La République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Saad Dahleb Blida 1

Institut d'Aéronautique et des études Spatiales (IAES)
Département Etudes Spatiales

Mémoire de fin d'études

En vue de l'obtention du diplôme **de Master**

En **: Aéronautique**

Option **: Télécommunications Spatiales**

Thème :

Détection intelligente de l'effet d'attaque malveillante dans un système de radio utilisant la distance Robuste et la distance Mahalanobis.

Proposé et encadré par :

Dr MOUMENA Ahmed.

Co-encadreur : Dr Sofiane TAHRAOUI.

Réalisé par :

BOUANZOUL Abdelkrim.

Année Universitaire : 2022 / 2023

Remerciement

Remerciement

En tout premier lieu, je remercie le bon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés, permis de mener à bien ce travail.

Je voudrais dans un premier temps remercier, mon encadreur Mr MOUMENA Ahmed je le remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant notre préparation de ce mémoire.

Nous remercions également le président et les membres du jury qui nous ferons l'honneur de juger notre travail.

Enfin, J'adresse mes sincères remerciements à tous les professeurs, intervenants et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté de me rencontrer et de répondre à mes questions durant mes recherches.

Dédicace

Dédicace :

Je dédie ce projet a :

À mes chers parents,

Et toute ma famille,

Grace à eux que je suis là, que dieu les bénisse et les protège

À tous ceux qui m'ont soutenu et m'ont encouragé depuis mon premier pas dans cette vie.

Ainsi mes amis

Et tous ceux qui m'aiment …….

BOUANZOUL Abdelkrim.

Résumé :

Les menaces à la sécurité de la couche physique ont évolué à partir d'attaques malveillantes dans les systèmes sans fil, en raison de leur nature furtive, rendant les systèmes de communication sans fil vulnérables. Dans ce travail, nous avons proposé un convertisseur modulé large bande centralisé (CMLB-C) combiné à un détecteur de classification basé sur la distance Mahalanobis (DM) et (distance robuste) DR. Le signal reçu par chaque récepteur radio dans chaque canal passe par différentes étapes pour obtenir un taux d'échantillonnage inférieur à Nyquist. Chaque récepteur fournit un taux minimum d'échantillons sur la base de la théorie d'écoute compressive (EC). Toutes les observations compressées de chaque canal sont rassemblées sous la forme d'une matrice appelée matrice de données compressées, qui est considérée directement comme l'entrée du classifieur proposé appelé DM-DR au niveau du centre de fusion (CF). Les résultats obtenus à partir de la simulation utilisant MATLAB et LIBRA montrent les effets d'attaque parfaitement dans le spectre en cas de présence de l'attaque basé sur deux hypothèses H_0 et H_1 pour faire la distinction. L'évaluation des performances est effectuée en termes de taux de détection d'anomalies basé sur la valeur seuil de chaque distance. En employant l'une des techniques d'apprentissage machine (ML) nommée ACP basée sur DD et DR, la performance de ce nouveau système proposé est bonne.

Mots clés : CMLB-C, DM, DR, EC, attaque, anomalie, classifieur.

Abstract :

Physical-layer security threats have evolved from malicious attacks in wireless systems, due to their furtive nature, make wireless communication systems vulnerable. In this work we proposed a centralized modulated wideband converter (C-MWC) combined with classifier detector based Mahalanobis distance (MD) and (robust distance) RD. The received signal at each radio receiver in each channel pass by different steps to realize sub-Nyquist sampling rate. Every receiver gives minimum rate of samples based compressed sensing (CS) theory. All compressed observations from each channel are collected in the form of matrix called compressed data matrix, which is considered directly as the input of the proposed classifier called MD-RD in the level of fusion center (FC). The results obtained from simulation using MATLAB and LIBRA show that this new proposed system the effects of attack perfectly in the presence of attack in the spectrum based on two hypotheses H_0 and H_1 to make distinction. Performance evaluation is performed in terms of anomaly detection rate-based threshold value of each distance. By employing one of the machine learning (ML) techniques-called PCA based MD and RD, the performance of this new proposed system shows good.

Key-words: C-MWC, MD, RD, CS, attack, anomaly, classifier.

Tables des matières

Table Des Matières :

Chapitre 2 : Sécurité de la couche physique au niveau la radio

Chapitre 3 : Théorie d'écoute compressive combinée avec un détecteur d'attaque basé sur la distance Mahalanobis et la distance Robuste

Chapitre 4 : Résultats de simulations et discussion

Liste Des Figures :

Liste des tableaux :

Liste des abréviations :

AWGN : Additive White Gaussian Noise.

CAI : Convertisseur Analogique Information.

CAN : Conversion Analogique-Numérique.

CF : Centre de Fusion.

CMLB : Convertisseur Modulé à Large Bande.

CMLB-C: Convertisseur Modulé à Large Bande Centralisé.

CWI : Continuous Wave Interference.

DM : Distance Mahalanobis.

DoS : Déni de service.

DS : Distance Robuste.

EC : Échantillonnage Compressif.

FFT : Fast Fourier Transform.

FPB : Filtre Passe-Bas.

GPS : Global Positioning System.

LIBRA : Library Robust Analysis.

LiDAR : Light Dtection and Ranging.

MCD : Modèle Conceptuel des données.

MCDCOV : Minimum Covariance Determinant.

OSI : Open Systems Interconnection.

PCA : Analyse des Composantes Principales.

PUE : Primary User Emulator.

RC : Radio Cognitive.

RCES : Réseau Coopératif d'Ecoute du Spectre

RRC : Réseaux de Radio Cognitive.

UP : Utilisateur Primaire.

US : Utilisateur Secondaire.

Introduction Générale

Introduction Générale :

Les systèmes de communication sont des cibles de choix pour les attaquants, car ils sont souvent déployés sur des réseaux largement accessibles et interconnectés. Les risques d'attaques peuvent inclure des interceptions, d'injection, de manipulation. De nombreux types d'attaques ont été expliqués dans littérature ces dernières années, mettant en évidence la nécessité de protéger les systèmes de communication contre les menaces potentielles [1].

A cet effet, l'écoute de spectre large bande est une méthode d'analyse des signaux électromagnétiques qui permet de détecter et d'identifier les différentes fréquences présentes dans un environnement donné [2]. Cette technique est utilisée dans de nombreux domaines, notamment en radiocommunication, en surveillance de l'environnement électromagnétique, en sécurité et en défense. Elle permet d'écouter les signaux émis par les équipements de communication, les radars, les émetteurs de télévision et de radio, ainsi que les interférences électromagnétiques indésirables.

En effet, la sécurité de la couche physique au niveau de la radio concerne la protection des communications sans fil contre les attaques malveillantes. Cette couche est la première couche du modèle OSI (Open Systems Interconnection) et assure la transmission des données entre les équipements de communication à travers l'air. Les menaces qui pèsent sur la sécurité de cette couche comprennent la captation des signaux, le brouillage, l'interception, la manipulation et la falsification [3].

L'Analyse des Composantes Principales (ACP) est une technique d'apprentissage automatique populaire utilisée pour réduire la dimensionnalité des données et la classification. Cette technique a montré des résultats prometteurs en termes de précision et de temps de calcul, ce qui en fait une méthode intéressante pour l'analyse de données volumineuses [4], [5].

Parmi les caractéristiques de l'ACP, La distance de Mahalanobis est une mesure statistique largement utilisée qui est sensible aux anomalies dans la matrice de données multidimensionnelles. Elle prend en compte la covariance des variables, ce qui permet de mieux comprendre les relations complexes entre celles-ci. Cette distance est souvent utilisée dans des domaines tels que la classification d'images, le traitement du signal et la reconnaissance de formes. En ce qui concerne l'autre caractéristique de l'ACP qui est la distance robuste, est une mesure robuste aux anomalies basé sur un estimateur statistique robuste MCD (Minimum

Covariance Determinant) par rapport à la distance Mahalanobis basé sur un estimateur classique S [6].

La combinaison entre le CMLB centralisé et le détecteur par classification est une approche intéressante pour résoudre le problème de la détection de l'effet d'attaque dans un environnement en évolution. Le travail proposé utilise le CMLB-C (Convertisseur Modulé Large Bande-centralisé) pour réaliser la théorie de sous Nyquist, tandis que le détecteur proposé est utilisé pour détecter l'effet d'attaque dans le spectre d'une façon intelligente basé sur la valeur de seuil pour faire une distinction entre le taux normal et le taux d'anomalie. Les résultats de l'étude du nouveau travail proposer ont montré de bons résultats peut détecter efficacement les effets d'attaque dans le spectre large bande.

Pour ce faire nous avons présenté ce travail de la façon suivante :
En plus d'une introduction générale et une conclusion générale, qui résume notre étude, le présent travail effectué en quatre chapitres comme suite :

- ✓ Dans le premier chapitre nous présentons l'écoute compressive.
- ✓ Dans le deuxième chapitre, nous allons mettre l'accent sur les différents types d'attaques qui peuvent perturber les systèmes de communication sans fil et comment on trouve une solution intelligente basé sur l'apprentissage automatique.
- ✓ Dans le troisième chapitre nous allons étudier le convertisseur modulé large bande centralisé combiné avec le détecteur de classification des données.
- ✓ Dans le dernier chapitre nous expliquons les résultats de simulation et discussion obtenus.

Contributions

Nous résumons les contributions et la nouveauté de ce travail comme suit :

- ❖ Sur le plan scientifique, il s'agit d'une nouvelle méthode par rapport aux autres techniques qui ont été développées ces dernières années dans littérature.
- ❖ Sur le plan technologique, notre travail se concentre sur la détection rapide et intelligente et dans un régime large bande du problème d'anomalie en présence des effets de l'attaqueur dans un système de communication sans fil.
- ❖ Utilisation d'un signal multi-bande analogique satellitaire GPS.
- ❖ CMLB centralisé basée sur la théorie d'EC est combiné au détecteur proposé sur la distance Mahalanobis et la distance Robuste pour la prise de décision.

- Utilisation de la structure CMLB pour réduire la consommation d'énergie de chaque RC, pour minimiser le prix et de réduire la complexité du calcul afin d'obtenir une détection plus rapide dans un régime large bande.
- Une solution intelligente est proposée pour améliorer la sécurité et la défense au niveau de la couche physique contre les attaques malveillantes dans le spectre.

Chapitre 1

Etat de l'art de l'écoute compressive

1.1 Introduction :

La radio cognitive est une technique de communication sans fil qui permet de détecter, analyser et adapter en temps réel l'utilisation des fréquences radio disponibles, en fonction des besoins des utilisateurs et des contraintes du réseau. Elle utilise des algorithmes sophistiqués et des systèmes intelligents pour exploiter les bandes de fréquences inutilisées, optimiser l'efficacité du spectre radio et améliorer la qualité de service. Dans ce chapitre, divers algorithmes d'écoute du spectre à large bande de sous Nyquist sont présentés. Une attention très particulière est accordée à l'utilisation des techniques de sous-Nyquist, y compris l'écoute compressive et les techniques d'échantillonnage de sous-Nyquist à canaux multiples [7].

1.2 Ecoute compressive :

L'écoute du spectre dans le régime à large bande est confrontée à des défis techniques considérables réside dans les taux d'échantillonnage très élevés requis par les méthodes conventionnelles d'estimation spectrale qui doivent fonctionner à la fréquence de Nyquist ou à une fréquence supérieure [8].

Le théorème d'échantillonnage de Nyquist nous dit que pour ne pas perdre d'information lors de l'échantillonnage uniforme d'un signal nous devons échantillonner au moins deux fois plus vite que sa largeur de bande, dans de nombreuses applications, le taux de Nyquist peut être si élevé que nous nous retrouvons avec trop d'échantillons et que nous devons les compresser pour les stocker ou les transmettre, il est très coûteux d'augmenter le taux ou la densité d'échantillonnage au-delà de l'état actuel de la technique.

Dans cette section, un nouveau domaine appelé échantillonnage compressif (EC) sera expliqué, l'échantillonnage compressif s'appuie sur les travaux de [9], [10], qui ont montré que si un signal a une représentation peu dense dans une base il peut être récupéré dans une autre base qui ont montré que si un signal a une représentation sparse dans une base il peut être récupéré à partir d'un petit nombre de projections sur une seconde base incohérente par rapport à la première.

