

Université du Québec
Institut national de la recherche scientifique
Centre Énergie Matériaux Télécommunications

**DÉTECTION À BASE D'ANN DANS DES SYSTÈMES MIMO-OFDM AVEC
DES ADC À BASSE RÉOLUTION**

Par

Shabnam Rezaei

Mémoire présenté pour l'obtention du grade de
Maître es Sciences, M.Sc.
en sciences de Télécommunications

Jury d'évaluation

Examineur externe	Prof. Claude D'Amours Université d'Ottawa
Examineur interne	Prof. Tayeb A. Denidni INRS-EMT
Directeur de recherche	Prof. Sofiène Affes INRS-EMT

Remerciements

Ce travail a été élaboré au sein de l'équipe Wireless Lab <www.wirelesslab.ca> du Centre Énergie, Matériaux et Télécommunications de l'Institut National de la Recherche Scientifique.

Je tiens à remercier toute personne qui a contribué de proche ou de loin à l'élaboration et au succès de ce travail.

Tout d'abord, je remercie mon directeur de recherche le Professeur Sofiène Affes qui m'a donné cette opportunité en premier lieu pour intégrer l'équipe Wireless Lab et qui m'a offert un environnement propice avec ses riches conseils et sa présence permanente afin d'assurer l'accomplissement de mes travaux de recherche.

Je suis aussi reconnaissant aux membres du Jury pour le temps et l'attention qu'ils ont portée à ce mémoire.

Finalement, je suis redevable à toute personne qui a contribué à ce travail par sa simple présence ou ses connaissances.

Résumé

Dans ce travail, nous proposons un réseau de neurones artificiels multicouches "Artificial Neural Network" (ANN), formé à l'algorithme de Levenberg-Marquardt, à utiliser dans la détection de signaux sur des systèmes de multiplexage par répartition orthogonale de la fréquence à plusieurs entrées "Multiple-Input and Multiple-Output-Orthogonal Frequency-Division Multiplexing" (MIMO-OFDM). Nous suggérons, en particulier, son usage avec les convertisseurs analogiques-numériques à faible résolution "Low Resolution Analog-to-Digital Converter" (LRADC).

Nous considérons un schéma de détection à l'aveugle dans lequel l'estimation de symbole de données est réalisée sans connaissance des informations d'état de canal sur le récepteur "Channel State Information at Receiver" (CSIR), contrairement aux algorithmes classiques.

La principale puissance du détecteur à base d'ANN réside dans son utilisation polyvalente avec n'importe quel schéma de modulation, à l'aveugle, mais sans modification de sa structure. Nous comparons, par simulations, ce nouveau récepteur avec des récepteurs classiques, afin de connaître le maximum de vraisemblance "Maximum-Likelihood" (ML), l'erreur quadratique moyenne minimale "Minimum Mean Squared Error" (MMSE) et le forçage à zéro "Zero Forcing" (ZF), en termes de performances du taux d'erreur de symbole "Symbol Error Rate" (SER).

Les résultats suggèrent que l'ANND approche le ML avec une complexité beaucoup plus faible, surpasse le ZF sur toute la plage de valeurs du rapport signal sur bruit "Signal-to-Noise Ratio" (SNR), et le fait également avec le MMSE, sur différentes gammes de SNR.

Termes indexés- réseaux de neurones artificiels; MIMO-OFDM; Convertisseurs analogique-numérique; Détection de signal.

Table des matières

Remerciements	iii
Résumé	v
Table des matières	vii
Liste des figures	ix
Liste des tableaux	xi
Liste des abréviations	xiii
Introduction générale	1
1 Techniques de détection dans les systèmes multi-antennes	5
1.1 Méthodes de détection linéaires	5
1.1.1 Méthode du maximum de vraisemblance (Maximum-Likelihood (ML))	6
1.1.2 Méthode de forçage à zéro (Zero-Forcing (ZF))	7
1.1.3 Méthode de l'erreur quadratique moyenne linéaire (Linear Minimum Mean Squared Error (L-MMSE))	7
1.2 Méthodes de détection à l'aveugle	7
1.3 Méthodes de détection semi-aveugles	8
1.4 Conclusion	9
2 Apprentissage Machine	11
2.1 Vue d'ensemble	11
2.2 Les Tâches d'apprentissage machine	11
2.3 Théorie	13
2.4 Approches	14
2.4.1 Types d'algorithmes d'apprentissage	14
2.4.2 Modèles	17
2.4.3 Modèles de formation	22
2.5 Techniques d'apprentissage machine pour les communications sans fil	23
2.6 Conclusion	24
3 Imperfections de ADC dans les systèmes sans fil multi-antennes	25
3.1 Présentation des ADC	25
3.1.1 Introduction	25
3.1.2 Types de convertisseurs ADC	26

3.1.3	Erreurs de quantification	28
3.2	ADC dans les systèmes MIMO	28
3.2.1	Modèle ADC	29
3.2.2	Résultats	30
3.3	Conclusion	31
4	Conception d'un algorithme de détection basé sur ANN et évaluation de ses performances	33
4.1	Introduction	33
4.2	Modèle du système	34
4.2.1	Systèmes à antennes multiples (MIMO)	34
4.2.2	Modulation de multiplexage par répartition orthogonale en fréquence (OFDM)	35
4.2.3	Modulation numérique sur les canaux sans fil	36
4.2.4	Système MIMO-OFDM	39
4.3	Modèle de réseaux de neurones artificiels	40
4.3.1	Introduction	40
4.3.2	Réseaux de Neurones de Perceptron Multicouches	41
4.3.3	Algorithme de Levenberg-Marquardt	42
4.4	Détection de signal basée sur ANN	43
4.4.1	Introduction	43
4.4.2	Détecteur basé sur ANN	43
4.4.3	Phase d'entraînement ANN	44
4.5	Résultats and discussion	44
5	Conclusion	49
	Références	49

Liste des figures

1.1	Schéma fonctionnel des systèmes de communication numériques (Goldsmith, 2005). . .	5
1.2	Système de bout en bout (Yan <i>et al.</i> , 2017).	6
1.3	Détecteur MMSE avec filtre anti-bruit (Wu <i>et al.</i> , 2014).	8
2.1	Un modèle de réseau neuronal artificiel (Kolman & Margaliot, 2005).	17
2.2	Un exemple de modèle d'arbre de décision (Estevez <i>et al.</i> , 2011).	20
3.1	Le modèle d'erreur ADC utilisé dans la simulation du système de communication (Choi <i>et al.</i> , 2016).	29
3.2	Le taux d'erreur sur les bits résultant pour les deux expériences.	31
4.1	Systèmes MIMO (He <i>et al.</i> , 2018).	35
4.2	Modèle de système MIMO-OFDM.	39
4.3	Réseau de neurones MLP.	41
4.4	Algorithme de formation de ANN.	44
4.5	Performances du SER des techniques de détection proposées et conventionnelles pour les systèmes 2×2 MIMO-OFDM pour BPSK et QPSK.	46
4.6	Performances du SER des techniques de détection proposées et conventionnelles pour les systèmes 4×4 MIMO-OFDM pour BPSK et QPSK.	46
4.7	Performances du SER des techniques de détection proposées et conventionnelles pour différents nombres de bits ADC pour 2×2 systèmes MIMO-OFDM pour QPSK. . .	47
4.8	Performances du SER des techniques de détection proposées et conventionnelles pour différents nombres de bits ADC pour 4×4 systèmes MIMO-OFDM pour QPSK. . .	48

Liste des tableaux

- 4.1 Paramètres du système MIMO-OFDM. 45
- 4.2 Paramètres ANN. 45
- 4.3 La Complexité de Temps Informatique. 47

Liste des abréviations

ADC	Analog-to-Digital Converter
ANN	Artificial Neural Network
ANND	Artificial Neural Network-based Detector
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase-Shift-Keying
CP	Cyclic Prefix
CSI	Channel State Information
CSIR	Channel State Information at Receiver
ENOB	Effective Number Of Bits
FF-MLP	Feed-Forward Multi-Layered Perceptron
FFT	Fast Fourier Transform
GA	Genetic Algorithm
IC	Integrated Circuits
ISI	Inter-Symbol Interference
LM	Levenberg-Marquardt
LRADC	Low Resolution Analog-to-Digital Converter
MDP	Markov Decision Process
MIMO	Multiple-Input and Multiple-Output
ML	Machine learning
ML	Maximum-Likelihood
MLP	Multi-Layered Perceptron
MLSE	Maximum-Likelihood Sequence Estimation
MMSE	Minimum Mean Squared Error
MPAM	M-ary Pulse Amplitude Modulation
MQAM	M-ary Quadrature Amplitude Modulation

OFDM	Orthogonal Frequency-Division Multiplexing
PSK	Phase-Shift-Keying
QPSK	Quadrature Phase-Shift-Keying
RBF	Radial Basis Function
RNN	Recurrent Neural Network
SER	Symbol Error Rate
SINAD	Signal-to-Noise And Distortion
SISO	Single Input Single Output
SNR	Signal-to-Noise Ratio
SQNR	Signal-to-Quantization-Noise Ratio
SVM	Support Vector Machine
ZF	Zero-Forcing

Introduction générale

La plupart des techniques de détection MIMO développées jusqu'à présent sont basées sur un CSIR parfait pour effectuer la détection. Par conséquent, l'estimation du canal pour de telles techniques est inévitable (Luo *et al.*, 2018). En fait, plus le CSIR est précis, meilleure sera la performance de détection. Les algorithmes d'estimation de canal les plus populaires sont : ML, ZF et MMSE. (Choi *et al.*, 2016) utilise un algorithme ZF pour effectuer une estimation de canal MIMO. (O'Shea *et al.*, 2017) envisage d'effectuer une estimation de canal basée sur ML pour prédire le CSIR. L'inconvénient de ces techniques est la nécessité d'effectuer des manipulations matricielles complexes qui prennent du temps (Seyman & Taşpınar, 2013), (Dörner *et al.*, 2018).

Par conséquent, de tels algorithmes classiques conviennent à une estimation CSIR à petite échelle (c'est-à-dire aux rares antennes) dans des systèmes de communication sans fil. D'autre part, les futurs systèmes sans fil tels que les systèmes mmWave ont pour objectif de fournir des liaisons de communication avec des débits de données élevés en utilisant de larges largeurs de bande.

À mesure que la bande passante du système augmente, le taux d'échantillonnage d'un convertisseur analogique-numérique (ADC) devrait également augmenter de façon linéaire. Malheureusement, des taux d'échantillonnage plus élevés entraînent une consommation d'énergie plus élevée des CAN (Walden, 1999).

L'utilisation de LRADC a été considérée comme une solution rentable pour réduire la consommation d'énergie des systèmes nécessitant des convertisseurs ADC à haute vitesse. En déployant des LRADC dans un système, le CSIR précis ne peut pas être obtenu à partir des signaux pilotes en raison de l'excès d'erreur de quantification. Pour résoudre ce problème, plusieurs méthodes d'estimation de canaux ont été développées.

Un problème commun à tous ces algorithmes est la nécessité d'une grande quantité de signaux pilotes pour surmonter l'erreur excessive. La détection aveugle dans les systèmes de communication MIMO a fait l'objet de nombreuses études. Par exemple, une technique de détection MIMO complètement aveugle a été développée dans (Tong *et al.*, 1994), et l'algorithme de classification K-means a été proposé dans (Liang *et al.*, 2013) pour la détection aveugle. Le principal inconvénient de ces algorithmes de détection à l'aveugle est la grande complexité de leur implémentation afin de produire un CSIR précis, qui peut ne pas être abordable pour des systèmes de communication pratiques. De plus, des méthodes de détection MIMO semi-aveugles, qui effectuent la détection de données et l'estimation de canaux en commun, ont été développées et ont démontré de meilleures performances que les méthodes de détection cohérentes.

Cependant, le problème clé lié à ces techniques aveugles est qu'elles ne sont disponibles que pour certains types de modulations, comme la modulation par décalage dans l'espace ou la modulation par décalage de phase "Phase-Shift-Keying" (PSK). Récemment, de nombreux chercheurs ont commencé à rechercher une solution innovante basée sur l'apprentissage automatique (ML). Par exemple, pour le problème principal de la détection aveugle discuté ci-dessus, les auteurs de Jeon *et al.* (2017) ont complètement repensé l'opération de détection en appliquant des algorithmes ML, sans recourir au CSIR.

Ils ont utilisé un classificateur KN-nearest neighbors (KNN) pour la détection de symboles sur des systèmes MIMO. Ainsi, ils ont démontré que KNN est comparable en termes de performance aux algorithmes de détection conventionnels. Cependant, sa précision peut diminuer dans les systèmes MIMO multi-utilisateurs en raison de ses limites inhérentes par rapport à d'autres algorithmes d'apprentissage tels que ANN. Dans cet article, nous proposons une approche de détection de signal aveugle basée sur ANN pour les systèmes MIMO-OFDM avec ADC parfaits et avec LRADC quand un récepteur CSIR explicite n'est pas connu.

Les réseaux de neurones ont offert des solutions à la pointe de la technologie dans de nombreux domaines autres que les classifications. Puisque nous pouvons considérer la détection comme un acte de classification, nous avons choisi cette technique comme classificateur. La formation d'un réseau de neurones implique de nombreux hyper-paramètres contrôlant la taille et la structure du réseau ainsi que la procédure d'optimisation permettant d'obtenir une meilleure détection sur des systèmes à grande échelle. L'approche proposée comprend deux phases. Dans la première phase,

l'émetteur envoie une séquence de symboles de données afin que le récepteur apprenne une fonction non linéaire qui décrit les relations entrée-sortie du système.

