  Soit $n_{12} = Max(n_9, n_{10}, n_{11})$ alors après le terme de rang n_{12}, les $k+2$ premiers termes de $\alpha_{(\mathsf{N},\mathsf{P})}((S''_p)^n)$ et $\alpha_{(\mathsf{N},\mathsf{P})}(S''_p)$, $\beta_\mathsf{P}((S''_c)^n)$ et $\beta_\mathsf{P}(S''_c)$, et $(S''_m)^n$ et S''_m, sont égaux.
  Les fonctions $\alpha_{(\mathsf{N},\mathsf{P})}$ et β_P sont toutes deux injectives donc, après le terme de rang n_{12}, les $k+2$ premiers termes de $(S''_p)^n$ et S''_p, $(S''_c)^n$ et S''_c, et $(S''_m)^n$ et S''_m, sont eux aussi égaux.

Par conséquent, les $k+1$ premiers termes des stratégies de $\mathcal{G}''_{f_0}((S''_p)^n, x^n, (S''_c)^n, m^n, (S''_m)^n)$ et $\mathcal{G}''_{f_0}(S''_p, x, S''_c, m, S''_m)$ sont les mêmes (modulo l'application des fonctions $\alpha_{(\mathsf{N},\mathsf{P})}$ et β_P, et en raison du décalage de stratégie) et d'après la définition $d_{\mathbb{S}_\mathsf{N}}$ et $d_{\mathbb{S}_P}$:

$$d_4\left(\mathcal{G}''_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n); \mathcal{G}''_{f_0}(S_p, x, S_c, m, S_m)\right)$$

est égale à :

$$d_{\mathbb{S}_\mathsf{N}}(\alpha_{(\mathsf{N},\mathsf{P})}((S_p)^n), \alpha_{(\mathsf{N},\mathsf{P})}(S_p)) + d_{\mathbb{S}_\mathsf{P}}(\beta_\mathsf{P}((S_c)^n), \beta_\mathsf{P}(S_c)) + d_{\mathbb{S}_\mathsf{P}}((S_m)^n, S_m)$$

qui est inférieur à $3.10^{-(k+1)} \leqslant 3.\frac{\varepsilon}{3} = \varepsilon$.

Soit $N_0 = max(n_8, n_{12})$. Nous pouvons affirmer que :

$$\forall \varepsilon > 0, \exists N_0 \in \mathbb{N}, \forall n \geqslant N_0,$$

$$d_2\left(\mathcal{G}''_{f_0}((S_p)^n, x^n, (S_c)^n, m^n, (S_m)^n); \mathcal{G}''_{f_0}(S_p, x, S_c, m, S_m)\right) \leqslant \varepsilon.$$

$\mathcal{G}''_{f_0}$ est par conséquent continue sur $(\mathcal{X}_4, d_4)$. □

Il nous reste à montrer que le $\mathcal{CIS}_4$ est un cas de chaos topologique au sens de Devaney. C'est-à-dire qu'il faudrait établir les deux propriétés de régularité et de transitivité pour la fonction $\mathcal{G}''_{f_0}$. Ce travail reste à réaliser.

IV

CONCLUSION

CHAPITRE 13

Bilan et perspectives

> La civilisation a pour but, non pas le progrès de la science et des machines, mais celui de l'homme.
>
> ALEXIS CARREL, CHIRURGIEN ET BIOLOGISTE, PRIX NOBEL (1873-1944)

Cet ultime chapitre est divisé en deux sections principales. Dans la première, nous donnons une synthèse générale des résultats obtenus dans cette thèse et nous présentons un bilan comparatif des propriétés des différents processus de dissimulation d'informations étudiés. Cette synthèse se poursuit par une section présentant diverses perspectives issues de nos travaux de recherches.

13.1/ BILAN

Durant cette thèse, nous avons commencé par approfondir les propriétés de sécurité du $\mathcal{CIW}_1$. Ce premier processus de dissimulation d'informations est basé sur des itérations chaotiques présentées à l'origine dans [11]. Elles sont rappelées dans l'état de l'art au chapitre 6. Jusqu'alors, seul le caractère chaotique (au sens de Devaney) de ces itérations avait été prouvé. Nous en avons établi la stégo-sécurité et nous avons prouvé qu'elles possédaient d'autres propriétés de sécurité, liées à leur comportement topologique, à savoir l'expansivité et le mélange topologique. Les propriétés topologiques de base avaient déjà été établies dans [18] pour certaines techniques d'étalement de spectre. Cependant, il avait été précédemment montré que ces dernières n'étaient pas expansives. De ce fait, contrairement à l'étalement de spectre, cette première méthode de dissimulation, basée sur les itérations chaotiques, permettait de faire face à des attaques de type KOA et KMA. Nous avons profité de cette première étude d'un algorithme de dissimulation d'informations, à base d'itérations chaotiques, pour mesurer combien ces processus impactaient l'image hôte en la dégradant. C'est le calcul du PSNR qui a

permis de quantifier cette dégradation.

Le problème avec le processus $\mathcal{CIW}_1$ est qu'il ne permet d'insérer qu'un seul bit dans une image donnée. Ce genre d'information binaire peut être satisfaisant en tatouage numérique (encore faudrait-il que le $\mathcal{CIW}_1$ soit robuste), mais la capacité de ce procédé est beaucoup trop faible pour satisfaire aux besoins de la stéganographie. C'est pourquoi, nous avons introduit un second algorithme de dissimulation d'informations.

Cette nouvelle méthode, notée $\mathcal{CIS}_2$, est basée sur des itérations chaotiques « améliorées ». Sa sécurité a été pleinement évaluée. Nous avons tout d'abord prouvé sa stégo-sécurité, avant d'établir les propriétés topologiques de base dont $\mathcal{CIS}_2$ faisait déjà preuve. Cette étude de sécurité du schéma $\mathcal{CIS}_2$, suivant l'approche topologique, a ensuite été étendue grâce à la mesure de son exposant de Lyapunov. Nous avons démontré que cet exposant était égal à $\ln(\mathrm{NP}^2)$, où N est le nombre de coefficients les moins significatifs du média de couverture et P la taille du message secret à embarquer.