L'échantillonnage à taux de Nyquist décrit complètement un signal en exploitant son caractère délimité par la bande, l'objectif de l'EC est de réduire le nombre de mesures nécessaires pour décrire complètement un signal en exploitant sa compressibilité et la différence réside dans le fait que les mesures ne sont plus des échantillons ponctuels mais des fonctionnels linéaires plus généraux du signal.

Considérons un signal à temps discret x, que nous considérons comme un vecteur de colonnes N x 1 avec des élément x[n], n = 1,2,, N. Tout signal peut être représenté en termes d'une base de vecteurs de N x 1 vecteurs $\{\psi_i\}_{i=1}^{N}$. Formation de la matrice de base (N x N)

$$\boldsymbol{\Psi} = [\psi^1|\psi^2|\dots|\psi^N] \qquad (1.1)$$

En empilant les vecteurs $\{\psi_i\}$ en colonnes, nous pouvons exprimer n'importe quel signal x sous la forme suivante

$$x = \sum_{i=1}^{N} s_i \psi_i \qquad (1.2)$$

$$x = \boldsymbol{\Psi} s$$

Où s est le vecteur colonne N x 1 des coefficients de pondération, il est clair que x et s sont des représentations équivalentes du même signal, avec x dans le domaine temporel et s dans le domaine numérique représentations équivalentes du même signal, avec x dans le domaine temporel et s dans le domaine $\boldsymbol{\Psi}$.

Nous nous concentrons sur les signaux qui ont une représentation sparse, où x est une combinaison linéaire de seulement K vecteurs de base, avec K ≪ N , c'est-à-dire que seulement K des s_i coefficients sont non nuls, la rareté est motivée par le fait que de nombreux signaux naturels et artificiels sont compressibles dans le sens où ils ne sont pas compressibles dans le sens où il existe une base Ψ où la représentation n'a que quelques grandes coefficients et de nombreux petits coefficients afin d'appliquer l'EC.

Considérons le processus de mesure linéaire qui calcule M $<$ N produits intérieurs entre x et un ensemble de vecteurs $\{\psi_i\}_{i=1}^{M}$ comme dans

$$y_j = \langle x, \phi_j \rangle \qquad (1.3)$$

L'empilement des mesures y_j dans M × 1 le vecteur y et les vecteurs de mesure ϕ^{T_j} en tant que lignes d'une matrice M × N et nous pouvons écrire :

$$y = \Phi x = \Phi\Psi s = \Theta s \qquad (1.4)$$

Où Θ est une matrice $M \times N$ il convient de noter que le processus de mesure est non adaptatif, c'est-à-dire que Φ ne dépend en aucune façon du signal x, La figure suivante illustre précédente :

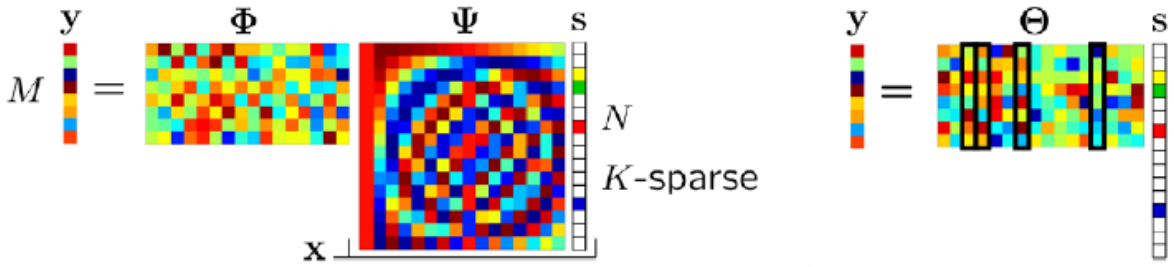

Figure 1.1: Processus de mesure de la compression [2].

La solution consiste en deux étapes : Dans la première étape, une matrice de mesure stable Φ stable qui garantit que l'information n'est pas endommagée par la réduction de la dimensionnalité de $x \in C^N$ à $y \in C^M$, dans la deuxième étape un algorithme de reconstruction est développé pour récupérer x à partir des mesures y.

1.3 Ecoute du spectre large bande :

Par rapport aux techniques à bande étroite existes dans la littérature, les techniques d'écoute du spectre à large bande visent à écouter une largeur de bande de fréquences qui dépasse la largeur de bande de cohérence du canal. Par exemple, pour exploiter les opportunités spectrales dans l'ensemble de la bande de télévision à ultra-haute fréquence (entre 300 MHz et 3 GHz), il faut utiliser des techniques d'écoute large bande devraient être utilisées. Nous notons que les techniques d'écoute à bande étroite ne peuvent pas être utilisées directement pour écouter le spectre à large bande, car elles prennent une seule décision binaire unique pour l'ensemble du spectre et ne peuvent pas identifier les opportunités spectrales individuelles qui se trouvent dans le spectre à large bande, l'écoute du spectre à large bande peut être classée en deux catégories : l'écoute large bande de Nyquist et l'écoute large bande sous-Nyquist. Le premier type traite les signaux numériques pris au taux de Nyquist ou au-dessus, tandis que le deuxième type acquiert des signaux en utilisant un taux d'échantillonnage inférieur au taux de Nyquist [11].

1.3.1 Ecoute large bande Sous-Nyquist :

En raison des inconvénients d'un taux d'échantillonnage élevé ou de la complexité d'implémentation systèmes Nyquist, les approches sous-Nyquist attirent de plus en plus d'attention dans le monde académique et que dans l'industrie [11]. L'écoute à large bande sous-Nyquist désigne la procédure d'acquisition de signaux à large bande en utilisant des taux d'échantillonnage inférieurs au Nyquist et de détecter les opportunités spectrales à l'aide de ces

mesures partielles. Deux types importants d'écoute à large bande sous-Nyquist sont l'écoute à large bande par l'écoute compressive et l'écoute à large bande sous-Nyquist multicanaux.

1.3.2 Ecoute large bande basée sur l'écoute compressive :

La technique d'écoute compressive (EC) est une technique qui permet d'acquérir efficacement d'acquérir un signal en utilisant relativement peu peu de mesures, ce qui permet de trouver une représentation unique du signal peut être trouvée sur la base de la compressibilité du signal dans un certain domaine. Le spectre à large bande étant intrinsèquement sparse en raison de la faible utilisation du spectre, il est possible de trouver une représentation unique du signal en raison de la faible utilisation du spectre, l'EC devient un candidat prometteur pour réaliser l'écoute de spectre à large bande en utilisant des taux d'échantillonnage inférieurs au taux de Nyquist. Tian et Giannakis ont d'abord introduit la théorie d'EC pour écouter le spectre à large bande dans [12]. Cette technique utilise moins d'échantillons plus proches du taux d'information, plutôt que l'inverse de la bande passante, pour effectuer l'écoute du spectre à large bande. Après la reconstruction du spectre à large bande, l'ondelette a été utilisée pour détecter les opportunités spectrales dans le spectre à large bande.

Cependant, l'EC s'est concentrée sur les signaux de longueur finie et à temps discret. Ainsi, des technologies innovantes sont pour étendre l'EC à l'acquisition de signaux à temps continu (c'est-à-dire pour mettre en œuvre l'EC dans le domaine analogique). Pour réaliser l'EC analogique, les auteurs dans [13], ont proposé un convertisseur analogique-information (CAI), qui pourrait être une bonne base pour la méthode d'EC décrite ci-dessus, le modèle basé sur le CAI se compose d'un générateur de nombres pseudo-aléatoires, d'un mélangeur, d'un accumulateur et d'un échantillonneur à faible taux, voir figure 1.2. Le générateur de nombres pseudo-aléatoires produit une séquence à temps discret qui démodule le signal $x(t)$ par un mélangeur. Puis, le signal démodulé est intégré par l'accumulateur, puis échantillonné à sous Nyquist à l'aide d'un facteur de décimation W. Ensuite, le signal sparse peut être reconstruit directement à partir de mesures partielles à l'aide de la méthode de reconstruction. Malheureusement, il a été identifié que la performance du CAI peut facilement être affectée par des imperfections de conception ou des inadéquations de modèles.

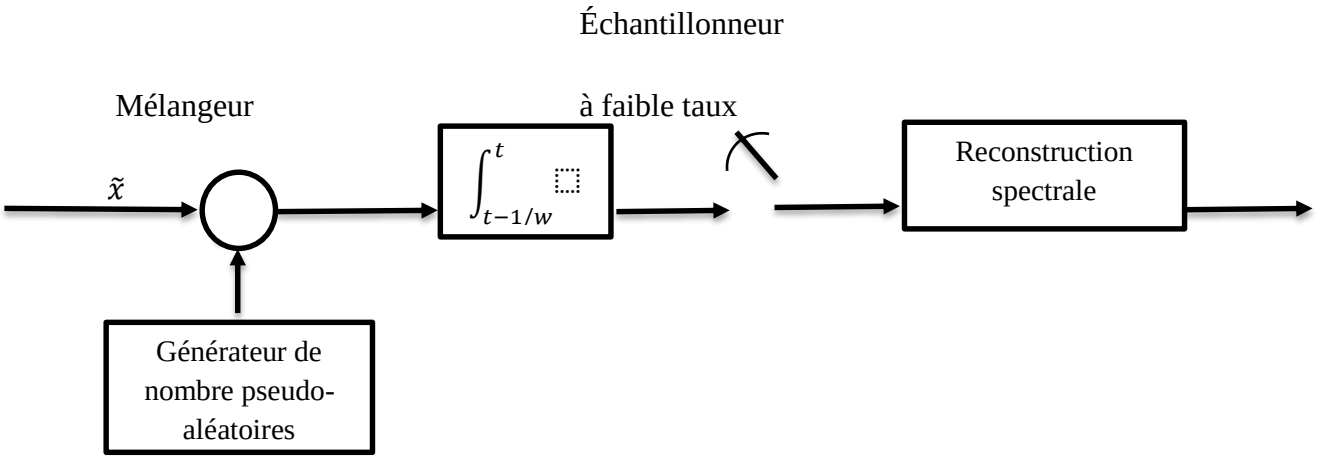

Figure 1.2: Ecoute à large bande basée sur un convertisseur analogique-information.

1.3.3 Ecoute à large bande sous-Nyquist à canaux multiples :

Pour contourner les incohérences des modèles, Les auteurs dans [14], ont proposé un modèle de convertisseur modulé à large bande (CMLB), voir figure 1.3 en modifiant le modèle CAI. La principale différence entre le CMLB et CAI est que le CMLB possède plusieurs canaux d'échantillonnage, l'accumulateur de chaque canal étant remplacé par un filtre passe-bas. L'un des avantages significatifs de l'introduction d'une structure de canaux parallèles dans la figure 1.3 est qu'elle offre une certaine robustesse face au bruit et aux inadéquations du modèle. En outre, la dimension de la matrice de mesure est réduite, ce qui rend la reconstruction spectrale plus efficace de manière computationnelle. Une autre méthode d'échantillonnage sous-Nyquistmulticanal est l'échantillonnage multi-coset, voir figure 1.4. L'échantillonnage multi-coset équivaut à choisir certains échantillons dans une grille uniforme, qui peut être obtenue à l'aide d'un taux d'échantillonnage f_s supérieur au taux de Nyquist. La grille uniforme est ensuite divisée en blocs de m échantillons consécutifs et, dans chaque bloc, v ($v < m$) échantillons sont conservés, tandis que le reste des échantillons est ignoré. Ainsi, l'échantillonnage multicoset est souvent mis en œuvre en utilisant v canaux d'échantillonnage avec un taux d'échantillonnage de f_s/m, les différents canaux d'échantillonnage ayant des décalages temporels différents. Pour obtenir une solution unique pour le spectre à large bande à partir de ces mesures partielles, le modèle d'échantillonnage doit être soigneusement conçu. Dans [15], il a été prouvé que certains motifs d'échantillonnage étaient valables pour une reconstruction unique du signal. L'avantage de l'approche multi-coset est que le taux d'échantillonnage dans chaque canal est m fois inférieure au taux de Nyquist. L'un des inconvénients de l'approche multi-coset est que la synchronisation des canaux doit être respectée de telle sorte que des

décalages temporels précis entre les canaux d'échantillonnage sont nécessaires pour satisfaire un modèle d'échantillonnage spécifique en vue d'une reconstruction spectrale robuste. Pour assouplir l'exigence de synchronisation multicanal, une approche d'écoute à large bande asynchrone à plusieurs canaux a été étudiée. Dans cette approche, un échantillonnage sous-Nyquist a été induit dans chaque canal d'échantillonnage pour envelopper la carte d'occupation du spectre sur elle-même ; le taux d'échantillonnage peut donc être réduit de manière significative. L'utilisation de différents taux d'échantillonnage dans différents canaux d'échantillonnage, comme le montre la figure 1.5 ci-dessous qui permet d'améliorer la performance de d'écoute du spectre à large bande. Plus précisément, dans le même temps d'observation, les nombres d'échantillons dans les canaux d'échantillonnage multiples sont choisis comme différents nombres premiers consécutifs. En outre, comme seules les amplitudes des spectres sous-Nyquist sont intéressantes, une telle approche d'écoute à large bande à plusieurs taux ne nécessite pas de synchronisation parfaite entre les canaux d'échantillonnage multiples, ce qui facilite la mise en œuvre.