Dans la deuxième phase, à l'aide du réseau de neurones formé, le récepteur détecte les symboles de données. Cette nouvelle approche peut être considérée comme un problème de classification dans l'apprentissage supervisé. Le classifieur cartographie le signal reçu sur l'un des vecteurs de symbole possibles. Le principal avantage de notre détecteur basé sur la ANN est que son fonctionnement est indépendant du schéma de modulation adopté par le système, contrairement aux techniques de détection classiques, dans lesquelles la structure du récepteur doit être adaptée au schéma de modulation.

Le travail est organisé comme suit. Dans le premier chapitre, nous présentons les techniques de détection les plus courantes adoptées dans un système multi-antennes. Dans le deuxième chapitre, nous présentons l'apprentissage automatique "Machine Learning" (ML) comme une alternative prometteuse pour la détection à faible coût par OFDM-MIMO. Dans le troisième chapitre, nous expliquons comment les imperfections du ADC peuvent affecter la précision du nouveau détecteur. Enfin, dans le quatrième chapitre, nous le simulons et l'évaluons juste avant de conclure ce travail dans le chapitre cinq.

Chapitre 1

Techniques de détection dans les systèmes multi-antennes

1.1 Méthodes de détection linéaires

Les systèmes MIMO sont définis comme des liaisons de communication point à point avec plusieurs antennes au niveau de l'émetteur et du récepteur (Yan *et al.*, 2017).

L'utilisation de plusieurs antennes au niveau de l'émetteur et du récepteur offre clairement des performances améliorées par rapport aux systèmes en diversité où l'émetteur ou le récepteur, mais pas les deux, possède plusieurs antennes. En particulier, des recherches récentes ont montré que les systèmes MIMO peuvent augmenter considérablement les débits de données des systèmes sans fil sans augmenter la puissance de transmission ou la bande passante (cf. Fig.1.2).

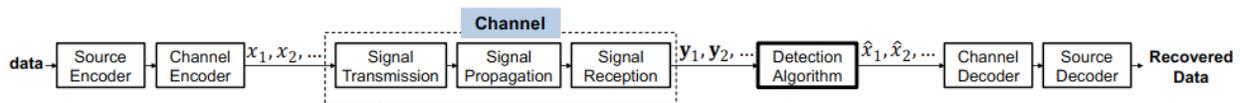


Figure 1.1 – Schéma fonctionnel des systèmes de communication numériques (Goldsmith, 2005).

Le coût de cette augmentation est le coût supplémentaire du déploiement de plusieurs antennes, l'encombrement de ces antennes supplémentaires (en particulier sur les petites unités portables) et la complexité supplémentaire requise pour le traitement du signal multidimensionnel.

La technologie MIMO à plusieurs antennes a été largement utilisée dans les systèmes de communication et dans d'autres domaines en raison de ses gains de performances considérables. Les chercheurs ont développé de nombreux types d'algorithmes de détection de signaux pour les systèmes MIMO. Les techniques de détection se divisent en deux grandes catégories: linéaire et non linéaire. Les techniques linéaires sont généralement les plus simples à implémenter et à comprendre conceptuellement. Les trois principaux algorithmes sont: le maximum de vraisemblance (ML), le forçage nul (ZF), l'erreur quadratique moyenne minimale linéaire (L-MMSE).

1.1.1 Méthode du maximum de vraisemblance (Maximum-Likelihood (ML))

Parmi les algorithmes de détection de signal MIMO, le meilleur est le détecteur de maximum de vraisemblance (ML) qui peut minimiser le taux d'erreur sur les bits pour les systèmes MIMO.

L'estimation de séquence de Maximum-likelihood évite le problème de l'amélioration du bruit car elle n'utilise pas de filtre d'égalisation, mais elle estime la séquence des symboles transmis (Yang & Hanzo, 2015).

La structure de l'entité MLSE est identique à celle de la Fig. 1.2 à la différence que l'égaliseur $H_{eq}(z)$ et le dispositif décisionnel sont remplacés par l'algorithme MLSE.

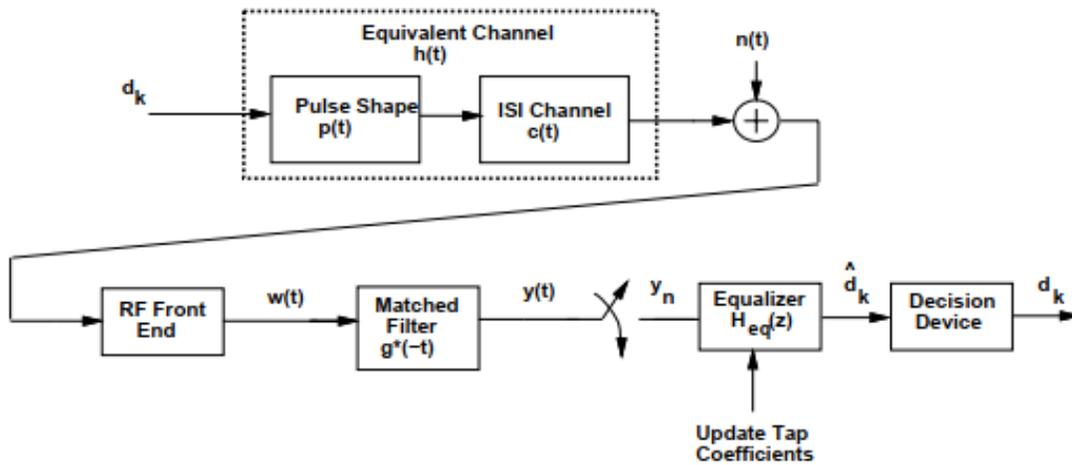


Figure 1.2 – Système de bout en bout (Yan *et al.*, 2017).

1.1.2 Méthode de forçage à zéro (Zero-Forcing (ZF))

Ce détecteur annule tous les ISI, mais peut entraîner une amélioration considérable du bruit. Le forçage à zéro peut provoquer une amplification du bruit si la valeur singulière minimale de \mathbf{H} est trop petite (Wang *et al.*, 2016).

Ceci peut être quantifié par la notion de numéro de condition de la matrice \mathbf{H} . Le nombre de condition de la matrice \mathbf{H} est une mesure des grandeurs relatives des valeurs singulières de \mathbf{H} . Il est défini comme le rapport entre la plus grande et la plus petite valeur singulière de \mathbf{H} . Lorsque le numéro de condition est égal à l'unité ou proche de l'unité, la matrice est dite bien conditionnée. Lorsque le numéro de condition est grand, la matrice est mal conditionnée.

Pour réduire la sensibilité des récepteurs linéaires au conditionnement de la matrice \mathbf{H} , nous pouvons ajouter un terme de régularisation à la fonction objectif selon la méthode de l'erreur quadratique moyenne linéaire.

1.1.3 Méthode de l'erreur quadratique moyenne linéaire (Linear Minimum Mean Squared Error (L-MMSE))

Dans la détection MMSE, l'objectif de la conception du détecteur est de minimiser l'erreur quadratique moyenne entre le symbole transmis d_k et son estimation \hat{d}_k à la sortie du détecteur (cf. Fig.1.3).

L'algorithme MMSE a une vitesse de détection supérieure, mais n'atteint pas les performances du détecteur ML (Kim *et al.*, 2008).

1.2 Méthodes de détection à l'aveugle

Les algorithmes de détection classiques brièvement présentés ci-dessus peuvent difficilement atteindre simultanément une précision de détection élevée et une complexité faible. Par conséquent, les chercheurs se sont montrés très intéressés par les nouvelles techniques ne nécessitant pas l'acquisition de CSI, notamment les méthodes de détection à l'aveugle.

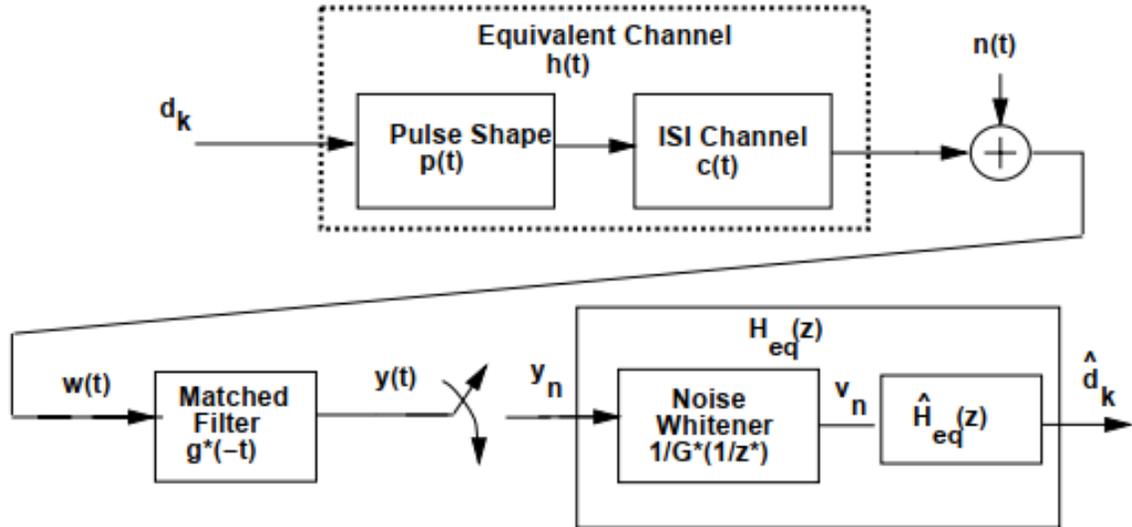


Figure 1.3 – Détecteur MMSE avec filtre anti-bruit (Wu *et al.*, 2014).

La détection aveugle estime le signal transmis sans recourir à des signaux pilotes, ce qui augmente le débit et simplifie la structure des données (Wu *et al.*, 2014).

La détection à l’aveugle sur les systèmes de communication MIMO en particulier a récemment fait l’objet d’études approfondies. Par exemple, une technique de détection MIMO complètement aveugle a été développée dans (Tong *et al.*, 1994), et l’algorithme de classification K-means a été proposé dans (Liang *et al.*, 2013) pour la détection aveugle. Cependant, le principal inconvénient de ces algorithmes de détection à l’aveugle est qu’ils impliquent des coûts de mise en œuvre élevés difficilement abordables dans les systèmes de communication pratiques.

1.3 Méthodes de détection semi-aveugles

Des méthodes de détection MIMO semi-aveugles, qui effectuent conjointement la détection de données et l’estimation de canaux, ont été développées et démontrent de meilleures performances que les méthodes de détection cohérentes. Cependant, leur principal problème est qu’ils ne fonctionnent bien que pour certains types de modulations, comme la modulation par déplacement d’espace ou la sélection par décalage de phase (PSK) (Wang *et al.*, 2016).

1.4 Conclusion

Les communications sans fil joueront un rôle majeur dans les futurs systèmes de communication. En fait, le besoin d'accès sans fil à l'Internet deviendra de plus en plus important et de nouveaux paradigmes de communication réseau, tels que les réseaux sans fil ad-hoc ou les systèmes intégrés mobiles / satellites, seront développés dans un proche avenir. Les algorithmes de détection joueront un rôle important dans la conception de systèmes de communication sans fil efficaces, car il sera impératif de respecter des contraintes spécifiques en termes de consommation d'énergie, de coût de calcul et de vitesse de détection. Les communications sans fil joueront un rôle majeur dans les futurs systèmes de communication. En fait, le besoin d'accès sans fil à l'Internet deviendra de plus en plus important et de nouveaux paradigmes de communication réseau, tels que les réseaux sans fil ad-hoc ou les systèmes intégrés mobiles / satellites, seront développés dans un proche avenir. Les algorithmes de détection joueront un rôle important dans la conception de systèmes de communication sans fil efficaces, car il sera impératif de respecter des contraintes spécifiques en termes de consommation d'énergie, de coût de calcul et de vitesse de détection.

Chapitre 2

Apprentissage Machine

2.1 Vue d'ensemble

L'appellation Apprentissage Machine a été inventée en 1959 par Arthur Samuel. Tom M. Mitchell a fourni une définition plus formelle et largement citée des algorithmes étudiés dans le domaine de l'apprentissage automatique: "Un programme informatique apprendrait de l'expérience E en ce qui concerne une classe de tâches T et la mesure de performance P si sa performance à des tâches dans T , mesurée par P , s'améliore avec l'expérience E ".

Cette définition de l'apprentissage automatique en termes de tâches offre une définition fondamentalement opérationnelle plutôt que de définir le champ en termes cognitifs. Cela fait suite à la proposition d'Alan Turing dans son article "Computing Machinery and Intelligence", dans lequel la question "Les machines peuvent-elles penser?" est remplacé par la question "Les machines peuvent-elles faire ce que nous (en tant qu'entités pensantes) pouvons faire?" (Ertekin *et al.*, 2011).

Dans la proposition de Turing, les diverses caractéristiques d'une machine à penser et ses diverses implications pour la construction de celle-ci sont exposées.

2.2 Les Tâches d'apprentissage machine

Les tâches d'apprentissage automatique sont classées en plusieurs grandes catégories.

Dans l'apprentissage supervisé, l'algorithme construit un modèle mathématique à partir d'un ensemble de données contenant à la fois les entrées et les sorties souhaitées. Par exemple, si la tâche déterminait si une image contenait un certain objet, les données d'apprentissage d'un algorithme d'apprentissage supervisé incluraient des images avec et sans cet objet (l'entrée), et chaque image aurait une étiquette (la sortie) indiquant s'il contenait l'objet (Chen & Wang, 2011). Dans des cas particuliers, la saisie peut n'être que partiellement disponible ou limitée à un retour d'informations spécial. Les algorithmes d'apprentissage semi-supervisés développent des modèles mathématiques à partir de données d'entraînement incomplètes, dans lesquelles une partie de l'échantillon n'a pas d'étiquette.