Nous avons identifié les contraintes applicatives liées à l'algorithme $\mathcal{CIS}_2$ afin de pouvoir en proposer une implémentation concrète utilisable en pratique. Cette implémentation a été développée et baptisée $\mathcal{CIS}_5$. Elle nous a permis de réaliser un dépôt logiciel auprès de l'Agence pour la Protection des Programmes [65].

Lors de nos travaux, pour intégrer les contraintes applicatives de $\mathcal{CIS}_2$ directement dans la modélisation formelle d'un nouveau processus, nous avons introduit un nouvel algorithme de dissimulation d'informations noté $\mathcal{DI}_3$. La rapidité de cette nouvelle version a été améliorée en réduisant le nombre de stratégies à prendre en compte. La stégo-sécurité du schéma $\mathcal{DI}_3$ a été étudiée. Nous avons alors pu produire des algorithmes concrets qui ont permis de vérifier la robustesse du procédé face à certaines attaques spatiales et fréquentielles. Enfin, nous avons utilisé le meilleur stéganalyseur du marché, *HugoBreakers*, pour estimer la capacité de ce stéganalyseur, basé sur l'intelligence artificielle, à détecter la présence d'informations secrètes insérées par le processus $\mathcal{DI}_3$. Les résultats, comparés avec quelques autres stéganographieurs réputés, se sont avérés plutôt encourageants.

Une certaine preuve de concept ayant été établie par la pratique, nous avons souhaité aller encore plus loin, en essayant d'améliorer les processus $\mathcal{CIS}_2$ et $\mathcal{DI}_3$. En effet, $\mathcal{CIS}_2$ requiert des contraintes applicatives supplémentaires pour être implémenté. L'algorithme $\mathcal{DI}_3$, pour sa part, n'est pas à base d'itérations chaotiques. Il est donc plus difficile d'établir sa sécurité topologique. Nous avons donc poursuivi nos investigations.

Nous avons alors introduit le schéma $\mathcal{CIS}_3$ et nous en avons établi sa stégo-sécurité. Ce processus améliore $\mathcal{CIS}_2$, car sa modélisation formelle tient directement compte des contraintes applicatives supplémentaires que nous avions du identifier pour implémenter $\mathcal{CIS}_2$. Un problème cependant demeurait, du fait que le système dynamique discret sur lequel est modélisé cet algorithme a un espace des phases qui ne permet pas d'atteindre la sécurité topologique du schéma de façon pertinente. Les preuves de sécurité, selon l'approche topologique, n'ont donc guère de sens pour le $\mathcal{CIS}_3$. Nous l'avons donc amélioré une dernière fois, ce qui a donné naissance au schéma $\mathcal{CIS}_4$.

Ce dernier processus a été pleinement défini, sa stégo-sécurité prouvée et il a, enfin, été modélisé sous la forme d'un système dynamique discret continu, sur un espace métrique, lui aussi, pleinement défini. La preuve de sécurité topologique de ce procédé reste à établir.

Le tableau 13.1 fourni une synthèse comparative des différents algorithmes mis au point.

Il présente les résultats des études conduites sur le plan de la sécurité. Ce tableau expose donc la capacité, de chacun des processus, à faire face aux attaques classiques du domaine de la science de l'information dissimulée.

	Attaques						
	KOA	KMA	WOA	CMA	EOA	Bits embarqués	Extraction
$\mathcal{NW}_{(\eta=1)}$	✗	✗	✓	✗		N	✓
$\mathcal{CIW}_1$	✓	✓	✓	✓	✓	1	✓
$\mathcal{CIS}_2$	✓	✓	✓	✓	✓	N	✓ (contraintes applicatives : $\mathcal{CIS}_5$)
$\mathcal{DI}_3$	✗	✗	✓	✗	✗	N	✓
$\mathcal{CIS}_3$	✗	✗	✓	✗	✗	N	✓
$\mathcal{CIS}_4$			✓			N	✓

TABLE 13.1 – Comparaison des processus de dissimulation d'informations étudiés durant la thèse.

13.2/ PERSPECTIVES

L'achèvement de cette thèse ne constitue pas une fin en soi, mais ouvre des perspectives. Certains travaux réalisés nécessitent, en effet, d'être poursuivis afin d'élargir le champ de leur portée et de générer d'autres champs de prospection scientifique.

Les différents processus de dissimulation d'informations, analysés tout au long de cette thèse, révèlent tous des points forts et des points faibles. Ils peuvent tous présenter un réel intérêt, soit dans un cadre théorique, soit dans un contexte applicatif. C'est pourquoi nous pensons que chacun de ces procédés nécessiterait une étude plus approfondie. Il serait notamment intéressant d'évaluer la complexité algorithmique du processus $\mathcal{CIS}_2$ afin de pouvoir comparer sa vitesse d'exécution avec celle de $\mathcal{DI}_3$.

Il est également prévu de réaliser un second dépôt logiciel à partir de $\mathcal{DI}_3$ afin que nous puissions l'exploiter dans la *plateforme de protection des documents numérique* du projet *Stégosécu ISIS*.

Les sections suivantes présentent quelques perspectives de recherches sur différents aspects.

13.2.1/ LA ROBUSTESSE

S'agissant de la robustesse, l'étude du schéma $\mathcal{DI}_3$, a tout juste été initiée. Cette étude mériterait d'être approfondie.