Presque toutes les techniques d'écoute à large bande sous-Nyquist exigent que le signal à large bande soit sparse sur une base appropriée. Compte tenu de la faible utilisation du spectre, la plupart des techniques d'écoute à large bande existantes signal à large bande est peu dense dans le domaine fréquence (c'est-à-dire que la base de rareté est une matrice de de Fourier). Cependant, à mesure que l'utilisation du spectre (par exemple, en raison de l'utilisation de techniques de radio dans les futurs réseaux cellulaires), le signal à large bande peut ne plus être clairsemé dans le domaine des fréquences.

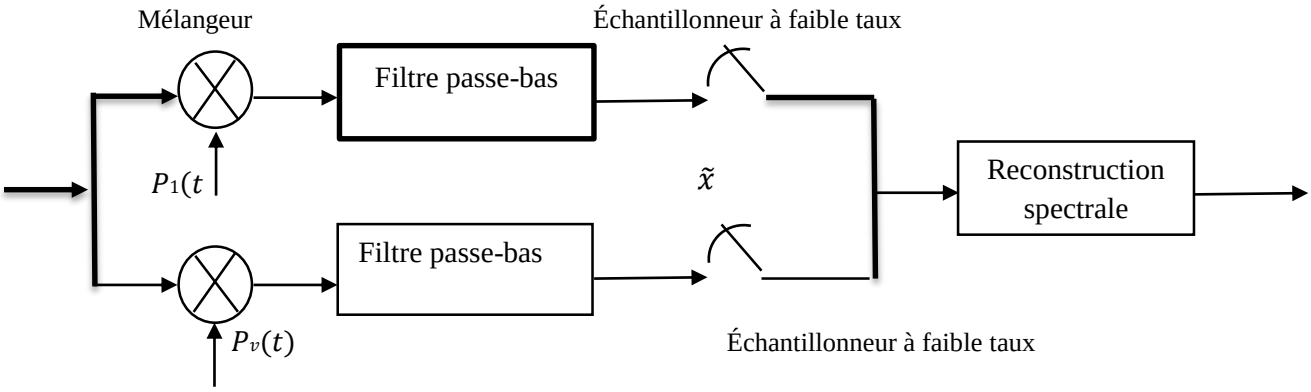

Figure 1.3 : Écoute à large bande basée sur CMLB.

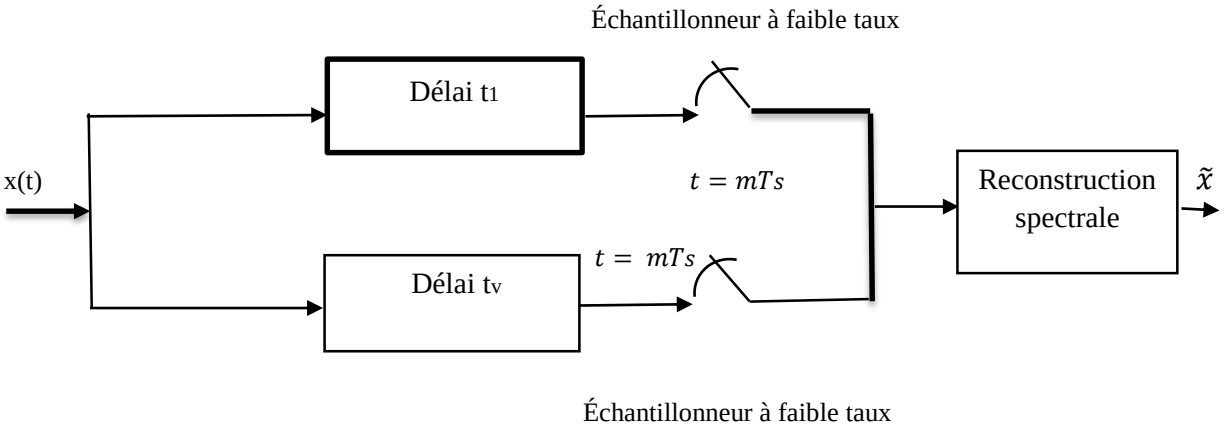

Figure 1.4: écoute à large bande basée sur un échantillonnage multi-coset

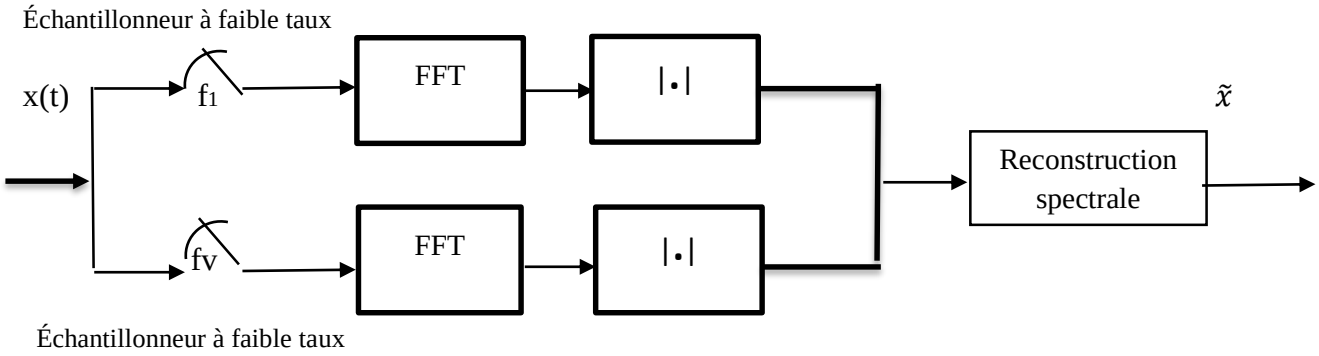

Figure 1.5 : écoute à large bande basée sur un échantillonnage à différents taux.

1.4 Ecoute coopérative à large bande :

L'écoute coopérative à large bande est une technique permettant aux utilisateurs de collaborer en écoutant ensemble un signal audio à travers une large plage de fréquences. Cette méthode a été développée pour améliorer la qualité de la transmission audio dans les réseaux de communication sans fil. Il a été démontré que cette méthode améliore considérablement la qualité de l'écoute pour les utilisateurs et peut également aider à réduire les taux d'erreur de transmission. Cette technique est donc largement utilisée dans les réseaux de communication sans fil. [16].

1.5 Applications d'écoute compressive :

Voici quelques applications de l'EC qui existe dans la littérature [17] :

1.5.1 Imagerie médicale :

L'écoute compressive (EC) a été utilisée en imagerie médicale pour réduire le nombre de scans IRM (Imagerie par Résonance Magnétique) nécessaires à la création d'images de haute qualité, réduisant ainsi le temps de scan et l'exposition du patient aux radiations.

1.5.2 Astronomie :

L'EC a été utilisée en astronomie pour reconstruire des images du ciel à haute résolution à partir de réseaux de capteurs comportant un nombre limité de pixels.

1.5.3 Communications sans fil :

Les techniques de l'EC peuvent être utilisées pour améliorer l'efficacité des communications sans fil en réduisant la quantité de données à transmettre

1.5.4 Traitement de l'audio et de la parole :

L'EC a été utilisée dans le traitement du son et de la parole pour compresser les données audios, réduire le bruit et améliorer la qualité de la parole.

1.5.5 Imagerie radar et sonar :

L'EC a été utilisée dans l'imagerie radar et sonar pour améliorer la détection des cibles et la résolution de l'imagerie.

1.6 Conclusion :

Dans ce chapitre, nous avons tout d'abord présenté la théorie d'échantillonnage compressive et ensuite, nous avons expliqué les techniques d'écoute du spectre à large bande de sous Nyquist. On a présenté quelques applications de l'EC qui existe dans la littérature. En outre, motivé par le fait que l'écoute du spectre à large bande est essentielle pour trouver de manière fiable les opportunités spectrales et réaliser un accès opportuniste au spectre pour les réseaux cellulaires de la prochaine génération, on a fait une petite étude de l'état de l'art des techniques d'écoute du spectre à large bande de sous Nyquist.

Chapitre 2

Sécurité de la couche physique au niveau la radio

2.1 Introduction :

Ce chapitre présente une revue de la littérature dans le domaine de la sécurité de la couche physique au niveau des radios. En outre, les principaux problèmes auxquels sont confrontés les réseaux de radio cognitive (RRCs) sont discutés, en mettant l'accent sur le défi de la sécurité, avec une référence spécifique aux menaces potentielles de la couche physique. Nous présentons les différents types d'attaques qui peuvent perturber les systèmes de communication sans fil et comment on trouve une solution intelligente basé sur l'apprentissage automatique pour faire face à ce genre de problème.

2.2 Réseaux de radio cognitive (RRC) et les menaces :

Les RRCs sont vulnérables à diverses attaques car ils sont généralement déployés dans des environnements non surveillés et utilisent des communications sans fil peu fiables [18]. Toutefois, il n'est pas simple de mettre en œuvre des mesures de sécurité dans les RRCs. L'un des principaux obstacles au déploiement de la sécurité sur les RRCs est que les RRCs actuels ont des capacités de calcul et de communication limitées. C'est pourquoi de nombreux chercheurs ont commencé à assurer la sécurité des RRCs à l'aide de différents mécanismes de sécurité. Les mécanismes de sécurité, y compris la gestion de la confiance, ont la capacité de sécuriser les RRCs contre les attaquants.

2.3 Attaques de la couche physique :

La couche physique gère la transmission et la réception des flux de bits entre un émetteur et un récepteur. Les exemples de fonctions assurées par la couche physique sont : la détection du signal, la sélection de la fréquence et la modulation, Comme toutes les autres couches, la couche physique de la radio cognitive peut faire l'objet de plusieurs attaques qui peuvent perturber son fonctionnement normal [19].

Ces attaques peuvent être classées en quatre catégories principales : (i) les attaques sur l'accès dynamique au spectre ; (ii) les attaques de manipulation ; (iii) les attaques d'écoute et (iv) les attaques d'injection de trafic malveillant. Cette classification est basée sur le processus de la radio cognitive qui est affecté par l'attaquant [21].

La catégorie des attaques sur l'accès dynamique au spectre comprend les attaques qui perturbent le processus d'écoute du spectre et empêchent les utilisateurs secondaires de détecter les canaux disponibles y compris les attaques d'émulation de UP (Utilisateur primaire) (PUE : primary user

emulator). La catégorie des attaques de manipulation vise à manipuler les paramètres radio conduisant à des décisions de détection erronées. Elle comprend la falsification des données de l'écoute du spectre et les attaques par fonction objective.

La catégorie des attaques par écoute clandestine vise à décoder les signaux pendant la transmission afin d'écouter des informations confidentielles.

Les attaquants peuvent être actifs ou passifs, les attaquants actifs visent à apporter des modifications au fonctionnement du système ou aux paquets échangés, tandis que les attaquants passifs visent à observer et à lire les informations du système à des fins ultérieures, sans affecter le système [20].

Les attaques classées dans chaque catégorie sont illustrées à la figure 2.1.