Les algorithmes de classification et les algorithmes de régression sont des types d'apprentissage supervisé. Les algorithmes de classification sont utilisés lorsque les sorties sont limitées à un ensemble limité de valeurs. Pour un algorithme de classification qui filtre les courriers électroniques, l'entrée serait un courrier entrant et le résultat serait le nom du dossier dans lequel le classer. Pour un algorithme qui identifie les spams, le résultat serait la prédiction de "spam" ou de "non spam", représentée par les valeurs booléennes vrai et faux. Les algorithmes de régression sont nommés pour leurs sorties continues, ce qui signifie qu'ils peuvent avoir n'importe quelle valeur dans une plage. Des exemples de valeur continue sont la température, la longueur ou le prix d'un objet.

Dans l'apprentissage non supervisé, l'algorithme construit un modèle mathématique à partir d'un ensemble de données qui ne contient que des entrées et aucune étiquette de sortie souhaitée. Des algorithmes d'apprentissage non supervisés sont utilisés pour trouver une structure dans les données, telle que le regroupement ou le regroupement de points de données. L'apprentissage non supervisé peut découvrir des modèles dans les données et peut regrouper les entrées en catégories, comme dans l'apprentissage par fonctions. La réduction de la dimensionnalité consiste à réduire le nombre de "fonctions", ou entrées, dans un ensemble de données.

Les algorithmes d'apprentissage actif accèdent aux sorties souhaitées (étiquettes de formation) pour un ensemble limité d'entrées basées sur un budget et optimisent le choix des entrées pour lesquelles des étiquettes de formation seront acquises. Lorsqu'ils sont utilisés de manière interactive, ils peuvent être présentés à un utilisateur humain pour étiquetage. Les algorithmes d'apprentissage du renforcement reçoivent une rétroaction sous forme de renforcement positif ou négatif dans un environnement dynamique et sont utilisés dans des véhicules autonomes ou pour apprendre à jouer à

un jeu contre un adversaire humain. D'autres algorithmes spécialisés en apprentissage automatique incluent la modélisation de sujets, où le programme d'ordinateur reçoit un ensemble de documents en langage naturel et trouve d'autres documents couvrant des sujets similaires.

Des algorithmes d'apprentissage automatique peuvent être utilisés pour trouver la fonction de densité de probabilité non observable dans les problèmes d'estimation de densité. Les algorithmes de méta-apprentissage apprennent leur propre biais inductif sur la base de l'expérience antérieure. Dans la robotique de développement, les algorithmes d'apprentissage par robot génèrent leurs propres séquences d'expériences d'apprentissage, également appelées programme, pour acquérir de manière cumulative de nouvelles compétences via l'exploration autoguidée et les interactions sociales avec les humains. Ces robots utilisent des mécanismes de guidage tels que l'apprentissage actif, la maturation, les synergies motrices et l'imitation.

2.3 Théorie

Un objectif essentiel de l'apprenant est de généraliser à partir de son expérience. La généralisation dans ce contexte est la capacité d'une machine à apprendre à exécuter avec précision des nouveaux exemples / tâches invisibles après avoir expérimenté un jeu de données d'apprentissage.

Les exemples de formation proviennent d'une distribution de probabilité généralement inconnue (considérée comme représentative de l'espace des occurrences) et l'apprenant doit construire un modèle général sur cet espace lui permettant de produire des prédictions suffisamment précises dans les nouveaux cas. L'analyse computationnelle des algorithmes d'apprentissage automatique et de leurs performances est une branche de l'informatique théorique appelée théorie de l'apprentissage par calcul. Comme les ensembles de formation sont finis et que l'avenir est incertain, la théorie de l'apprentissage ne fournit généralement aucune garantie quant à la performance des algorithmes (Ren *et al.*, 2018).

Au lieu de cela, les limites probabilistes sur la performance sont assez courantes. La décomposition biais-variance est un moyen de quantifier l'erreur de généralisation. Pour obtenir les meilleures performances possibles dans le contexte de la généralisation, la complexité de l'hypothèse doit correspondre à la complexité de la fonction sous-jacente aux données. Si l'hypothèse est moins complexe que la fonction, le modèle sous-exploite les données. Si la complexité du modèle augmente en ré-

ponse, l'erreur d'apprentissage diminue. Cependant, si l'hypothèse est trop complexe, le modèle est sujet à un sur-ajustement et la généralisation sera moins applicable.

Outre les limites de performance, les théoriciens de l'apprentissage étudient la complexité temporelle et la faisabilité de l'apprentissage. Dans la théorie de l'apprentissage informatique, un calcul est considéré comme réalisable s'il peut être effectué en temps polynomial. Il existe deux types de résultats en termes de complexité temporelle. Les résultats positifs montrent qu'une certaine classe de fonctions peut être apprise en temps polynomial; ceux négatifs montrent le contraire.

2.4 Approaches

Les types d'algorithmes d'apprentissage automatique diffèrent par leur approche, le type de données qu'ils font entrer et sortir et le type de tâche ou de problème qu'ils sont censés résoudre.

2.4.1 Types d'algorithmes d'apprentissage

Enseignement supervisé

Les algorithmes d'apprentissage supervisé construisent un modèle mathématique d'un ensemble de données contenant à la fois les entrées et les sorties souhaitées. Les données sont appelées " données d'apprentissage " et consistent en un ensemble d'exemples d'apprentissage. Chaque exemple d'apprentissage comporte une ou plusieurs entrées et une sortie souhaitée, également appelée " signal de supervision ".

Dans le modèle mathématique, chaque exemple d'apprentissage est représenté par un tableau ou un vecteur, parfois appelé " vecteur de caractéristiques ", et les données d'apprentissage sont représentées par une matrice. Grâce à l'optimisation itérative d'une fonction objective, les algorithmes d'apprentissage supervisé apprennent une fonction qui peut être utilisée pour prédire la sortie associée à de nouvelles entrées (Ledesma *et al.*, 2018).

Une fonction optimale permettra à l'algorithme de déterminer correctement la sortie pour les entrées qui ne faisaient pas partie des données d'apprentissage. Un algorithme qui améliore la

précision de ses sorties ou de ses prévisions au fil du temps est réputé avoir appris à effectuer cette tâche.

Les algorithmes d'apprentissage supervisé incluent la classification et la régression. Les algorithmes de classification sont utilisés lorsque les sorties sont limitées à un ensemble de valeurs, et les algorithmes de régression sont utilisés lorsque les sorties peuvent avoir une valeur numérique quelconque dans un intervalle. L'apprentissage par similarité est un domaine de l'apprentissage automatique supervisé étroitement lié à la régression et à la classification, mais l'objectif est d'apprendre à partir d'exemples en utilisant une fonction de similarité qui mesure la similitude ou la corrélation entre deux objets.

Il y a des applications dans le classement, les systèmes de recommandation, le suivi de l'identité visuelle, la vérification faciale et la vérification du locuteur.

Dans le cas des algorithmes d'apprentissage semi-supervisés, certains exemples d'apprentissage manquent d'étiquettes d'apprentissage, mais ils peuvent néanmoins être utilisés pour améliorer la qualité d'un modèle.

Dans l'apprentissage faiblement supervisé, les étiquettes de formation sont bruyantes, limitées ou imprécises; cependant, ces étiquettes sont souvent moins chères à obtenir, ce qui permet d'obtenir un plus grand nombre de formations efficaces.

Apprentissage non supervisé

Les algorithmes d'apprentissage non supervisés prennent un ensemble de données contenant uniquement des entrées et trouvent une structure dans les données, telle que le regroupement ou la mise en cluster de points de données. Les algorithmes tirent donc des enseignements à partir des données de test qui n'ont pas été étiquetées, classées ou catégorisées. Au lieu de répondre aux commentaires, des algorithmes d'apprentissage non supervisés identifient les points communs des données et réagissent en fonction de la présence ou de l'absence de tels points communs dans chaque nouvel élément de données.

Une application centrale de l'apprentissage non supervisé concerne l'estimation de la densité en statistique, bien que l'apprentissage non supervisé englobe d'autres domaines impliquant la synthèse et l'explication des caractéristiques de données.

L'analyse en grappes est l'assignation d'un ensemble d'observations en sous-ensembles (appelés " grappes ") afin que les observations d'une même grappe soient similaires en fonction d'un ou de plusieurs critères prédéfinis, tandis que les observations tirées de grappes différentes sont différentes.

Différentes techniques de regroupement supposent différentes hypothèses sur la structure des données, souvent définies par une métrique de similarité et évaluées, par exemple, par leur compacité interne ou par la similarité entre les membres d'un même cluster, et par la séparation entre les clusters. D'autres méthodes sont basées sur la densité estimée et la connectivité graphique (Jan, 2018).

Apprentissage par renforcement

L'apprentissage par renforcement est un domaine de l'apprentissage automatique qui concerne la manière dont les agents logiciels doivent agir dans un environnement afin de maximiser la notion de récompense cumulative.

En raison de sa généralité, le domaine est étudié dans de nombreuses autres disciplines, telles que la théorie des jeux, la théorie du contrôle, la recherche opérationnelle, la théorie de l'information, l'optimisation par simulation, les systèmes multi-agents, l'intelligence en essaim, les statistiques et les algorithmes génétiques.

Dans l'apprentissage automatique, l'environnement est généralement représenté sous la forme d'un processus de décision de Markov " Markov Decision Process" (MDP). De nombreux algorithmes d'apprentissage par renforcement utilisent des techniques de programmation dynamique. Les algorithmes d'apprentissage par renforcement ne supposent pas la connaissance d'un modèle mathématique exact du MDP et sont utilisés lorsque des modèles exacts sont impossibles. Les algorithmes d'apprentissage par renforcement sont utilisés dans les véhicules autonomes ou pour apprendre à jouer à un jeu contre un adversaire humain.

2.4.2 Modèles

Pour effectuer un apprentissage automatique, il faut créer un modèle, formé à partir de certaines données d'apprentissage, capable de traiter des données supplémentaires pour effectuer des prévisions.

Différents types de modèles ont été utilisés et étudiés pour les systèmes d'apprentissage automatique.

Réseaux de neurones artificiels (ANN)

Les réseaux de neurones artificiels (ANN), ou systèmes connexionnistes, sont des systèmes informatiques vaguement inspirés par les réseaux de neurones biologiques qui constituent le cerveau des animaux. De tels systèmes " apprennent " à effectuer des tâches en prenant en compte des exemples, généralement sans être programmés avec des règles spécifiques à une tâche (cf. Fig.2.1).

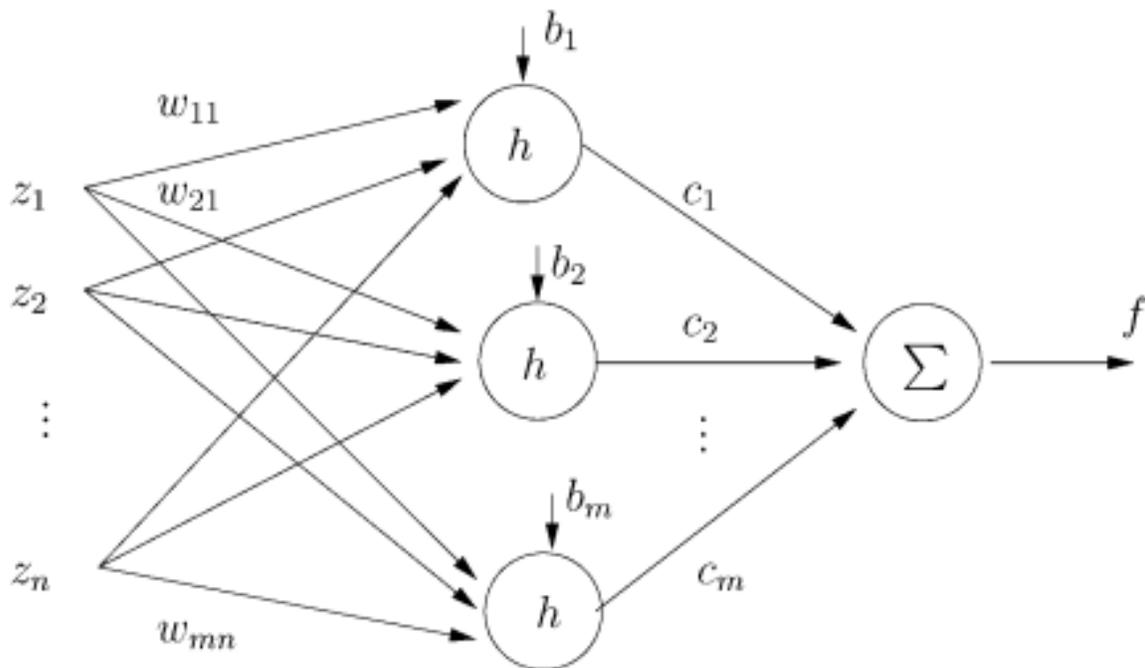


Figure 2.1 – Un modèle de réseau neuronal artificiel (Kolman & Margaliot, 2005).

ANN est un modèle basé sur un ensemble d'unités ou de nœuds connectés appelés "neurones artificiels", qui modélisent de manière lâche les neurones d'un cerveau biologique. Chaque connexion,

comme les synapses dans un cerveau biologique, peut transmettre des informations, un "signal", d'un neurone artificiel à un autre (Jan, 2018).

Un neurone artificiel qui reçoit un signal peut le traiter puis signaler d'autres neurones artificiels qui y sont connectés. Dans les implémentations ANN courantes, le signal entre des neurones artificiels est un nombre réel et la sortie de chaque neurone artificiel est calculée par une fonction non linéaire de la somme de ses entrées. Les connexions entre les neurones artificiels sont appelées " bords ".