Rappelons que le programme $\mathcal{CIS}_5$ est une implémentation concrète de l'algorithme $\mathcal{CIS}_2$ qui tient compte des contraintes applicatives liées à sa modélisation formelle. Sur la base de ce programme, nous pourrions mener une étude de robustesse du schéma $\mathcal{CIS}_2$ afin de pouvoir le comparer avec les autres algorithmes de dissimulation d'informations.

Par ailleurs, pour obtenir des résultats encore plus probants, il faudrait définir quels seraient les bons LSC et MSC à sélectionner. Comme il existe de multiples manières de définir les LSC et les MSC, du fait de la diversité des méthodes de description de médias (domaine spatial, domaine fréquentiel, contour de formes, etc.), il paraît donc judicieux d'étudier l'impact de toutes ces méthodes sur la robustesse de chacun des différents processus. Il se pourrait que le domaine de description fréquentiel (pour les médias hôte) fournisse de meilleurs résultats sur le plan de la robustesse. Dans ce cas, le domaine de description fondé sur les ondelettes semble, *a priori*, être davantage pertinent que celui fondé sur les cosinus discrets. L'étude des contours de formes dans les images pourrait également permettre de fournir de bons LSC.

Nous pourrons aussi mener de nouvelles études de robustesse sur des attaques qui n'ont pas encore été étudiées, afin d'améliorer davantage notre compréhension des processus à ce niveau.

13.2.2/ LA STÉGANALYSE

Rappelons que la stéganalyse vise à étudier l'impact des méthodes de dissimulation d'informations du point de vue de la détection du filigrane.

Seul $\mathcal{DI}_3$ a été stéganylisé. De la même manière que nous avons prévu d'étudier la robustesse de l'algorithme $\mathcal{CIS}_2$ à partir de son implémentation $\mathcal{CIS}_5$, nous prévoyons également de stéganalyser $\mathcal{CIS}_2$. Cela nous permettra de pouvoir comparer $\mathcal{CIS}_2$, ou plus précisément son implémentation $\mathcal{CIS}_5$, avec les autres algorithmes de dissimulation d'informations de la littérature.

Comme pour la robustesse mentionnée à la section précédente, il faudrait étendre l'étude de la stéganalyse des différents processus à l'ensemble des méthodes de description de médias. À cette fin, d'autres stéganalyseurs que *HugoBreaker* devraient être testés. Les résultats pourraient ainsi être comparés à un plus grand ensemble de stéganographieurs concurrents. Pour contrecarrer ces outils, à base d'intelligence artificielle, une étude pratique et systématique du bon choix des LSC et des MSC apparaît, à nouveau, comme étant nécessaire.

Il faudrait encore étudier d'autres stéganographieurs, tant du point de vue de la stégosécurité que du point de vue du chaos topologique. Il serait important, à ce niveau, de produire une définition de détectabilité qui devrait utiliser des notions reliées à la théorie de la complexité.

13.2.3/ LA SÉCURITÉ

Nous avons vu qu'il y a deux manières complémentaires d'étudier formellement la sécurité des algorithmes de dissimulation d'informations : l'approche probabiliste (stégo-sécurité) et l'approche topologique (chaos topologique).

Afin d'approfondir nos connaissances des processus de dissimulation d'informations mis au point, il importera d'étudier les nouvelles propriétés topologiques les concernant. Lors de cette étude, nous pourrons analyser la propriété d'expansivité de $\mathcal{CIS}_2$, en tentant de fournir une évaluation de sa constante d'expansivité. Il sera également possible d'étudier son entropie topologique.

À l'issue de ces travaux complémentaires, nous dresserons un nouveau comparatif, plus complet, des processus étudiés face aux autres processus existants. Ce comparatif tiendra compte des nouvelles propriétés topologiques mises en exergue. Une telle synthèse est essentielle pour pouvoir juger de la meilleure adéquation de tel ou tel processus avec des contextes d'applications spécifiques, eu égard à la performance de ces processus ainsi qu'à leur niveau de sécurité et de robustesse. Ce type d'étude a été initié dans [15] pour des applications relatives à l'anonymisation des données sur Internet et dans [16] pour le cas du Web sémantique.

Sur le plan théorique, le lien entre la stégo-sécurité et la sécurité topologique nécessiterait d'être étudié, afin de savoir s'il existe une relation de cause à effet entre chacune de ces deux propriétés. Au début de cette thèse, nous avions commencé à explorer cette piste, en ramenant ces deux notions à leur plus petit dénominateur commun : la topologie. Pour la sécurité topologique ce dénominateur est évident. Concernant la stégo-sécurité, qui est basée sur des notions de probabilités, ce dénominateur correspond bien, lui aussi, à la topologie. En effet, la théorie des probabilités est fondée sur la théorie de la mesure et la théorie des intégrales de Lebesgue, toutes deux s'appuyant sur la théorie des espaces topologiques. Même si dans cette première approche nous n'avons pas réussi à extirper un lien évident, nous pensons que cette piste d'étude reste intéressante et qu'elle doit être explorée davantage.

Pour finir, l'étude théorique de notre ultime algorithme de dissimulation d'informations ($\mathcal{CIS}_4$) nécessiterait d'être menée à son terme. La preuve de sa sécurité topologique nous permettra de le comparer aux autres algorithmes qui sont à la fois stégo-sûrs et topologiquement sûrs. Une fois cette conjecture démontrée rigoureusement, nous pourrons évaluer la constante de sensibilité du schéma ainsi que son exposant de Lyapunov. Comme nous l'avons fait pour $\mathcal{CIS}_2$, nous pourrons alors étudier la complétude et la perfection topologiques de l'espace métrique $\mathcal{X}_4$.