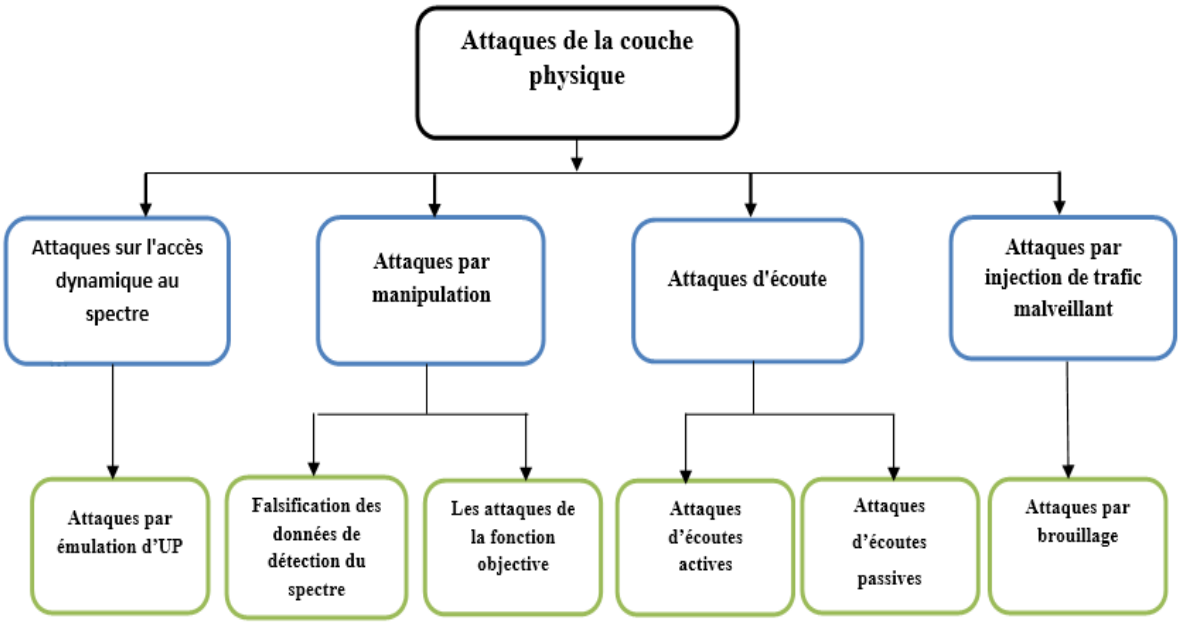

Figure 2.1 : Classification des attaques ciblant la couche physique de RC [21].

2.4 Exemples de quelques attaques au niveau de la couche physique des radios

Voici exemples de quelques attaques qui existent dans la littérature :

2.4.1 Attaque par brouillage intentionnel :

Il s'agit de l'un des types d'attaque les plus fondamentaux pouvant être perpétré par des USs dans les RRCs. L'US malveillant brouille les utilisateurs primaires et secondaires en émettant intentionnellement et continuellement dans une bande sous licence [18] cette attaque peut être plus grave et plus dangereuse dans les RRCs si un US malveillant mobile agit dans une zone géographique et si l'utilisateur secondaire se déplace vers une autre zone avant que l'attaque ne soit détectée.

2.4.2 Attaque par émulation d'utilisateur sous licence :

Un RRC utilise le spectre sous licence lorsqu'il est libre ; sinon, il utilise la bande sans licence. L'attaquant peut brouiller la bande sous licence et émuler l'UP (Utilisateur Primaire), limitant ainsi le RRC à fonctionner dans les bandes sans licence et limitant la capacité du RRC. Ce problème n'a pas encore de solution [18], [22].

2.4.3 Brouillage du canal de contrôle commun :

Dans ce cas, l'attaquant transmet des impulsions périodiques dans le spectre du canal de contrôle. Le brouillage d'un canal bloque la communication probable entre tous les nœuds de radio cognitive. L'ultra large bande, en tant que déploiement du canal de contrôle commun, peut résoudre le problème du brouillage [18], [22].

2.4.4 Attaque par déni de service :

L'attaque par déni de service empêche les USs d'utiliser la bande de fréquences des UPs. Les attaquants génèrent des résultats d'écoute montrant que la bande de fréquences des UPs est occupée par des UPs. Si leurs résultats d'écoute sont agrégés dans le processus de prise de décision finale sans filtrage approprié, ils pourraient influencer négativement la décision finale, ce qui entraînerait des erreurs de fausse alarme et la perte de la possibilité d'utiliser les bandes de fréquences UPs lorsqu'elles sont réellement disponibles. Si les attaques réussissent, la performance du système se dégradera fortement [23].

2.4.5 Attaque par brouillage du récepteur primaire :

Dans les RRCs, une entité malveillante peut profiter d'un manque de connaissances sur l'emplacement des récepteurs primaires pour causer intentionnellement des interférences

nuisibles à un récepteur primaire victime. Cette attaque se produit lorsqu'une entité malveillante plus proche du récepteur primaire victime participe à un protocole de collaboration et demande que les transmissions des autres USs soient dirigées vers l'utilisateur malveillant [23].

2.4.6 Attaque par amplification de la sensibilité :

Dans les RRCs, certaines techniques d'écoute des UPs sont plus sensibles aux transmissions primaires afin d'éviter les interférences avec le réseau primaire. Les USs sont vulnérables à cet événement, car cela entraîne de fréquentes fausses détections et des occasions manquées. Une entité malveillante peut amplifier la sensibilité et donc le nombre d'occasions manquées en rejouant les transmissions primaires [23].

2.4.7 Attaque des utilisateurs secondaires qui se chevauchent :

Dans les architectures centralisées et distribuées des RRCs, plusieurs réseaux secondaires peuvent coexister dans la même région géographique. Les transmissions d'entités malveillantes peuvent nuire aux utilisateurs primaires et secondaires, non seulement dans un réseau, mais aussi dans d'autres RRCs [23].

2.4.8 Attaque de falsification des données d'écoute du spectre :

Dans cette attaque, un attaquant peut envoyer de faux résultats d'écoute du spectre local à un collecteur de données, ce qui amène ce dernier à prendre une mauvaise décision en matière d'écoute du spectre dans les RRCs [24].

2.5 Contre-mesures pour quelques attaques qui existent dans la littérature :

2.5.1 Contre-mesures de brouillage

Le brouillage des contre-mesures fait référence à l'utilisation de techniques de brouillage pour perturber le fonctionnement des systèmes de communication et de radar ennemis ou hostiles. L'objectif est de désactiver ou de dégrader l'efficacité de ces systèmes, en les rendant moins efficaces, voire inutilisables. Le brouillage des contre-mesures peut être réalisé par différents moyens, tels que la transmission d'un signal puissant sur la même fréquence que le système cible, l'émission de bruit dans la bande de fréquence du système cible ou la transmission d'un signal qui imite le signal du système cible afin de le confondre.

Le brouillage des contre-mesures peut être un outil efficace, car il peut perturber la capacité de l'ennemi à communiquer et à naviguer, entraver sa capacité à lancer des missiles ou d'autres armes, et créer la confusion et le chaos sur le champ de bataille.

Toutefois, le brouillage des contre-mesures peut également avoir des conséquences inattendues, telles que la perturbation des communications amies et des dommages collatéraux à l'électronique civile. Il est donc important d'étudier soigneusement l'utilisation du brouillage de contre-mesures et de l'employer judicieusement et avec précision [25].

Des techniques de détection d'anomalies intelligentes basées sur l'apprentissage automatique et la reconnaissance de patrons ont à la capacité d'identifier les anomalies et de les supprimer basé sur la classification. Ces techniques sont donc considérées comme un outil important pour détecter les attaques dans des scénarios en temps réel.

2.5.2 Contre-mesures d'attaque par émulation d'utilisateur primaire :

L'une des contre-mesures pour prévenir une attaque par émulation d'utilisateur primaire au niveau de la couche physique du radio serait de mettre en place un mécanisme de détection et de suivi des signaux émis par les émetteurs primaires légitimes. Ce mécanisme consiste à établir un profil de signal des émetteurs primaires et à comparer les signaux reçus à ce profil afin de détecter toute émulation ou tentative de compromettre le système en émettant des signaux similaires.

En outre, il est possible de mettre en place des méthodes d'authentification et de chiffrement pour les communications entre les émetteurs primaires et le système de contrôle, afin de s'assurer que seuls les signaux provenant de sources légitimes soient autorisés à pénétrer dans le système.

Enfin, l'utilisation de technologies de brouillage peut également être envisagée pour perturber les signaux émis par les émulateurs potentiels et empêcher leur utilisation dans des attaques potentielles [21].

2.6 Conclusion :

La sécurité de la couche physique dans les RRCs est d'une importance capitale pour assurer une communication sécurisée entre les utilisateurs primaires et secondaires. Les techniques de sécurité de la couche physique dans les RRCs visent à assurer la sécurité en utilisant des techniques de traitement du signal telles que l'étalement du spectre et ses techniques de transmission sécurisée. Ces techniques permettent de réduire la capacité d'un adversaire à écouter ou à interférer dans les communications des utilisateurs légitimes. En outre, les techniques de sécurité de la couche physique peuvent également être utilisées pour détecter et prévenir les attaques de brouillage en se basant l'apprentissage automatique et la reconnaissance de patrons d'anomalies.

Chapitre3

Théorie d'écoute compressive combinée avec un détecteur d'attaque basé sur la distance Mahalanobis et la distance Robuste

3.1 Introduction :

La détection des anomalies dans les données multi variées est une tâche importante en statistique, car ce type de données peut fausser toute procédure statistique. Dans les contextes d'exploration de données et d'apprentissage automatique, de nombreuses techniques standard telles que l'analyse en composantes principales et l'analyse discriminante linéaire sont intrinsèquement sensibles aux observations atypiques [26].

Il existe de nombreuses méthodes pour détecter les anomalies dans les données à haute dimension. Parmi les études antérieures, citons : Shi et al [27] ont proposé un algorithme de détection d'anomalies en haute dimension basé sur un algorithme génétique, etc. La distance classique de Mahalanobis est une méthode courante de détection d'anomalies. Cependant, il s'agit d'une méthode basée sur le vecteur moyen de l'échantillon et la matrice de covariance de l'échantillon. Étant donné que les algorithmes classiques basés sur le vecteur moyen et la matrice de covariance sont sensibles aux anomalies, la distance de Mahalanobis classique est également sensible aux anomalies. De nombreux auteurs ont proposé des méthodes d'estimation robustes pour le vecteur moyen et la matrice de covariance, comme l'estimateur S [28], et l'estimateur MCD (Minimum Covariance Determinant) [29].

Dans le même temps, la distance de Mahalanobis robuste est proposée dans la littérature sur la base du vecteur moyen et de la matrice de covariance robustes. En outre, l'estimateur MCD rapide est largement appliqué car sa computation est simple et rapide [30], et qu'il présente une grande robustesse. Par rapport à la distance de Mahalanobis classique, on constate une bonne amélioration de la robustesse. En 2014, Feng et al [31] ont appliqué la distance de Mahalanobis robuste basée sur l'estimateur MCD rapide à l'analyse des données de nuages de points LiDAR (Light Detection and Ranging).

3.2 Modèle du signal et du système :

3.2.1 Modèle de signal :

On désigne par $y(t)$ le signal analogique multibande reçu correspondant à la i^{th} transmission$1 \leq i \leq n_B$, reçu au p^{th} RC, $1 \leq p \leq P$. Soit $y(t)$ un signal à temps continu

à valeurs réelles, supporté par $\mathcal{F} = [-f_{max}\,;\,+f_{max}]$ et composé d'un maximum de n_B transmissions, tel que :

$$y(t) = \sum_{i=1}^{n_B} \propto_i (t) \tag{3.1}$$

Où $\propto_i (t)$ est un processus passe-bande. La largeur de bande unilatérale de chaque transmission est censée ne pas dépasser $\mathcal{B}$. Formellement, la transformée de Fourier de $y(t)$ définie par [8] :

$$Y(f) = \int_{-\infty}^{+\infty} y(t)\, e^{-j2\pi ft} dt \tag{3.2}$$

est nul pour tout $f \notin \mathcal{F}$. Nous désignons par $f_{nyq} = 1/T_{nyq}$ le taux de Nyquist de $y(t)$. Seulement n_B et $\mathcal{B}$ ou au moins une borne supérieure pour chaque, sont censés être connus. $K = 2n_B$, sa sparsité (le facteur 2 provient du fait que chaque signal contribue à deux bandes de fréquences symétriques). Les fréquences porteuses et les modulations de $\propto_i (t)$ sont inconnues. Le signal est reçu par P récepteurs, avec $P \geq 2n_B$.