Les neurones et les arêtes artificiels ont généralement un poids qui s'ajuste au fur et à mesure de l'apprentissage. Le poids augmente ou diminue l'intensité du signal lors d'une connexion. Les neurones artificiels peuvent avoir un seuil tel que le signal ne soit envoyé que si le signal agrégé dépasse ce seuil. Généralement, les neurones artificiels sont agrégés en couches. Différentes couches peuvent effectuer différents types de transformations sur leurs entrées. Les signaux voyagent du premier calque (le calque d'entrée) au dernier calque (le calque de sortie), après avoir traversé les calques plusieurs fois.

L'objectif initial de l'approche ANN était de résoudre les problèmes de la même manière que le ferait un cerveau humain. Cependant, avec le temps, l'attention s'est portée sur l'exécution de tâches spécifiques, conduisant à des déviations par rapport à la biologie humaine.

Les réseaux de neurones artificiels ont été utilisés pour diverses tâches, notamment la vision par ordinateur, la reconnaissance de la parole, la traduction automatique, le filtrage des réseaux sociaux, le jeu de plateau et les jeux vidéo et le diagnostic médical.

L'apprentissage en profondeur consiste en plusieurs couches cachées dans un réseau de neurones artificiels. Cette approche tente de modéliser la façon dont le cerveau humain transforme la lumière et le son en vision et en audition. Certaines applications réussies d'apprentissage en profondeur sont la vision par ordinateur et la reconnaissance de la parole. Certains réseaux de neurones artificiels sont des systèmes adaptatifs et sont utilisés, par exemple, pour modéliser des populations et des environnements qui changent constamment (Frénay & Verleysen, 2016).

Les réseaux de neurones peuvent être matériels (les neurones sont représentés par des composants physiques) ou basés sur des logiciels (modèles informatiques), et peuvent utiliser diverses topologies et algorithmes d'apprentissage. Voici trois types principaux de ANN.

Réseau de neurones Feed-Forward

Ce type de réseau de neurones est l'une des formes les plus simples d'ANN, où les données ou l'entrée sont acheminées dans une direction. Les données passent par les nœuds d'entrée et sortent par les nœuds de sortie. Ce réseau de neurones peut avoir ou ne pas avoir des couches cachées. En termes simples, il a une onde à propagation frontale et aucune propagation en arrière en utilisant une fonction d'activation de classification (Dörner *et al.*, 2018).

Les réseaux de neurones à feed-forward sont de deux types:

-Perceptron à couche unique: le type de réseau de neurones le plus simple est un réseau à couche unique de perceptron, qui consiste en une couche unique de nœuds de sortie : les entrées sont alimentées directement vers les sorties via une série de poids.

-Perceptron multicouche: cette classe de réseaux est constituée de plusieurs couches d'unités de calcul, généralement interconnectées de manière rétroactive. Chaque neurone d'une couche possède des connexions dirigées vers les neurones de la couche suivante. Dans de nombreuses applications, les unités de ces réseaux appliquent une fonction sigmoïde comme fonction d'activation.

Réseau de neurones à fonction radiale (Radial Basis Function (RBF))

Les fonctions de base radiales considèrent la distance d'un point par rapport au centre. Les fonctions RBF ont deux couches, les fonctions étant combinées à la fonction de base radiale de la couche intérieure, puis la sortie de ces fonctions est prise en compte lors du calcul de la même sortie dans le pas de temps suivant, qui est essentiellement une mémoire (Frénay & Verleysen, 2016).

Réseau de neurones récurrents (Recurrent Neural Network (RNN))

Le réseau de neurones récurrents fonctionne sur le principe de la sauvegarde de la sortie d'une couche et de la restituer à l'entrée afin d'aider à prévoir le résultat de la couche

Ici, la première couche est formée de la même manière que le réseau de neurones à réaction directe avec le produit de la somme des poids et des caractéristiques. Le processus de réseau de neurones récurrent commence une fois que cela est calculé. Cela signifie que, d'une étape à l'autre,

chaque neurone se souviendra de certaines informations qu'il avait précédemment. Cela fait en sorte que chaque neurone se comporte comme une cellule de mémoire lorsqu'il effectue des calculs (Frénay & Verleysen, 2016).

Dans ce processus, nous devons laisser le réseau de neurones travailler sur la propagation frontale et mémoriser les informations dont il a besoin pour une utilisation ultérieure. Ici, si la prédiction est fautive, nous utilisons le taux d'apprentissage ou la correction d'erreur pour effectuer de petits changements, de sorte qu'elle s'efforce progressivement de faire la bonne prédiction lors de la propagation en retour.

Arbres de décision

L'apprentissage de l'arbre décisionnel utilise un arbre décisionnel comme modèle prédictif pour passer des observations relatives à un élément (représentées dans les branches) aux conclusions relatives à la valeur cible de l'élément représentée dans les feuilles, comme illustré à la Fig.2.2.

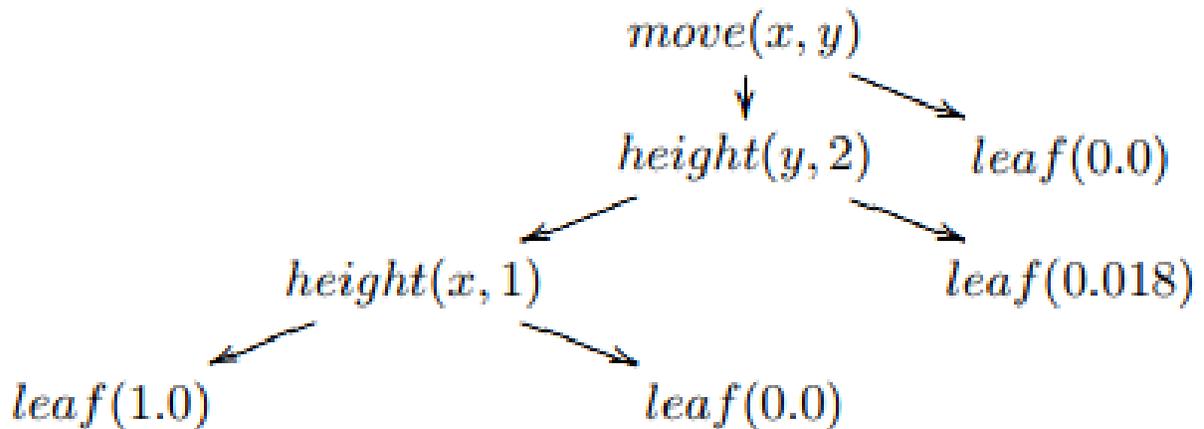


Figure 2.2 – Un exemple de modèle d'arbre de décision (Estevez *et al.*, 2011).

C'est l'une des approches de modélisation prédictive utilisées en statistique, en fouille de données et en apprentissage machine. Les modèles d'arborescence dans lesquels la variable cible peut prendre un ensemble discret de valeurs sont appelés des arbres de classification; dans ces arborescences, les feuilles représentent les étiquettes de classe et les branches représentent les conjonctions d'entités menant à ces étiquettes de classe.

Les arbres de décision où la variable cible peut prendre des valeurs continues (généralement des nombres réels) sont appelés des arbres de régression. Dans l'analyse de décision, un arbre de décision peut être utilisé pour représenter visuellement et explicitement les décisions et la prise de décision.

Dans l'exploration de données, un arbre de décision décrit les données, mais l'arbre de classification résultant peut être une entrée pour la prise de décision.

Machines à vecteurs de support

Les machines à vecteurs de support "Vector Support Machine" (SVM), également appelées « réseaux de vecteurs de support », constituent un ensemble de méthodes d'apprentissage supervisé associées utilisées pour la classification et la régression. À partir d'un ensemble d'exemples d'apprentissage, chacun étant classé comme appartenant à l'une des deux catégories, un algorithme d'apprentissage SVM crée un modèle qui prédit si un nouvel exemple appartient à l'une ou à l'autre catégorie

Un algorithme de formation SVM est un classifieur linéaire binaire non probabiliste, bien qu'il existe des méthodes telles que la mise à l'échelle Plat pour utiliser SVM dans un paramètre de classification probabiliste (Estevez *et al.*, 2011).

En plus d'effectuer une classification linéaire, les SVM peuvent effectuer efficacement une classification non linéaire en utilisant ce que l'on appelle " le truc du noyau ", mappant implicitement leurs entrées dans des espaces de fonctions de grande dimension.

Réseau Bayésien

Un réseau bayésien, un réseau de croyances ou un modèle graphique cyclique dirigé est un modèle graphique probabiliste qui représente un ensemble de variables aléatoires et leur indépendance conditionnelle avec un graphe cyclique dirigé (DAG) (Kolman & Margaliot, 2005).

Par exemple, un réseau bayésien pourrait représenter les relations probabilistes entre les maladies et les symptômes. Compte tenu des symptômes, le réseau peut être utilisé pour calculer les probabilités de présence de diverses maladies. Il existe des algorithmes efficaces qui effectuent l'inférence et l'apprentissage

Les réseaux bayésiens modélisant des séquences de variables, tels que les signaux de parole ou les séquences de protéines, sont appelés réseaux bayésiens dynamiques. Les généralisations de réseaux bayésiens capables de représenter et de résoudre des problèmes de décision dans des conditions d'incertitude sont appelées " diagrammes d'influence ".

Algorithmes génétiques

Un algorithme génétique " Genetic Algorithm " (GA) est un algorithme de recherche et une technique heuristique qui imite le processus de sélection naturelle en utilisant des méthodes telles que la mutation et le croisement pour générer de nouveaux génotypes dans l'espoir de trouver de bonnes solutions à un problème donné.

Dans l'apprentissage automatique, des algorithmes génétiques ont été utilisés dans les années 1980 et 1990. À l'inverse, des techniques d'apprentissage automatique ont été utilisées pour améliorer les performances d'algorithmes génétiques et évolutifs (Estevez *et al.*, 2011).

2.4.3 Modèles de formation

Habituellement, les modèles d'apprentissage automatique requièrent beaucoup de données pour fonctionner correctement. Ainsi, lors de la formation d'un modèle d'apprentissage automatique, il est souvent nécessaire de collecter un grand échantillon représentatif de données provenant d'un ensemble de formation (Frénay & Verleysen, 2016).

Les données de l'ensemble de formation peuvent être aussi variées qu'un corpus de texte, une collection d'images ou des données collectées auprès d'utilisateurs individuels d'un service. Le sur-ajustement est une chose à surveiller lors de la formation d'un modèle d'apprentissage automatique.

2.5 Techniques d'apprentissage machine pour les communications sans fil

Un émetteur, un canal et un récepteur constituent un système de communication sans fil typique. Le modèle de canal décrit le phénomène physique sous-jacent et dicte la conception de l'émetteur et du récepteur (Dörner *et al.*, 2018).

Les schémas de signalisation sont conçus en fonction des objectifs du système, tels que la robustesse face aux dégradations du canal, la facilité de mise en œuvre et la minimisation de la puissance d'émission. De même, les algorithmes de détection sont conçus pour optimiser les performances du système en fonction des contraintes choisies.

La plupart des systèmes de communication sans fil déployés utilisent des signaux électromagnétiques à des fréquences inférieures à 10 GHz. Dans ces systèmes, les canaux sont conformes à des modèles mathématiques traitables décrivant la propagation électromagnétique avec une précision raisonnable.

Cependant, dans d'autres systèmes, les canaux ne disposent pas de modèles de canaux faciles à décrire. Dans cette catégorie, nous mettons en évidence deux sous-catégories de canaux: les canaux décrivant la propagation des signaux électromagnétiques à ondes millimétriques et les modèles de canaux inconnus décrivant la propagation des signaux non électromagnétiques. La première offre de grandes quantités de bande passante pour prendre en charge des débits de données élevés dans les systèmes de prochaine génération. Dans cette catégorie, les modèles de canaux peuvent être connus, mais ils sont trop complexes et / ou changent trop rapidement pour être estimés avec une précision raisonnable. Cette estimation d'informations d'état de canal (CSI), cependant, est estimée pour de nombreux algorithmes de détection actuels. La deuxième sous-catégorie comprend les canaux moléculaires (c'est-à-dire les canaux par lesquels les signaux moléculaires sont envoyés), qui offrent un moyen de communiquer lorsque la propagation électromagnétique est impossible ou peu pratique. Dans cette catégorie, les modèles de canaux sont inconnus ou difficiles à obtenir. Par conséquent, les algorithmes de détection traditionnels ne peuvent pas être utilisés. Ces considérations, dans les deux sous-catégories, ont incité les chercheurs à adopter des outils d'estimation provenant d'autres domaines pour identifier le canal ou contourner sa déduction explicite.

Les techniques d'apprentissage automatique ont récemment connu un succès sans précédent dans les problèmes de classification pour lesquels il n'existe pas de modèle mathématique bien défini. Par exemple, la correspondance des pixels en objet ou du langage parlé en signification.

2.6 Conclusion

La diversité et la complexité croissantes des architectures de réseau mobile ont rendu la surveillance et la gestion de la multitude d'éléments de réseau difficiles à gérer. Par conséquent, l'intégration d'une intelligence machine polyvalente dans les futurs réseaux mobiles suscite un intérêt de recherche sans précédent. Cette tendance se reflète dans les solutions d'apprentissage automatique (ML) aux problèmes allant de la sélection de la technologie d'accès radio à la détection de programmes malveillants, ainsi que dans le développement de systèmes en réseau prenant en charge les pratiques d'apprentissage automatique. ML permet d'extraire systématiquement des informations précieuses à partir de données de trafic et de découvrir automatiquement des corrélations qui auraient autrement été trop complexes pour être extraites par des experts humains.

Chapitre 3

Imperfections de ADC dans les systèmes sans fil multi-antennes

3.1 Présentation des ADC

3.1.1 Introduction

En électronique, un convertisseur analogique-numérique (ADC, A/N ou A-à-D) est un système qui convertit un signal analogique, tel qu'un son capté par un microphone ou une lumière entrant dans un appareil photo numérique, en un signal numérique. Un ADC peut également fournir une mesure isolée, telle qu'un dispositif électronique, qui convertit une tension ou un courant analogique d'entrée en un nombre numérique représentant l'amplitude de la tension ou du courant (He *et al.*, 2018). En général, la sortie numérique est un nombre binaire à complément de deux qui est proportionnel à l'entrée. Toutefois, il existe d'autres possibilités.