13.2.4/ DISSIMULATION D'INFORMATIONS DANS D'AUTRES MÉDIAS

Afin de pouvoir élargir les perspectives d'exploitation économique de nos technologies, fondées sur la dissimulation d'informations dans des médias numériques, nous avons prévu d'élargir nos expérimentations, études et comparaisons, en utilisant de nouveaux types de médias, autres que les images (vidéos, sons, mails, PDF [1], fichiers Windows Office, fichiers OpenOffice, etc.). Ces nouvelles perspectives laissent présager d'autres possibilités d'exploitation scientifique et économique, en raison de la large utilisation de

1. Portable Document Format

ces nouveaux médias dans le domaine informatique, sur Internet et sur le marché. À la différence des images, qui ne sont utilisées que dans un champ d'application relativement restreint, ces autres médias offrent des possibilités applicatives bien plus larges. Il en va de même pour les débouchés économiques qui en découleront.

Pour pouvoir dégager de réelles perspectives d'exploitations économiques de la théorie de l'information dissimulée, il est impératif que nous allions plus loin dans l'étude scientifique de ces nouveaux médias. Nous prévoyons donc d'analyser le comportement de nos solutions de dissimulation d'informations face à ces nouveaux médias, en mettant en évidence les nouvelles contraintes décelées. Même si l'objectif d'exploitation économique continuera à guider une partie de nos travaux, la réussite de cette phase de valorisation dépendra fortement de notre capacité à mener des recherches scientifiques fondamentales dans ce domaine. Les études de robustesse et de stéganalyse, décrites dans les sections précédentes, pourront alors être conduites de façon analogue en s'intéressant à ces nouveaux médias.

Le cas des documents PDF s'avère prometteur. En effet, dans le monde économique, le travail avec des images est bien souvent à la marge, sauf peut-être pour des domaines d'activités très spécialisés, comme par exemple le traitement d'images dans le secteur de la photographie ou le traitement vidéo. Grâce aux études de terrain que nous avons réalisées auprès de différents prospects, nous avons pu constater que le format le plus utilisé dans le monde économique est le format PDF (outre les fichiers Windows Office utilisés dans le domaine administratif). Nous proposons donc, dans un contexte scientifique et applicatif, de focaliser nos efforts sur l'étude de la dissimulation d'informations dans les documents PDF.

De plus, en tenant compte des premiers retours du marché, obtenus lors des études marketing préalablement menées (*Stégosécu ISIS*) en collaboration avec des cabinets externes spécialisés, nous avons pu constater que le format de document PDF était, en effet, celui le plus largement utilisé dans le domaine informatique, sur Internet et dans bien d'autres secteurs du marché. Nous avons pu constater que les documents PDF sont en meilleure adéquation avec les processus de gestion interne des entreprises (B2B [85]), avec les habitudes de travail de leur personnel et avec les utilisateurs (B2C [86]) de leurs produits informatiques. Ces tiers (personnes physiques ou personnes morales) constitueront les prospects vers lesquels s'orientera la *plateforme de protection des documents numériques* que la future société *Stégosécu ISIS* sera à même de leur proposer.

Ainsi, comme nous l'avons signalé précédemment, est-il d'une importance cruciale (pour garantir la compétitivité et l'avance technologique de la future société *Stégosécu ISIS*), de continuer à approfondir nos recherches scientifiques dans ce domaine. À cette fin, nous mettrons, notamment, en application les résultats mis en évidence par plusieurs membres de notre équipe de recherche [3]. Ces travaux concernent la dissimulation d'informations dans des documents PDF, avec une perspective d'intégration directe de cette technologie au sein de la *plateforme de protection des documents numériques*, produit phare de *Stégosécu ISIS*.

Les premiers résultats de ces recherches ont été prometteurs. C'est pourquoi il est primordial que nous puissions améliorer notre connaissance des techniques de protection des documents PDF et de dissimulation d'informations dans ce type de médias. Il s'agira alors de mettre en évidence des protocoles de protection, dont nous devrons garantir la sécurité par des modèles mathématiques précis, comme nous l'avons fait, jusqu'ici dans cette thèse.

Cette démarche de recherches ultérieures, s'inscrit évidemment dans le programme de Recherche et Développement de la future société *Stégosécu ISIS*. Ce programme pourra, le cas échéant, être conduit par l'équipe de recherche.

13.2.5/ AUTRES CONTEXTES D'ÉTUDE DE LA DISSIMULATION D'INFORMATIONS

Grâce à nos multiples contributions, nous disposons, à présent, d'une meilleure visibilité et d'une meilleure connaissance des contraintes liées à la dissimulation d'informations, à la sécurisation des données et à la protection des documents numériques.

Cela nous permet d'envisager l'étude d'autres domaines de recherches périphériques dans lesquels nous pourrons mettre à profit ces connaissances, à partir du recul que nous avons aujourd'hui.

Nous pourrons notamment analyser comment nous pourrions appliquer nos résultats dans le domaine des réseaux de capteurs, domaine qui est un axe de recherche lui aussi développé au sein de notre équipe.

Nous pourrons aussi étudier dans quelle mesure la dissimulation d'informations peut être appliquée au domaine de la signature numérique, en proposant de nouveaux protocoles de signatures. Ces protocoles pourraient être fondés sur des techniques de tatouage numérique. En fonction des résultats obtenus lors de ces futures études, nous pourront envisager le transfert de ces technologies à la future société *Stégosécu ISIS*.

13.2.6/ GÉNÉRATEURS DE NOMBRES PSEUDO-ALÉATOIRES

Dans un domaine connexe à celui de la dissimulation d'informations, mais faisant lui aussi partie du contexte général de la sécurité informatique, il sera également possible de mener de nouvelles études sur les générateurs de nombres pseudo-aléatoires chaotiques. Nous pourrons alors comparer ces générateurs chaotiques avec d'autres générateurs, après leur avoir fait passer les tests statistiques classiquement utilisés pour quantifier la performance des PRNG. Nous utiliserons alors les suites de tests bien connues comme NIST [2] [67], la batterie DieHARD [56] ou le test strict TestU01 [52], qui est aujourd'hui le test le plus pointu pour l'évaluation de la qualité des PRNG.

