

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

**LOCALISATION EN INTÉRIEUR BASÉE SUR L'APPRENTISSAGE AUTOMATIQUE:
EXPLOITATION DES INFORMATIONS SUR L'ÉTAT DES CANAUX WI-FI**

**MÉMOIRE PRÉSENTÉ
COMME EXIGENCE PARTIELLE DE LA
MAÎTRISE EN GÉNIE ÉLECTRIQUE**

**PAR
LOTFI BENCHARIF**

DÉCEMBRE 2020

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

MAÎTRISE EN GÉNIE ÉLECTRIQUE (M. Sc. A.)

Direction de recherche :

Messaoud Ahmed Ouameur

Prénom et nom

directeur de recherche

Daniel Massicotte

Prénom et nom

codirecteur de recherche

Jury d'évaluation

Usef Faghihi

Prénom et nom

Professeur UQTR

Abdellah Chehri

Prénom et nom

Professeur UQAC

Résumé

La localisation en intérieur a récemment reçu une grande attention en raison de l'utilisation potentielle d'une large gamme de services intelligents. L'exploitation du signal Wi-Fi, en tant que technologie de la localisation d'un appareil cible, est l'une des approches les plus étendues et les plus recherchées. En particulier, les points d'accès Wi-Fi sont largement déployés dans n'importe quel environnement intérieur. En analysant différentes technologies et algorithmes, nous avons conclu que la méthode basée sur le Wi-Fi s'est avérée être l'approche la plus prometteuse pour déterminer l'emplacement d'un appareil. Les méthodes de suivi intérieures existantes basées sur le Wi-Fi obtiennent principalement des informations sur l'état des canaux, ou obtiennent une force de signal reçu approximative via un ordinateur personnel pour faire le suivi.

Dans ce mémoire, nous implémentons le système sur une carte réseau 802.11 commerciales de type Intel NIC 5300. Nous fournissons un guide d'installation et de mise en œuvre du matériel et des logiciels nécessaires à un système comme celui-ci, et soulignons les points clés qui permettront aux futurs chercheurs de mettre en place rapidement des équipements.

Des approches d'apprentissages profonds et récurrents sont rigoureusement étudiées pour améliorer la précision des systèmes de localisation. Une approche basée sur le réseau récurrent à mémoire court et long terme (LSTM) a été comparée à d'autres approches de pointe. Une explication complète pour sélectionner les meilleurs hyper paramètres est présentée puis nous évaluons ses performances dans différents scénarios intérieurs typiques.

Avant-propos

Tout d'abord je remercie ALLAH le tout puissant de m'avoir donné de l'effort et la patience pour achever ce mémoire.

Je tiens tout d'abord à exprimer ma profonde gratitude à mon directeur de recherche, Monsieur Messaoud Ahmed Ouameur, qui s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu me consacrer.

Je voudrais remercier également mon co-directeur de recherche, Monsieur Daniel Massicotte pour le suivi régulier et l'aide qu'il a pu m'apporter tout au long de l'élaboration de ce mémoire.

Mes remerciements s'étendent aussi aux membres des évaluateurs pour avoir accepté de me consacrer une partie de leur temps afin d'examiner et de juger mon travail.

Enfin, que tous les enseignants ayant participé à ma formation tout au long du cycle de mes études voient ici ma profonde reconnaissance.

Un grand merci à mes parents, mon frère, mes sœurs ainsi que toute la famille et tous mes amis pour leurs soutiens et encouragements.

Table des matières

Résumé.....	iii
Avant-propos.....	iv
Table des matières.....	v
Liste des tableaux.....	viii
Liste des figures	ix
Liste des symboles	xii
Chapitre 1 - Introduction.....	16
1.1 Contexte du projet de recherche.....	17
1.2 Problématique.....	18
1.3 Objectifs	22
1.4 Méthodologie.....	23
1.5 Organisation du mémoire	24
Chapitre 2 - La localisation en intérieur.....	25
2.1 Introduction	25
2.2 Les systèmes et technologies de localisation.....	25
2.2.1 Satellites GPS.....	29
2.2.2 Réseau cellulaire GSM	31
2.2.3 Wi-Fi.....	33

2.2.4	Autre système de localisation	35
2.3	Techniques de localisation	41
2.3.1	L'instant d'arrivée (ToA).....	41
2.3.2	La différence des instants d'arrivées ou TDoA	43
2.3.3	Le temps d'aller-retour du signal ou RToF.....	45
2.3.4	L'angle d'arrivée ou AoA.....	46
2.3.5	RSSI	47
2.4	Algorithme d'empreinte digitale.....	49
2.4.1	Algorithme voisin le plus proche (Near Neighbors, NN).....	50
2.4.2	Algorithme des k plus proches voisins (k Nearest Neighbors, kNN)	51
2.4.3	Algorithme pondéré des k-voisins les plus proches.....	52
2.5	Conclusion.....	53
Chapitre 3 - Implémentation de Intel NIC 5300		54
3.1	Introduction	54
3.2	Information d'état de canal.....	54
3.3	Le choix du matériel.....	56
3.4	Traitement de ses données sur Matlab.....	63
3.5	Modèle de localisation en intérieur basé sur l'apprentissage automatique	72

3.5.1	K- Voisin le plus proche (kNN).....	72
3.5.2	Machine à vecteurs de support (SVM)	72
3.5.3	Réseau récurrent à mémoire court et long terme LSTM.....	73
3.5.4	Arbres de décision (DT).....	74
3.6	Conclusion.....	75
Chapitre 4 - Résultats et discussion		76
4.1	Introduction	76
4.2	Zone d'essai.....	76
4.3	Mesure de RSS et CSI signal.....	78
4.4	Équipements et outils	79
4.4.1	Point d'accès	79
4.4.2	Outil de mesure.....	80
4.5	Environnement expérimental et procédure.....	81
4.6	Évaluation de différentes méthodes de classification.....	85
4.6.1	Premiers scénarios	85
4.6.2	Deuxièmes scénarios.....	96
4.1	Conclusion.....	107
Chapitre 5 - Conclusion		108
Bibliographies		110

Liste des tableaux

Tableau 2-1 Résumé des différentes technologies sans fil pour la localisation [9]	40
Tableau 4-1 Ensemble finalisé des meilleurs hyper paramètres pour le scénario 1	88
Tableau 4-2 Ensembles finalisés des meilleurs hyper paramètres pour le scénario 2.....	97
Tableau 4-3 Comparaisons de la précision de localisation (%) pour différentes méthodes d'apprentissage.....	105
Tableau 4-4 Comparaisons de la précision de localisation (%) pour différentes méthodes de mesure basées sur RSS.	106

Liste des figures

Figure 2-1 Aperçu des systèmes de positionnement sans fil actuels [24].....	26
Figure 2-2 Visibilités des satellites GPS à Melbourne le 3 juillet 2013 (source : http://www.n2yo.com/).....	30
Figure 2-3 Architectures typiques des systèmes basés sur iBeacon [9].....	36
Figure 2-4 ToA à trois positions [49].....	42
Figure 2-5 Localisation et détection de proximité basées sur TDoA [9].....	44
Figure 2-6 Principes de localisation par la méthode d'AoA [24].....	47
Figure 2-7 Localisations basées sur RSSI [9].....	48
Figure 3-1 a) Carte Wi-Fi et b) Adaptateur PCI-Express	57
Figure 3-2 Script pour le début d'enregistrement du CSI.....	62
Figure 3-3 Scripts pour envoyer des pings entre le Tx et Rx.....	63
Figure 3-4 Données collectées au point 1 au LSSI.....	64
Figure 3-5 Les informations collectées de la première cellule	64
Figure 3-6 Scripts pour la normalisation du point A au LSSI.....	67
Figure 3-7 Le CSI obtenu avant la normalisation.....	68
Figure 3-8 Le CSI obtenu normalisé.....	68
Figure 3-9 Exemple où le $N_{tx} = 1$	69
Figure 3-10 Scripts pour analyser les trois chemins spatiaux différents.....	70
Figure 3-11 Exemple où le $N_{tx} = 2$	70
Figure 3-12 Tracer du SNR pour les 3 antennes.....	71
Figure 3-13 Calculs des valeurs SNR effectives.....	72
Figure 3-14 L'unité de base du LSTM	73
Figure 4-1 a) Deuxième étage du bâtiment L-Léon-Provancher (LP) et b) maison.....	77

Figure 4-2 Formes d'onde RSS mesurées à différents points de référence.....	78
Figure 4-3 Formes d'onde CSI mesurées à différents points de référence.....	79
Figure 4-4 Photos du modèle du AP.....	80
Figure 4-5 Schémas de la collecte des données.....	80
Figure 4-6 Données collectées.....	81
Figure 4-7 Scripts utilisés pour la réorganisation des données.....	82
Figure 4-8 Matrices de données obtenues après la réorganisation.....	82
Figure 4-9 Classifications des données pour chaque point de mesure.....	83
Figure 4-10 Schémas de l'organisation de nos données.....	84
Figure 4-11 Scripts pour la réorganisation de nos données d'entraînement et de validation.....	84
Figure 4-12 Dispositions du pavillon Léon-Provenche au LSSI 2488LP et les placements du point d'accès.....	85
Figure 4-13 Cartes d'environnement expérimental au LSSI (2488LP).....	86
Figure 4-14 Organigrammes du processus de sélection des hyperparamètres.....	87
Figure 4-15 Résultats de la simulation au LSSI.....	89
Figure 4-16 a) Matrices de confusion pour l'apprentissage et b) la validation au LSSI.....	90
Figure 4-17 Matrices de confusion: (Model 1 : SVM quadratique) (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) au LSSI.....	94
Figure 4-18 ROC et AUC courbes pour (Model 1 : SVM quadratique) , (Model 2 :kNN pondéré) , (Model 3 :Arbre de décision) , (Model 4 :kNN) , (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) au LSSI.....	96
Figure 4-19 Architectures de la maison et les placements du point d'accès.....	97
Figure 4-20 Résultats de la simulation à la maison.....	98
Figure 4-21 a) Matrices de confusion pour l'apprentissage et b) la validation à la maison.....	100

Figure 4-22 Matrices de confusion: (Model 1 : SVM quadratique), (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) à la maison	103
Figure 4-23 ROC et AUC courbes pour (Model 1 : SVM quadratique), (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) à la maison	104

Liste des symboles

AoA	Angle d'arrivée (<i>Angle of arrival</i>)
AUC	Aire sous la courbe ROC (<i>Area Under The Curve</i>)
BLE	<i>Bluetooth Low Energy</i>
BLPA	<i>Bluetooth local position application</i>
BS	Station de base (<i>Base station</i>)
BTS	Station de transmission de base (<i>Base transceiver station</i>)
CDF	<i>Cumulative distribution function</i>
CDMA	Accès multiple par répartition en code (<i>Code division multiple access</i>)
CFR	Channel frequency response
CID	<i>Cell identity</i>
CNN	Réseau neuronal convolutif (<i>Convolutional Neural Networks</i>)
CSI	Informations d'état de canal (<i>Channel state information</i>)
DBL	Appareils basés localement (<i>Devices based local</i>)
dBm	<i>Decibel-milliwatt</i>

DF	Radiogoniométrie (<i>Direction finding</i>)
DL	Apprentissage profondi (<i>Deep learning</i>)
DT	Arbre de décision (<i>Decision tree</i>)
E-OTD	<i>Enhanced Observed Time Difference</i>
GPS	<i>Global Positioning System</i>
GSM	<i>Global System for Mobile Communications</i>
HMM	<i>Hidden Markov Model</i>
IEEE	Institut des ingénieurs électriciens et électroniciens (<i>Institute of Electrical and Electronics Engineers</i>)
IoT	Internet des objets (<i>Internet of things</i>)
IPS	Système de positionnement en intérieur (<i>Indoor positioning system</i>)
ISM	Bande industrielle, scientifique et médicale (<i>Industrial, scientific and medical</i>)
kNN	Méthode des k plus proches voisins (<i>k-nearest neighbors</i>)
LOS	<i>Line-of-sight</i>
LPWA	<i>Low Power Wide Area Network</i>
LSTM	<i>Long Short-Term Memory</i>

MAC	<i>Media Access Control</i>
ML	Apprentissage automatique (<i>Machine learning</i>)
MNN	<i>Multilayer neural network</i>
MTC	Communication de type machine <i>machine type communication</i>
NLOS	<i>Non-line-of-sight</i>
NN	<i>Near Neighbors</i>
ns	<i>Nanosecond</i>
OFDM	<i>Orthogonal frequency-division multiplexing</i>
OSS	<i>One step secant</i>
PDA	Assistant personnel numérique (<i>Personal digital assistant</i>)
RF	Radio fréquence (<i>Radio frequency</i>)
ROC	<i>Receiver Operating Characteristics</i>
RSS	<i>Received Signal Strength</i>
RTof	<i>Retour time of flight</i>
Rx	Receveur
STFT	<i>Short-Time Fourier Transform</i>

SVM	Machine à vecteurs de support (<i>support vector machine</i>)
TDoA	<i>Time Difference Of Arrival</i>
ToA	<i>Time Of Arrival</i>
Tx	Transmetteur
UWB	<i>Ultra wideband</i>
WLAN	<i>Wireless Local Area Network</i>
WSN	<i>Wireless Sensor Networks</i>

Chapitre 1 - Introduction

Actuellement, la localisation fait l'actualité dans le secteur High Tech et de nouvelles technologies. Grâce à ces techniques de géolocalisation, nous pouvons maintenant positionner des personnes ou du matériel avec une précision variante pouvant aller de 10 km à 10 cm.

Localisé est l'un des besoins fondamentaux de la vie depuis le tout début de ce monde. Au début, les informations de position des cibles sont estimées en sentant, en sondant, en traçant ou en voyant. Au fur et à mesure que le monde se développe, tout devient de plus en plus compliqué, de même que l'espace dans lequel nous vivons. Les chercheurs accordent une grande attention à cette technologie et ils ont maintenant affaire à un endroit avec beaucoup d'informations. Ce qui est important pour eux est d'obtenir autant de détails que possible sur l'environnement autour et la position d'eux-mêmes.

Les téléphones intelligents et d'autres appareils sans fil ont beaucoup évolué ces dernières années et ont abouti à une large gamme de services, y compris la localisation en intérieur.

La localisation en intérieur est le processus d'obtention de l'emplacement d'un utilisateur ou d'un appareil dans un cadre ou un environnement intérieur. La localisation des appareils en intérieur a été largement étudiée au cours des dernières années, principalement dans des environnements industriels et pour les réseaux de capteurs sans fil et la robotique. Cependant, il y a à peine dix ans, la prolifération à grande échelle des téléphones intelligents et des appareils portables dotés de capacités de communication sans fil a rendu la localisation et le

suivi de ces appareils synonyme de localisation et de suivi des utilisateurs correspondants et a permis un large éventail des applications et services associés.

1.1 Contexte du projet de recherche

La localisation des utilisateurs et des équipements connaît un essor fulgurant d'applications dans de différents domaines comme exemple l'industrie, de la gestion des catastrophes, de la santé [1-3], de la gestion des bâtiments, de la surveillance et de nombreuses autres applications. Cette technologie peut également profiter à de nombreux systèmes novateurs tels que l'Internet des objets (IoT) [4], les architectures intelligentes (telles que les villes intelligentes[5], les bâtiments intelligents [6], les réseaux intelligents [7]) et la communication de type machine (MTC) [8].

Aujourd'hui, l'émergence de IoT a apporté une nouvelle révolution économique dont les entreprises espèrent profiter. L'un des services que ces objets peuvent fournir est la localisation précise, d'où l'on peut dire que l'IoT est une fusion de nombreuses technologies hétérogènes et de normes de communication conçues pour fournir des connexions de bout en bout à des milliards d'appareils.