Soit $r(t)$ est le signal reçu par chaque utilisateur RC pour chaque canal p. Nous supposons une écoute coopérative centralisée du spectre à large bande dans laquelle un groupe de P RC communique avec le CF (Centre de Fusion) par des canaux de contrôle. Pour les canaux radio entre les utilisateurs primaires (UPs) et les radios cognitives, nous supposons un évanouissement plat avec une distribution de Rayleigh et un effet d'ombre, et nous supposons des canaux de contrôle entre les radios cognitives et la station de base. L'écoute par chaque utilisateur RC est effectuée périodiquement et les mesures compressées sont envoyées au CF. Considérons une représentation continue du signal reçu $r(t)$ à chaque RC en présence du bruit et d'un signal malveillant s'il est présent d'après les deux hypothèses H_0 et H_1 :

Hypothèse H_0 : le signal source est corrompu uniquement avec le bruit.

$$r_p(t) = \sum_{i=1}^{n_B} H_{ip} \propto_{ip} (t) + w_p(t), \tag{3.3}$$

Hypothèse H_1 : le signal source est corrompu avec le bruit et le signal d'attaque.

$$r_p(t) = \sum_{i=1}^{n_B} H_{ip} \propto_{ip} (t) + H'_p J_p(t) + w_p(t), \tag{3.4}$$

$t = 1, \dots, n_s, p = 1, \dots, P,$

n_s est le nombre de points qui représentent le signal spectral de la source multi-bande dans le domaine temporel discret. n_B est le nombre de sous-bandes générées par les UPs. P est le nombre de RCs utilisateurs secondaires (USs).

Où, H_{ip} représentent les canaux radio entre les UPs à bande étroite et les radios, $r_p(t)$ est le signal reçu par chaque $p$utilisateur radio, $\propto_i (t)$ est le signal sparse multi-bande lorsque les i^{th} UPs sont présentes et $w_p(t)$ est le bruit blanc additif gaussien (AWGN :Adding White Gaussian Noise), J_p : est le signal malveillant d'attaque, H'_p représente les canaux radio entre UA (Utilisateur d'attaque) et (USs).

Nous supposons seulement un simple RC qui observe et échantillonne le signal reçu $r(t)$ à la fréquence d'échantillonnage de Nyquist. Ensuite, le RC simple envoie ses observations compressées au CF pour qu'il les traite. Nous pouvons généraliser cette méthode pour tous les autres USs via collaboration centralisée.

Nous considérons deux hypothèses pour faire une décision d'après la valeur du seuil pour chaque valeur de la distance décrite dans ce travail s'il s'agit d'une valeur d'anomalie ou pas.

Les deux distances : distance Mahalanobis (DM) et distance robuste (DR) sont capable de détecter les anomalies basant sur la valeur du seuil. La matrice de données compressées obtenue via collaboration est considérée comme l'entrée du détecteur intelligent proposé. Afin de rendre possible la détection des anomalies avec un petit nombre d'échantillons (n_c) par rapport au taux d'échantillonnage de Nyquist, nous considérons dans ce travail la théorie de l'échantillonnage sous-Nyquist. Dans ce cas et à chaque utilisateur radio, on observe le signal analogique multi-bande reçu x de manière compressive avec un taux très inférieur au taux d'échantillonnage de Nyquist basé sur la théorie d'EC tel que :

$$x = \Phi y + w \tag{3.5}$$

Où, $x \in \mathbb{R}^{n_c}$ est la mesure compressée, Φ est la matrice d'écoute, $y \in \mathbb{R}^{n_s}$ est le signal reçu, w est le bruit gaussien additif, et $n_c \ll n_s$.

Nous aimerions utiliser les mesures compressées directement pour détecter le problème d'anomalie, sans passer par l'étape de récupération du signal multi bande. Les récepteurs radios peuvent utiliser la même matrice d'écoute Φ ou des matrices différentes pour chaque utilisateur

radio. Nous supposons ici que tous les utilisateurs radios utilisent la même matrice Φ par chaque radio.

Après l'application de la technique d'échantillonnage sous-Nyquist dans un régime large bande par un convertisseur large bande modulé (CMLB), chaque radio envoie toutes ses mesures compressées via coopération par une matrice d'observations au CF. Il est clair que l'envoi de toutes les mesures compressées par chaque récepteur radio nécessite une consommation d'énergie moins, moins de calculs.

3.2.2 Modèle du signal d'attaque par interférence d'ondes continues (CWI : Continuous Wave Interference)

Le brouillage CWI peut être écrit comme suit [32] :

$$j(t) = \sum_{k=1}^{K} \sqrt{2P_{J_k}} \cos\left[2\pi\left(f_c \mp f_{\Delta_k}\right)t + \varphi_k\right] \qquad (3.6)$$

Ou P_{J_k}, f_c, f_{Δ_k} et φ_k représentent la puissance, la fréquence centrale (identique à la fréquence L1), la fréquence de réglage etla phase aléatoire de la k ième tonalités, respectivement. Nous choisissons la fréquence de réglage comme étant : $f_{\Delta_k} = (0, \pm 0{,}1, \pm 0{,}2) f_{CA}$ avec f_{CA} étant la fréquence du code C/A.

3.2.3 Modèle d'écoute coopérative centralisée du spectre à large bande :

Dans ce travail, nous proposons un schéma d'écoute coopérative centralisé qui collecte les données compressées sous forme d'une matrice au niveau du centre de fusion. Nous expliquons tout d'abord le schéma d'échantillonnage compressif proposé dans ce travail.

Nous expliquons ensuite le détecteur intelligent basé sur la DM et DR juste avant le processus de récupération du signal multi bande dans un régime large bande.

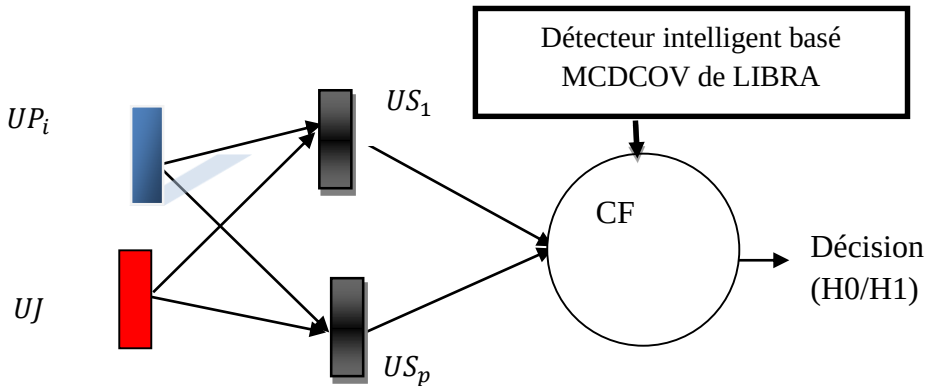

Figure.3.1 : Schéma d'échantillonnage compressif coopératif centralisé combiné avec le détecteur proposé basé sur LIBRA (Library Robust Analysis).

3.3 Convertisseur modulé à large bande centralisé basé sur l'écoute coopérative du spectre Système CMLB Centralisé :

Le frontal CMLB prétraite un signal analogique multi bande reçu $y(t)$ en utilisant des p^{th} canaux. Le terme multi-bande fait référence à un signal analogique dont le spectre de fréquences est concentré sur N_B bandes de fréquences dont la largeur de bande individuelle $\mathcal{B}$**.** Il faut que les conditions suivantes soient remplies $2n_B\mathcal{B} < f_{nyq}$. Pour éviter les effets de bord $f_s \geq f_p \geq \mathcal{B}$ doit être garanti [33], [34].

f_s: Fréquence d'échantillonnage. f_p: Fréquence du pseudo-bruit.

Nous utilisons le cadre CMLB centralisé pour effectuer un échantillonnage sous-Nyquist dans le réseau coopératif d'écoute du spectre (RCES). La seule différence avec le CMLB est que chaque RC n'effectue qu'un échantillonnage à un canal avec un mélangeur, un filtre passe-bas (FPB) et une conversion analogique-numérique (CAN). D'une part, nous pouvons réduire considérablement le coût matériel d'un simple RC. D'autre part, il est facile d'ajuster et de satisfaire le nombre de canaux d'échantillonnage dont nous avons besoin, en profitant du fait que de nombreux RCs sont inclus dans un RCES.

Le CMLB basé sur le convertisseur analogique information (CAI) via collaboration est proposé par les auteurs dans [35] pour le signal multi bande analogique plus pratique. Comme l'illustre la figure 3.2 ci-dessous, le support de fréquence d'un signal analogique multi bande réside dans

de nombreux intervalles continus répartis sur un spectre à large bande. Un exemple simple de signal analogique multi-bande à temps continu est le suivant

Le signal GPS L1CA quittant l'antenne du ième satellite peut être représenté analytiquement par [32]

$$x_{L1}(t) = \sqrt{2P}.D_{L1}(t).C_{L1}(t)\cos(2\pi f_{L1}t + \varphi_{L1}) \quad (3.7)$$

x_{L1} : Signal GPS L1CA transmis.

P_{L1} : Puissance totale L1.

$D_{L1}(t)$: Données de navigation L1.

$C_{L1}(t)$: Code C/A L1.

φ_{L1} : Phase de la porteuse L1.

f_{L1} : Fréquence de porteuse L1.

3.4 Description du système :

Le CMLB centralisé (CMLB-C) est similaire au CMLB, mais nous prenons en compte le déphasage et la perte de transmission. θ_p et la perte de transmission γ_p dans chaque canal d'échantillonnage p comme le montre la figure ci-dessous. Il est évident que lorsque α_p est plus petit, l'atténuation est plus importante. Maintenant, nous pouvons représenter le signal analogique multi bande reçu dans le p^{th} RC comme suit : $x_p(t) = \gamma_p e^{j\theta_p} y(t)$ [33].

La figure.3.2 montre la structure du système CMLB, où le signal analogique multi bande d'entrée $y_p(t)$ entre P canaux simultanément. Dans le canal p^{th} canal, $y_p(t)$ est multiplié par un mélange d'une fonction de séquence de pseudo-bruit $pn_p(t)$ qui est $T_{\mathbb{p}}$ périodique. Plus précisément, $pn_p(t)$ est une fonction de pseudo-aléatoire qui alterne entre les niveaux $\{\pm 1\}$ pour chaque D intervalle de temps égaux.

Après le mélange, le spectre du signal analogique multi bande est tronqué par un filtre passe-bas (FPB) avec une coupure de $1/2T_s$ et le signal filtré est échantillonné à la fréquence $1/T_s$.

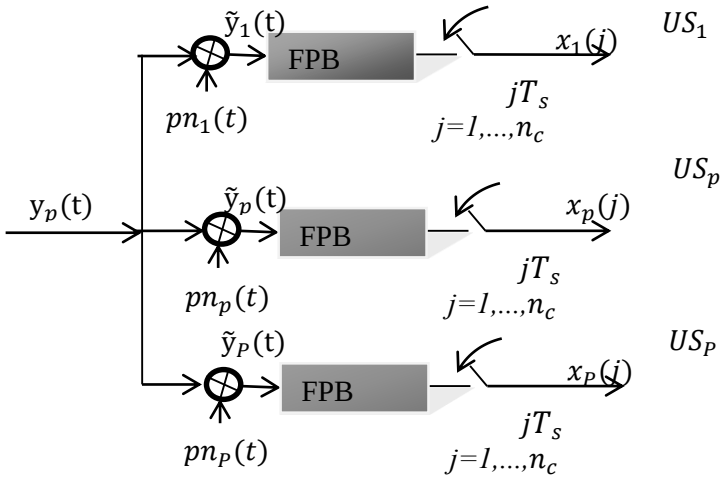

Figure 3.2 : schéma du CMLB.

3.5 Distances :

Les observations compressées de chaque utilisateur radio sont collectées sous forme d'une matrice de données compressée au niveau du CF et servent d'entrée au détecteur intelligent proposé dans ce travail. Nous utilisons cette technique pour détecter les problèmes d'anomalie à l'aide de deux estimateurs robustes qui sont la moyenne et la covariance [36].

3.5.1 Distance de Mahalanobis et la distance robuste :

La distance robuste d'une observation i est utilisée pour détecter s'il s'agit d'une anomalie ou non.