Il existe plusieurs architectures d'ADC. En raison de la complexité et de la nécessité de composants parfaitement adaptés, tous les convertisseurs ADC les plus spécialisés sont implémentés en tant que circuits intégrés "Integrated Circuits" (IC). Les performances d'un ADC se caractérisent principalement par sa bande passante et son rapport signal sur bruit (SNR). La bande passante d'un ADC se caractérise principalement par son taux d'échantillonnage. Le SNR d'un ADC est influencé par de nombreux facteurs, notamment la résolution, la linéarité et la précision (dans quelle mesure

les niveaux de quantification correspondent au vrai signal analogique), le repliement du spectre et la gigue. Le SNR d'un ADC est souvent résumé en termes de nombre de bits effectif "Effective Number Of Bits" (ENOB), soit le nombre de bits de chaque mesure renvoyés qui ne sont en moyenne pas de bruit (Rizvi *et al.*, 2009). Un ADC idéal a un ENOB égal à sa résolution. Les ADCs sont choisis pour correspondre à la bande passante et au rapport signal sur bruit requis du signal à numériser. Si un ADC fonctionne à un taux d'échantillonnage supérieur à deux fois la largeur de bande du signal, alors, selon le théorème d'échantillonnage de Nyquist-Shannon, une reconstruction parfaite est possible.

La présence d'une erreur de quantification limite même le SNR d'un ADC idéal. Toutefois, si le SNR du ADCs dépasse celui du signal d'entrée, ses effets peuvent être négligés, ce qui donne une représentation numérique essentiellement parfaite du signal d'entrée analogique.

3.1.2 Types de convertisseurs ADC

Convertisseur ADC à double pente

Composants fondamentaux:

- . Intégrateur
- . Interrupteurs à commande électronique
- . Compteur
- . Horloge
- . Logique de contrôle
- . Comparateur

Un ADC à double pente (DS-ADC) intègre une tension d'entrée inconnue V_{IN} pendant une durée déterminée T_{INT} , puis " désintègre " T_{DEINT} en utilisant une tension de référence connue V_{REF} pendant une durée variable. Le principal avantage de cette architecture par rapport à la pente unique est que le résultat de la conversion finale est insensible aux erreurs dans les valeurs

des composants. En d'autres termes, toute erreur introduite par une valeur de composant pendant le cycle d'intégration sera annulée pendant la phase de désintégration (Zhang *et al.*, 2016).

Lors de la simulation dans ce travail, nous utilisons ce type de ADC.

Convertisseur ADC par approximation successive

Un convertisseur approximatif successif est un type de convertisseur analogique-numérique qui convertit une forme d'onde analogique continue en une représentation numérique discrète via une recherche binaire à travers tous les niveaux de quantification possibles avant de finalement converger vers une sortie numérique pour chaque conversion.

Convertisseur ADC Flash

Un convertisseur analogique-numérique flash (également appelé convertisseur analogique-numérique à conversion directe) est un type de convertisseur analogique-numérique qui utilise une échelle de tension linéaire avec un comparateur à chaque " échelon " de l'échelle pour comparer la tension d'entrée aux tensions de référence successives. Souvent, ces échelles de référence sont constituées de nombreuses résistances; cependant, les implémentations modernes montrent que la division de tension capacitive est également possible. La sortie de ces comparateurs est généralement introduite dans un codeur numérique, qui convertit les entrées en une valeur binaire.

Convertisseur ADC Delta-Sigma

L'une des technologies ADC les plus avancées est ce que l'on appelle le « delta-sigma ». Dans ce convertisseur, le signal de tension d'entrée analogique est connecté à l'entrée d'un intégrateur, produisant à la sortie un taux de variation de la tension, ou pente, correspondant à la magnitude d'entrée. Cette tension de rampe est ensuite comparée au potentiel de la masse (0 volt) par un comparateur. Le comparateur agit comme une sorte de ADC 1 bit produisant 1 bit de sortie (" haut " ou " bas ") selon si la sortie de l'intégrateur est positive ou négative. La sortie du comparateur est ensuite verrouillée par une bascule de type D cadencée à une fréquence élevée, puis renvoyée vers

un autre canal d'entrée de l'intégrateur, afin de piloter l'intégrateur dans la direction d'une sortie 0 volt.

3.1.3 Erreurs de quantification

Les erreurs de quantification sont les erreurs les plus importantes pouvant affecter le fonctionnement des systèmes de communication sans fil MIMO. Ces erreurs sont introduites dans un ADC idéal par quantification avec une erreur d'arrondissement entre la tension d'entrée analogique du ADC et la valeur numérisée de sortie. L'erreur est non linéaire et dépend du signal (Yuan *et al.*, 2017).

Les erreurs de quantification sont distribuées du courant continu à la fréquence de Nyquist. Par conséquence, si une partie de la bande passante de l'ADC n'est pas utilisée, comme c'est le cas avec le sur-échantillonnage, certaines erreurs de quantification se produiront hors bande, ce qui améliorera effectivement le rapport signal à canal de quantification-bruit "Signal-to-Quantization-Noise Ratio" (SQNR) pour la bande passante utilisée (Sun *et al.*, 2019). Dans un système sure-échantillonné, la mise en forme du bruit peut être utilisée pour augmenter encore le SQNR en forçant davantage d'erreurs de quantification hors bande.

3.2 ADC dans les systèmes MIMO

Les imperfections des convertisseurs analogique-numérique côté récepteur affectent les performances globales du système. Un système de communication numérique sans fil se compose de plusieurs parties résumées ci-dessus. Il est très difficile d'analyser théoriquement l'influence des imperfections de l'ADC, telles que la gigue et la distorsion non linéaire, ainsi que de la quantification idéale, sur les différentes parties et sur les performances globales d'un système de communication.

Par conséquence, un modèle ADC avec des paramètres variables a été simulé dans les systèmes de communication MIMO. L'objectif du modèle est d'obtenir des résultats exemplaires de la manière dont l'ADC peut affecter les performances de liaison d'un système de communication (Choi *et al.*, 2019).

3.2.1 Modèle ADC

Le modèle consiste en un certain nombre de modèles d'erreur additifs non linéaires suivis d'une quantification idéale. La Fig.3.1 montre la structure du modèle.

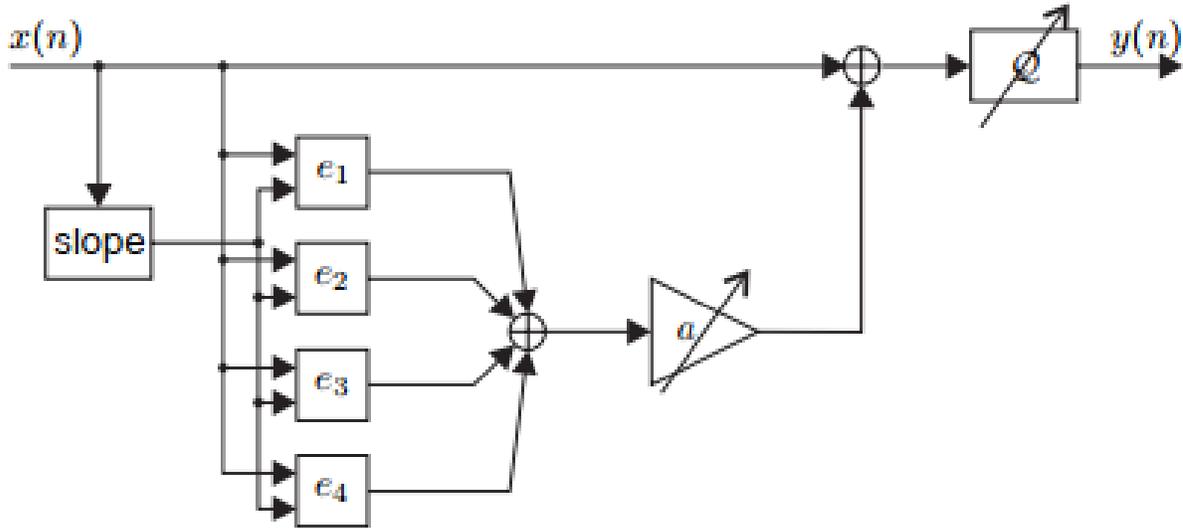


Figure 3.1 – Le modèle d'erreur ADC utilisé dans la simulation du système de communication (Choi et al., 2016).

L'entrée $X(n)$ du modèle est mise à l'échelle pour être comprise dans l'intervalle $[-1, 1]$. Les fonctions d'erreur e_1 à travers e_4 sont définies comme:

$$\begin{aligned}
 e_1(x, s) &= 2\Delta s^2, \\
 e_2(x, s) &= 3\Delta |s|x, \\
 e_3(x, s) &= 5\Delta |s|(x + 10\Delta - 20\Delta(r - 1))/(20\Delta)g - g/2, \\
 e_4(x, s) &= 2\Delta x^3.
 \end{aligned} \tag{3.1}$$

Les index de temps sont omis pour simplifier la notation. Ici, s est la pente du signal d'entrée, estimée en tant que différence arrière:

$$s(n) = (x(n) - x(n - 1))/2. \tag{3.2}$$

et r et g sont calculés comme

$$r = \lfloor (x/20\Delta) \rfloor + 1, \quad (3.3)$$

$$g = \begin{cases} +1 & \text{if } r \text{ is even;} \\ -1 & \text{if } r \text{ is odd.} \end{cases}$$

Le paramètre Δ est la taille du pas de quantification et le résultat est quantifié en b bits par le paramètre quantificateur Q .

$$\Delta = 2/2^b. \quad (3.4)$$

3.2.2 Résultats

1) Quantification variable : Dans le premier cas de test, le gain d'erreur est fixé à $a = 1$ et le nombre de bits b varie de 6 à 16 par incréments de 2. Chaque réglage est testé 150 fois. Réduire le nombre de bits en dessous de 10 augmentera considérablement le BER (cf. Fig.3.2(a)). Cependant, la différence entre le modèle ADC avec et sans fonctions d'erreur ('o' et 'x', respectivement) est négligeable. Ce résultat quelque peu surprenant mène au deuxième cas de test.

2) Gain d'erreur variable: Dans le deuxième cas de test, le nombre de bits était fixé à $b = 10$ et le gain d'erreur a été augmenté. Le rapport signal sur bruit et distorsion "Signal-to-Noise And Distortion" (SINAD) résultant a été évalué pour chaque réglage. Il ressort des résultats que ce système de communication est assez robuste contre les erreurs de non-linéarité dans l'ADC de récepteur (cf. Fig.3.2(b)).

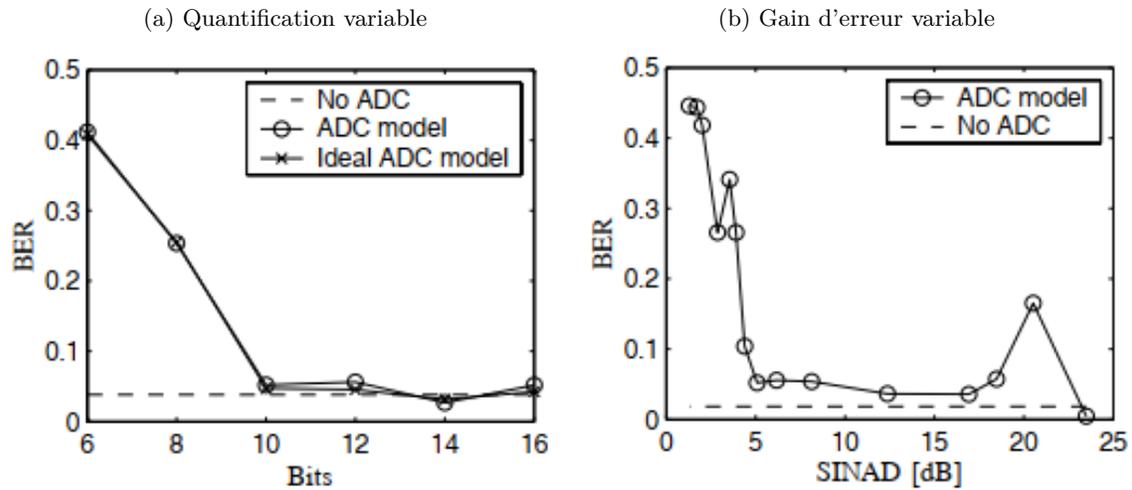


Figure 3.2 – Le taux d'erreur sur les bits résultant pour les deux expériences.

3.3 Conclusion

Les résultats indiquent que, pour ce système spécifique, le nombre de bits ADC a eu une incidence sur les performances du système de communication, en termes de taux d'erreur sur les bits. Toutefois, les modèles d'erreur non linéaires testés n'affectaient pas le système avec des niveaux de distorsion raisonnables. L'une des raisons de l'impact relativement faible est qu'il n'y a ni émetteur brouilleur ni canal adjacent dans la simulation. Les interféreurs et les canaux adjacents peuvent créer des composants parasites dans le signal souhaité lorsque des non-linéarités ADC sont présentes.

Chapitre 4

Conception d'un algorithme de détection basé sur ANN et évaluation de ses performances

4.1 Introduction

Nous proposons une approche de détection de signal aveugle basée sur ANN pour les systèmes MIMO-OFDM avec ADC parfaits et avec LR-ADC lorsque le récepteur n'indique pas le CSIR explicite.

Les réseaux de neurones ont atteint l'état de la technique dans plus de domaines que les autres algorithmes de classification. Puisque nous considérons la détection comme un acte de classification, nous avons choisi cet algorithme comme classificateur.