Bien que la recherche actuellement se concentre sur les technologies émergentes liées à la communication machine à machine à distance, les technologies existantes à court et moyen terme, tel que Bluetooth, Zigbee, Wi-Fi, Ultra Large Bande (UWB), etc., resteront des parties inséparables du réseau IoT. C'est pourquoi un grand nombre d'appareils IoT utiliseront plus d'une interface de communication, une pour la communication courte et une pour la communication longue portée. Alors que les technologies IoT longue portée soient conçues pour fournir des solutions de communication à couverture élevée et à faible consommation

d'énergie, elles ne peuvent pas prendre en charge les débits de données élevés requis par diverses applications locales.

D'un autre côté, bien que les technologies IoT à longue portée n'aient pas été conçues avec une disposition de localisation à l'intérieur, de nombreuses applications IoT nécessiteront une localisation et/ou une navigation intérieure/extérieure transparente et omniprésente des appareils tant statiques que mobiles.

Les technologies de communication à courte portée traditionnelle peuvent estimer assez précisément l'emplacement intérieur relatif d'un dispositif IoT par rapport à certains points de référence, mais l'emplacement global (c'est-à-dire les coordonnées géographiques longitude-latitude) de ces appareils reste inconnu. À moins que la position globale du point de référence ne soit également connue. Les technologies IoT à distance émergentes peuvent estimer l'emplacement global d'un appareil (car l'emplacement exact de son point d'accès est généralement connu), mais sa précision est considérée comme très faible, en particulier dans les environnements intérieurs [9].

1.2 Problématique

Les services basés sur la localisation ont attiré beaucoup d'attention au cours de ces dernières années. Contrairement au scénario extérieur, la localisation en intérieur reste une tâche difficile en raison de l'absence de signaux du système de positionnement global (GPS) [10]. Les informations de localisation en l'intérieur gagnent en importance dans les services et applications de communication modernes [11], tels que la surveillance de la santé des personnes âgées et handicapées en situation d'urgence [12], la reconnaissance des activités

pour l'efficacité énergétique dans les maisons intelligentes [13] et de nombreuses autres applications liées à IoT [14, 15].

À titre d'exemple, les auteurs dans [16] ont discuté des réseaux de capteurs de faible puissance clairsemée utilisant des capteurs de potentiel électrique passif pour les bâtiments intelligents sensibles à la foule où le système de détection passive peut surveiller à distance les signaux physiologiques humains. Ces systèmes utilisent des signaux radio, des champs magnétiques ou d'autres informations sensorielles collectées par différents types de capteurs avec diverses technologies (GPS, Wi-Fi, Bluetooth, ZigBee, ultrasons et infrarouge) [14]. Ces technologies varient considérablement en termes de précision, de couverture, d'efficacité, de sécurité et de fiabilité, qui restent un obstacle impératif dans la localisation en intérieur [15] [14]. Différentes méthodes de localisation à l'intérieur peuvent généralement être classées en trois groupes principaux en fonction des informations de modélisation utilisées: méthodes basées sur la force du signal reçu (RSS), l'angle d'arrivée (AoA) ou le temps d'arrivée (ToA) [17]. S'appuyant sur des techniques d'empreintes digitales Wi-Fi, le RSS a été utilisé avec succès pour la localisation active des appareils sans fil [10, 18].

Les auteurs de [18] sont les premiers à adopter le traçage de rayons comme prédicteur de canal pour aider l'empreinte digitale en intérieur en utilisant des mesures de réponse impulsionnelle de canal pour obtenir une réduction des erreurs de localisation de 25% [18].

Les auteurs de [15] et [11] ont présenté un système UWB personnalisé et un système Wi-Fi modifié pour acquérir d'autres métriques que RSS. Après avoir transformé le signal 802.11 large bande en impulsions à bande étroite, le système WiSee [13] utilise un prétraitement par transformée de Fourier à court terme (STFT) pour déduire différentes caractéristiques discriminantes.

D'autre part, les auteurs dans [11] ont présenté une approche de la localisation en intérieur en utilisant un réseau neuronal convolutif (CNN) pour la classification des canaux et la régression d'erreur de télémétrie sur les traces d'informations d'état de canal brut (CSI) unidimensionnel.

Dans [15], les auteurs ont utilisé LSTM pour la reconnaissance d'activité plutôt que d'autres techniques classiques d'apprentissage automatique (Machine learning ou ML), telles que le modèle de Hidden Markov caché (HMM). Même si l'approche LSTM souffre d'un long temps d'apprentissage, elle surpasse les techniques de ML. Outre que les signaux Wi-Fi, l'adoption à grande échelle attendue des appareils IoT utilisant des réseaux étendus à basse consommation (Low Power Wide Area Network ou LPWAN) telle que LoRaWAN, ouvre de nouvelles opportunités pour explorer la quantité massive de données générées pour permettre divers services et applications qui reposent sur la localisation précise des personnes, biens et actifs, allant de la domotique et de la vie assistée à l'automatisation accrue des processus de production et de logistique et à l'optimisation du réseau sans fil. Les réseaux LPWA sont différents, car ils font des compromis uniques par rapport aux technologies traditionnelles prédominantes dans la scène IoT telles que Zig-Bee et Bluetooth.

Les auteurs de [19] ont développé une infrastructure de localisation en intérieur basé sur le LPWAN à l'aide d'un réseau LoRaWAN personnalisé. Les résultats montrent qu'ils peuvent prédire l'emplacement intérieur avec une précision de plus de 98% à l'aide d'un réseau neuronal multicouche (MNN).

Battiti *et al.* dans [20] ont proposé une méthode pour déterminer l'emplacement à l'aide d'un classificateur basé sur un réseau neuronal. Ils ont adopté une architecture de perceptron multicouche (MLP) et une méthode d'apprentissage *one step secant* (OSS). Ils ont choisi une

architecture à trois couches avec trois unités d'entrée, huit unités de couche cachée et deux sorties, car l'architecture peut obtenir les erreurs d'apprentissage et de test les plus faibles et n'est pas très sensible aux effets "*overfitting*". Ils rapportent que seuls cinq échantillons de puissance du signal à différents endroits sont suffisants pour obtenir une erreur de distance moyenne de 3 m. En augmentant le nombre d'exemples d'entraînement, l'erreur de distance moyenne est réduite à 1,5 m.

Dans [21], Xian et *al.* ont proposé un schéma d'apprentissage de la distribution de la propagation du signal basé sur le modèle et l'algorithme de positionnement de l'assistant de suivi, dans lequel une machine d'état est utilisée pour obtenir de manière adaptative une plus grande précision dans les états de suivi et de non-suivi. Le système atteint une précision de détermination de position statique de 90% avec une précision de 2m. Pour les appareils mobiles ambulants, une précision de 5 m est obtenue.

Ladd *et al.* [22] ont proposé un algorithme de localisation de robot bayésien basé sur une grille utilisant l'infrastructure IEEE 802.11. Dans la première étape de l'algorithme, l'hôte utilise le modèle de probabilité pour calculer ses probabilités de localisation dans plusieurs emplacements différents en fonction du RSS des 9 points d'accès. La deuxième étape utilise la vitesse maximale limitée de l'utilisateur mobile pour optimiser les résultats (à partir de la première étape) et rejette les solutions qui ont des changements significatifs dans la localisation de l'hôte mobile. Selon que le deuxième étage est utilisé (83% et 77% du temps), l'hôte peut prédire que sa position est à moins de 1,5 m.

Dans [23], les auteurs présentent une approche basée sur l'apprentissage en profondeur (Deep learning ou DL) pour la localisation en intérieur en utilisant des mesures de la qualité du canal de transmission, y compris le RSS et les informations d'état du canal (Chanel state

information ou CSI). Ils ont divisé un plan de pièce rectangulaire en blocs bidimensionnels. Chaque bloc est traité comme une classe et définit la localisation comme un problème de classification. En utilisant RSS et CSI, ils ont développé quatre réseaux de neurones profonds mis en œuvre avec des MLP et des réseaux de neurones convolutifs unidimensionnels (1-D-CNN) pour estimer la position des objets dans la pièce. Les résultats expérimentaux montrent que le réseau 1-D-CNN utilisant les informations CSI peut obtenir d'excellentes performances de positionnement.

Dans ce mémoire, nous allons voir l'étude des systèmes de localisation basés sur le Wi-Fi pour une utilisation en intérieur.

Le système Wi-Fi présente l'avantage de disposer d'une infrastructure de plus en plus présente grâce à son déploiement dans les environnements internes. Le système Wi-Fi, comme toute technologie sans fil, a des signaux qui peuvent surmonter des obstacles multiples dans les bâtiments (murs, meubles, etc.). La localisation basée sur le Wi-Fi est principalement obtenue en mesurant le RSS et le CSI. Après ses mesures de RSS et CSI sont ensuite utilisées par des techniques de ML pour localiser des objets. Mais, les mesures des données sont affectées par les caractéristiques de l'environnement et cela présente plusieurs défis à cause des interférences causées par d'autres activités sans fil et l'atténuation causée par les obstacles.

1.3 Objectifs

La localisation en intérieur présente plusieurs axes de recherche, dans ce projet, l'axe principal de notre recherche est la mise en œuvre d'une architecture de système de localisation en intérieur basé sur le Wi-Fi, qui nous permet de faire une campagne de mesure

pour les données Wi-Fi et l'application des modèles d'apprentissages automatiques pour traiter le problème de localisation comme un problème de classification. Pour ce faire, les sous-objectifs suivants seront réalisés :

- Étudiez d'abord les technologies, techniques et algorithmes pouvant être utilisés dans la conception de notre projet.
- La mise au point d'un système de localisation en intérieur performant avec une précision améliorée, qui nous permet de faire la collecte des données dans deux scénarios différents (une maison et à l'université), nous évaluerons l'adéquation des intensités du signal reçu aux signaux Wi-Fi pour le problème de localisation en intérieur et exploiter toutes les données qu'on a acquises comme le CSI et RSS.
- Appliquer différents types d'apprentissages automatiques pour la classification de ses résultats en plusieurs classes, travailler sur notre algorithme pour obtenir une meilleure précision et faire une comparaison entre les différentes méthodes de réseau de neurones pour voir la technique qui nous donne de meilleurs résultats.

1.4 Méthodologie

Ce travail est réalisé dans le laboratoire des signaux et systèmes intégrés (LSSI), il est fait en plusieurs étapes afin de pouvoir obtenir de meilleurs résultats.

- Analyse préliminaire des technologies ainsi que techniques existantes de la localisation en intérieur, avec des exemples d'application et dans la littérature
- Choisir le matériel pour faire notre setup. Il existe plusieurs types de cartes Wi-Fi, dans ce projet le choix de cette carte est important. Donc nous avons cherché une

carte spécifique qui nous a permis d'acquies plusieurs informations lors de la collecte de données.

- Créer plusieurs scénarios différents pour les mesures des données dans le LSSI, l'université et une maison. Cette partie va nous permettre d'effectuer des tests de notre setup dans différents endroits et en présence des différents facteurs qui peuvent affecter les signaux radiofréquence.
- Applications des différentes techniques d'apprentissage automatique sur les données collectées dans différents scénarios, pour effectuer la classification en plusieurs classes et comparer les résultats des différentes techniques de ML.
- Optimisation de notre LSTM qui est la dernière phase de notre projet pour pouvoir avoir une meilleure précision dans différents scénarios.

1.5 Organisation du mémoire

La structure de ce mémoire est comme suit. Ce chapitre présente une introduction au contexte du projet ainsi que différentes problématiques qu'il y a dans la littérature et nos principaux objectifs qu'on veut atteindre. Dans le deuxième chapitre, nous résumerons la revue de littérature sur les différentes techniques, technologies, algorithmes et systèmes utilisés pour la localisation en intérieur. Après, le chapitre III est divisé en deux sections, la première est dédiée aux algorithmes de ML et DL utilisés dans le problème de localisation en intérieur sur les données Wi-Fi, et la deuxième décrira notre setup de collecte de données Wi-Fi. Ensuite, le chapitre IV montre les résultats obtenus dans différents scénarios de localisation en intérieur avec la comparaison des résultats obtenue. Finalement, le chapitre V présentera une conclusion générale de notre travail avec les travaux futurs.

Chapitre 2 - La localisation en intérieur

2.1 Introduction

Dans ce chapitre, nous allons nous focaliser sur les définitions des systèmes, technologies et techniques générales relatives à la localisation en intérieur. Tout d'abord, on va commencer par les systèmes et technologies de la localisation. Par la suite, les techniques basiques de localisation seront présentées. La dernière section va être consacrée aux algorithmes d'empreinte digitale.

2.2 Les systèmes et technologies de localisation

Les technologies de localisation ont connu un essor important avec le développement de l'émission radio, plusieurs systèmes de localisation ont été développés pour des besoins militaires, la navigation maritime et aérienne. Actuellement, la maîtrise des technologies de l'espace offre à l'homme un GPS extrêmement efficace et précis. L'intégration de cette technologie avec les infrastructures fonctionnelles dans plusieurs domaines de la vie moderne constitue une source d'information de grand intérêt pour suivre en temps réel les trajectoires.

Les techniques de localisation évoluent continuellement depuis des années. Sa précision en extérieur a été grandement améliorée, et il est si courant que le GPS a presque banni l'utilisation des cartes. Actuellement, des millions d'utilisateurs ont des navigateurs qui peuvent les guider dans n'importe quel endroit lors d'un voyage en voiture ou des applications de téléphone intelligent en temps réel qui peuvent les diriger partout vers n'importe quel endroit.

Toutefois, les utilisateurs utilisent de la localisation à l'intérieur des bâtiments: à la maison, au travail, au centre commercial, les endroits où les signaux GPS ne sont pas disponibles. Des applications qui permettent, par exemple les gens peuvent avoir besoin de naviguer dans un bâtiment public comme un centre commercial ou un hôpital. Il peut être difficile de trouver son chemin dans un tel immeuble, de guider une personne aveugle à l'intérieur d'un bâtiment jusqu'à l'emplacement souhaité et d'aider les visiteurs à trouver leur chemin vers une certaine destination.

Plusieurs types de technologies sans fil sont utilisées pour la localisation à l'intérieur. La figure 2.1 représente un aperçu des systèmes de localisation sans fil actuels.

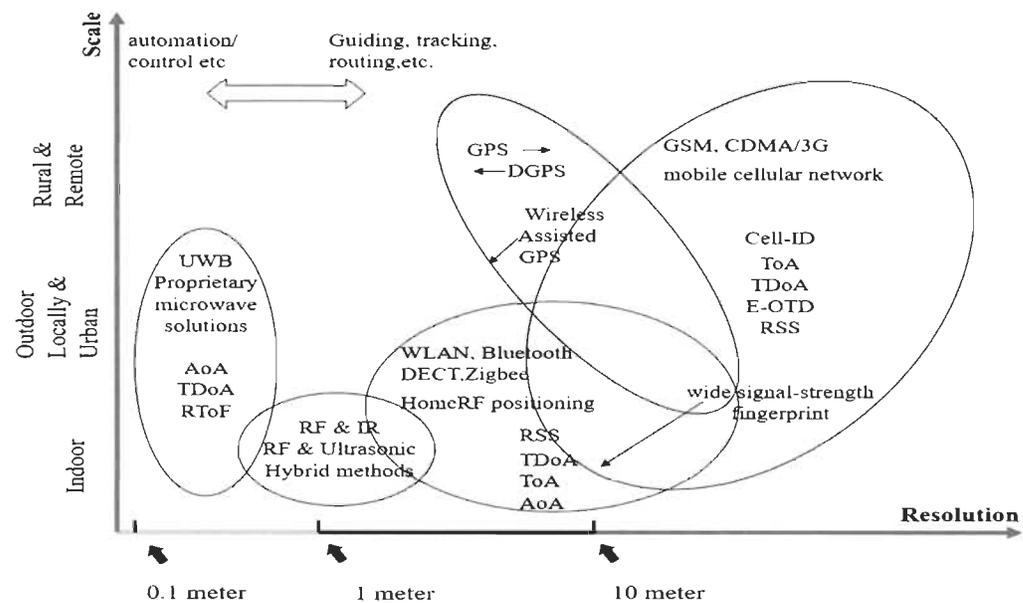


Figure 2-1 Aperçu des systèmes de positionnement sans fil actuels [24]

Les performances des systèmes de positionnement en intérieur peuvent être analysées en fonction de six propriétés différentes [24]:

- **Exactitude** : L'exactitude (ou erreur de localisation) est la condition la plus importante des systèmes de positionnement. Elle mesure la taille de l'écart entre

la position de point de consigne réelle dans cette zone considérée et l'emplacement du même point de consigne estimé par le système de positionnement intérieur qui donne l'erreur de performance du système de positionnement en intérieur (IPS). Certains compromis entre une précision « appropriée » et d'autres caractéristiques sont nécessaires.