Elle est définie comme suit :

$$DR_i = ((\mathrm{x_i} - \hat{\mu}_{MCD})^T \hat{\Sigma}_{MCD}^{-1} (\mathrm{x_i} - \hat{\mu}_{MCD})^{1/2} \tag{3.8}$$

Avec $\hat{\mu}_{MCD}$ et $\hat{\Sigma}_{MCD}$ les estimations de l'emplacement et de la dispersion du MCD. Cette distance robuste est la robustification directe de la distance de Mahalanobis

$$DM_i = ((\mathrm{x_i} - \bar{x})^T S^{-1} (\mathrm{x_i} - \bar{x}))^{1/2} \tag{3.9}$$

Qui utilise la moyenne classique $\bar{x}$ et la matrice de covariance empirique S comme estimations de l'emplacement et de la dispersion.

Sous l'hypothèse de normalité, les observations d'anomalies sont celles dont la distance robuste est supérieure à la valeur seuil $\sqrt{\chi^2_{p,0.975}}$ [36].

Les observations régulières dont la distance robuste ne dépasse la valeur du seuil.

3.6 Algorithme proposé :

$X\ (n_c\ , P)$: Matrice de données compressées

Début

Calculer la moyenne (μ) et la covariance (S)

Pour $\boldsymbol{i\ =\ 1 : n_c}$

$DM_i\ = ((\mathrm{x_i} - \bar{x})^T S^{-1} (\mathrm{x_i} - \bar{x}))^{1/2}$

Fin pour

Calculer le seuil de la distance Mahalanobis

$Seuil_{\ DM} \longleftarrow \chi^2_{p,1-\alpha}$

Si $DM\ (i) > Seuil_{DM}$ **Alors**

Déclaré l'observation (i) comme anomalie

Sinon

Déclaré l'observation (i) normale

Fin si

Calculer la moyenne $(\hat{\mu}_{MCD})$ et la covariance $\hat{\Sigma}_{MCD}$

Pour $i\ =\ 1 : n_c$

$DR_i = ((\mathrm{x_i} - \hat{\mu}_{MCD})^T \hat{\Sigma}^{-1}_{MCD} (\mathrm{x_i} - \hat{\mu}_{MCD})^{1/2}$

Fin pour

Calculer le seuil de la distance robuste

$Seuil_{DR} \leftarrow \chi^2_{p,1-\alpha}$

Si $DR\ (i)\ > Seuil_{DR}$ **Alors**

Déclaré l'observation (i) comme anomalie

Sinon

Déclaré l'observation (i) normale

Fin si

End

3.7 Conclusion :

En conclusion, la théorie d'écoute compressive combinée avec un détecteur d'anomalies basé sur la distance Mahalanobis et la distance robuste est une approche efficace pour la détection d'anomalies dans le spectre. En utilisant cette méthode basée sur l'apprentissage automatique de classification, il est possible de détecter s'il y a la présence d'un signal d'attaque dans le spectre ou pas. Deux mesures ont été montrées qui est sensibles aux anomalies qui sont la moyenne et la covariance basé sur l'estimateur MCD par rapport à l'estimateur S classique. L'algorithme proposé qui permet de réaliser notre objectif a été montré dans ce chapitre.

Chapitre 4

Résultats de simulations et discussion

4.1 Introduction :

Matlab est un langage de programmation de haut niveau et un environnement de calcul numérique largement utilisé dans les domaines de l'ingénierie, des sciences et des mathématiques. Il fournit un environnement de programmation et de développement interactif avec de nombreuses fonctions intégrées, des boîtes à outils et des applications pour l'analyse de données, la visualisation, la modélisation et la simulation. Matlab peut être utilisé pour un large éventail de tâches telles que le traitement d'images, l'analyse de signaux, la conception de commandes, la modélisation financière, etc. Le langage est couramment utilisé dans l'enseignement supérieur et les instituts de recherche pour le calcul numérique et la simulation [37].

LIBRA est une bibliothèque MATLAB pour l'analyse robuste est développée à ROBUST@Leuven, le groupe de recherche sur les statistiques robustes à la KU Leuven. Elle contient des implémentations conviviales de plusieurs procédures robustes. Ces méthodes sont résistantes aux anomalies dans les données. Actuellement, la bibliothèque contient des fonctions pour la localisation uni variée, l'échelle et l'asymétrie, la localisation multi variée et l'estimation de la covariance (MCD), la régression, l'analyse en composantes principales, la régression en composantes principales, Régression par moindres carrés partiels, classification, regroupement, détection des erreurs pour les données, détection des anomalies. À titre de comparaison, plusieurs fonctions non robustes sont également incluses. De nombreux outils graphiques sont fournis pour la vérification des modèles et la détection des anomalies [38].

4.2 Paramètres de simulation :

Dans ce travail, un détecteur intelligent basé sur la distance Mahalanobis (DM) et la distance robuste (DR) proposé basé sur l'apprentissage automatique tel que décrit dans ce travail est combiné avec un CMLB centralisé via coopération de radios. Nous générons un signal analogique multi bande x(t) à partir d'utilisateurs primaires (UPs) satellitaires GPS L1CA à bande étroite modulés par une modulation BPSK avec une longueur de points N_s et N est le nombre de sous-bandes supposé dans ce travail. La présence d'un utilisateur attaquant malveillant, appelé brouillage par sweep, est notée $J(t)$. Le signal analogique multi-bande GPS L1CA noté $x(t)$ reçu est observé par chaque récepteur radio à large bande basé sur l'écoute coopérative du spectre. Nous donnons le nombre de récepteurs P RC. La fréquence f_{L1} Après l'application d'un taux d'échantillonnage inférieur à Nyquist, chaque récepteur radio fournit un

nombre minimum d'échantillons compressés noté n_c ce nombre minimum d'échantillons est suffisant pour savoir s'il y a un problème d'anomalies sur la base des caractéristiques des deux types de distances DM et DR. Le seuil de décision pour les deux types de distances est donné par $seuil = \chi^2_{P,1-\alpha}$. Les paramètres de simulation utilisés dans ce travail sont indiqués dans le tableau 4.1:

Notation	Description	Valeur
P	Nombre de récepteurs radio	50
f_{L1}	Fréquence de GPS L1	1.57 GHz
n_s	Nombre de points	462500
n_B	Nombre de bandes	2
n_c	Nombre d'échantillons compressés	2500
$\mathcal{B}$	Largeur de bande de chaque bande	2 MHz
e_i	Énergie aléatoire de chaque bande	[1,2]
SNR	rapport signal-bruit	43dB
SJR	Rapport signal sur attaque	-50dB
L	Facteur de sou-échantillonnage	185
$C(t)$	Nombre d'intervalles ±1 dans chaque période de $pn_p(t)$	1023
f_c	Fréquence central	f_{L1}
f_{CA}	Fréquence code C/A	1,023MHz
f_1 φ_1	Première fréquence du CWI Première phase du CWI	f_c 0
f_2 φ_2	Deuxième fréquence du CWI Deuxième phase du CWI	$f_c + 0{,}1f_{CA}$ $\pi/2$

f_3	Troisième fréquence du CWI	$f_c + 0{,}2f_{CA}$
φ_3	Troisième phase du CWI	π

Tableau.4.1 : présente les paramètres de simulation considérés dans ce travail.

4.3 Résultats :

Hypothèse H_0 : le signal source est corrompu uniquement avec AWGN.

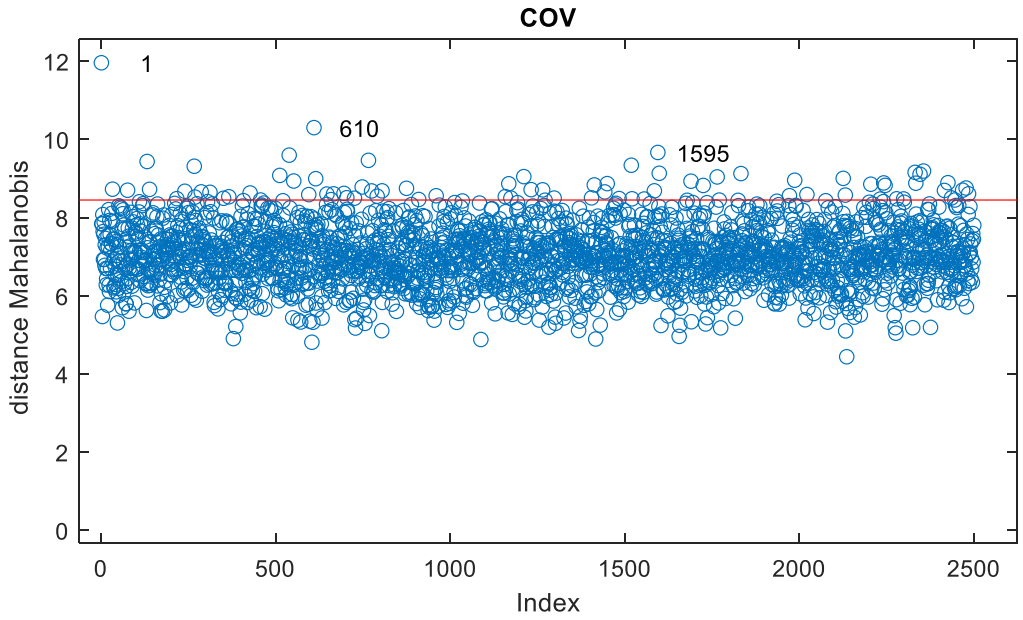

Figure 4.1.a : distance Mahalanobis obtenue par la covariance classique selon l'hypothèse H_0.

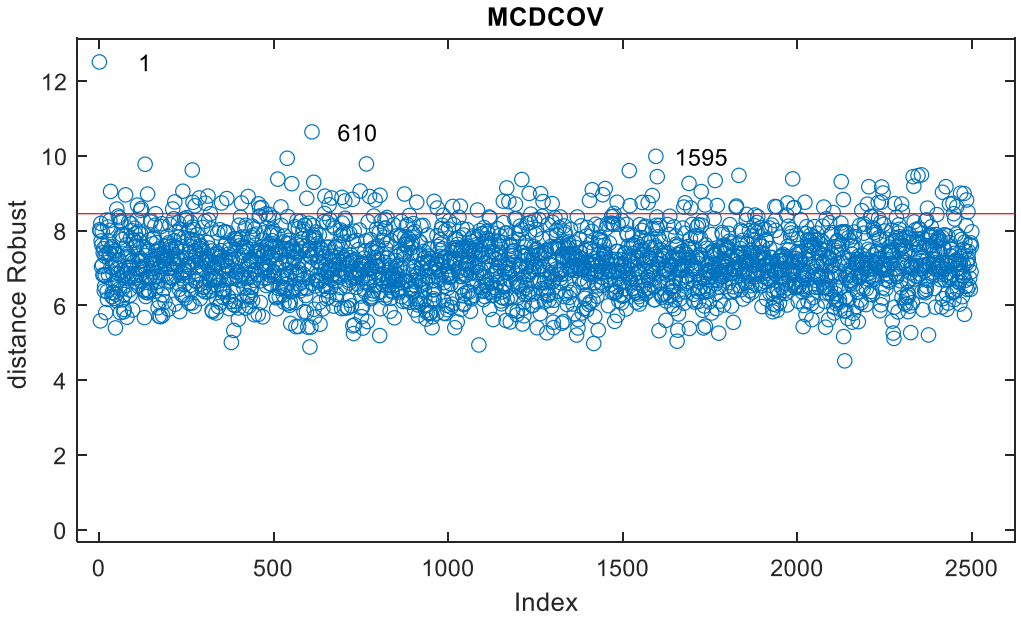

Figure 4.1.b : distance robuste obtenue par la covariance MCD d'après l'hypothèse H_0.