La formation d'un réseau neuronal implique de nombreux hyper-paramètres contrôlant la taille et la structure du réseau ainsi que la procédure d'optimisation permettant de mieux détecter les systèmes à grande échelle. L'approche proposée comprend deux phases. Au cours de la première phase, l'émetteur envoie une séquence de symboles de données afin que le récepteur apprenne une fonction non linéaire qui décrit les relations entrée-sortie du système. Dans la deuxième phase, à

l'aide du réseau de neurones formé, le récepteur détecte les symboles de données. Cette nouvelle approche peut être considérée comme un problème de classification dans l'apprentissage supervisé.

Le classifieur mappe le signal reçu sur l'un des vecteurs de symbole possibles. L'avantage principal de nos détecteurs basés sur la ANN est que son fonctionnement est indépendant du schéma de modulation utilisé dans le système, contrairement aux techniques de détection classiques dans lesquelles la structure du récepteur doit être adaptée au schéma de modulation.

4.2 Modèle du système

En combinant les technologies de multiplexage par répartition en fréquence orthogonale et multi-entrées à entrées multiples, les systèmes sans fil ont réussi à obtenir des débits de données élevés et une efficacité spectrale élevée, inatteignables pour les systèmes SISO conventionnels.

Ces améliorations du débit de données et de l'efficacité spectrale tiennent au fait que les schémas MIMO et OFDM sont bien des technologies de transmission parallèles dans les domaines spatial et fréquentiel (Cimini, 1985).

Un autre avantage de la combinaison de la technique OFDM avec le système MIMO est la robustesse du système dans les canaux d'évanouissement sélectifs en fréquence.

4.2.1 Systèmes à antennes multiples (MIMO)

Plusieurs antennes au niveau de l'émetteur et/ou du récepteur d'un système mobile peuvent augmenter les débits de données et les performances (systèmes à plusieurs entrées et plusieurs sorties) ou réduire l'ISI et les interférences d'autres utilisateurs (antennes intelligentes).

Les systèmes MIMO sont définis comme des liaisons de communication point à point avec plusieurs antennes au niveau de l'émetteur et du récepteur. L'utilisation de plusieurs antennes au niveau de l'émetteur et du récepteur offre clairement des performances améliorées par rapport aux systèmes en diversité où l'émetteur ou le récepteur, mais pas les deux, possède plusieurs antennes.

En particulier, des recherches récentes ont montré que les systèmes MIMO peuvent augmenter considérablement les débits de données des systèmes sans fil sans augmenter la puissance de

transmission ou la bande passante. Le coût de cette augmentation est le coût supplémentaire du déploiement de plusieurs antennes, l'encombrement de ces antennes supplémentaires (en particulier sur les unités portables de petite taille) et la complexité supplémentaire requise pour le traitement du signal multidimensionnel.

Le travail récent dans les systèmes MIMO inclut la capacité de ces systèmes sous différentes hypothèses concernant la connaissance du canal, le codage et le décodage optimaux pour ces systèmes et les stratégies de transmission pour les systèmes non codés. Un système de communication point à point à bande étroite (évanouissements plats) utilisant n antennes d'émission et de réception est illustré à la Fig.4.1

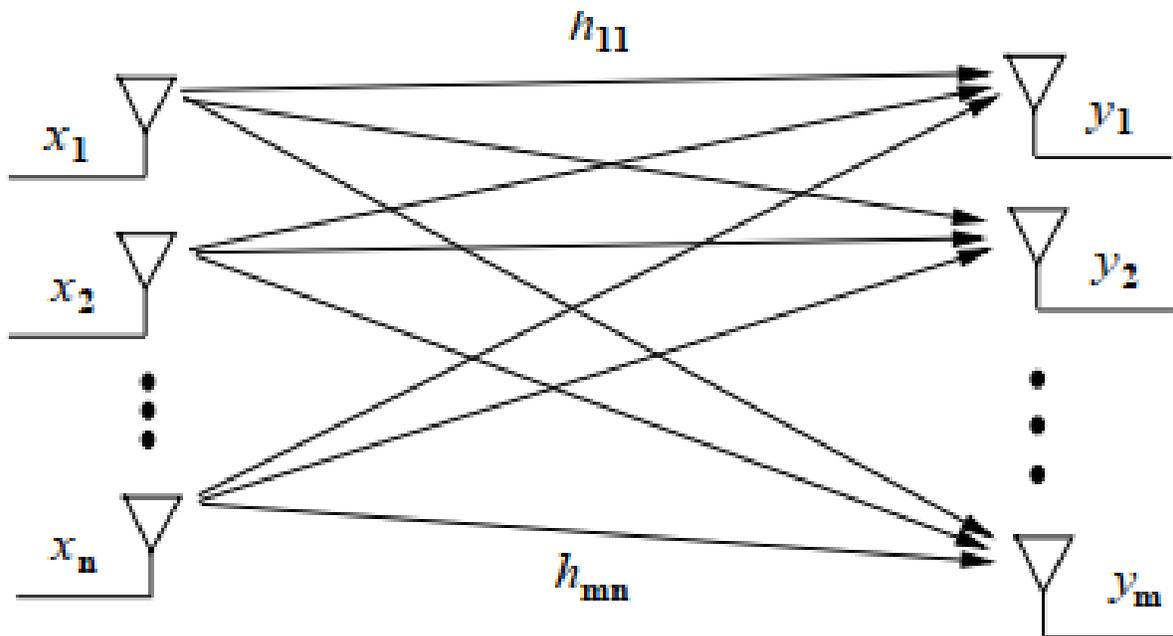


Figure 4.1 – Systèmes MIMO (He *et al.*, 2018).

4.2.2 Modulation de multiplexage par répartition orthogonale en fréquence (OFDM)

L'idée de base de la modulation multi-porteuse est de diviser le flux binaire transmis en plusieurs sous-flux différents et de les envoyer sur de nombreux sous-canaux différents. Typiquement, les sous-canaux sont orthogonaux dans des conditions de propagation idéales, auquel cas la modulation à porteuses multiples est souvent appelée multiplexage par répartition orthogonale de fréquence (OFDM).

Le débit de données sur chacun des sous-canaux est nettement inférieur au débit total et la bande passante correspondante du sous-canal est bien inférieure à la largeur de bande totale du système. Le nombre de sous-flux est choisi pour assurer que chaque sous-canal a une largeur de bande inférieure à la largeur de bande de cohérence du canal, de sorte que les sous-canaux subissent un évanouissement relativement plat. Ainsi, l'ISI sur chaque sous-canal est petit.

De plus, dans l'implémentation discrète d'OFDM, souvent appelée multiton discrète (DMT), l'ISI peut être complètement éliminé par l'utilisation d'un préfixe cyclique. Les sous-canaux dans OFDM n'ayant pas besoin d'être contigus, un large bloc de spectre continu n'est donc pas nécessaire pour les communications multi-porteuses à haut débit.

La forme la plus simple de modulation multi-porteuses divise le flux de données en plusieurs sous-flux à transmettre sur différents sous-canaux orthogonaux centrés sur différentes fréquences de sous-porteuses. Le nombre de sous-flux est choisi pour rendre le temps de symbole sur chaque sous-flux beaucoup plus grand que le délai de propagation du canal ou, de manière équivalente, pour rendre la largeur de bande du sous-flux inférieure à la largeur de bande de cohérence du canal. Cela garantit que les sous-flux ne connaîtront pas un ISI significatif.

4.2.3 Modulation numérique sur les canaux sans fil

Les progrès réalisés au cours des dernières décennies dans le traitement du matériel et du signal numérique ont rendu les émetteurs-récepteurs numériques beaucoup moins chers, plus rapides et plus éco-énergétiques que les émetteurs-récepteurs analogiques.

Plus important encore, la modulation numérique offre de nombreux autres avantages par rapport à la modulation analogique, notamment des débits de données plus élevés, des techniques de correction d'erreur performantes, la résistance aux dégradations de canal, des stratégies d'accès multiples plus efficaces et une sécurité et une confidentialité accrues.

Plus précisément, les techniques de modulation de haut niveau telles que MQAM permettent des débits de données beaucoup plus élevés en modulation numérique par rapport à la modulation analogique avec la même largeur de bande de signal. Les progrès du codage et de la modulation codée appliqués à la signalisation numérique rendent le signal beaucoup moins sensible au bruit et aux évanouissements, et des techniques d'égalisation ou multi-porteuses peuvent être utilisées

pour atténuer l'ISI. Les techniques à spectre étalé appliquées à la modulation numérique peuvent éliminer ou combiner des trajets multiples, résister aux interférences et détecter plusieurs utilisateurs simultanément.

Enfin, la modulation numérique est beaucoup plus facile à chiffrer, ce qui accroît le niveau de sécurité et de confidentialité des systèmes numériques. Pour toutes ces raisons, les systèmes en cours de construction ou proposés pour des applications sans fil sont tous des systèmes numériques. La modulation et la détection numériques consistent à transférer des informations sous forme de bits sur un canal de communication.

Les bits sont des chiffres binaires prenant les valeurs 1 ou 0. Ces bits d'information sont dérivés de la source d'information, qui peut être une source numérique ou une source analogique ayant transité par un convertisseur analogique-numérique. Les sources analogiques numériques et converties A/D peuvent être compressées pour obtenir la séquence de bits d'information. La modulation numérique consiste à mapper les bits d'information en un signal analogique pour la transmission sur le canal.

La détection consiste à déterminer la séquence de bits d'origine en fonction du signal reçu sur le canal. Les principales considérations lors du choix d'une technique de modulation numérique particulière sont les suivantes:

- Haut débit.
- Efficacité spectrale élevée (occupation de la bande passante minimale).
- Efficacité énergétique élevée (puissance d'émission minimale requise).
- Robustesse aux dégradations du canal (probabilité d'erreur binaire minimale).
- Mise en œuvre à faible coût / puissance: il s'agit souvent d'exigences contradictoires et le choix de la modulation repose sur la recherche de la technique qui permet le meilleur compromis entre ces exigences.

Il existe deux catégories principales de modulation numérique: la modulation d'amplitude/ phase et la modulation de fréquence. Comme la modulation de fréquence a généralement une enveloppe de signal constante et est générée à l'aide de techniques non linéaires, cette modulation est éga-

lement appelée modulation à enveloppe constante ou modulation non linéaire, et la modulation amplitude/phase est également appelée modulation linéaire.

La modulation linéaire a généralement de meilleures propriétés spectrales que la modulation non linéaire.

Modulation d'amplitude et de phase

En modulation d'amplitude et de phase, le train de bits d'information est codé dans l'amplitude et/ou la phase du signal transmis. Spécifiquement, sur un intervalle de temps de T_s , $K = \log_2 M$ bits sont codés dans l'amplitude et/ou la phase du signal transmis $s(t)$, $0 < t < T_s$.

Le signal transmis sur cette période peut être écrit en fonction de sa représentation dans l'espace du signal. Ces composantes de signal en phase et en quadrature sont des signaux en bande de base ayant des caractéristiques spectrales déterminées par la forme de l'impulsion $g(t)$.

$$s(t) = s_I(t) \cos(2\pi f_c t + \phi_0) - s_Q(t) \sin(2\pi f_c t + \phi_0) \quad (4.1)$$

Il existe trois types principaux de modulation d'amplitude / phase:

- Modulation d'amplitude d'impulsion "M-ary Pulse Amplitude Modulation" (MPAM): informations codées en amplitude uniquement.
- Phase Shift Keying (MPSK): informations codées en phase uniquement.
- Modulation d'amplitude en quadrature "M-ary Quadrature Amplitude Modulation" (MQAM): information codée à la fois en amplitude et en phase.

Dans ce travail, nous adoptons les conceptions BPSK et QPSK.

4.2.4 Système MIMO-OFDM

Nous considérons le système MIMO-OFDM avec N antennes d'émission et M antennes de réception, ayant ainsi une matrice de canaux $M \times N$ et N_s sous-porteuses. La structure de l'émetteur-récepteur est illustrée à la Fig.4.2. Le signal transmis est formé en symboles OFDM en appliquant

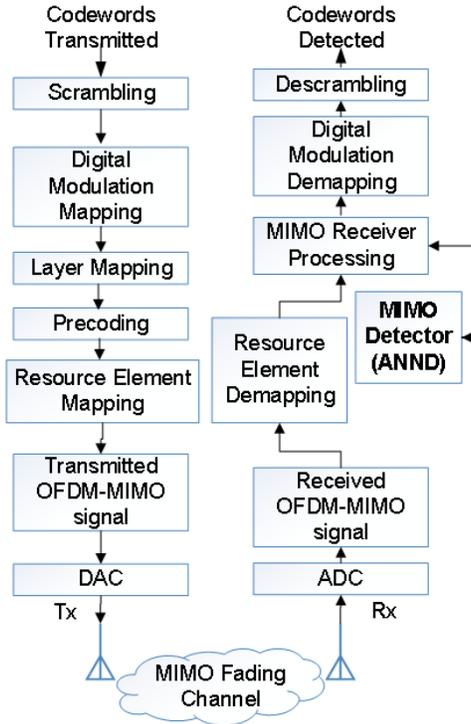


Figure 4.2 – Modèle de système MIMO-OFDM.

la transformée de Fourier rapide inver $IFFT$ comme suit:

$$\begin{aligned}
 x_n(l) &= IFFT\{X_n(k)\} \\
 &= \sum_{k=0}^{N_s-1} X_n(k)e^{j2\pi kl/N_s}, \quad l = 0, 1, \dots, N_s - 1,
 \end{aligned} \tag{4.2}$$

où $X_n(k)$ est le signal transmis par la $n^{\text{ième}}$ antenne d'émission sur la $k^{\text{ième}}$ sous-porteuse. De plus, pour éviter les interférences entre symboles (ISI), le préfixe cyclique (CP) est inséré. Après l'insertion du CP, les signaux OFDM sont prêts à être transmis. Au niveau du récepteur, juste après le ADC, la transformée de Fourier rapide (FFT) fonctionne comme un moyen de démodulation OFDM. Ainsi, la relation entrée-sortie du modèle de système MIMO-OFDM peut être exprimée sous forme de

matrice comme suit:

$$\begin{aligned}
 \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n}, \\
 \mathbf{y} &= [y_1, \dots, y_M]^T, \\
 \mathbf{x} &= [x_1, \dots, x_N]^T,
 \end{aligned} \tag{4.3}$$

où \mathbf{y} est le vecteur du signal reçu, \mathbf{H} désigne la matrice de canaux $M \times N$, \mathbf{x} est le vecteur du signal transmis et \mathbf{n} est le vecteur $M \times 1$ du bruit gaussien additif (AWGN) dont les éléments ont une moyenne nulle et une variance de σ^2 ; et T désigne l'opération de transposition.