- **Complexité** : La complexité du système IPS peut être attribuée au matériel, aux logiciels et aux facteurs de fonctionnement. Elle peut être donnée par le temps de calcul requis pour l'algorithme de positionnement. Le taux de localisation est plus élevé si la complexité est plus élevée. Le taux de localisation est un indicateur important de la complexité. Le double du taux de localisation est le décalage de localisation, qui est le délai entre le déplacement d'une cible mobile vers un nouvel emplacement et le signalement du nouvel emplacement de cette cible par le système.
- **Précision** : La précision est la mesure entre les écarts des points d'estimation. Une fonction de distribution cumulative précise (CDF) peut être utilisée pour calculer les forces et les faiblesses, où plus la concentration des courbes est élevée, plus la précision est élevée. La précision ne tient compte que de la valeur des erreurs de distance moyenne. Cependant, la précision de localisation tient compte de la cohérence du fonctionnement du système, c'est-à-dire qu'il s'agit d'une mesure de la robustesse de la technique de positionnement, car elle révèle la variation de ses performances sur de nombreux essais.
- **Robustesse** : La bonne robustesse du système est pendant le processus de positionnement, même si l'une des sources de signal agit anormalement ou ne peut

pas être reçue, le système de positionnement intérieur doit encore fonctionner avec un certain niveau de précision. Parfois, le signal d'un émetteur est totalement bloqué, de sorte que le signal ne peut pas être obtenu à partir de certaines unités de mesure. La seule information permettant d'estimer l'emplacement est le signal des autres unités de mesure. Quelquefois, certaines unités de mesure peuvent ne pas fonctionner ou être endommagées dans un environnement difficile. Les techniques de positionnement doivent utiliser ces informations incomplètes pour calculer l'emplacement.

- **Évolutivité:** L'évolutivité est la plage prise en charge par le système de positionnement intérieur, qui comprend la zone d'installation expérimentale, la couverture du signal et le positionnement spatial 2-Dimension (2-D) ou 3-Dimension (3-D). Certains systèmes peuvent prendre en charge les espaces 2-D et 3-D. Le caractère évolutif d'un système assure la fonction de positionnement normal lorsque la portée de positionnement devient grande. Habituellement, les performances de positionnement se dégradent lorsque la distance entre l'émetteur et le récepteur augmente.
- **Coût:** Le coût d'un système de localisation peut dépendre de nombreux facteurs. Les coûts de construction du système comprennent le temps passé, la consommation d'énergie, l'espace, le poids et l'argent. Le facteur temps est lié à l'installation et à la maintenance et parfois, nous devons tenir compte de certains coûts irrécupérables.

2.2.1 *Satellites GPS*

Les systèmes satellitaires sont largement utilisés dans le monde. Ils se composent de plusieurs ensembles de satellites qui fournissent des signaux de positionnement géospatial autonomes avec une couverture mondiale. Le système de navigation global qui existe aujourd'hui permet à tout appareil électronique équipé d'un récepteur de signal satellite de calculer sa position 3-D (longitude, latitude et altitude).

Les services de localisation GPS peuvent être utilisés dans toutes les conditions météorologiques et partout sur la planète. La seule exigence est qu'il ne doit y avoir aucun obstacle entre l'appareil mobile et le ciel. Le système de satellites GPS est à l'origine composé de 24 satellites placés sur six orbites.

De nos jours, le GPS est considéré comme un outil / utilitaire de positionnement populaire pour localiser les objets mobiles nécessitant des services de base de position et de navigation. Par exemple, les puces GPS intégrées dans les téléphones intelligents deviennent des composants standard pour permettre d'obtenir des informations de localisation. Pour la détermination de la position 2-D, par ex. latitude et longitude (ouest et nord), un récepteur GPS doit suivre les signaux GPS d'au moins trois satellites. Avec quatre satellites ou plus en vue, le récepteur peut déterminer la position 3-D de l'utilisateur. De plus, le système de localisation GPS est basé sur la mesure du temps nécessaire pour localiser les récepteurs et comme les ondes reçues souffrent de réflexions, le système ne sera pas adapté aux zones fermées ou aux espaces où le signal ne pénètre pas correctement [25].

La figure 2.1 est un instantané de la visibilité des satellites GPS à Melbourne le 3 juillet 2013. La seule condition requise est qu'il ne doit pas y avoir d'obstacle entre l'appareil mobile et le ciel.

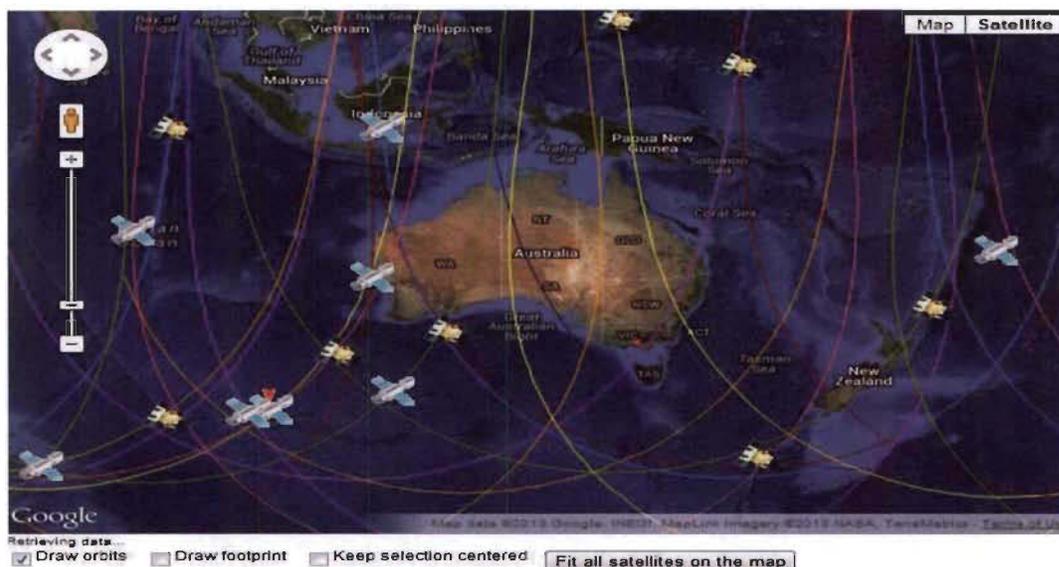


Figure 2-2 Visibilités des satellites GPS à Melbourne le 3 juillet 2013 (source : <http://www.n2yo.com/>)

Le système GPS a inspiré certains chercheurs pour la localisation en intérieur en utilisant différents matériaux [26].

Récemment, Atmel et U-blox ont annoncé le lancement d'un nouveau type de technologie de suivi des signaux faibles par GPS appelé SuperSense. Avec ce nouveau logiciel GPS, la navigation GPS devient possible dans les intérieurs de bâtiments et les profonds canyons urbains en raison de sa sensibilité de suivi au-delà de -158 dBm. Ses performances ne sont pas rapportées pour l'instant.

Locata Corporation a inventé une nouvelle technologie de localisation appelée Locata [27], pour un positionnement de précision aussi bien à l'intérieur qu'à l'extérieur. Une partie de la « technologie Locata » consiste en un émetteur-récepteur synchronisé dans le temps

appelé LocataLite. Un réseau de LocataLites forme un LocataNet, qui transmet des signaux de type GPS qui permettent un positionnement en un seul point en utilisant des mesures de phase de porteuse pour un appareil mobile (une Locata). Le groupe de navigation et de localisation par satellite (SNAP) de l'Université de New South Wales a aidé au développement d'une locata et à l'essai de la nouvelle technologie. Les expériences de test démontrent la preuve de concept de la «technologie Locata» et montrent que le positionnement du point de phase de la porteuse (sans liaisons de données par modem radio) est possible avec une précision sous-centimètre [27].

2.2.2 Réseau cellulaire GSM

Le GPS n'est pas la seule manière de résoudre les problèmes de géolocalisation, il est possible d'utiliser des technologies habituellement utilisées pour le transfert de données comme le système mondial de communications mobiles (GSM). Cette technique permet le positionnement d'un terminal GSM en se basant sur certaines informations relatives aux antennes GSM auxquelles le terminal est connecté.

Le GSM est une norme mondiale pour les communications mobiles. Jusqu'à présent, le GSM est une technologie de pointe dans le domaine des téléphones mobiles et offre une bonne couverture dans les zones urbaines et rurales. En Amérique du Nord, le GSM utilise les bandes de fréquences 850 MHz et 1900 MHz et en Europe, Moyen-Orient et Afrique le GSM utilise les bandes de fréquences 900 MHz et 1800.

Le réseau GSM est divisé en plusieurs cellules. Chaque cellule contient une station de base de transmission (BTS) appelée station d'émetteur-récepteur de base. Les stations de base GSM sont généralement équipées de plusieurs antennes directionnelles, qui définissent la

couverture de la cellule. Les cellules sont identifiées par un ID de cellule unique (CID) et selon leur emplacement, elles peuvent couvrir environ 35 kilomètres [28]. Chaque appareil mobile est associé à une cellule. La sélection de cellules est basée sur la force du signal BTS reçu par la station mobile. En général, celui avec le signal le plus fort est sélectionné. Cependant, dans certains cas, par exemple pour éviter la saturation, d'autres unités peuvent être sélectionnées.

De nombreux systèmes ont utilisé des systèmes de réseau cellulaire mobile à accès multiple mobile / par répartition par code (GSM / CDMA) pour estimer la localisation de clients mobiles externes. Cependant, en fonction de la taille de la cellule, la précision de la méthode utilisant l'identifiant de cellule ou la différence de temps d'observation améliorée (E-OTD) est généralement faible (environ 50-200 m). Généralement, dans les zones densément peuplées (comme les zones urbaines), la précision est plus élevée, tandis que dans les environnements ruraux, la précision est plus faible [29].

Otsason *et al.* [30] ont proposé le premier système de suivi intérieur précis basé sur le GSM intérieur et l'utilisation d'empreintes digitales de grande puissance de signal. Les données utilisées sont collectées à partir des 6 cellules GSM les plus puissantes détectées et en utilisant des lectures de jusqu'à 29 canaux GSM supplémentaires. Ils ont présenté les résultats d'expériences utilisant la technologie des k plus proches voisins (kNN) pondérés sur les empreintes digitales de la force du signal recueillies dans trois bâtiments à plusieurs étages. Les résultats montrent que leur système de suivi intérieur peut distinguer la différence entre les étages et atteindre une précision moyenne aussi basse que 2,5 m.

Pour l'extérieur, le travail d'utilisation du GSM pour localiser le téléphone mobile a été effectué. Goetz *et al.* [31] recommandent d'utiliser la méthode de différence des instants

d'arrivées (TDoA) pour localiser les appareils mobiles. Ils rapportent une précision de cinq mètres. Habituellement, le nombre de gratte-ciel et de bâtiments dans la zone affectera la précision du système.

2.2.3 *Wi-Fi*

Notre société fait face à un boom technologique dans l'industrie du sans-fil et les applications connexes. Ces dernières années, la technologie des réseaux locaux sans fil (WLAN) a connu une croissance sans précédent. La recherche a révélé que presque tous les derniers ordinateurs portables, assistants numériques personnels (PDA) et téléphones portables sont désormais équipés de cartes d'interface réseau sans fil. Cela montre que la demande de WLAN augmente, il convient non seulement aux entreprises, mais aussi très populaire dans les universités, les bureaux et les foyers [32].

Wi-Fi est l'un des standards de l'Institut des ingénieurs électriciens et électroniciens ou IEEE pour le WLAN. Le dernier amendement de la norme originale (802.11) est le 802.11n qui a été ratifié à la fin de 2009 [33]. Cependant, des normes plus anciennes comme 802.11a / b / g sont encore couramment utilisées. La différence entre ces variations réside dans la fréquence, la modulation et la vitesse de transfert maximale. Les bandes de fréquences de la norme Wi-Fi pour toutes les variations sont 2,4 GHz et 5 GHz.

Cette norme de réseau local sans-fil de milieu de gamme fonctionnant dans la bande de 2,4 GHz pour l'industrie, la science et la médecine (ISM), est devenue très populaire dans les points d'accès publics et les entreprises ces dernières années. Avec 11, 54 ou 108 Mb/s, IEEE 802.11 est actuellement le principal standard de WLAN. Par conséquent, il est donc intéressant d'utiliser également une infrastructure WLAN existante pour la localisation en

intérieur, en ajoutant un serveur de localisation. La précision d'un système de positionnement WLAN typique utilisant RSS est d'environ 3 à 30 m, et le taux de mise à jour est d'environ quelques secondes [24].

La norme IEEE 802.11 fonctionne dans la bande de fréquences ISM est principalement utilisée pour fournir des fonctions de réseau et des connexions internet avec divers appareils dans des environnements privés, publics et commerciaux. Au départ, la portée de réceptions du Wi-Fi est d'environ 100 mètres [24], mais dans IEEE 802.11ah (principalement optimisé pour les services IoT), elle est maintenant passée à 1 km [34, 35].

Bahl *et al.* [35] ont proposé un système de localisation et de suivi des utilisateurs dans les bâtiments RADAR, qui utilise la méthode des plus proches voisins dans la technologie spatiale du signal, qui est la même que la méthode de kNN. Les auteurs proposent deux types de méthodes pour déterminer la localisation de l'utilisateur. La première dépend de la mesure empirique de la force du signal du point d'accès en phase hors ligne, la deuxième est la modélisation de la propagation du signal. La précision du système RADAR est d'environ 2 à 3 m. Dans leur travail de suivi [36], le RADAR a été amélioré par l'algorithme de type Viterbi. Par conséquent, le 50 centiles du système RADAR est d'environ 2,37 à 2,65 m et son 90 centiles est d'environ 5,93 à 5,97 m.

Dans [37] les auteurs ont proposés un système de positionnement intérieur Wi-Fi combine les fonctions des méthodes de mesure trilatérale et d'analyse de scène. Le système peut déterminer à quel étage l'utilisateur peut se trouver. L'expérience montre que l'algorithme peut atteindre une précision de 100% dans leur expérience.

2.2.4 Autre système de localisation

D'autres techniques et méthodes ont également été exploitées telles que :

2.2.4.1 Bluetooth

Bluetooth (ou IEEE 802.15.1) inclut des spécifications pour les couches physiques et contrôles d'accès au support (MAC) qui connectent divers appareils sans fil fixes ou mobiles dans un espace personnel spécifique. La dernière version de Bluetooth, Bluetooth Low Energy (BLE), également connu sous le nom de Bluetooth Smart, peut fournir un débit de données amélioré de 24 Mb/s et une plage de couverture de 70 à 100 mètres, et a une efficacité énergétique supérieure par rapport à l'ancienne version [38].