Hypothèse H_1. : le signal source est corrompu avec AWGN et le signal d'attaque proposé :

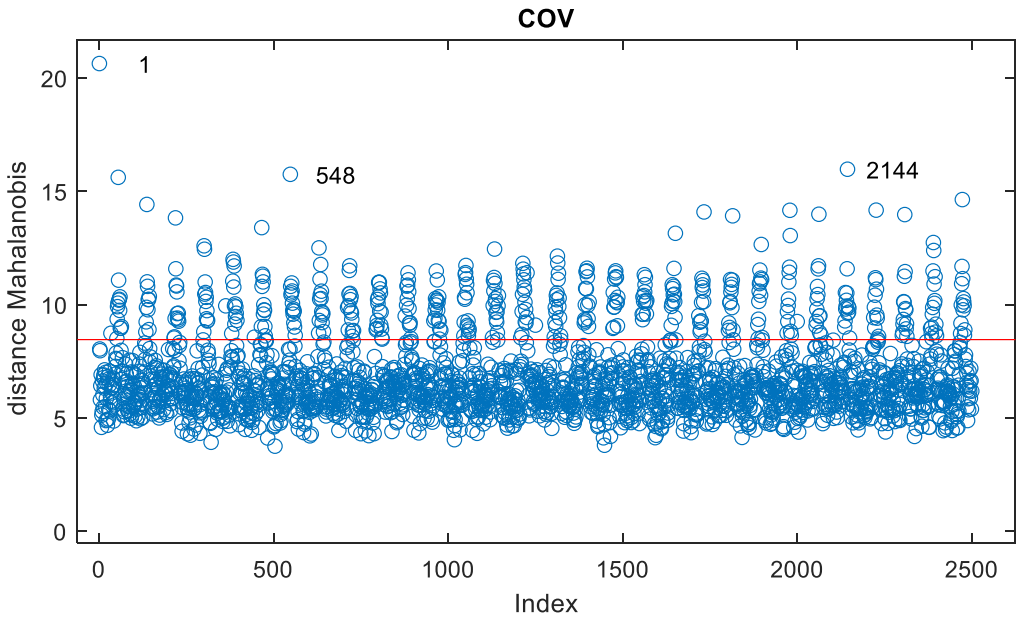

Figure 4.2.a : distance Mahalanobis obtenue d'après l'hypothèse H_1.

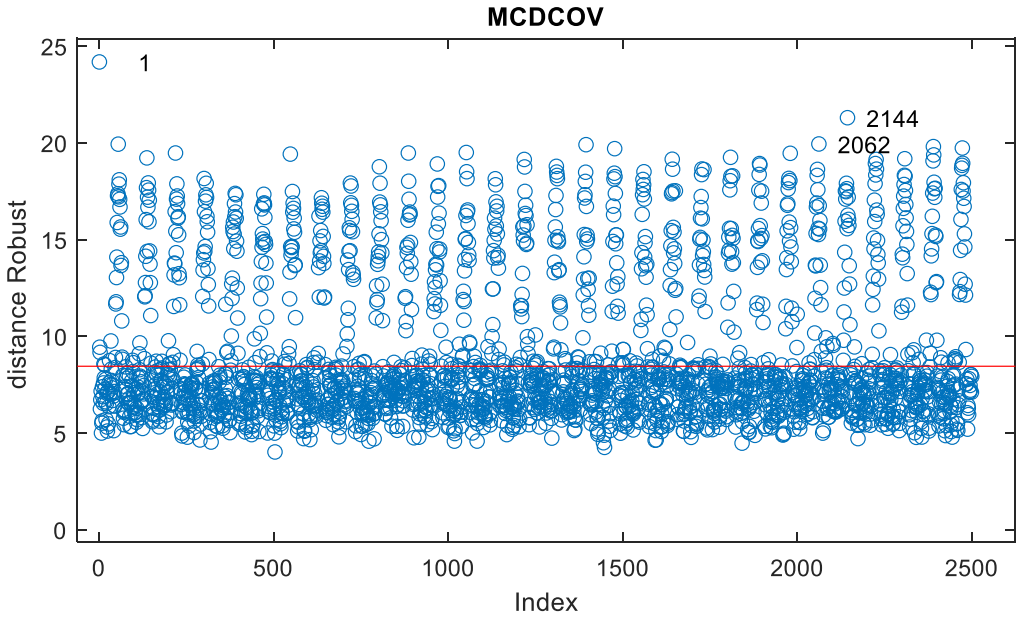

Figure 4.2.b : distance robuste obtenue d'après l'hypothèse H_1.

La ligne rouge représente la valeur seuil de chaque distance, pour prendre une décision sur la présence d'anomalies dans l'ensemble de données multi-variées du schéma proposé, pour *P*RCs, avec 1-α=97,5%, la valeur seuil est définie comme suit

Seuil_DM	8.451
Seuil_DR	8.451

Tableau.4.2 : valeurs du seuil des deux types de distances

	Taux d'anomalies	%	Taux normales	%
DM cas (H_0)	67	2,68%	2433	97,32%
DR cas (H_0)	105	4,2%	2395	95,8%
DM cas (H_1)	442	17,68%	2058	82,32%
DR cas (H_1)	791	31,64%	1709	68,36%

Tableau 4.3 : représente le taux d'anomalies et normales des deux types de distances.

4.4 Discussion :

On remarque clairement d'après les résultats obtenus dans les deux figures des deux types de distance et dans le cas de l'hypothèse H0 : Figure.4.1.a (DM) et Figure.4.1.b (DR)que le taux d'anomalies obtenu pour la distance Mahalanobis (DM=67, estimation en pourcentage2,68%), et pour la distance Robuste (DR=105, estimation en pourcentage4,2%) est trop faible signifie clairement que le signal source est corrompu uniquement avec le bruit. Par contre les résultats obtenus dans les deux figures suivantes des deux types de distances et dans le cas de l'hypothèse H1 : Figure.4.2.a (DM) et Figure.4.2.b (DR), on remarque clairement que le taux d'anomalies obtenu pour la distance Mahalanobis (DM=442, estimation en pourcentage 17,68%), et pour la distance Robuste (DR=791, 31,64%) est trop élevé par rapport au cas $\boldsymbol{H_0}$ signifie bien clairement qu'il s'agit d'une attaque malveillante dans le spectre. On remarque aussi que la distance DR a été plus affectée par l'attaque que la distance DM dans le cas $\boldsymbol{H_1}$et ça due à l'estimateur MCD qui plus robuste à détecter les anomalies par rapport à l'estimateur S classique de l'ACP. Ces résultats obtenus par le détecteur proposé de classification ont réussi avec succès de faire la distinction entre la présence et l'absence de l'attaque dans le spectre. Dans ces cas la valeur de seuil de chaque distance est importante pour faire la distinction basée sur des hypothèses.

Ces résultats obtenus et par comparaison d'après les deux hypothèses supposées montrent que le nouveau schéma proposé fonctionne avec une grande performance dans l'identification des anomalies en utilisant LIBRA. Ces résultats obtenus par la classification sont performants, bons si on les compare avec les résultats obtenus par les auteurs dans [39] qui étudie le cas de classification des scores de l'APC utilisant l'estimateur classique S. La mesure de distance est suffisamment sensible pour détecter les anomalies sur la base de leurs caractéristiques comme la moyenne, la covariance basée sur l'estimateur MCD, et la covariance basée sur l'estimateur S classique de l'ACP (Analyse des composantes Principales). Un autre avantage du détecteur de classification proposé est qu'au cours de l'étape de détection, les statistiques peuvent être calculées en moins de temps, ce qui permet d'utiliser cette technique dans des scénarios en temps réel.

Avantages du détecteur proposé pour la prise de décision en apprentissage automatique :

- Apprentissage automatique basé sur la classification des patrons/observations.
- Détection intelligente par classification.
- Complexité du système est minimisée grâce à l'EC.

Étapes à suivre pour extraire nos résultats obtenus ci-dessus

- **Cas hypothèse H_1**

Premièrement on exécute main_demo.m, ce dernier fais appel à la fonction : mcdcov de LIBRA : [out]=mcdcov(compressed data matrix), pour avoir les résultats qu'on a obtenus ci-dessus dans le cas de l'hypothèse H_1.

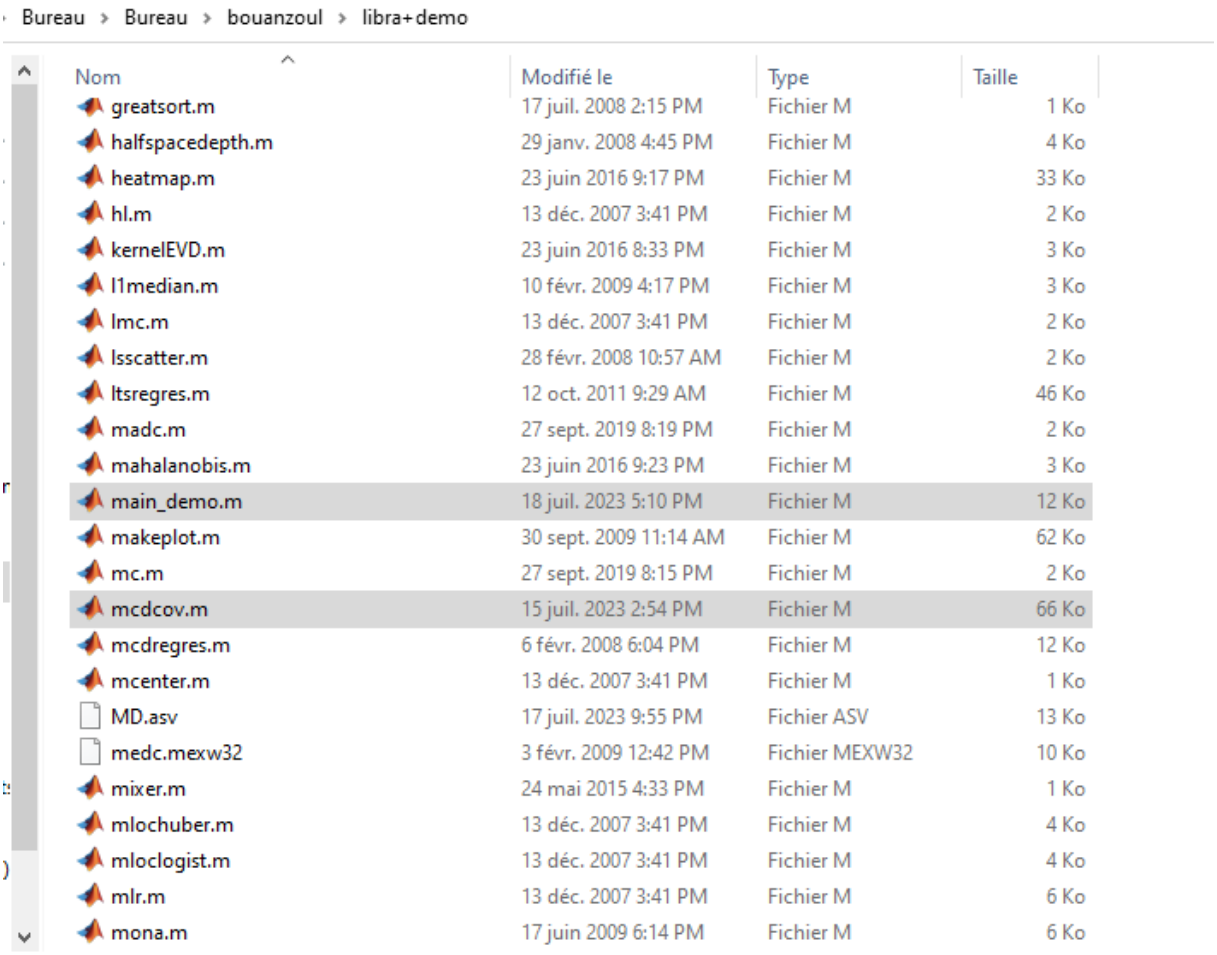

Figure 4.3 : LIBRA.