Après ADC, la FFT exploite les symboles OFDM reçus et le symbole de la $m^{\text{ième}}$ antenne sur la $k^{\text{ième}}$ sous-porteuse peut être écrit ainsi:

$$\begin{aligned}
 Y_m(k) &= FFT\{y_m(l)\} \\
 &= \frac{1}{N_s} \sum_{l=0}^{N_s-1} y_m(l) e^{-j2\pi kl/N_s}, \\
 k &= 0, 1, \dots, N_s - 1,
 \end{aligned} \tag{4.4}$$

ci-après, si des détecteurs classiques étaient déployés, les symboles du pilote pourraient être extraits.

Cependant, bien que la détection avec ces méthodes ne soit pas très efficace pour des valeurs de SNR faibles ou un CSIR faible, cela signifie que l'application de ces algorithmes n'est pas le choix optimal dans cette situation. Nous allons donc trouver une nouvelle méthode pour obtenir de meilleures performances de détection. Nous allons donc à la suite proposer un nouveau détecteur basé sur ANN basé sur la classification par algorithme de Levenberg-Marquardt.

4.3 Modèle de réseaux de neurones artificiels

4.3.1 Introduction

Le réseau neuronal artificiel est un système informatique inspiré du réseau neuronal biologique. Une ANN est basée sur une collection d'unités connectées appelées neurones artificiels, par poids

synaptiques, qui comporte plusieurs classes en termes de mode de connexion entre neurones et fonctions d'activation, ainsi que plusieurs algorithmes d'apprentissage du réseau. Dans cet article, nous utilisons les réseaux de neurones de perceptrons multicouches (MLP) comme réseau et l'algorithme de Levenberg-Marquardt pour la méthode d'apprentissage.

4.3.2 Réseaux de Neurones de Perceptron Multicouches

En fonction du type de connexion entre les neurones, plusieurs types de ANN sont définis. Le type de ANN le plus courant est le perceptron multicouche à feed-forward (FF-MLP).

Un MLP se compose de plusieurs couches qui comportent au moins une couche d'entrée, une couche cachée et une couche de sortie. Chaque nœud est un neurone utilisant une fonction d'activation non linéaire (Hagan & Menhaj, 1994). Dans un FF-MLP, le signal d'entrée est transmis à travers une fonction d'activation pour produire la sortie du neurone. Il peut y avoir des neurones cachés qui jouent un rôle interne dans le réseau (Demuth *et al.*, 2014). La Fig. 4.3 montre un modèle simple d'un réseau d'annuaire composé de couches d'entrée, cachées et de sortie.

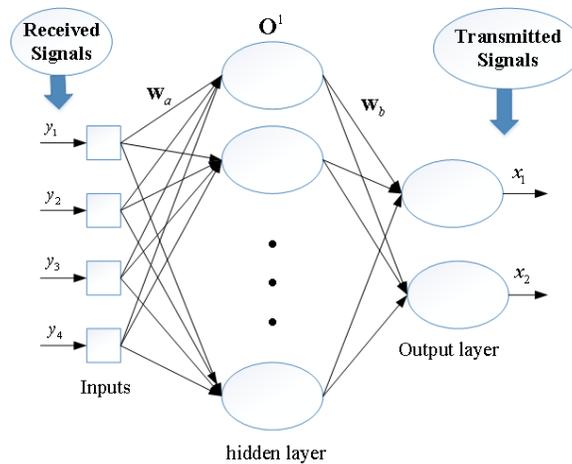


Figure 4.3 – Réseau de neurones MLP.

Une fois qu'un réseau de neurones est formé pour une tâche spécifique, ses poids peuvent être utilisés comme initialiseurs pour un autre ANN pour une tâche similaire.

4.3.3 Algorithme de Levenberg-Marquardt

Pour former une ANN, nous avons besoin d'un algorithme et la signification de la formation est l'ajustement des poids entre neurones et algorithmes.

L'algorithme de Levenberg-Marquardt (LM) est utilisé pour résoudre des problèmes d'optimisation des moindres carrés non linéaires. Dans cette formation, afin d'ajuster les poids des couches cachée et de sortie, une fonction d'erreur $\mathbf{E}(\mathbf{w})$ est définie et doit être minimisée pendant la phase d'apprentissage comme suit:

$$E(w) = \frac{1}{2} \sum_{r=1}^R (d_r - x_r)^2, \quad (4.5)$$

Où d_r est la $r^{\text{ième}}$ sortie désirée, x_r est la $r^{\text{ième}}$ sortie réelle et R le nombre de points de sortie. A chaque itération, les poids des réseaux doivent être mis à jour en fonction de:

$$\Delta \mathbf{w} = -(\mathbf{J}^T \mathbf{J} + \phi \mathbf{I})^{-1} \mathbf{J}^T E, \quad (4.6)$$

Où $\Delta \mathbf{w}$ est le vecteur de différence de poids, \mathbf{J} est la matrice jacobienne contenant les dérivées des erreurs de réseau par rapport aux poids et aux biais, ϕ est le taux d'apprentissage qui détermine la vitesse de changement des poids \mathbf{w} à chaque étape.

Le choix d'un taux d'apprentissage élevé entraîne toutefois une formation plus rapide, avec le coût d'un temps de convergence plus long pour l'algorithme. Il y a donc un compromis entre le taux d'apprentissage et le taux de convergence de l'algorithme en cours d'exécution. Bien qu'une valeur trop grande de ϕ accélère le processus d'apprentissage, cela peut provoquer une oscillation, ce qui signifie que l'algorithme ne converge pas. D'autre part, un rythme d'apprentissage trop faible entraîne la longue convergence de l'algorithme.

Nous pouvons le modifier intelligemment pendant la phase d'apprentissage en fonction du temps nécessaire à la convergence de notre réseau.

4.4 Détection de signal basée sur ANN

4.4.1 Introduction

Dans cette section, nous proposons une technique de détection à l'aveugle pour notre modèle de système MIMO-OFDM, basée sur l'approche de classification dans les réseaux ANN.

4.4.2 Détecteur basé sur ANN

Notre structure ANND proposée pour notre système comporte deux couches: une couche cachée et une couche de sortie. Nous considérons la fonction hyperbolique tangente comme la fonction d'activation dans chaque neurone de la couche cachée avec sept neurones et la fonction d'activation à limite dure dans la couche de sortie.

La structure neuronale à quatre entrées et deux sorties sert à adapter les signaux OFDM au réseau neuronal. Les symboles reçus sont constitués de signaux complexes alors que le réseau de neurones utilise des signaux réels.

Pour adapter le réseau de neurones au système, chaque signal complexe est séparé en parties réelle et imaginaire. Les opérations mathématiques dans les couches du réseau pendant les phases de travail et d'entraînement sont les suivantes:

$$\begin{aligned}
 net_j^1 &= \sum_{i=1}^{l_1} (y_i \cdot w_{aji}), \quad j = 1, 2, \dots, 7, \\
 o_j^1 &= f(net_j^1) = (e^{2net_j^1} - 1) / (e^{2net_j^1} + 1), \\
 net_r^2 &= \sum_{j=1}^{l_2} (o_j^1 \cdot w_{brj}), \quad r = 1, 2, \\
 x_r &= f(net_r^2) = net_r^2,
 \end{aligned} \tag{4.7}$$

Où w_{aji} est le poids de l'entrée dans la couche cachée au $j^{\text{ième}}$ nœud, l_1 est le nombre de nœuds d'entrée qui est 4 dans ce projet, w_{brj} est le poids de la couche cachée dans la couche en sortie au $r^{\text{ième}}$ nœud et l_2 est le nombre de neurones cachés dans la couche qui est ici 7.

4.4.3 Phase d'entraînement ANN

L'algorithme d'apprentissage pour le réseau neuronal à feed-forward est présenté ci-dessous:

- 1) Initialisez le vecteur de pondération \mathbf{w} et le taux d'apprentissage ϕ ;
- 2) Prenez le signal reçu comme entrée et estimez le signal transmis en utilisant le détecteur ML comme cible conformément à la Fig.4.4;
- 3) Calculez Fonction d'erreur selon (4.5);
- 4) Calculez la différence de poids $\Delta\mathbf{w}$ selon (4.6);
- 5) Recalculez la fonction d'erreur après avoir utilisé les nouveaux poids conformément $\mathbf{w} + \Delta\mathbf{w}$;
- 6) Si l'erreur est plus petite que celle calculée à l'étape 3, réduisez le taux d'apprentissage (ϕ) de 0,1 fois. Sinon, augmentez-le 2 fois et ensuite allez à l'étape 2;
- 7) Terminez la phase d'apprentissage si la fonction d'erreur est inférieure à la valeur prédéfinie.

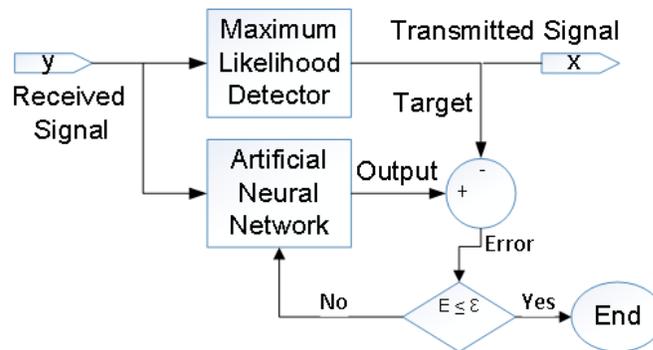


Figure 4.4 – Algorithme de formation de ANN.

4.5 Résultats and discussion

Dans cette section, en raison de la simulation extensive, nous comparons les performances de notre détecteur ANND, basé sur la technologie ANN, en termes de SER, avec les techniques de détection classiques ML, MMSE et ZF pour notre modèle de système MIMO-OFDM. Nous faisons une telle comparaison en ce qui concerne le SNR et les résultats sont montrés dans les Fig. 4.5, Fig. 4.6, Fig. 4.7 et Fig. 4.8 et nous comparons également les complexités de temps de calcul de ces détecteurs dans le tableau 4.3 pour montrer la plus faible complexité de ANND en comparaison à l'algorithme ML, respectivement. Les paramètres de simulation pour notre modèle de système MIMO-OFDM et le réseau de neurones à perceptrons multicouches sont répertoriés dans les Tableaux 4.1 et 4.2.

Tableau 4.1 – Paramètres du système MIMO-OFDM.

Paramètre	Valeur
Fréquence porteuse f_c	5 GHz
Fréquence d'échantillonnage f_s	3 MHz
Taille FFT	64
Type de modulation	BPSK, QPSK
Type de canal	canal de Rayleigh plat
Nombre d'antennes	2×2 , 4×4
Portée du SNR	0–30 dB
Égaliseur récepteur MIMO	ML, MMSE, ZF, ANND
Les ADCs	ADC parfait, différents nombres de bits ADC

Tableau 4.2 – Paramètres ANN.

Paramètre	Valeur
Nombre d'entrées	4
Nombre de sorties	2
Nombre de couches cachées	1
Nombre de neurones dans la couche cachée	7
Numéro d'époque / Itération	1000
Fonction de formation / algorithme	Levenberg-Marquardt
Métrique de performance	Erreur quadratique moyenne
Fonction d'erreur cible E	10^{-3}
Taux d'apprentissage initial ϕ	0.35

Les Fig. 4.5 et Fig. 4.6 représentent la comparaison des performances du SER des quatre techniques de détection ANND, ML, MMSE et ZF pour deux schémas de modulation de BPSK et QPSK lorsqu'un ADC parfait, c'est-à-dire à résolution infinie, est déployé dans le récepteur.

Les Fig. 4.5 et Fig.4.6 concernent un système MIMO 2×2 et 4×4 , respectivement. La Fig. 4.5, montre que ANND surpasse ZF pour toutes les valeurs de SNR pour BPSK et QPSK. De plus, il est montré qu'il ne sera jamais meilleur que le détecteur optimal, c'est-à-dire ML.

Par rapport au MMSE, la supériorité de ANND dépend à la fois du schéma de modulation et de la valeur du rapport SNR. Pour BPSK, ANND surpasse le MMSE à des valeurs de faible SNR; il est inférieur pour QPSK avec des valeurs de SNR élevé.