Désormais, la plupart des appareils mobiles disposent d'un module Bluetooth, il fonctionne dans la bande ISM 2,4 GHz. Par rapport au WLAN, le débit binaire brut est inférieur (1 Mbps).

Lors de l'utilisation de Bluetooth pour le positionnement, le problème le plus important est l'incertitude du temps de réponse. Il faut 2 à 11 secondes pour découvrir un appareil à proximité de l'utilisateur. Dans un environnement mobile, cela peut gravement affecter la précision de la position. Rendre les appareils Bluetooth plus détectables augmentera le pourcentage de temps passé dans ce que l'on appelle « l'état d'interrogation », qui nécessite plus d'énergie. L'absence de temps de réponse prévisible reste un inconvénient majeur des systèmes de positionnement Bluetooth.

Marco *et al.* [39] ont proposé un système de localisation utilisant Bluetooth qui peut fournir des estimations de position avec 90% de précision et 0,5m de précision lors d'une

marche. Cependant, ce système présente l'inconvénient d'avoir une couverture limitée en raison de sa courte portée.

Antti *et al.* ont présenté la conception et la mise en œuvre d'une application de positionnement local Bluetooth (BLPA) [40]. Leurs méthodes sont qu'ils convertissent le niveau de puissance du signal reçu en une estimation de distance basée sur un modèle de propagation simple, puis utilisent un filtre de Kalman étendu (EKF) basé sur la distance. La précision de BLPA sera de 3,76 m.

Apple a annoncé iBeacon lors de la World Developers Conference (WWDC) de 2013 [41]. Le protocole est spécialement conçu pour la détection de proximité et les services basés sur la proximité. Ce protocole permet aux appareils compatibles BLE (également appelés iBeacon ou balises) d'envoyer des balises ou des signaux à intervalles périodiques.

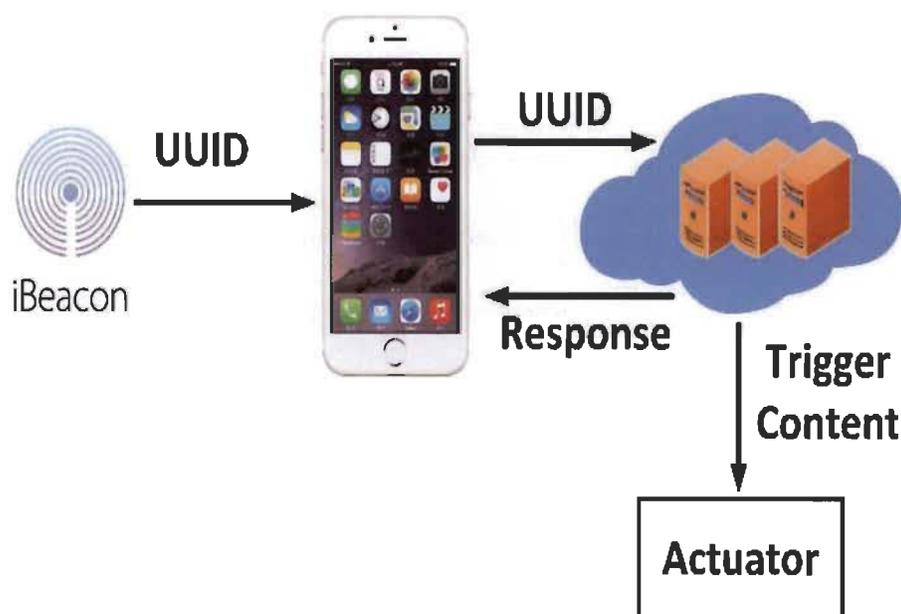


Figure 2-3 Architectures typiques des systèmes basés sur iBeacon [9]

2.2.4.2 ZigBee

ZigBee est basé sur la norme IEEE 802.15.4, qui couvre les couches physiques et MAC pour les réseaux personnels à faible coût, à faible débit et à faible consommation d'énergie [42]. ZigBee définit la couche supérieure de la pile de protocoles et est principalement utilisé dans les réseaux de capteurs sans fil. La couche réseau de ZigBee est responsable du routage multisauts et de l'organisation du réseau, tandis que la couche application est responsable de la communication distribuée et du développement d'applications. Bien que ZigBee ait tendance à localiser les capteurs dans les réseaux de capteurs sans fil (WSN), il n'est pas facile d'utiliser ZigBee sur la plupart des appareils utilisateur, ce qui n'est pas propice à la localisation des utilisateurs à l'intérieur.

Les auteurs dans [43] proposent un système de positionnement basé sur un réseau de capteurs basé sur IEEE802.15.4 / ZigBee. Le système utilise des empreintes digitales et utilise SVM pour estimer l'emplacement des nœuds. L'expérience a été menée dans une galerie minière souterraine et les résultats ont montré que leurs systèmes pouvaient localiser le nœud mobile avec une erreur de distance moyenne de 0,6 à 1 mètre.

2.2.4.3 Dispositif d'identification par radiofréquence (RFID)

Dispositif d'identification par radiofréquence est une méthode de stockage et de récupération de données par transmission électromagnétique vers un circuit intégré compatible au radio fréquence (RF) et il est maintenant considéré comme une méthode pour améliorer le processus de traitement des données [44]. Le système RFID comprend plusieurs composants de base, notamment plusieurs lecteurs RFID, des étiquettes RFID et la communication entre eux. Le lecteur RFID peut lire les données envoyées par l'étiquette

RFID. Les lecteurs et étiquettes RFID utilisent des RF et des protocoles définis pour envoyer et recevoir des données. Les étiquettes RFID sont classées comme passives ou actives.

Les étiquettes RFID passives peuvent fonctionner sans piles. Ils sont principalement utilisés pour remplacer la technologie traditionnelle des codes à barres et sont plus légers, plus petits et moins chers que les étiquettes actives. Les étiquettes RFID actives sont de petits émetteurs-récepteurs qui peuvent envoyer activement leur ID (ou d'autres données supplémentaires) en réponse à des requêtes. À l'exception des gammes de fréquences basses et hautes, la gamme de fréquences utilisée est semblable au boîtier RFID passif [24].

2.2.4.4 Ultra large bande (UWB)

UWB est basé sur l'envoi d'impulsions ultra-courtes avec une période de temps < 1 nanoseconde (ns) sont transmises sur une large bande passante (> 500 MHz), dans la gamme de fréquences de 3,1 à 10,6 GHz, en utilisant un rapport cyclique de service très bas [24] peut réduire la consommation d'énergie. Cette technologie est principalement utilisée dans les systèmes de communication à courte portée, tels que les périphériques PC et d'autres applications intérieures.

UWB est devenu une technologie particulièrement intéressante pour le suivi en intérieur, car elle n'est pas perturbée par d'autres signaux (en raison de sa différence fondamentale de type de signal et de spectre radio), tandis que les signaux UWB (en particulier les fréquences contenues dans le spectre UWB) peuvent pénétrer à travers divers matériaux, y compris les murs (bien que les métaux et les liquides puissent interférer avec les signaux UWB) [9].

2.2.4.5 *Autres systèmes*

Un certain nombre d'autres systèmes, comme signal acoustique, ultrason et infrarouge ont également été étudiés pour leur utilisation potentielle et leur capacité de localisation.

L'infrarouge est un rayonnement électromagnétique avec de longues longueurs d'onde, qui est souvent utilisé dans les communications à courte portée entre des appareils intégrés d'infrarouges. Les communications infrarouges sont utiles pour les environnements intérieurs. Cependant, l'infrarouge ne peut pas pénétrer les murs, ce qui rend son utilisation très difficile entre les pièces.

L'ultrason est une onde de pression sonore oscillante avec une fréquence supérieure à la limite supérieure de la plage d'audition humaine. Les appareils à ultrasons peuvent être utilisés pour détecter des objets et mesurer des distances.

Tous les systèmes ci-dessus ne peuvent fonctionner que dans des mesures à courte portée. Le tableau suivant montre un résumé des différentes technologies sans fil utilisé pour la localisation.

Tableau 2-1 Résumé des différentes technologies sans fil pour la localisation [9]

Technologie	Portée maximale	Débit maximum	Consommation d'énergie	Avantage	Inconvénient
IEEE 802.11 n 802.11 ac 802.11 ad	250m exte 35m intér Quelques mètres	600 Mbps 1.3 Gb/s 4.6 Mbps	Modéré Modéré Modéré	Largement disponible, haute précision, ne nécessite pas de matériel supplémentaire complexe	Sujette au bruit, nécessite des algorithmes de traitement complexes
UWB	10-20 m	460 Mbps	Modéré	Immunité aux interférences, offre une grande précision	Portée plus courte, nécessite du matériel supplémentaire sur différents appareils utilisateur, coût élevé
Ultrason	Quelques dizaines de mètres	30 Mbps	Faible modérée	Comparativement moins d'absorption	Dépendance élevée au placement du capteur
RFID	200 m	1.67 Gb/s	Faible	Consommez une faible puissance, a une large portée	La précision de localisation est faible
Bluetooth	100 m	24 Mbps	Faible	Débit élevé, plage de réception, faible consommation d'énergie	Faible précision de localisation, sujette au bruit
LoRa	15 km	37.5 kb/s	Extrêmement bas	Large plage de réception, faible consommation d'énergie	Longue distance entre la station de base et l'appareil, atténuation du signal extérieur-intérieur du serveur en raison des murs du bâtiment
SigFox	50 km	100 Bps	Extrêmement bas	Large plage de réception, faible consommation d'énergie	Longue distance entre la station de base et l'appareil, atténuation du signal extérieur-intérieur du serveur en raison des murs du bâtiment
IEEE 802.11 ah	1 km	100 kb/s	Extrêmement bas	Large plage de réception, faible consommation d'énergie	Pas complètement exploré pour la localisation, les performances à voir dans l'environnement intérieur

2.3 Techniques de localisation

Le calcul d'une position repose principalement sur trois métriques différentes, la multilatération, l'angulation et la prise d'empreinte, ce sont les trois techniques les plus couramment utilisées pour résoudre les problèmes de localisation. Ces techniques peuvent utiliser différentes technologies (GPS, GSM, Wi-Fi et Bluetooth) pour les espaces intérieurs et extérieurs.

On différencie de nombreuses techniques qu'on examine une à une.

2.3.1 L'instant d'arrivée (ToA)

Un signal électromagnétique a une vitesse de propagation constante. Ces méthodes sont basées sur l'estimation du délai de propagation entre l'émetteur et le récepteur. Le retard est déterminé par l'inter corrélation croisée entre le signal reçu et le signal émis. La position du plus grand pic dans cette corrélation indique l'arrivée du signal de *ligne of sight* (LoS) et donne la valeur de retard à partir du temps de référence.

Puisque ToA est basé sur la différence entre l'heure d'arrivée et l'heure de transmission, tous les récepteurs et émetteurs doivent être synchronisés au niveau de l'horloge. Ils nécessitent généralement un matériel plus complexe et des références temporelles absolues. Compte tenu de la vitesse de propagation du signal et du besoin de synchronisation, cela est devenu un axe important de la recherche dans ce domaine [45]. ToA peut être mesuré par différentes techniques de signalisation, telles que la mesure à spectre étalé à séquence directe [46, 47] ou à ultra-large bande (UWB) [48].

Comme le montre la figure 2-4, la distance entre le capteur et la cible peut être déterminée lors de la mesure du ToA. Ensuite, l'intersection du cercle et la distance déterminée par le

capteur sont la position de la cible mesurée. Afin d'obtenir un certain point d'articulation unique, il doit y avoir au moins 3 cercles, ce qui signifie qu'au moins 3 capteurs sont nécessaires dans un système de positionnement 2-D.

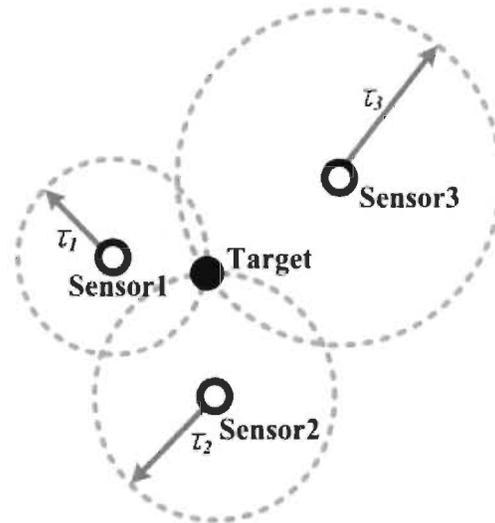


Figure 2-4 ToA à trois positions [49]

La technique ToA présente deux inconvénients principaux [50]. Le premier inconvénient est qu'il nécessite des matériaux supplémentaires très spéciaux pour la mesure. La précision de mesure de ToA a une influence importante sur la précision de l'estimation de position. La vitesse de propagation des signaux radio dans l'air est d'environ 3×10^8 m/s, ce qui équivaut presque à la vitesse de la lumière. L'erreur de mesure de ToA est d'une microseconde, ce qui entraîne une erreur de distance estimée d'environ 300 mètres. Le deuxième inconvénient concerne tous les émetteurs et récepteurs du système. Les appareils doivent être synchronisés avec précision. Le signal transmis doit être horodaté afin que l'unité de mesure puisse identifier la distance parcourue par le signal.

2.3.2 La différence des instants d'arrivées ou TDoA

La différence de temps d'arrivée (TDoA) est une variation de ToA, mais elle ne nécessite que les stations fixes d'avoir des horloges internes synchronisées. Il implique plusieurs récepteurs stationnaires qui ont des horloges synchronisées qui collaborent pour trouver l'emplacement de la source du signal [51]. TDoA exploite la différence des temps de propagation des signaux de différents émetteurs, mesurés au niveau du récepteur. Les mesures TDoA ($T_{D(i,j)}$ - des émetteurs i et j) sont converties en valeurs de distance physique $L_{D(i,j)} = c \cdot T_{D(i,j)}$, où c est la vitesse de la lumière. Le récepteur est maintenant situé sur l'hyperboloïde donné par l'équation (2.1).

$$L_{D(i,j)} = \sqrt{(X_i - x)^2 + (Y_i - y)^2 + (Z_i - z)^2} - \sqrt{(X_j - x)^2 + (Y_j - y)^2 + (Z_j - z)^2} \quad 2.1$$

Où (X_i, Y_i, Z_i) sont les coordonnées du nœud émetteur / référence i et (x, y, z) sont les coordonnées du récepteur / utilisateur.

La TDoA d'au moins trois émetteurs est nécessaire pour calculer l'emplacement exact du récepteur à l'intersection de trois (ou plus) hyperboloïdes. Les équations hyperboliques peuvent être résolues par régression linéaire [52] ou en utilisant l'expansion des séries de Taylor pour linéariser les équations.

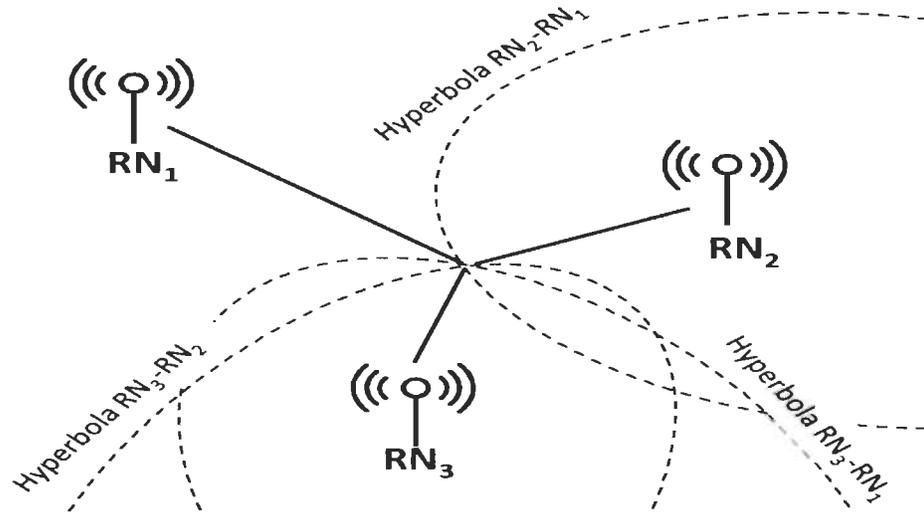


Figure 2-5 Localisation et détection de proximité basées sur TDoA [9]

La figure 2-5 montre comment utiliser trois nœuds de références différents pour obtenir la position 2-D de n'importe quelle cible. Elle montre aussi les hyperboles formées par les résultats de la mesure de la position de l'utilisateur (point noir) obtenu à partir des nœuds de références.