Nous pourrons également réaliser de nouveaux tests statistiques sur le générateur "Blum-Blum-Shub" [46, 81], dans le but de juger du lien qui existe entre la performance de ce type de générateur et la taille des nombres premiers dont il dépend.

En fonction des résultats obtenus, la perspective d'une valorisation économique de ces travaux pourra aussi être envisagée.

2. National Institute of Standards and Technology of the U.S. Government

Bibliographie

[1] Delicious social bookmarking, http ://delicious.com/. Retrieved June, 2012 from http: //delicious.com/.

[2] The frick collection, http ://www.frick.org/. Retrieved June, 2012 from http://www.frick. org/.

[3] Jean-François Couchot Ahmad Bitar, Rony Darazi and Raphaël Couturier. Steganography in pdf documents using spread transform dither modulation. page 5p. Université Antonine, Baabda, Lebanon - www.upa.edu.lb, Université de Franche Comté, Belfort, France - www.univ-fcomte.fr, 2014.

[4] M. R. Anderberg. *Cluster Analysis for Applications*. Academic Press, 1973.

[5] ANR. Astrid. http://www.agence-nationale-recherche.fr/financer-votre-projet/ appel-detail/accompagnement-specifique-des-travaux-de-recherches-et-d-innovation-defense-astr 2014. [En ligne ; Page disponible le 11 mai 2014].

[6] J. M. Bahi and S. Contassot-Vivier. Stability of fully asynchronous discrete-time discrete state dynamic networks. *IEEE Transactions on Neural Networks*, 13(6) :1353–1363, 2002.

[7] Jacques Bahi and Sylvain Contassot-Vivier. Basins of attraction in fully asynchronous discrete-time discrete-state dynamic networks. *IEEE Transactions on Neural Networks*, 17(2) :397–408, 2006.

[8] Jacques Bahi, Nicolas Friot, and Christophe Guyeux. Lyapunov exponent evaluation of a digital watermarking scheme proven to be secure. In *IIH-MSP'2012, 8-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages 359–362, Piraeus-Athens, Greece, July 2012. IEEE Computer Society.

[9] Jacques Bahi, Nicolas Friot, and Christophe Guyeux. Topological study and lyapunov exponent of a secure steganographic scheme. In Javier Lopez and Pierangela Samarati, editors, *SECRYPT'2013, Int. Conf. on Security and Cryptography. SECRYPT is part of ICETE - The International Joint Conference on e-Business and Telecommunications*, pages ***–***, Reykjavik, Iceland, July 2013. SciTePress. 8 pages. To appear.

[10] Jacques Bahi and Christophe Guyeux. Hash functions using chaotic iterations. *Journal of Algorithms & Computational Technology*, 4(2) :167–181, 2010.

[11] Jacques Bahi and Christophe Guyeux. A new chaos-based watermarking algorithm. In *SECRYPT 2010, International conference on security and cryptography*, Athens, Greece, 2010. To appear.

[12] Jacques Bahi and Christophe Guyeux. A new chaos-based watermarking algorithm. In *SECRYPT'10, Int. conf. on security and cryptography*, pages 455–458, Athens, Greece, July 2010. SciTePress.

[13] Jacques M. Bahi. *Algorithmes asynchrones pour des systèmes différentiels-algébriques. Simulation numérique sur des exemples de circuits électriques*. PhD thesis, Université de Franche-Comté, 1991.

[14] Jacques M. Bahi. *Méthodes itératives dans des espaces produits. Application au calcul parallèle*. Habilitation Ã diriger des recherches, Université de Franche-Comté, 1998.

[15] Jacques M. Bahi, François Couchot, Nicolas Friot, and Christophe Guyeux. Application of steganography for anonymity through the internet. In *IHTIAP'2012, The First Workshop on Information Hiding Techniques for Internet Anonymity and Privacy*, pages ***–***, Venice, Italy, June 2012. To appear.

[16] Jacques M. Bahi, François Couchot, Nicolas Friot, and Christophe Guyeux. A robust data hiding process contributing to the development of a semantic web. In *INTERNET'2012, The Fourth International Conference on Evolving Internet*, pages ***–***, Venice, Italy, June 2012. To appear.

[17] Jacques M. Bahi, Jean-François Couchot, and Christophe Guyeux. Steganography : A class of secure and robust algorithms. *The Computer Journal*, 55(6) :653–666, 2012.

[18] Jacques M. Bahi and Christophe Guyeux. A chaos-based approach for information hiding security. arXiv N^o 0034939, April 2010.

[19] J. Banks, J. Brooks, G. Cairns, and P. Stacey. On devaney's definition of chaos. *Amer. Math. Monthly*, 99 :332–334, 1992.

[20] P. Bas, T. Filler, and T. Pevný. Break our steganographic system — the ins and outs of organizing boss. In T. Filler, editor, *Information Hiding, 13th International Workshop*, Lecture Notes in Computer Science, Prague, Czech Republic, May 18–20, 2011. Springer-Verlag, New York.

[21] Dimitri P. Bertsekas and John N. Tsitsiklis. Parallel and distributed iterative algorithms : a selective survey, 1988.

[22] Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and distributed computation : numerical methods*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[23] Christian Cachin. An information-theoretic model for steganography. *Information and Computation*, 192 :41 – 56, 2004.

[24] Francois Cayre, Caroline Fontaine, and Teddy Furon. Kerckhoffs-based embedding security classes for woa data hiding. *IEEE Transactions on Information Forensics and Security*, 3(1) :1–15, 2008.

[25] François Cayre, Caroline Fontaine, and Teddy Furon. Watermarking security : theory and practice. *IEEE Transactions on Signal Processing*, 53(10) :3976–3987, 2005.