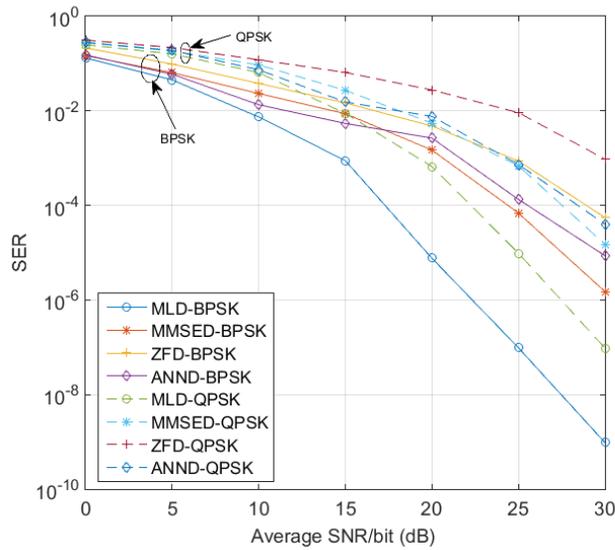


Figure 4.5 – Performances du SER des techniques de détection proposées et conventionnelles pour les systèmes 2×2 MIMO-OFDM pour BPSK et QPSK.

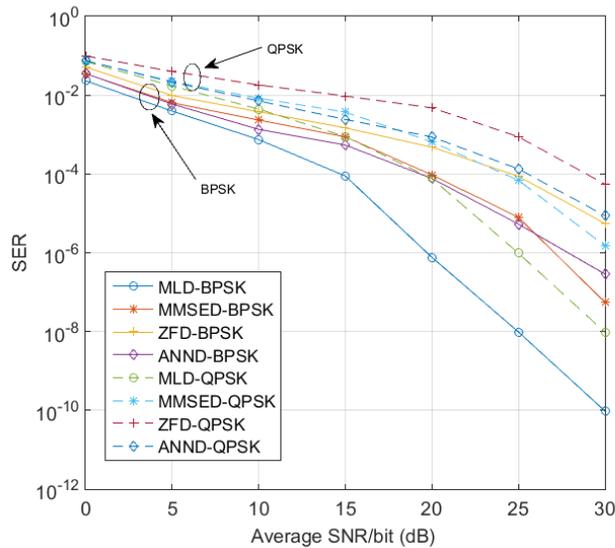


Figure 4.6 – Performances du SER des techniques de détection proposées et conventionnelles pour les systèmes 4×4 MIMO-OFDM pour BPSK et QPSK.

La différence de performance est due au nombre limité de temps d'entraînement et de couches cachées. Il convient de noter qu'il est bien connu que le MMSE converge vers ZF de manière asymptotique pour les valeurs de faible SNR. Les Fig. 4.7 et Fig. 4.8 montrent les SER de la technique de détection proposée pour un ADC parfait pour un détecteur ML afin de définir une bande inférieure et divers nombres de bits ADC pour d'autres détecteurs, en utilisant une modulation QPSK avec des antennes 2×2 et 4×4 . Il est observé que la technique de détection proposée permet

une réduction du SER dans le système par rapport à l'algorithme MMSE pour tous les différents nombres de bits ADC, en particulier dans le cas d'une taille de 4×4 antennes.

On peut en déduire que les performances du détecteur neuronal sont meilleures que celles des algorithmes MMSE et ZF pour les systèmes MIMO-OFDM avec convertisseurs ADC à faible résolution, en particulier dans la plage de faible rapport SNR, avec non seulement une taille d'antenne de 4×4 mais également une taille d'antenne de 2×2 . En conséquence, on peut s'attendre à une performance significative de la technique de détection ANN dans d'autres types de modulation pour un système avec des convertisseurs ADC à faible résolution.

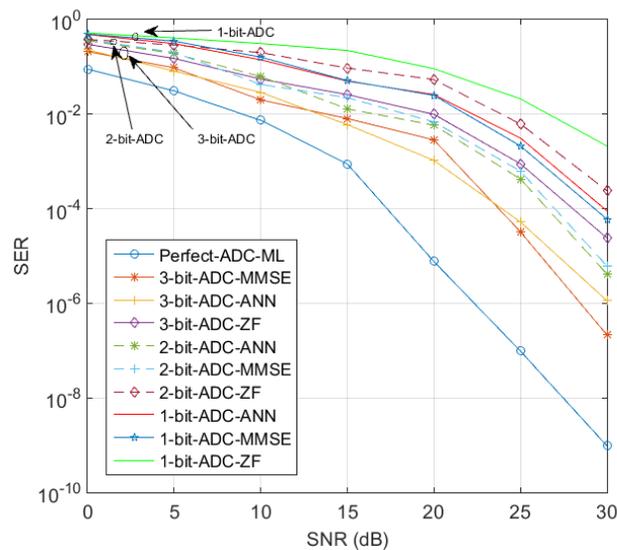


Figure 4.7 – Performances du SER des techniques de détection proposées et conventionnelles pour différents nombres de bits ADC pour 2×2 systèmes MIMO-OFDM pour QPSK.

Comme nous pouvons le voir au tableau 4.3, notre détecteur ANND proposé a une complexité temporelle inférieure à celle de l'algorithme ML et l'impact de cet avantage sera plus important dans un système à grande échelle (système comportant un grand nombre d'antennes).

Tableau 4.3 – La Complexité de Temps Informatique.

Algorithms	Complexité temporelle (/ML)
MLD	1.000
MMSED	0.0511
ZFD	0.0425
ANND	0.3429

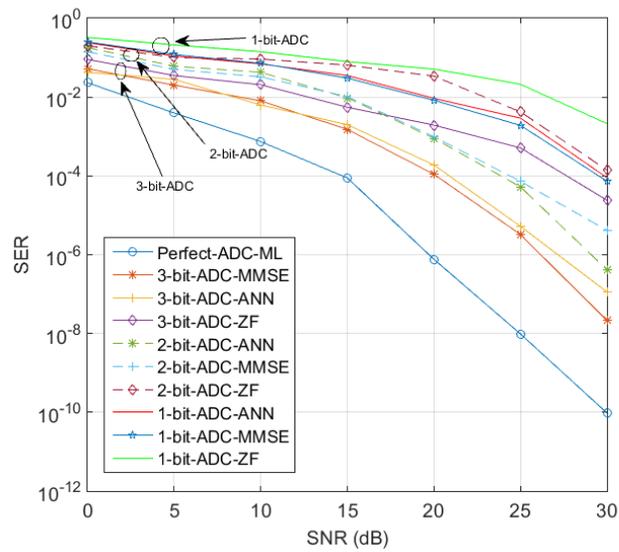


Figure 4.8 – Performances du SER des techniques de détection proposées et conventionnelles pour différents nombres de bits ADC pour 4×4 systèmes MIMO-OFDM pour QPSK.

Chapitre 5

Conclusion

Dans cet thèse, nous avons proposé une nouvelle méthode de détection de signal basée sur un réseau neuronal artificiel. Nous avons considéré un système MIMO-OFDM avec des convertisseurs analogiques/numériques parfaits et imparfaits, c'est-à-dire un bit.

Nous avons comparé le détecteur à base d'ANN proposé avec les méthodes de détection classiques du maximum de vraisemblance, de l'erreur quadratique moyenne minimale et du forçage nul pour les deux cas d'ADC et pour différents schémas de modulation. Nous avons montré que ANND suit bien les performances du MMSE en termes de SER et peut être considéré comme une alternative prometteuse pour les applications nécessitant un grand nombre d'antennes, telles que l'immense MIMO.

Plusieurs travaux de recherche et développement peuvent être considérés comme des extensions de ce travail. L'étape suivante consiste à concevoir, avec le même principe, un algorithme pour une système de communication massif MIMO. En second lieu, la mise en œuvre de notre solution nous permettra de mettre en œuvre ce nouvel algorithme et, sur le plan personnel, cela nous donnera l'occasion de toucher plusieurs aspects de la recherche scientifique et du développement.

Références

- Chen K & Wang S (2011). Semi-supervised learning via regularized boosting working on multiple semi-supervised assumptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1):129–143. DOI:10.1109/TPAMI.2010.92.
- Choi J, Lee G & Evans BL (2019). Two-stage analog combining in hybrid beamforming systems with low-resolution adcs. *IEEE Transactions on Signal Processing*, 67(9):2410–2425. DOI:10.1109/TSP.2019.2904931.
- Choi J, Mo J & Heath RW (2016). Near maximum-likelihood detector and channel estimator for uplink multiuser massive mimo systems with one-bit adcs. *IEEE Transactions on Communications*, 64(5):2005–2018. DOI:10.1109/TCOMM.2016.2545666.
- Cimini L (1985). Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Transactions on Communications*, 33(7):665–675. DOI:10.1109/TCOM.1985.1096357.
- Demuth HB, Beale MH, De Jess O & Hagan MT (2014). *Neural network design*. Martin Hagan.
- Dörner S, Cammerer S, Hoydis J & t. Brink S (2018). Deep learning based communication over the air. 12(1):132–143. DOI:10.1109/JSTSP.2017.2784180.
- Ertekin S, Bottou L & Giles CL (2011). Nonconvex online support vector machines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2):368–381. DOI:10.1109/TPAMI.2010.109.
- Estevez J, Toledo P & Alayon S (2011). Using an induced relational decision tree for rule injection in a learning classifier system. *2011 IEEE Congress of Evolutionary Computation (CEC)*, pages 647–654.
- Frénay B & Verleysen M (2016). Reinforced extreme learning machines for fast robust regression in the presence of outliers. *IEEE Transactions on Cybernetics*, 46(12):3351–3363. DOI:10.1109/TCYB.2015.2504404.
- Goldsmith A (2005). *Wireless Communications Systems Design and Considerations*. Wiley.
- Hagan MT & Menhaj MB (1994). Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993. DOI:10.1109/72.329697.
- He H, Wen C & Jin S (2018). Bayesian optimal data detector for hybrid mmwave mimo-ofdm systems with low-resolution adcs. *IEEE Journal of Selected Topics in Signal Processing*, 12(3):469–483. DOI:10.1109/JSTSP.2018.2818063.

- Jan T (2018). Ensemble of semi-supervised models for iot resource scheduling and sharing. *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, pages 376–378.
- Jeon Y, Hong S & Lee N (2017). Blind detection for MIMO systems with using supervised learning. *2017 IEEE International Conference on Communications (ICC)*, pages 1–6.
- Kim N, Lee Y & Park H (2008). Performance analysis of mimo system with linear mmse receiver. *IEEE Transactions on Wireless Communications*, 7(11):4474–4478. DOI:10.1109/TWC.2008.070785.
- Kolman E & Margaliot M (2005). Are artificial neural networks white boxes? *IEEE Transactions on Neural Networks*, 16(4):844–852. DOI:10.1109/TNN.2005.849843.
- Ledesma S, Ibarra-Manzano M, Cabal-Yepez E, Almanza-Ojeda D & Avina-Cervantes J (2018). Analysis of data sets with learning conflicts for machine learning. *IEEE Access*, 6:45062–45070. DOI:10.1109/ACCESS.2018.2865135.
- Liang H, Chang RY, Chung W & Kuo S (2013). A reduced-complexity blind detector for mimo system using k-means clustering algorithm. *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pages 1–5.
- Luo C, Ji J, Wang Q, Chen X & Li P (2018). Channel state information prediction for 5G Wireless Communications: A Deep learning approach. *IEEE Transactions on Network Science and Engineering*, pages 1–1. DOI:10.1109/TNSE.2018.2848960.
- O’Shea TJ, Erpek T & Clancy TC (2017). Physical layer deep learning of encodings for the mimo fading channel. *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 76–80.
- Ren Y, Wang Y & Zhu J (2018). Spectral learning for supervised topic models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3):726–739. DOI:10.1109/TPAMI.2017.2682085.
- Rizvi UH, Janssen GJM & Weber JH (2009). Ber analysis for mpam signal constellations in the presence of fading and adc quantization noise. *IEEE Communications Letters*, 13(10):733–735. DOI:10.1109/LCOMM.2009.091586.
- Seyman MN & Taşpınar N (2013). "channel estimation based on neural network in space time block coded MIMO-OFDM system". *Digital Signal Processing*, 23(1):275 – 280. DOI:https://doi.org/10.1016/j.dsp.2012.08.003.
- Sun P, Wang Z, Heather RW & Schniter P (2019). Joint channel-estimation/decoding with frequency-selective channels and few-bit adcs. *IEEE Transactions on Signal Processing*, 67(4): 899–914. DOI:10.1109/TSP.2018.2887189.
- Tong L, Xu G & Kailath T (1994). Blind identification and equalization based on second-order statistics: a time domain approach. *IEEE Transactions on Information Theory*, 40(2):340–349. DOI:10.1109/18.312157.
- Walden RH (1999). Analog-to-digital converter survey and analysis. *IEEE Journal on Selected Areas in Communications*, 17(4):539–550. DOI:10.1109/49.761034.

- Wang S, Zhang L, Li Y, Wang J & Oki E (2016). Multiuser mimo communication under quantized phase-only measurements. *IEEE Transactions on Communications*, 64(3):1083–1099. DOI:10.1109/TCOMM.2016.2527641.
- Wu M, Yin B, Wang G, Dick C, Cavallaro JR & Studer C (2014). Large-scale mimo detection for 3gpp lte: Algorithms and fpga implementations. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):916–929. DOI:10.1109/JSTSP.2014.2313021.
- Yan X, Long F, Wang J, Fu N, Ou W & Liu B (2017). Signal detection of mimo-ofdm system based on auto encoder and extreme learning machine. *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 1602–1606.
- Yang S & Hanzo L (2015). Fifty years of mimo detection: The road to large-scale mimos. *IEEE Communications Surveys Tutorials*, 17(4):1941–1988. DOI:10.1109/COMST.2015.2475242.
- Yuan J, Jin S, Wen C & Wong K (2017). The distributed mimo scenario: Can ideal adcs be replaced by low-resolution adcs? *IEEE Wireless Communications Letters*, 6(4):470–473. DOI:10.1109/LWC.2017.2704604.
- Zhang T, Wen C, Jin S & Jiang T (2016). Mixed-adc massive mimo detectors: Performance analysis and design optimization. *IEEE Transactions on Wireless Communications*, 15(11):7738–7752. DOI:10.1109/TWC.2016.2606592.