Les méthodes conventionnelles de calcul des estimations TDoA sont basées sur la corrélation. TDoA peut être estimée à partir de l'inter corrélation entre les signaux reçus à une paire d'unités de mesure. Admettons que pour le signal transmis est $s(t)$ et que le signal reçu à la première station de base est $x_1(t)$ mais altéré par le bruit $n_1(t)$ [53]. Supposons aussi que $s(t)$ est retardé par D et altéré par le bruit $n_2(t)$, , alors on aura :

$$x_1(t) = s(t) + n_1(t) \quad 2.2$$

$$x_2(t) = As(t - D) + n_2(t) \quad 2.3$$

Où A est le rapport d'amplitude. On veut estimer D , le décalage de temps d'arrivée (TDOA) et $s(t)$ entre les deux récepteurs. Il peut également être nécessaire d'estimer l'amplitude de mise à l'échelle A , et en estimant la mise à l'échelle d'amplitude, un récepteur approprié peut être sélectionné [53]. Par conséquent, la fonction d'inter corrélation des signaux reçus par les deux récepteurs est donnée par :

$$\hat{R}_{x_1x_2}(\tau) = \frac{1}{T} \int_0^T x_1(t)x_2(t - \tau)dt \quad 2.4$$

La TDoA estimée est la valeur de τ qui maximise $\hat{R}_{x_1x_2}(\tau)$. Cette méthode nécessite que le récepteur partage une référence temporelle précise, mais n'impose pas d'exigences à la station mobile. Habituellement, des techniques de traitement dans le domaine fréquentiel sont utilisées pour trouver τ .

2.3.3 *Le temps d'aller-retour du signal ou RToF*

RToF mesure le temps de propagation du signal aller-retour (émetteur-récepteur) pour estimer la distance entre Tx et Rx [54]. Pour RToF, la condition de synchronisation d'horloge n'est pas aussi stricte que les exigences ToA, mais leurs mécanismes de mesure de distance sont les mêmes. ToA et RToF ont des mécanismes de télémétrie similaires. Après avoir reçu le signal de l'émetteur, le récepteur répond à l'émetteur, puis l'émetteur calcule le total aller-retour ToA. Le principal avantage de RToF est que, par rapport à ToA, une synchronisation d'horloge relativement modérée est requise entre Tx et Rx. Cependant, la précision de l'estimation RToF est affectée par les mêmes facteurs que ToF (c'est-à-dire la fréquence d'échantillonnage et la largeur de bande du signal) qui dans ce cas est plus strict puisque le signal est transmis et reçu deux fois [9].

Un autre problème important pour les systèmes basés sur RToF est le temps de réponse du récepteur, qui dépend fortement de l'électronique du récepteur et des frais généraux de protocole. Si le temps de propagation entre l'émetteur et le récepteur est important par rapport au temps de réponse, ce dernier peut être ignoré, mais dans les systèmes à courte portée comme pour le suivi en intérieur, le retard ne peut pas être ignoré. Soit t_1 le moment où Tx_i envoie un message à Rx_j , et Rx_j reçoit le message à l'instant $t_2 = t_1 + t_p$. j renvoie le signal à Tx_i au temps t_3 , puis le reçoit à t_4 et v est la vitesse de propagation des signaux radio dans l'air. Par conséquent, la distance entre i et j peut être calculée comme suit [54] :

$$D_{ij} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2} \times v \quad 2.5$$

2.3.4 L'angle d'arrivée ou AoA

La mesure de l'angle d'arrivée (AoA) détermine la direction d'origine du signal reçu. L'estimation AoA *direction finding* (DF) peut être effectuée à l'aide d'antennes directionnelles ou de réseaux d'antennes. Afin de mesurer l'angle de réception, la station de base (BS) du système appliquant l'AoA est équipée des antennes sensibles à la direction, qui est généralement composée d'un réseau d'éléments d'antenne pouvant diviser sa direction [55].

En AoA, la position de la cible est à l'intersection de plusieurs paires de lignes de direction angulaire, dont chacune va de la station de base ou de la station balise à la cible en mouvement [45]. Comme le montre la figure 2.6, AoA peut utiliser au moins deux points de référence connus A et B et deux angles (θ_1, θ_2) pour obtenir la position 2-D de la cible P [56].

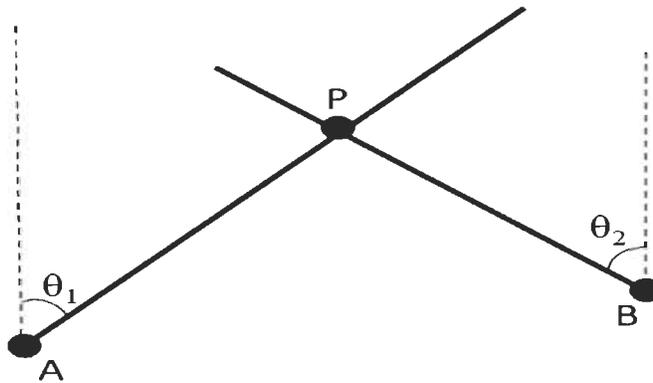


Figure 2-6 Principes de localisation par la méthode d'AoA [24]

L'avantage de l'AoA est qu'il ne peut utiliser que 3 unités de mesure pour le positionnement 3-D ou 2 unités pour le positionnement 2-D pour déterminer l'estimation de position, et il n'y a pas de synchronisation d'horloge entre les unités de mesure. Les inconvénients comprennent une configuration matérielle relativement complexe et une diminution significative de la précision de position lorsque la cible en mouvement s'éloigne de l'unité de mesure. L'utilisation d'AoA pour un positionnement précis doit prendre en compte la réflexion de directions trompeuses, ou en mesurant la directivité de l'ouverture de l'antenne [57, 58].

2.3.5 RSSI

RSSI indique la force du signal radio reçu, qui est principalement mesurée en dBm (1 dBm = 1.259 milliwatts). Plus le récepteur est proche de l'émetteur, plus le signal est fort, car la force du signal est atténuée lors de sa propagation dans l'air, et l'atténuation est proportionnelle à la distance [59]. Par exemple, le chipset Atheros Wi-Fi utilise des valeurs RSSI comprises entre 0 et 60, tandis que Cisco utilise une plage entre 0 et 100. En utilisant

le RSSI et un modèle de propagation de perte de chemin simple [60], la distance d entre Tx et Rx peut être estimée comme :

$$RSSI = -10n \log_{10}(d) + A \quad 2.6$$

Où n est l'exposant de perte de trajet (qui varie de 2 dans l'espace libre à 4 dans les environnements intérieurs) et A est la valeur RSSI à une distance de référence du récepteur.

Dans le cas de l'emplacement de base de l'appareil (DBL), la localisation basée sur RSS nécessite une trilatération ou une latération en N points, c'est-à-dire que le RSS au niveau de l'appareil est utilisé pour estimer la distance absolue entre l'appareil utilisateur et au moins trois points de référence, puis la géométrie / trigonométrie de base est appliquée à l'équipement utilisateur pour obtenir sa position par rapport à un point de référence, comme le montre la figure 2-7 [9].

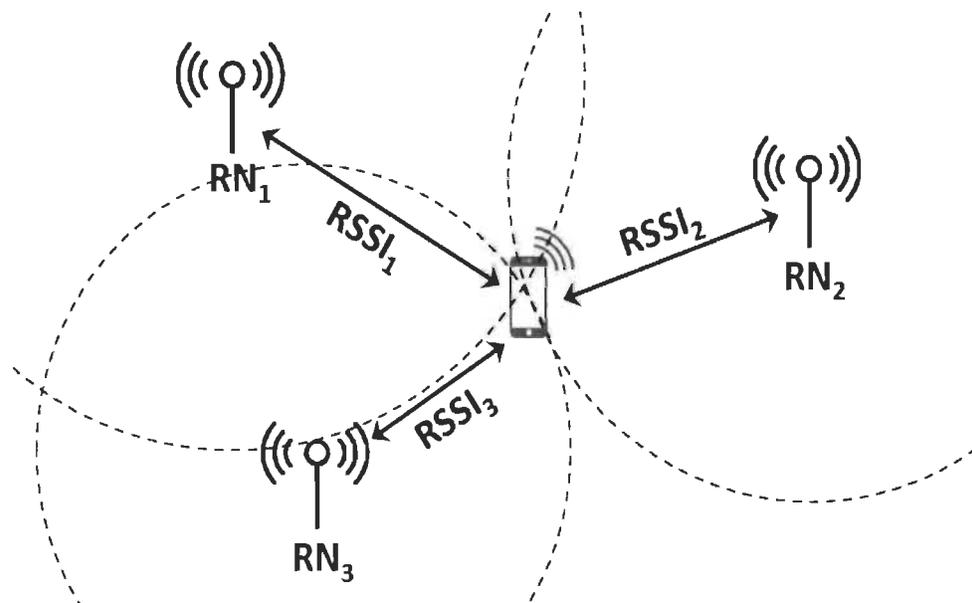


Figure 2-7 Localisations basées sur RSSI [9]

Bien que la méthode basée sur RSS soit simple et économique, elle souffre d'une mauvaise précision de positionnement (en particulier dans des conditions de *non-line-of-*

sight) en raison d'une atténuation supplémentaire du signal causée par la transmission à travers le mur. D'autres obstacles importants et fluctuations RSS dus à l'évanouissement par *multipath fading* et au bruit interne sont importants [61] [62]. Différents filtres ou mécanismes de calcul de la moyenne peuvent être utilisés pour atténuer ces effets. Cependant, il est impossible d'obtenir une précision de positionnement plus élevée sans utiliser des algorithmes complexes.

2.4 Algorithme d'empreinte digitale

Le mot vient du mot « *fingerprint* », qui fait référence aux empreintes digitales. Pour fonctionner, cette technologie nécessite une base de données qui associe un ensemble d'éléments qui caractérisent l'emplacement à certains endroits de l'environnement considéré. Ces éléments doivent être capables de distinguer chaque emplacement des autres emplacements de l'environnement. Si cette condition n'est pas remplie dans ce cas, les éléments qui composent "l'empreinte" ne sont pas importants. La sélection des caractéristiques du signal qui constituent l'empreinte digitale a évidemment une grande influence sur la précision de positionnement.

Chaque empreinte digitale a les informations RF ou RSS de chaque point d'accès et l'empreinte digitale de l'emplacement donné a une combinaison unique par rapport à toutes les empreintes digitales de la carte [63]. En ce qui concerne d'algorithmes d'empreintes digitales, il peut être classé en deux groupes de base, algorithme déterministe et probabiliste [64]

- **Algorithme déterministe :** Il essaie de trouver la distance de signal statistique minimale entre le vecteur de positionnement RSSI détecté et les vecteurs de

localisations de divers points de référence d'étalonnage. Il peut être égal ou non à la distance physique minimale entre l'emplacement physique réel de l'appareil mobile et l'emplacement enregistré de l'échantillon d'étalonnage. Le point d'échantillonnage avec la plus petite distance de signal statistique entre lui-même et le vecteur de position détecté est généralement considéré comme la meilleure estimation de position d'origine incluse dans la base de données d'étalonnage.

- **Algorithme probabiliste :** Puisqu'un tableau vectoriel d'emplacements spécifiques a été détecté, il utilise l'inférence de probabilité pour déterminer la probabilité d'emplacements spécifiques. L'algorithme traite la base de données d'étalonnage elle-même comme une distribution de probabilité conditionnelle a priori pour déterminer la probabilité d'un emplacement spécifique [65].

2.4.1 Algorithme voisin le plus proche (*Near Neighbors, NN*)

L'algorithme déterministe le plus ancien et le plus simple est l'algorithme NN utilisé dans RADAR, qui est un système de localisation et de suivi des utilisateurs basé sur RF. Il est développé en deux phases: phase hors ligne et phase de positionnement. Cet algorithme fonctionne par le calcul de la distance euclidienne entre le RSS actuel d'un emplacement inconnu et la valeur moyenne de la force du signal de chaque point d'accès enregistré dans la base de données, puis de trouver le voisin le plus proche (distance euclidienne minimale), en utilisant l'estimation des coordonnées d'un emplacement inconnu [66]. Prenons exemple le cas que $RSS_m = \{RSS_{m1}, RSS_{m2}, \dots, \dots, RSS_{mn}\}$ est le tableau des valeurs moyennes de la force du signal en temps réel mesuré à un emplacement inconnu et n est le nombre de points d'accès (Access points, APs). $RSS_{m1} \dots RSS_{mn}$ sont les valeurs moyennes de la force du signal de AP_1 , AP_2 et AP_3 respectivement.

$RSS_i = \{RSS_{i1}, RSS_{i2}, \dots, \dots, RSS_{in}\}$ est le tableau des valeurs moyennes de la force du signal de tous les points de référence de la base de données, où RSS_{i1} est la valeur moyenne de la force du signal à partir de AP_1 et RSS_{in} est la valeur moyenne de la force du signal AP_n respectivement, la distance euclidienne peut alors être définie comme:

$$d = \sqrt{(RSS_{m1} - RSS_{i1})^2 + (RSS_{m2} - RSS_{i2})^2 + \dots + (RSS_{mn} - RSS_{in})^2} \quad 2.7$$

2.4.2 Algorithme des k plus proches voisins (k Nearest Neighbors, kNN)

L'algorithme kNN est une extension de l'algorithme NN présenté ci-dessus. k est le nombre de la distance euclidienne minimale entre chaque point de référence et l'emplacement de l'appareil [67]. L'algorithme kNN mesure la distance entre un scénario de requête et un ensemble de scénarios dans l'ensemble de données, elle calcule simplement les distances euclidiennes entre la lecture $RSSI$ en direct et chaque empreinte de point de référence.

Lors de la première identification de k ($k \geq 2$) voisins les plus proches sont d'abord identifiés à partir de la base de données d'empreintes digitales en fonction des distances euclidiennes calculées entre la force du signal de chaque point de référence et celle obtenue à l'emplacement de l'appareil, puis, sur la base de toute la moyenne des coordonnées (x, y) des k voisins les plus proches détermine la meilleure position estimée. Si $k = 1$, alors kNN devient NN . Dans le cas où la distance euclidienne et le vecteur de force du signal de l'utilisateur sont $\{RSS_{i1}, RSS_{i2}, \dots, RSS_{in}\}$, où RSS_{in} se réfère à RSS de AP_n au point de référence i , alors la distance euclidienne peut être calculée par la formule suivante :

$$d_i = \sqrt{\sum_{k=1}^n (RSS_{ik} - RSS_k)^2} \quad 2.8$$

Et la mesure de distance absolue est donnée comme suit :

$$d_i = \sum_{k=1}^n |RSS_{ik} - RSS_k| \quad 2.9$$

Étant donné que l'atténuation du signal dans un réseau sans fil 802.11 est affectée par de nombreux facteurs, la valeur RSS obtenue à deux emplacements différents peut être presque la même. En général, un algorithme NN qui ne considère que le point de référence avec la valeur RSS la plus proche obtiendra généralement une position estimée loin de la position réelle. Cependant, kNN considérera plus de voisins proches et de meilleurs résultats peuvent être obtenus [68].