[26] D. Chazan and W. Miranker. Chaotic relaxation. *Linear algebra and its applications*, pages 199–222, 1969.

[27] Jeremy Clark, P. C. van Oorschot, and Carlisle Adams. Usability of anonymous web browsing : an examination of tor interfaces and deployability. In *Proceedings of the 3rd symposium on Usable privacy and security*, SOUPS '07, pages 41–51, New York, NY, USA, 2007. ACM.

[28] Ingemar J. Cox, Senior Member, Joe Kilian, F. Thomson Leighton, and Talal Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6 :1673–1687, 1997.

[29] Ministère de l'Enseignement Supérieur et de la Recherche. Pepite - tremplin pour l'entrepreneuriat Étudiant. http://www.enseignementsup-recherche.gouv.fr/cid77179/ouverture-des-inscriptions-au-prix-pepite-tremplin-pour-l-entrepreneuriat-etudiant.html, 2014. [En ligne ; Page disponible le 11 mai 2014].

[30] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems, 2nd Edition*. Westview Pr., March 2003.

[31] ESIEA. Perseus technology for anonymity through the internet, 02 2012. [On line - 2012.02.22].

[32] ESIEA. Perseus technology for anonymity through the internet - firefox plugin, 02 2012. [On line - 2012.02.22].

[33] Eric Filiol. Passive and active leakage of secret data from non networked computer. In *Black Hat*, 2008.

[34] Enrico Formenti. *Automates cellulaires et chaos : de la vision topologique à la vision algorithmique*. PhD thesis, École Normale Supérieure de Lyon, 1998.

[35] Enrico Formenti. *De l'algorithmique du chaos dans les systèmes dynamiques discrets*. PhD thesis, Université de Provence, 2003.

[36] Jessica J. Fridrich, Tomás Pevný, and Jan Kodovský. Statistically undetectable jpeg steganography : dead ends challenges, and opportunities. In Deepa Kundur, Balakrishnan Prabhakaran, Jana Dittmann, and Jessica J. Fridrich, editors, *MM&Sec*, pages 3–14. ACM, 2007.

[37] Nicolas Friot, Christophe Guyeux, and Jacques Bahi. Chaotic iterations for steganography - stego-security and chaos-security. In Javier Lopez and Pierangela Samarati, editors, *SECRYPT'2011, Int. Conf. on Security and Cryptography. SECRYPT is part of ICETE - The International Joint Conference on e-Business and Telecommunications*, pages 218–227, Sevilla, Spain, July 2011. SciTePress.

[38] T. Furon. Security analysis, 2002. European Project IST-1999-10987 CERTIMARK, Deliverable D.5.5.

[39] Teddy Furon. A survey of watermarking security. In Mauro Barni, Ingemar J. Cox, Ton Kalker, and Hyoung Joong Kim, editors, *IWDW*, volume 3710 of *Lecture Notes in Computer Science*, pages 201–215, Siena, Italy, September 15-17 2005. Springer.

[40] Jean-Pierre Goedgebuer, Laurent Larger, and Henri Porte. Optical cryptosystem based on synchronization of hyperchaos generated by a delayed feedback tunable laser diode. *Phys. Rev. Lett.*, 80 :2249–2252, Mar 1998.

[41] Christophe Guyeux. *Le désordre des itérations chaotiques et leur utilité en sécurité informatique*. PhD thesis, Université de Franche-Comté, 2010.

[42] Christophe Guyeux and Jacques Bahi. An improved watermarking algorithm for internet applications. In *INTERNET'2010. The 2nd Int. Conf. on Evolving Internet*, pages 119–124, Valencia, Spain, September 2010.

[43] Christophe Guyeux, Nicolas Friot, and Jacques Bahi. Chaotic iterations versus spread-spectrum : chaos and stego security. In *IIH-MSP'10, 6-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages 208–211, Darmstadt, Germany, October 2010.

[44] Stégosécu ISIS. Site web de la future sas. http://www.stegosecu.fr, 2014. [En ligne ; Page disponible le 11 mai 2014].

[45] Robert Jenkins. Isaac. In Dieter Gollmann, editor, *Fast Software Encryption*, volume 1039 of *Lecture Notes in Computer Science*, pages 41–49. Springer Berlin / Heidelberg, 1996. 10.1007/3-540-60865-6_41.

[46] P. Junod. *Cryptographic secure pseudo-random bits generation : The Blum-Blum-Shub generator*. August, 1999.

[47] T. Kalker. Considerations on watermarking security. In *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*, pages 201–206, 2001.

[48] Stefan Katzenbeisser and Fabien A. Petitcolas, editors. *Information Hiding Techniques for Steganography and Digital Watermarking*. Artech House, Inc., Norwood, MA, USA, 2000.

[49] Auguste Kerckhoffs. La cryptographie militaire. *Journal des sciences militaires*, IX :5–83, January 1883.

[50] J. Kodovský and J. Fridrich. Steganalysis in high dimensions : fusing classifiers built on random subspaces. In *Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics XIII, San Francisco, CA*, January 2011.

[51] J. Kodovský, J. Fridrich, and V. Holub. Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics and Security*, PP Issue :99 :1 – 1, 2011. To appear.

[52] P. L'Ecuyer and R. Simard. Testu01 : A software library in ansi c for empirical testing of random number generators. *Laboratoire de simulation et d'optimisation. Université de Montréal IRO*, 2009.

[53] Marie-Jeanne Lesot and Maria Rifqi. Order-based equivalence degrees for similarity and distance measures. In *IPMU*, pages 19–28, 2010.

[54] H.S. Malvar and D. Florêncio. Improved spread spectrum : A new modulation technique for robust watermarking. *IEEE Trans. Signal Proceeding*, 53 :898–905, 2003.