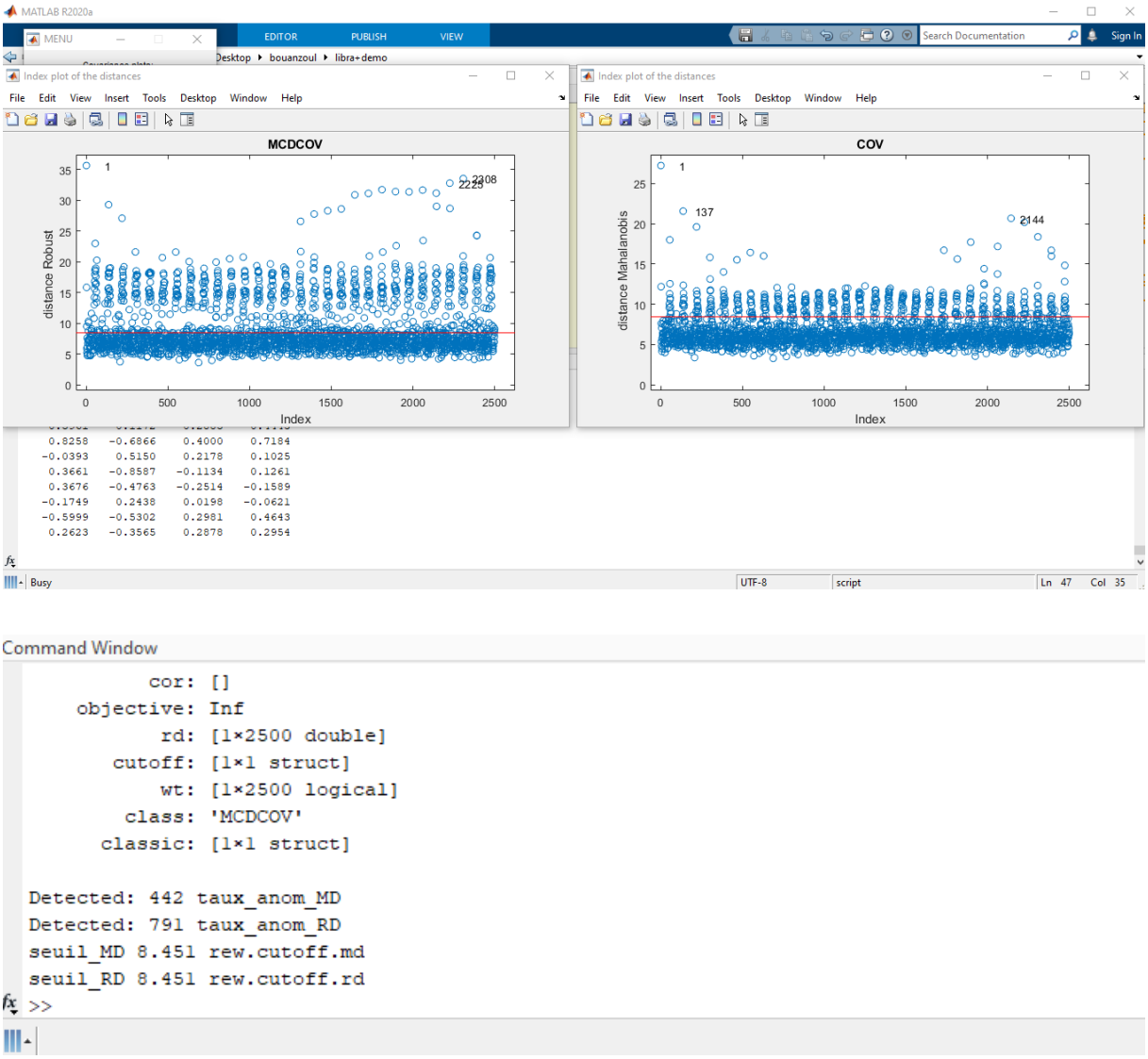

Figure 4.4 : Capture des résultats de simulation par matlab.

4.5 Conclusion :

Nous concluons que les résultats obtenus par simulation et d'après les deux types d'hypothèses supposées ont montrés une bonne efficacité et performance d'avoir une détection intelligente et robuste par classification utilisant LIBRA. On a pu réaliser l'objectif de ce travail de minimiser la complexité des calculs et la consommation d'énergie par les radios et une détection rapide et intelligente a été réalisé.

Conclusion générale et perspectives

Conclusion générale :

Un nouveau schéma intelligent a été proposé. Le CMLB centralisé combiné au détecteur d'anomalies basé sur DM et DR basé sur la librairie d'analyses des données robuste LIBRA en cas de présence d'attaque malveillante. L'un des principaux objectifs de ce travail est de concevoir une technique de détection intelligente par classification. Des hypothèses sont montrées dans ce travail pour distinguer la présence d'attaque ou pas. Nous obtenons une matrice de données compressées qui conserve les propriétés linéaires du signal multi-bande échantillonné à large bande de sous Nyquist. Chaque utilisateur radio via une stratégie de coopération centralisée envoie ses mesures compressées au CF pour obtenir une détection intelligente. Le CF traite les observations compressées de chaque utilisateur radio sous la forme d'une matrice de données compressée utilisant le détecteur proposé. Le CF prend une décision basée sur la valeur du seuil de chaque distance décrite dans le chapitre 3 pour savoir s'il s'agit d'une attaque présente dans spectre ou pas. La performance de ce schéma donne de bons résultats, la complexité du système est considérablement réduite, la consommation d'énergie de chaque radio est réduite, on a réalisé une détection intelligente sans passer par l'étape de récupération. Ce nouveau schéma est considéré comme une solution prometteuse dans les systèmes de communication cognitifs à l'avenir et peut être utilisé dans les scénarios en temps réel.

Perspectives:

- Étude d'autres techniques de classification basée sur l'apprentissage automatique.
- Etude dans le cas de classifieur Hybride.

Références

Références

[1] EsraAltulaihan. Cybersecurity Threats, Countermeasures and Mitigation Techniques on the IoT: Future Research Directions. Electronics 2022, 11(20),3330; https://doi.org/10.3390/electronics11203330.

[2] YounessArjoune, and Naima Kaabouch. Wideband Spectrum Sensing: A Bayesian Compressive Sensing Approach. Sensors 2018, 18(6), 1839 https://doi.org/10.3390/s18061839.

[3] Yulong Zou. 2015. Securing Physical-Layer Communications for Cognitive Radio Networks. IEEE Communications Magazine, 53(9).doi 10.1109/MCOM.2015.7263345

[4] Tom Howley et al. The Effect of Principal Component Analysis on Machine Learning Accuracy with High Dimensional Spectral Data. Conference: Applications and Innovations in Intelligent Systems XIII, Proceedings of AI-2005, the Twenty-fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, Cambridge, UK, December 2005. doi: 10.1007/1-84628-224-1_16.

[5] Reem Melki et al. Machine Learning for Physical Layer Security: Limitations, Challenges and Recommendation. Conference: 16TH INTERNATIONAL CONFERENCE ON SIGNAL IMAGE TECHNOLOGY & INTERNET BASED SYSTEMS (SITIS 2022), Dijon, France. DOI: 10.1109/SITIS57111.2022.00017.

[6] Tai-Ning Yang, Sheng-De Wang. 1999. Robust algorithms for principal component analysis. Pattern Recognition Letter journal.DOI:10.1016/S0167-8655(99)00060-4.

[7] Mitola III, J. (1999). Cognitive radio: An integrated agent architecture for software defined radio. PhD thesis, KTH, Royal Institute of Technology, Stockholm, Sweden.

[8] E. Candes, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," IEEE Trans. on Information Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[9] Lamelas Polo, Y. Compressive Wideband Spectrum Sensing for Cognitive Radio Applications. Master thesis, 2008, delft yniversity, pages: Lamelas Polo, Y.

[10] D. L. Donoho, Compressed sensing," IEEE Trans. on Information Theory, vol. 52, no. 4, pp. 1289-1306, April 2006.

[11] J. A. Tropp et al., "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," IEEE Trans. Info. Theory, vol. 56, no. 1, Jan. 2010, pp. 520–44.

[12] Z. Tian and G. Giannakis, "Compressive Sensing for Wideband Cognitive Radios," Proc. IEEE Int'l. Conf. Acoustics, Speech, and Sig. Proc., Honolulu, HI, April 2007, pp. 1357–60.

[13] Hongjyan Sun et al. WIDEBAND SPECTRUM SENSING FOR COGNITIVE RADIO NETWORKS: A SURVEY. IEEE Wireless Communications • April 2013, pp:1-8.

[14] Vivek Upadhyaya and Dr. Mohammad Salim. Compressive Sensing: Methods, Techniques, and Applications. Conference: International Conference on Applied Scientific Computational Intelligence using Data Science (ASCI 2020), pp:1-23, DOI: 10.1088/1757-899X/1099/1/012012

[15] R. Venkataramani and Y. Bresler, "Perfect Reconstruction Formulas and Bounds on Aliasing Error in sub–Nyquist Nonuniform Sampling of Multiband Signals," IEEE Trans. Info. Theory, vol. 46, no. 6, Sept. 2000, pp. 2173–83.

[16] Nemanja Milosevic et al. Wide-band cooperative spectrum sensing method. CogART '11: Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management, October 2011, No.28, Pages 1–5

[17] M. Mishali and Y. C. Eldar, "Blind Multiband Signal Reconstruction: Compressive Sensing for Analog Signals," IEEE Trans. Sig. Proc., vol. 57, no. 3, March 2009, pp. 993–1009.

[18] Saziaparvin et al. Cognitive radio network security: A survey. Journal of Network and Computer Applications, 2012, 35(6):1691–1708

[19] Forouzan, Behrouz A. "Data Communications and Networking." McGraw-Hill, 2012.

[20] K. Anbukkarasi, Countermeasure against physical layer attack in cognitiveradio networks, Int. J. Electron. Commun. Eng. (March) (2017) 195–200.

[21] Fatima Salahdine, KaabouchNaima. Security Threats, Detection, and Countermeasures for Physical Layer in Cognitive Radio Network: A Survey. Physical Communication, April 2020,DOI: 10.1016/j.phycom.2020.101001

[22] Clancy TC, Goergen N. Security in cognitive radio networks: threats and mitigation. In: Proc. of 3rd international IEEE conference on in cognitive radio oriented wireless networks and communications. Singapore; 2008. p. 1–8.

[23] Qin TYH, Leung C, Shen Z, Miao C. Towards a trust aware cognitive radio architecture. Newsletter ACM SIGMOBILE Mobile Computing and Communica- tions Review 2009;13(2):86–95.

[24] Chen R, Park J-M, Hou YT, Reed JH. Toward secure distributed spectrumsensing in cognitive radio networks. IEEE Communications Magazine, Special Issue on Cognitive Radio Communications 2008b:50–5.

[25] Cheng, Chi-Hao; Tsui, James (2021). An Introduction to Electronic Warfare; from the First Jamming to Machine Learning Techniques. Oxon: CRC Press. p. 47. ISBN 978-87-7022-435-2.

[26] Tarr, S. M¨uller, and N. C. Weber. Robust estimation of precision matrices under cellwise contamination. Computational Statistics & Data Analysis, 93:404–420, 2016.

[27] Shi, D.D. Jia, R.Y. and Huang, Y.T. (2009) Improvement of Outlier Detection Algorithm in High Dimension Based on Genetic Algorithm. Journal of Computer Technology and Development, 19, 141-143.

[28] Donoho, D.L. (1982) Breakdown Properties of Multivariate Location Estimators. Technical Report, Harvard University, Boston.

[29] Grübel, R. (1988) A Minimal Characterization of the Covariance Matrix. Metrika, 35, 49-52. https://doi.org/10.1007/BF02613285

[30] Rousseeuw, P.J. and Driessen, K.V. (1999) A Fast Algorithm for the Minimum Co-variance determinant Estimator. Technometrics, 41, 212-223. https://doi.org/10.1080/00401706.1999.10485670

[31] Feng, L., Li, B. and Huang, L. (2014) Detection and Analysis of Lidar Point Cloud Gross Error Based on Robust Mahalanobis Distance. Geodesy and Geodynamics,34, 168-173.

[32] Chengzhen Wang et al. Machine learning-based approach to GPS antijamming. GPS Solutions, July 2021, 25(3), pp :1-13, doi:10.1007/s10291-021-01154-7.

[33] Ahmed Moumena. Quickest physical-layer MGD anomaly detection for jamming attacks in centralized modulated wideband converter-based ROC curve. Int. J. Commun. Syst. 32(18) (2019)]

[34] Moshe Mishali, Yonina C. Eldar. (2011), Wideband Spectrum Sensing at Sub-Nyquist Rates, IEEE Signal Process. Mag, Vol.28, 4, pp.102-135.

[35] Alberts, D. S., J. J. Garstka, et F. P. Stein. (1999), Network Centric Warfare Developing and Leveraging Information Superiority. 2e édition, Washington, D.C., DoD C4ISR Cooperative Research Program.

[36] Sabine Verboven et al. LIBRA: a Matlab library for robust analysis, February 2005, Chemometrics and Intelligent Laboratory Systems 75(2):127-136, doi: 10.1016/j.chemolab.2004.06.003.

[37] MATLAB: A Practical Introduction to Programming and Problem Solving de Stormy Attaway

[38] https://github.com/mwgeurts/libra

[39] Ahmed Moumena. Centralized Wideband Cooperative CRs Combined with PCA Anomaly Detector in the Presence of malicious Attacks Conference: The IEEE Seventh International Conference on Image and Signal Processing and their Applications (ISPA22); Mostaghanem, Algeria. DOI: 10.1109/ISPA54004.2022.9786282].