2.4.3 Algorithme pondéré des k-voisins les plus proches

En raison du NLOS dans un environnement intérieur, nous recevons des empreintes digitales avec la même force de signal à différents endroits. Dans ce cas, si nous utilisons l'algorithme de moyenne de la force du signal, nous obtiendrons une position inexacte. Donc on utilise la méthode pondérée des k plus proches voisins génère le résultat final par la moyenne pondérée des k positions qui ont k distances euclidiennes minimales [66]. Dans le cas où $RSS_m = \{RSS_{m1}, RSS_{m2}, \dots, \dots, RSS_{mn}\}$ est le tableau des valeurs moyennes de la force du signal en temps réel à des emplacements inconnus et n est le nombre des points d'accès, $RSS_j = \{RSS_{j1}, RSS_{j2}, \dots, \dots, RSS_{jn}\}$, ont les distances euclidiennes minimales pour RSS_m , L_j est le lieu correspondant à RSS_j , $j = 1 \dots k$. L'emplacement estimé (EL) peut-être calculé comme suit:

$$EL = \sum_{j=1}^k \frac{1}{d(RSS_j, RSS_m)} \times L_j / \sum_{j=1}^k \frac{1}{d(RSS_j, RSS_m)} \quad 2.10$$

Les algorithmes déterministes peuvent fournir une précision de positionnement raisonnable. Cependant, les performances de ces algorithmes sont généralement liées à la taille de la base de données. Par conséquent, l'établissement de grandes bases de données est la garantie de base pour une grande précision. De plus, la variation de l'intensité du signal mesurée à chaque point est importante, ainsi la moyenne des intensités du signal, plutôt qu'une mesure individuelle, est utilisée pour estimer les emplacements.

2.5 Conclusion

Ce chapitre donne un aperçu des technologies de positionnement générales, notamment le GPS, GSM, le Wi-Fi, Bluetooth, Zigbee, RFID, UWB et d'autres systèmes. Comparés à d'autres systèmes de positionnement, les systèmes de positionnement Wi-Fi sont plus adaptés et performants aux environnements intérieurs de plus de leur omniprésente partout. Cinq types de mesures pour le positionnement (à savoir ToA, TDoA, AoA, RTof et RSS) ont été développés. Trois algorithmes d'empreintes digitales sont introduits, dont l'algorithme NN, kNN et Weighted-kNN, et leurs caractéristiques spécifiques sont étudiées. Une étude approfondie d'un système de localisations à base de la technologie Wi-Fi sera détaillée dans les prochains chapitres.

Chapitre 3 - Implémentation de Intel NIC 5300

3.1 Introduction

Dans ce chapitre, nous allons présenter la configuration et l'installation du matériel et expliquer la procédure pour la mise en œuvre du logiciel. Dans la section matérielle, nous décrirons les caractéristiques et les points critiques pour l'installation du matériel choisi, nous indiquerons aussi les points essentiels et les problèmes rencontrés au cours du processus de l'implémentation de notre système, ainsi que les solutions apportées. Dans la section des logiciels, nous passerons par les étapes d'installation du logiciel pour la collecte des informations d'état du canal (ou *channel state information*, CSI). À la fin nous allons présenter les différents modèles de ML que nous allons utiliser pour traiter nos résultats.

3.2 Information d'état de canal

Les informations d'état de canal peuvent fournir des informations plus détaillées et plus fines sur les sous-porteuses par rapport au RSSI. Dans la communication sans fil, CSI décrit les facteurs de transmission du signal sur chaque trajet de transmission, tels que la diffusion du signal, l'atténuation de la distance et la dégradation de l'environnement. Lorsqu'un signal *Orthogonal frequency-division multiplexing* (OFDM) traverse le canal à trajets multiples, le signal reçu dans une sous-porteuse peut être écrit comme suit:

$$y = Hx + n \quad 3.1$$

Où x est le vecteur de signal d'émission ou on assume pour simplicité une seule antenne de transmission, y est le vecteur de signal de réception, H est la matrice d'état de canal et n est les bruits blancs gaussiens.

Ainsi, sur la base de (3.1), les informations de canal de chaque sous-porteuse peuvent être estimées comme suit:

$$\hat{H} = \frac{y}{x} \quad 3.2$$

Où \hat{H} est, dans le contexte de transmission OFDM, la réponse du canal dans le domaine fréquentiel (ou channel frequency response, CFR). Les informations CSI correspondantes peuvent être extraites d'un adaptateur sans fil Wi-Fi. Pour la sous-porteuse k de fréquence centrale f_k , son CSI est défini comme $H(f_k)$. Compte tenu de la différence des différentes sous-porteuses, la matrice H du CSI peut être écrite comme suit:

$$H = [H_1, H_2, H_3, \dots, H_N] \quad 3.3$$

Où N est le nombre de sous-porteuses. Dans ce travail, le nombre de N est fixé à 30. H_i est le CSI de i^{th} sous-porteuse. La signature CSI contient à la fois des informations d'amplitude et de phase, qui sont définies dans l'équation suivante:

$$H_i = |H_i|e^{j\angle H_i} \quad 3.4$$

Où $|H_i|$ est son amplitude, et $\angle H_i$ est sa phase [69]. Pour les systèmes à entrées multiples (MIMO) avec plusieurs antennes de réception et de transmission, le CSI de chaque sous-porteuse peut être exprimé sous la forme d'une matrice $p \times q$, où p est le nombre d'antennes de transmission et q le nombre d'antennes de réception. Ainsi, pour chaque paquet de données, on peut dériver une matrice $p \times q \times N$, où N est le nombre total de sous-porteuses.

3.3 Le choix du matériel

La première étape c'est de choisir un jeu de composants (*chipset*) qui doit respecter deux conditions importantes qui sont :

- Le fonctionnement en protocole Wi-Fi 802.11n.
- L'existence d'un protocole spécialisé pouvant accéder à la couche physique depuis l'espace utilisateur pour récupérer le CSI.

D'après la littérature, il n'y a que deux outils logiciels connus pour ce type de chipsets Wi-Fi. Daniel Halperin *et al.* en 2011, ont développé un outil qui fonctionne exclusivement pour la carte NIC Intel Wireless Link 5300 (IWL5300) [69]. En 2015, Yaxiong Xie *et al.* [70] ont développé un deuxième outil conçu pour fonctionner pour n'importe quel appareil avec un chipset du fabricant Atheros, installé soit dans un point d'accès OpenWRT OS ou une carte NIC d'un ordinateur portable.

Dans notre cas, nous avons décidé d'utiliser l'IWL5300, donc nous allons expliquer le processus d'installation de la carte Intel et le logiciel pour récupérer les données que nous avons collectées et nous identifierons tous les problèmes rencontrés lors de l'installation du matériel.

L'IWL5300 fournit des informations d'état de canal 802.11n dans un format qui rapporte les matrices de canal pour 30 groupes de sous-porteuses, soit environ un groupe pour 2 sous-porteuses à 20 MHz ou une sur 4 à 40 MHz. Chaque entrée de matrice de canal est un nombre complexe, avec une résolution signée de 8 bits pour les parties réelle et imaginaire. Il spécifie le gain et la phase du trajet du signal entre une seule paire d'antennes d'émission-réception.

L'outil logiciel Linux pour la carte IWL5300 calcule le CSI au côté récepteur lorsque les paquets qui arrivent sont envoyés aux débits 802.11n et les transmettent aux couches supérieures du système d'exploitation, où l'utilisateur peut y accéder.

Pour ce faire, il existe deux principaux endroits où appliquer les modifications logicielles: le système d'exploitation et le micrologiciel de la carte sans fil. Le changement dans le micrologiciel de la carte est transparent pour nous, car nous ne pouvons pas voir le code, mais nous sommes conscients de son objectif. Le nouveau micrologiciel de la carte active le mode de débogage dans la carte. Ce mode a été conçu par les développeurs du chipset pour vérifier le fonctionnement du calcul CSI, donc le chipset le calculera pour tout paquet qui arrive avec un débit 802.11n.

Notre setup se compose de composants importants qui sont la carte sans fil Intel NIC 5300 (IWL530), d'un ordinateur où la brancher, adaptateur externe pour étendre la carte réseau sans fil mini-PCI vers PCI-Express, trois antennes Wi-Fi et 3 fils mini-PCI vers antenne Wi-Fi.

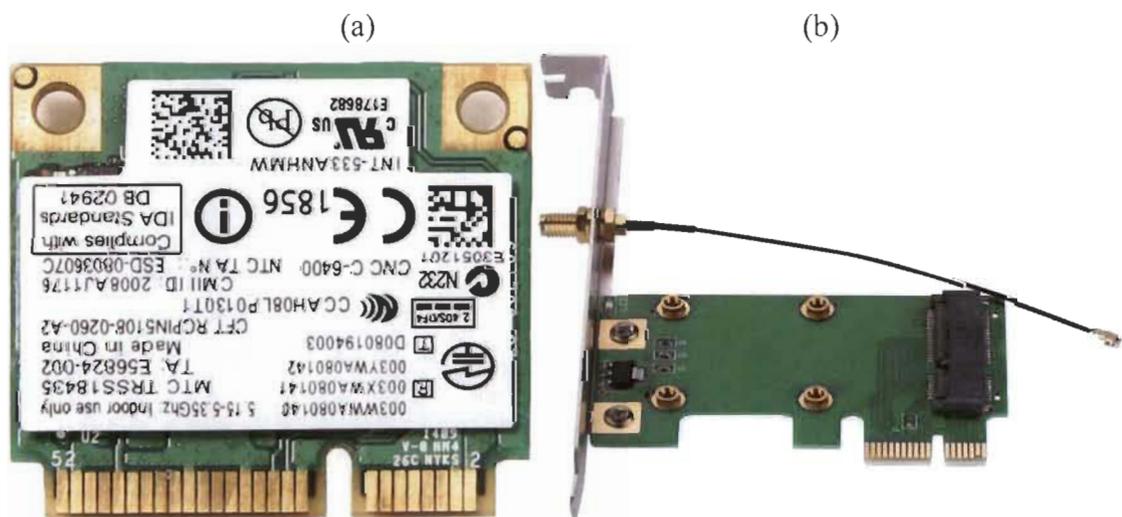


Figure 3-1 a) Carte Wi-Fi et b) Adaptateur PCI-Express

Sur le marché il y a une carte pleine taille et une demi-taille. Le petit a été conçu pour être intégré aux ordinateurs portables les plus récents et les plus compacts, là où le gros ne l'est pas. La deuxième caractéristique importante de l'une des cartes est qu'elles utilisent un connecteur PCI-Express [71].

Important : La carte IWL530 n'est pas compatible avec les adaptateurs mini PCI-e vers USB elle doit être branché directement avec la carte mère, dans notre cas nous l'avons essayé avec cet adaptateur et ça n'a pas marché, mais nous avons réussi le montage avec un adaptateur mini PCI-e vers PCI-Express.

Pour ce projet, nous avons eu accès au Laboratoire des Signaux et Systèmes Intégrés LSSI à l'Université du Québec à Trois-Rivières, où on a de nouveaux ordinateurs très puissants avec un processeur Xeon E5-2620 v3 @2.0 GHz, Carte graphique Nvidia RTX-2080ti et une RAM de 32GB DDR4.

Après le branchement de notre carte et les antennes, l'ordinateur va démarrer, et nous pouvons installer l'outil CSI pour Linux [69].

Nous suivons les instructions d'installations dont la première est de choisir les systèmes d'exploitation Linux basés sur une version du noyau Linux en amont entre 3.2 (par exemple Ubuntu 12.04) et 4.2 (par exemple Ubuntu 14.04.4).

Pour cette application, nous avons essayé la version Ubuntu 12.04 et nous avons eu quelques soucis, donc après plusieurs recherches sur GitHub, nous avons choisi une ancienne version d'Ubuntu 14.04 avec des noyaux de version 3.13 qui a marché parfaitement.

Une fois le système d'exploitation configuré et mis en marche, nous pouvons commencer par l'installation de l'outil CSI en suivant les instructions de [69]. Comme indiqué, ce

processus comporte cinq étapes à effectuer à partir de la ligne de commande et peut être résumé comme suit:

A. Conditions préalables

Installation des outils de compilation, des en-têtes de développement Linux selon la version du noyau et le client Git.

Après cela, le guide d'installation recommande de désactiver le gestionnaire de réseau pour éviter les conflits dans la gestion des connexions sans fil, mais dans mon cas ça ne m'a pas aidé dans les étapes suivantes parce que j'ai plus accès à internet et cela m'empêche sur les prochaines étapes pour le téléchargement du micrologiciel modifié.

```
echo iface wlan0 inet manual | sudo tee -a /etc/network/interfaces  
sudo restart network-manager
```

Pour éviter de perdre accès à internet, nous conseillons de ne pas exécuter les deux lignes ci-dessus et ça va nous faciliter de continuer la suite de l'installation

B. Créer et installer le pilote sans fil modifié

Obtenir l'arborescence des sources Linux de l'outil CSI, avec les modifications appropriées du pilote sans fil correspondant à la version du noyau. Lors de ce processus de clonage direct du référentiel, CSI Tool a pris beaucoup de temps pour cet ordinateur (plus de 12h). Pour gagner du temps, il est préférable de créer un clonage du référentiel sur un autre ordinateur pour nous faciliter la tâche si nous voulons refaire notre setup sur une autre machine.

Il est indispensable de fusionner ces modifications de pilote avec l'arborescence des sources Linux pour la version du noyau fournie par la distribution.

Enfin, nous devons construire et installer le pilote modifié comme expliqué dans les instructions [69]. Pour cet ordinateur, nous obtiendrons le message impossible de lire la clé privée après le chargement du nouveau pilote, mais cela n'affectera pas le fonctionnement de l'outil.

C. Installer le micrologiciel modifié

Pour l'installation du micrologiciel modifié, il faut tout d'abord, se procurer le matériel supplémentaire de l'outil CSI et remplacer le micrologiciel d'origine de la carte IWL 5300 par le micrologiciel modifié [69].

D. Créer l'outil d'enregistrement de l'espace utilisateur

Cette étape consiste à créer l'outil qui enregistre dans l'espace utilisateur le CSI qui provient du pilote [69].

E. Activer l'enregistrement et le test

Enfin, nous devons télécharger le pilote et le recharger en activant les capacités d'enregistrement. Pour activer les différentes possibilités d'enregistrement, un paramètre appelé *connector_log* est chargé en même temps que le pilote. Le paramètre *connector_log* chargé avec le pilote indique le type d'informations que nous voulons enregistrer.

En principe, cela suffirait à configurer l'outil si nous voulons calculer le CSI pour les paquets provenant d'un modem Wi-Fi. Nous devons établir une connexion avec un point d'accès sans mot de passe et exécuter la commande suivante pour démarrer l'utilitaire d'enregistrement CSI:

```
sudo linux-80211n-csitool-supplementary/netlink/log_to_file csilssi.dat
```

En exécutant cette commande, nous avons remarqué qu'un nouveau fichier s'est créé en nom de **csilssi.dat** sur notre ordinateur, mais le problème est que ce fichier est vide pour cela nous avons dû refaire les étapes suivantes :

Le Wi-Fi est activé (l'option activée le Wi-Fi est cochée), mais les réseaux Wi-Fi disponibles ne sont pas affichés, alors nous essayons d'exécuter la commande suivante pour chercher les réseaux Wi-Fi

```
sudo iw dev wlan0 scan
```

mais nous obtenons l'erreur suivante : «*command failed: Network is down (-100)*»

Pour cette erreur il y a plusieurs choses à vérifier, parce que quand on a désactivé le gestionnaire de réseau pour qu'il ne contrôle pas l'interface sans fil, cela signifie que nous ne pouvons pas utiliser l'interface graphique de gestionnaire de réseau pour nous connecter à un réseau sans fil. Si nous mettons sur liste noire le module du noyau, nous devons charger explicitement le module chaque fois que nous redémarrons notre ordinateur, en utilisant la commande de l'étape E (sudo modprobe iwlmwifi connector_log=0x1).