[55] Yongyi Mao and Xiang Chen. An encryption algorithm of chaos based on sine square mapping. In *Proceedings of the 2011 Fourth International Symposium on Computational Intelligence and Design - Volume 01*, ISCID '11, pages 131–134, Washington, DC, USA, 2011. IEEE Computer Society.

[56] G. Marsaglia. Diehard : a battery of tests of randomness. *http ://stat.fsu.edu/ geo/diehard.html*, 1996.

[57] J. A. Martínez-Ñonthe, A. Díaz-Méndez, M. Cruz-Irisson, L. Palacios-Luengas, J. L. Del-Río-Correa, and R. Vázquez-Medina. Cryptosystem with one dimensional chaotic maps. In *Proceedings of the 4th international conference on Computational intelligence in security for information systems*, CISIS'11, pages 190–197, Berlin, Heidelberg, 2011. Springer-Verlag.

[58] J.-C. Miellou. Algorithmes de relaxation chaotique à retards. *Rairo*, R1 :148–162, 1975.

[59] J.-C. Miellou. Itérations chaotiques à retards, étude de la convergence dans le cas d'espaces partiellement ordonnés. *C.R.A.S. Paris*, 280 :233–236, 1975.

[60] Thomas Mittelholzer. An information-theoretic approach to steganography and watermarking. In Andreas Pfitzmann, editor, *Information Hiding*, volume 1768 of *Lecture Notes in Computer Science*, pages 1–16, Dresden, Germany, September 29 - October 1. 1999. Springer.

[61] Jacques Bahi Nicolas Friot and Christophe Guyeux. CIS-5, programme sécurisé de stéganographie basé sur les itérations chaotiques, jan 2012.

[62] Luis Perez-Freire, Pedro Comesanñ, Juan Ramon Troncoso-Pastoriza, and Fernando Perez-Gonzalez. Watermarking security : a survey. In *LNCS Transactions on Data Hiding and Multimedia Security*, 2006.

[63] Luis Perez-Freire, F. Pérez-gonzalez, and Pedro Comesañ. Secret dither estimation in lattice-quantization data hiding : A set-membership approach. In Edward J. Delp and Ping W. Wong, editors, *Security, Steganography, and Watermarking of Multimedia Contents*, San Jose, California, USA, January 2006. SPIE.

[64] Tomás Pevný, Tomás Filler, and Patrick Bas. Using high-dimensional image models to perform highly undetectable steganography. In Rainer Böhme, Philip W. L. Fong, and Reihaneh Safavi-Naini, editors, *Information Hiding*, volume 6387 of *Lecture Notes in Computer Science*, pages 161–177. Springer, 2010.

[65] Agence pour la Protection des Programmes. Agence pour la protection des programmes (app). http://www.app.asso.fr, 2014. [En ligne ; Page disponible le 11 mai 2014].

[66] Tor Project. Tor : Anonymity online - protect your privacy. defend yourself against network surveillance and traffic analysis, 02 2012. [On line - 2012.02.22].

[67] NIST Special Publication 800-22 rev. 1. A statistical test suite for random and pseudorandom number generators for cryptographic applications. NIST, August 2008.

[68] Maria Rifqi, Vincent Berger, and Bernadette Bouchon-Meunier. Discrimination power of measures of comparison. *Fuzzy Sets and Systems*, 110 :189–196, 2000.

[69] Maria Rifqi, Marcin Detyniecki, and Bernadette Bouchon-Meunier. Discrimination power of measures of resemblance. In *IFSA'03*, 2003.

[70] François Robert. *Discrete Iterations : A Metric Study*, volume 6 of *Series in Computational Mathematics*. Springer-Verlag, 1986.

[71] Sylvie Ruette. *Chaos en dynamique topologique, en particulier sur l'intervalle, mesures d'entropie maximale*. PhD thesis, Université d'Aix-Marseille II, 2001.

[72] Laurent Schwartz. *Analyse : topologie générale et analyse fonctionnelle*. Hermann, 1980.

[73] Claude E. Shannon. Communication theory of secrecy systems. *Bell Systems Technical Journal*, 28 :656–715, 1949.

[74] Li Shujun, Li Qi, Li Wenmin, Mou Xuanqin, and Cai Yuanlong. Statistical properties of digital piecewise linear chaotic maps and their roles in cryptography and pseudorandom coding. *Proceedings of the 8th IMA International Conference on Cryptography and Coding*, 1 :205–221, 2001.

[75] Gustavus J. Simmons. The prisoners' problem and the subliminal channel. In *Advances in Cryptology, Proc. CRYPTO'83*, pages 51–67, 1984.

[76] Kaushal Solanki, Anindya Sarkar, and B. S. Manjunath. Yass : Yet another steganographic scheme that resists blind steganalysis. In Teddy Furon, François Cayre, Gwenaël J. Doërr, and Patrick Bas, editors, *Information Hiding*, volume 4567 of *Lecture Notes in Computer Science*, pages 16–31. Springer, 2007.

[77] MahlonC. Stacy, KurtE. Augustine, and RichardA. Robb. Image boss : An image database system designed for research. *Journal of Digital Imaging*, 10(1) :56–59, 1997.

[78] K. Tanaka, Y. Nakamura, and K. Matsui. Embedding secret information into a dithered multi-level image. In *IEEE Military Communications Conference*, 1990.

[79] CNIT GAUSS-CNRS UVIGO, UNIGE. First summary report on fundamentals. Technical report, ECRYPT, European Network of Excellence in Cryptology, march 2005. Retrieved from www.ecrypt.eu.org/ecrypt1/documents/D.WVL.1-2.0p.pdf.

[80] Tirkel Rankin Van. Electronic water mark, 1993.

[81] Umesh Vazirani and Vijay Vazirani. Efficient and secure pseudo-random number generation (extended abstract). In George Blakley and David Chaum, editors, *Advances in Cryptology*, volume 196 of *Lecture Notes in Computer Science*, pages 193–202. Springer Berlin / Heidelberg, 1985. 10.1007/3-540-39568-7_17.