Mais dans notre cas le problème est que l'erreur que nous voyons est *Network is down (-100)* et cela apparaît parce que l'interface wlan0 n'a pas été activée, nous devons le faire manuellement en utilisant les utilitaires de ligne de commande pour configurer l'interface sans fil par l'exécution de cette ligne:

```
sudo ip link set wlan0 up
```

Cela nous a aidés à nous connecter au Wi-Fi en exécutant la ligne suivante :

```
sudo iw wlan0 connect -w nom-de-notre-Wi-Fi
```

Mais maintenant, le problème est si nous regardons du côté du routeur, les statistiques du routeur montrent que l'adresse IP 0.0.0.0 est attribuée à notre appareil. Si nous nous connectons à un autre appareil, le routeur attribue correctement l'adresse IP. Pour cela nous avons dû faire la configuration de l'adresse IP manuellement en changeant le fichier qui se trouve dans ce dossier '/etc/network/interfaces' en ajoutant une adresse IP statique.

Ça se peut aussi qu'on obtienne aussi l'erreur suivante *Network is unreachable* qu'on peut corriger avec cette commande :

```
sudo dhclient wlan0
```

Maintenant la figure 3-2 montre une partie du script pour commencer à enregistrer les données collectées par la carte NIC

```
lotfi@lotfiben: ~
lotfi@lotfiben:~$ sudo modprobe iwlmwifi connector_log=0x1
lotfi@lotfiben:~$ sudo ifconfig wlan0 up
lotfi@lotfiben:~$ sudo iw wlan0 connect -w lotfi
wlan0 (phy #0): connected to f8:1a:67:2e:e5:07
lotfi@lotfiben:~$ sudo linux-80211n-csitool-supplementary/netlink/log_to_file cs
otfiiii.dat
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
wrote 393 bytes [msgcnt=0]
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 213 bytes: id: 26 val: 1 seq: 0 clen: 213
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
received 393 bytes: id: 26 val: 1 seq: 0 clen: 393
```

Figure 3-2 Script pour le début d'enregistrement du CSI

En même temps dans un autre terminal on va commencer à envoyer des pings au point d'accès en utilisant la commande suivante

```
Ping 192.168.0.1
```

```

lotfi@lotfiben: ~
64 bytes from 192.168.0.1: icmp_seq=5 ttl=64 time=1.29 ms
64 bytes from 192.168.0.1: icmp_seq=6 ttl=64 time=1.14 ms
64 bytes from 192.168.0.1: icmp_seq=7 ttl=64 time=1.85 ms
64 bytes from 192.168.0.1: icmp_seq=8 ttl=64 time=1.29 ms
64 bytes from 192.168.0.1: icmp_seq=9 ttl=64 time=1.29 ms
^C
--- 192.168.0.1 ping statistics ---
9 packets transmitted, 9 received, 0% packet loss, time 8011ms
rtt min/avg/max/mdev = 1.145/2.706/8.690/2.418 ms
lotfi@lotfiben:~$ ping 192.168.0.1
PING 192.168.0.1 (192.168.0.1) 56(84) bytes of data.
64 bytes from 192.168.0.1: icmp_seq=1 ttl=64 time=1.34 ms
64 bytes from 192.168.0.1: icmp_seq=2 ttl=64 time=2.29 ms
64 bytes from 192.168.0.1: icmp_seq=3 ttl=64 time=1.28 ms
64 bytes from 192.168.0.1: icmp_seq=4 ttl=64 time=1.32 ms
64 bytes from 192.168.0.1: icmp_seq=5 ttl=64 time=1.30 ms
64 bytes from 192.168.0.1: icmp_seq=6 ttl=64 time=1.30 ms
64 bytes from 192.168.0.1: icmp_seq=7 ttl=64 time=1.29 ms
64 bytes from 192.168.0.1: icmp_seq=8 ttl=64 time=1.29 ms
64 bytes from 192.168.0.1: icmp_seq=9 ttl=64 time=1.28 ms
64 bytes from 192.168.0.1: icmp_seq=10 ttl=64 time=2.18 ms
64 bytes from 192.168.0.1: icmp_seq=11 ttl=64 time=3.64 ms
64 bytes from 192.168.0.1: icmp_seq=12 ttl=64 time=1.29 ms

```

Figure 3-3 Scripts pour envoyer des pings entre le Tx et Rx

3.4 Traitement de ses données sur Matlab

Pour le traitement de ses données nous allons procéder avec un ordinateur portable Dell Inspiron 15 7000 Gaming, ses spécifications les plus intéressantes sont:

- Intel(R) Core(TM) i7-7700HQ CPU @2.80GHz
- NVIDIA GeForce GTX 1050 Ti 4 Go de vRAM
- RAM DE 16 Go

Grâce à Matlab 2020a nous allons pouvoir procéder au traitement de ses données que nous avons collectées, maintenant pour lire le fichier de trace CSI nous devons avoir le matériel supplémentaire de l'outil CSI qui se trouve sur le GitHub [69], et on exécute cette commande pour notre fichier :

```
>> csi_trace = read_bf_file('sample_data/log.all_csiwifilssiA1.dat');
```

Name	Value
csi_trace	1971x1 cell

Figure 3-4 Données collectées au point 1 au LSSI

Dans notre exemple de fichier, `csi_trace` est un tableau de cellules 1x1971, qui contient 1971 structures. Celui-ci contient les informations CSI pour 1971 paquets reçus. Inspectons l'une des entrées:

```
>> csi_entry = csi_trace{1}

csi_entry =

    struct with fields:

        timestamp_low: 1.4924e+09
        bfee_count: 2416
            Nrx: 3
            Ntx: 2
        rssi_a: 38
        rssi_b: 40
        rssi_c: 32
        noise: -78
        agc: 36
        perm: [2 1 3]
        rate: 2316
        csi: [2x3x30 double]
```

Figure 3-5 Les informations collectées de la première cellule

On va expliquer chaque ligne [69] :

- **timestamp_low** : L'heure d'arrivée du paquet Wi-Fi, qui est représentée par les 32 bits inférieurs de l'horloge de la carte réseau. Cet horodatage représente également l'heure d'arrivée des valeurs RSSI et CSI comprises dans le paquet Wi-Fi.

- **bfee_count** : est simplement un décompte du nombre total de mesures de formation de faisceaux qui ont été enregistrées par le pilote et envoyées à l'espace utilisateur.
- **Nrx** : représente le nombre d'antennes utilisées pour recevoir le paquet par ce NIC, et sa valeur est fixée à 3.
- **Ntx** : représente le nombre d'antennes utilisées du côté de l'émetteur (c'est-à-dire le point d'accès) et sa valeur est définie sur 2.
- **perm** nous indique comment le NIC permutait les signaux des 3 antennes de réception dans les 3 chaînes RF qui traitent les mesures. La valeur d'échantillon de [2 1 3] implique que l'antenne B a été envoyée à la chaîne RF A, l'antenne A à la chaîne B et l'antenne C à la chaîne C. Cette opération est effectuée par un module de sélection d'antennes dans le NIC et correspond généralement à la commande des antennes par ordre décroissant de RSSI.
- **rate** est le débit auquel le paquet a été envoyé. Il faut savoir que les bits d'antenne sont omis, car il n'y a aucun moyen pour le récepteur de savoir quelles antennes d'émission ont été utilisées.
- **noise** : Le bruit mesuré sur le canal.
- **agc** : Représente le paramètre de contrôle automatique du gain du NIC mesuré en dB. La valeur de ce champ ainsi que la valeur du champ de bruit sont nécessaires pour convertir l'unité des valeurs RSSI de dB en dBm.
- **RSSI_a** : RSSI_a représente la valeur RSSI reçue à la première antenne du NIC mesurée en dB.

- **RSSI_b** : RSSI_b représente la valeur RSSI reçue à la deuxième antenne du NIC mesurée en dB.
- **RSSI_c** : RSSI_c représente la valeur RSSI reçue au niveau de la troisième antenne du NIC mesurée en dB.
- **csi** est le CSI lui-même, normalisé à une référence interne. Il s'agit d'une matrice 3-D $N_{tx} \times N_{rx} \times 30$ où la troisième dimension est à travers 30 sous-porteuses dans le canal OFDM. Pour un canal de 20 MHz de large, celles-ci correspondent à environ la moitié des sous-porteuses OFDM, et pour un canal de 40 MHz, c'est environ une sur 4 sous-porteuses. Les sous-porteuses mesurées sont définies par la norme IEEE 802.11n-2009 (dans le tableau 7-25f à la page 50).

Maintenant que nous avons décrit toutes les lignes de cette structure, il va falloir que nous les mettions toutes ensemble pour calculer le CSI en unités absolues, plutôt que le niveau de référence interne d'Intel. En particulier, la combinaison des valeurs RSSI et AGC va nous permettre d'obtenir le RSS en dBm, et inclusion du bruit pour obtenir le SNR.

La matrice CSI rapportée par Intel ne répond pas à la norme. Il est rapporté sans tenir compte de l'amplification. Cependant, on s'attend à ce que le CSI soit en dBm, et pour obtenir le dB par rapport au bruit, on doit ajouter le RSSI et l'amplification et le bruit. Grâce à ce script *getscaled_csi* [69], nous avons pu procéder à cette normalisation. Les calculs ne sont que des expressions relativement logiques de ce qui précède:

1. Les entrées de la matrice CSI sont en amplitude.
2. L'amplitude au carré est la puissance; la moyenne est la puissance moyenne; la racine carrée de la moyenne la remet en unités d'amplitude.
3. RSSI - AGC - autres constantes donne les unités absolues de la matrice CSI en dB.

4. À la fin il ajoute du bruit de fond.

Donc, pour résumer le script, la matrice CSI est divisée par la racine carrée de sa puissance et multipliée par la racine carrée du RSS ($\text{total_RSSI} - (\text{AGC} + \text{Magic Number})$) exprimé en mW), les résultats seront le CSI matrice en unité d'amplitude tenant compte de la puissance reçue, et après la normalisation, le CSI de chaque canal est exprimé en dB par rapport au bruit.

La trace CSI est une cellule qui contient beaucoup d'informations. Nous n'avons besoin que des informations CSI. Par conséquent, nous devons utiliser la fonction *get_scaled_csi* pour extraire les informations CSI dans une nouvelle matrice afin de pouvoir traiter les données CSI. La fonction *get_scaled_csi* est également incluse dans le supplément linux-80211n-csitool. Les parties de codes utilisant cette fonction sont illustrées à la figure 3-6.

```

1  clear all
2  clc
3
4  csi_trace1 = read_bf_file('csiwifi1ssiA1.dat');
5  |
6  for i=1:1952
7  | csi_entry(i) = csi_trace1{i};
8  | csi1(:, :, :, i) = get_scaled_csi(csi_entry(i));
9  | end
10 LSSIA1= squeeze(num2cell(permute(csi1, [1,2,3,4]),1:3)
11 LSSIA1= cell2mat(LSSIA1);
12 LSSIA1(:, :, :) = LSSIA1(1:3000, :, :);
13 LSSIA1=permute(LSSIA1, [3 2 1]);
14 save('csilssiAA1.mat', 'LSSIA1', 'LSSIA1')
--

```

Figure 3-6 Scripts pour la normalisation du point A au LSSI

Comme nous pouvons voir ici nous avons le CSI qui n'est pas normalisé

```

csi_trace1(1, 1).csi
-----
val(:,:,1) =
    28.0000 +17.0000i    1.0000 - 3.0000i    4.0000 -10.0000i
    1.0000 - 1.0000i    0.0000 -13.0000i    -3.0000 + 3.0000i

val(:,:,2) =
    23.0000 -39.0000i    -4.0000 - 6.0000i    -14.0000 - 4.0000i
    0.0000 + 3.0000i    -16.0000 - 4.0000i    -4.0000 + 2.0000i

val(:,:,3) =
   -40.0000 -25.0000i   -12.0000 + 8.0000i    -4.0000 +16.0000i
    4.0000 + 0.0000i   -10.0000 +13.0000i    4.0000 + 4.0000i

val(:,:,4) =
   -24.0000 +35.0000i    10.0000 +16.0000i    18.0000 + 3.0000i
    3.0000 - 3.0000i    10.0000 +19.0000i    6.0000 - 2.0000i

```

Figure 3-7 Le CSI obtenu avant la normalisation

Après la normalisation nous obtenons le nouveau CSI

```

2x3x30 complex double
-----
val(:,:,1) =
    15.8060 + 9.5965i    0.5645 - 1.6935i    2.2580 - 5.6450i
    0.5645 - 0.5645i    0.0000 - 7.3385i    -1.6935 + 1.6935i

val(:,:,2) =
    12.9835 -22.0156i   -2.2580 - 3.3870i   -7.9030 - 2.2580i
    0.0000 + 1.6935i   -9.0320 - 2.2580i   -2.2580 + 1.1290i

val(:,:,3) =
   -22.5801 -14.1125i   -6.7740 + 4.5160i   -2.2580 + 9.0320i
    2.2580 + 0.0000i   -5.6450 + 7.3385i    2.2580 + 2.2580i

val(:,:,4) =
   -13.5480 +19.7576i    5.6450 + 9.0320i   10.1610 + 1.6935i
    1.6935 - 1.6935i    5.6450 +10.7255i    3.3870 - 1.1290i

```

Figure 3-8 Le CSI obtenu normalisé

Lors de la collecte des données Ntx des fois change (de 2 à 1, généralement 2, rarement 1). Après une recherche approfondie sur le forum du GitHub, nous avons trouvé que le pilote Intel n'est pas responsable du changement de Ntx. D'après les recherches effectuées, nous avons trouvé que la faible diversité spatiale entre les antennes modifie la taille de la matrice CSI renvoyée par le pilote. Par exemple, si nous sommes dans une condition pour laquelle $N_{tx} = 1$ et que nous essayons d'augmenter la distance entre les antennes N_{rx} , N_{tx} peut passer de 1 à 2, en fonction de la distance. Comme nous pouvons le voir sur la figure suivante un exemple où le N_{tx} est de 1.