[82] Yong Wang, Kwok-Wo Wong, Xiaofeng Liao, and Tao Xiang. A block cipher with dynamic s-boxes based on tent map. *Communications in Nonlinear Science and Numerical Simulation*, 14(7) :3089 – 3099, 2009.

[83] Wikipédia. Espace complet — wikipédia, l'encyclopédie libre, 2010. [En ligne ; Page disponible le 9-août-2010].

[84] Wikipédia. Analyse constructive. http://fr.wikipedia.org/wiki/Analyse_constructive, 2014. [En ligne ; Page disponible le 11 mai 2014].

[85] Wikipédia. Business to business (b2b). http://fr.wikipedia.org/wiki/Business_to_business, 2014. [En ligne ; Page disponible le 11 mai 2014].

[86] Wikipédia. Business to consumer (b2c). http://fr.wikipedia.org/wiki/Business_to_consumer, 2014. [En ligne ; Page disponible le 11 mai 2014].

[87] Wikipédia. Discrete cosine transform. http://en.wikipedia.org/wiki/Discrete_cosine_transform, 2014. [En ligne ; Page disponible le 11 mai 2014].

[88] Wikipédia. Démonstration constructive. http://fr.wikipedia.org/wiki/D%C3%A9monstration_constructive, 2014. [En ligne ; Page disponible le 11 mai 2014].

[89] Wikipédia. Microcontrôleur. http://fr.wikipedia.org/wiki/Microcontr%C3%B4leur, 2014. [En ligne ; Page disponible le 11 mai 2014].

[90] Wikipédia. Wavelet. http://en.wikipedia.org/wiki/Wavelet, 2014. [En ligne ; Page disponible le 11 mai 2014].

[91] Araminta Wordsworth. News in National Post.com, February 2012. Availlable at http://news.nationalpost.com/2012/02/13/hamza-kashgari/.

[92] G.U. Yule and MG Kendall. An introduction to the theory of. *Statistics*, 1950.

Résumé :

Les systèmes dynamiques discrets, œuvrant en itérations chaotiques ou asynchrones, se sont avérés être des outils particulièrement intéressants à utiliser en sécurité informatique, grâce à leur comportement hautement imprévisible, obtenu sous certaines conditions. Ces itérations chaotiques satisfont les propriétés de chaos topologique et peuvent être programmées de manière efficace. Dans l'état de l'art, elles ont montré tout leur intérêt au travers de schémas de tatouage numérique. Toutefois, malgré leurs multiples avantages, ces algorithmes existants ont révélé certaines limitations. Cette thèse a pour objectif de lever ces contraintes, en proposant de nouveaux processus susceptibles de s'appliquer à la fois au domaine du tatouage numérique et au domaine de la stéganographie. Nous avons donc étudié ces nouveaux schémas sur le double plan de la sécurité topologique et de la sécurité dans un cadre probabiliste . L'analyse de leur niveau de sécurité respectif a permis de dresser un comparatif avec les autres processus existants comme, par exemple, l'étalement de spectre. Des tests applicatifs ont été conduits pour stéganalyser les processus proposés et pour évaluer leur robustesse. Grâce aux résultats obtenus, nous avons pu juger de la meilleure adéquation de chaque algorithme avec des domaines d'applications ciblés comme, par exemple, l'anonymisation sur Internet, la contribution au développement d'un web sémantique, ou encore une utilisation pour la protection des documents et des données numériques. Parallèlement à ces travaux scientifiques fondamentaux, nous avons proposé plusieurs projets de valorisation avec pour objectif la création d'une entreprise de technologies innovantes.

Mots-clés : *dissimulation d'informations; étalement de spectre; exposant de Lyapunov; Internet; itérations chaotiques; modèles mathématiques; preuves de sécurité; protection des documents numériques; robustesse; sécurisation des données; sécurité informatique; sécurité topologique; stéganalyse; stéganographie; stégo-sécurité; systèmes dynamiques discrets; tatouage numérique; tests applicatifs; théorie du chaos; topologie; vie privée.*

Abstract:

Discrete dynamical systems by chaotic or asynchronous iterations have proved to be highly interesting tools in the field of computer security, thanks to their unpredictible behavior obtained under some conditions. More precisely, these chaotic iterations possess the property of topological chaos and can be programmed in an efficient way. In the state of the art, they have turned out to be really interesting to use notably through digital watermarking schemes. However, despite their multiple advantages, these existing algorithms have revealed some limitations. So, these PhD thesis aims at removing these constraints, proposing new processes which can be applied both in the field of digital watermarking and of steganography. We have studied these new schemes on two aspects: the topological security and the security based on a probabilistic approach. The analysis of their respective security level has allowed to achieve a comparison with the other existing processes such as, for example, the spread spectrum. Application tests have also been conducted to steganalyse and to evaluate the robustness of the algorithms studied in this PhD thesis. Thanks to the obtained results, it has been possible to determine the best adequation of each processes with targeted application fields as, for example, the anonymity on the Internet, the contribution to the development of the semantic web, or their use for the protection of digital documents. In parallel to these scientific research works, several valorization perspectives have been proposed, aiming at creating a company of innovative technology.

Keywords: *application tests; chaotiques iterations; chaos theory; computer security; data hiding; data security; digital Watermarking; discret dynamycal systems; Internet; lyapunov exponent; mathematics models; privacy; protection of digital documents; robustness; security proofs; spread spectrum; steganalysis; steganography; stego-security; topological security; topology.*

SPIM

■ École doctorale SPIM 16 route de Gray F - 25030 Besançon cedex
■ tél. +33 (0)3 81 66 66 02 ■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr

UFC
UNIVERSITÉ DE FRANCHE-COMTÉ