Field	Value
timestamp_low	1.7088e+09
bfee_count	2839
Nrx	3
Ntx	1
rss_i_a	38
rss_i_b	39
rss_i_c	29
noise	-82
agc	37
perm	[2,1,3]
rate	2308
csi	1x3x30 complex double

Figure 3-9 Exemple où le $N_{tx} = 1$

Sinon la majorité du temps de notre N_{tx} est de 2 comme le montre la figure suivante

csi_trace1{2, 1}	
Field ^	Value
timestamp_low	1.4924e+09
bfee_count	2417
Nrx	3
Ntx	2
rss_i_a	38
rss_i_b	40
rss_i_c	32
noise	-78
agc	36
perm	[2,1,3]
rate	2315
csi	2x3x30 complex double

Figure 3-11 Exemple où le Ntx =2

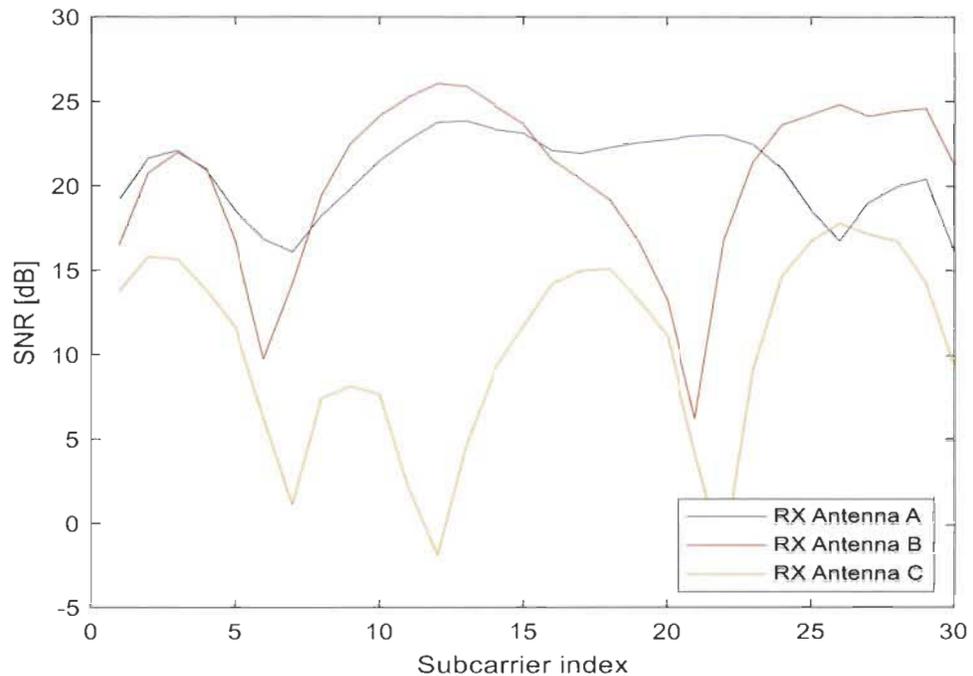
Regardons les trois chemins spatiaux différents sur le lien 1×3 que nous avons mesuré:

```
>> csi_entry = csi_trace1{424};
>> csi = get_scaled_csi(csi_entry);
>> plot(db(abs(squeeze(csi).')))
legend('RX Antenna A', 'RX Antenna B', 'RX Antenna C', 'Location', 'SouthEast' );
xlabel('Subcarrier index');
ylabel('SNR [dB]');
>> db(get_eff_SNRs(csi), 'pow')
```

Figure 3-10 Scripts pour analyser les trois chemins spatiaux différents

Dans la commande plot, squeeze() transforme CSI en une matrice 3×30 en supprimant la première dimension singleton. dB() convertit de l'espace linéaire (tension) en espace logarithmique (base 10, puissance), abs() convertit chaque nombre complexe en sa grandeur. Finalement le .' est l'opérateur transpose qui transforme le CSI compressé de la matrice $3 \times$

30 en une matrice 30×3 . Combinés, nous obtenons la figure 3-12. Un exemple de collecte CSI est illustré à la figure 3-12, utilisant 3 antennes de réception, elle décrit l'amplitude d'un paquet collecté pour un échantillon de 30 sous-porteuses.



Nous terminerons par calculer le SNR effectif à partir de nos matrices CSI. Pour ce faire, nous utilisons le script *get_eff_SNRs* [69], qui prend en entrée une matrice CSI et renvoie une matrice 7×4 de valeurs SNR effectives dans l'espace linéaire (puissance). Les 4 colonnes correspondent au SNR effectif en utilisant les quatre schémas de modulation 802.11, à savoir BPSK / QPSK / 16QAM / 64QAM.

Les 7 lignes correspondent aux sept sélections d'antennes possibles avec 3 antennes et 1, 2 ou 3 flux spatiaux. En particulier, les 3 premières lignes correspondent à des émissions mono-flux avec l'antenne A, B ou C. Les 3 lignes suivantes correspondent à des émissions double flux avec les antennes AB, AC ou BC. La dernière ligne correspond à une transmission à 3 flux utilisant toutes les antennes.

```

-> db(get_eff_SNRs(csi), 'pow')

ans =

    18.1916    18.4004    19.5788    21.9611
   -156.5356  -156.5356  -156.5356  -156.5356
   -156.5356  -156.5356  -156.5356  -156.5356
   -156.5356  -156.5356  -156.5356  -156.5356
   -156.5356  -156.5356  -156.5356  -156.5356
   -156.5356  -156.5356  -156.5356  -156.5356
   -156.5356  -156.5356  -156.5356  -156.5356

```

Figure 3-13 Calculs des valeurs SNR effectives

3.5 Modèle de localisation en intérieur basé sur l'apprentissage automatique

Ces dernières années, avec l'essor de l'intelligence artificielle, du *cloud computing* et des technologies de big data, les technologies de navigation et de localisation autonomes sont rapidement mises à jour. En particulier, le développement de ML a apporté de nouvelles solutions à la localisation basée sur les signaux RF.

3.5.1 K- Voisin le plus proche (kNN)

K-plus proche voisin est une méthode simple qui tente d'effectuer une classification en calculant la distance entre les caractéristiques. L'algorithme kNN considère k points d'étalonnage. La sélection de ces points basée sur la sélection des points k les plus proches dans l'espace des fonctionnalités pour approximer la position de l'utilisateur.

3.5.2 Machine à vecteurs de support (SVM)

Les machines vectorielles de support [72] sont une technique puissante utilisée pour la classification et la régression des données. Ils sont utilisés comme classificateur non paramétrique superviser pour résoudre les problèmes de reconnaissance de formes. Les SVM

sont utilisés dans le système de localisation en entraînant des vecteurs de support sur une carte radio constituée de points de maillage. Les SVM analysent la relation entre les empreintes digitales entraînées et leurs points de grille en traitant chaque point de grille comme une classe. Les empreintes digitales RSSI testées sont prises en entrée dans le SVM qui prédit à quelle classe le test appartient.

3.5.3 Réseau récurrent à mémoire court et long terme LSTM

LSTM [73] est un réseau récurrent à mémoire court et long terme et un réseau de neurones récurrent dans le temps qui a été utilisé dans de nombreux domaines de la technologie. LSTM est une excellente variante de RNN qui hérite des caractéristiques de la plupart des modèles RNN et résout le problème du gradient de fuite causée par la réduction progressive du processus de rétropropagation du gradient. Dans le traitement de tâches spécifiques, LSTM est bien adapté pour traiter les problèmes qui sont fortement corrélés avec les séries temporelles, telles que la traduction automatique, la génération de dialogues, l'encodage / décodage, etc. Par conséquent, nous choisissons LSTM comme modèle d'apprentissage en profondeur pour notre application de localisation par Wi-Fi.

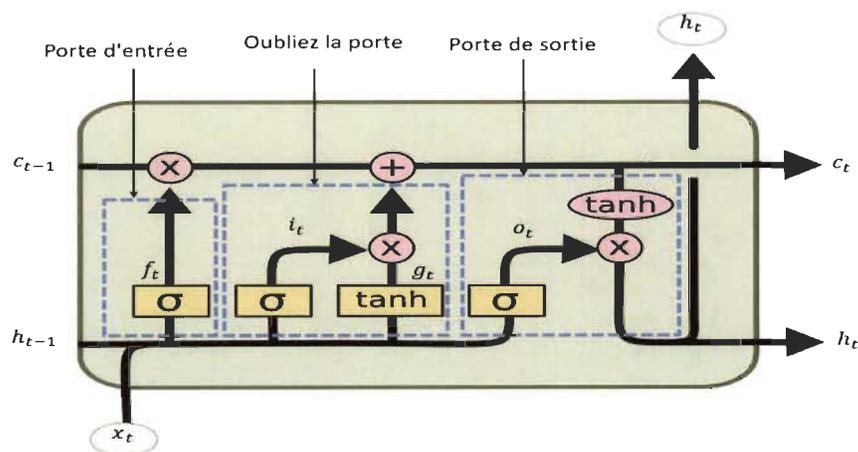


Figure 3-14 L'unité de base du LSTM

L'unité de base du modèle LSTM est représentée sur la figure 3-14, qui comprend la porte d'oubli, la porte d'entrée et la porte de sortie. Dans la porte oubliée, la partie oubliée de l'état de cellule c_t est apportée par l'entrée x_t , l'état de cellule c_{t-1} et l'état caché h_{t-1} . Dans la porte d'entrée, le vecteur resté de l'état de cellule c_t est contribué par la valeur que la sortie de x_t par la fonction sigmoïde et tanh. L'état caché h_t est contribué par la mise à jour c_t et de la sortie ensemble.

3.5.4 Arbres de décision (DT)

L'arbre de décision (DT) est une méthode d'apprentissage supervisé non paramétrique pour la régression et la classification. L'algorithme DT est basé sur un arbre de décision binaire construit à partir de l'ensemble de données d'apprentissage. L'algorithme commence par apprendre des règles de décision simples dérivées des caractéristiques des données. Il est construit en trois nœuds: nœud racine, nœud interne et nœud terminal. Le nœud racine n'a pas de bords entrants, zéro ou plusieurs bords sortants. L'intérieur a exactement un bord entrant et deux bords sortants ou plus. Le terminal a des fronts d'entrée et aucun front de sortie. Chaque nœud terminal de l'arbre de décision se voit attribuer une étiquette de catégorie. Chacun de ces nœuds correspond à une décision prise sur l'un des paramètres d'entrée. Le nœud est ensuite divisé en nouveaux sous-ensembles, un pour chacun des sous-arbres du nœud, de telle sorte que le même emplacement cible se trouve dans les mêmes sous-ensembles [74].

L'algorithme s'arrête lorsqu'il y a une décision pure et que l'incertitude est inefficace. Une décision pure signifie que le sous-ensemble de données de chaque nœud contient un et un seul emplacement cible.

3.6 Conclusion

Dans ce chapitre nous avons présenté un système de mesure et de localisation basé sur un réseau Wi-Fi que nous avons mis au point, qui nous a permis de faire une collecte de données et d'acquérir de différentes informations telles que CSI, RSS. Nous avons aussi expliqué la procédure comment exploiter ses données sur Matlab. À la fin nous avons vu quelques méthodes de ML utilisées dans la littérature pour la localisation en intérieur que nous allons appliquer dans le prochain chapitre.

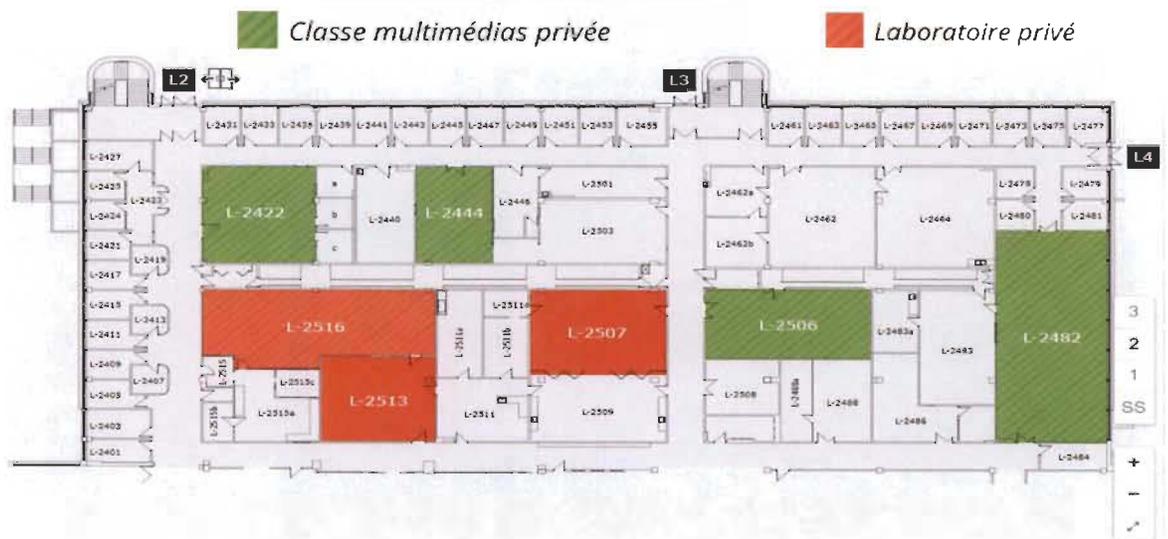
Chapitre 4 - Résultats et discussion

4.1 Introduction

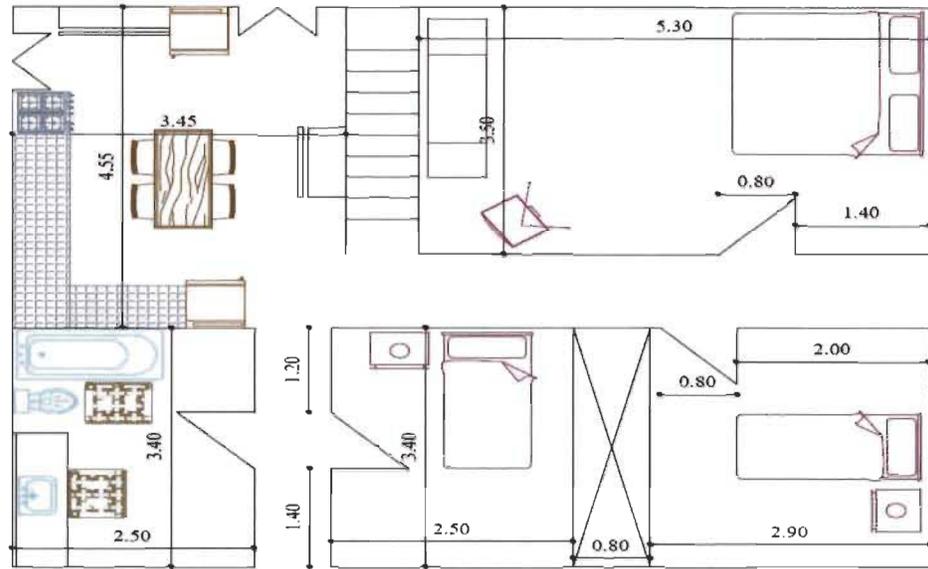
Le but de ce chapitre est d'évaluer les différentes techniques d'apprentissage automatique des données collectées par notre setup de Wi-Fi. La collecte de ses données va se faire sur deux différents scénarios. Une étude comparative de performance de différentes techniques qui sont LSTM, SVM quadratique, kNN pondéré, arbre de décision, classificateur d'ensemble et SVM optimisable pour voir quelle méthode nous obtenons une meilleure précision avec nos données collectées.

4.2 Zone d'essai

Dans cette section, nous donnons tous d'abord un aperçu d'architecture du système qui collecte les données. L'emplacement choisi pour les expériences est reparti sur deux emplacements, la première partie se fait au LSSI à l'Université du Québec à Trois-Rivières (Figure 4-1) et la deuxième partie c'est dans un milieu résidentiel (Maison).



a)



b)

Figure 4-1 a) Deuxième étage du bâtiment L-Léon-Provancher (LP) et b) maison

L'expérience a été menée pendant le soir lors de la période de la pandémie du Covid-19, le trafic humain dans cette zone n'était donc pas très dense.

La prise d'empreinte digitale est une technique populaire de localisation en intérieur par radiofréquence. Il faut souvent une étude environnementale pour obtenir des empreintes de l'environnement où le système de localisation va être utilisé. Nous allons effectuer cette technique en deux phases : apprentissage hors ligne et la validation en ligne. Au cours de la phase hors ligne, une étude de l'environnement est effectuée pour collecter RSS et CSI à partir de point d'accès à différent point de référence, pour former notre LSTM. Ainsi chaque point de référence est représenté par son empreinte digitale. Toutes ses empreintes vont former des vecteurs de signaux que nous allons stocker dans une base de données. Après dans la phase en ligne, les mesures de validation correspondant aux empreintes stockées pour estimer l'emplacement de l'utilisateur à l'aide des réseaux de neurone

4.3 Mesure de RSS et CSI signal

Les figures 4-2 et 4-3 montrent les formes d'onde de puissance de signal d'échantillons RSS et d'échantillons CSI, respectivement. Les différentes couleurs de la figure indiquent les signaux mesurés par différentes antennes de réception. Les mesures ont été effectuées sur quatre points différents à l'université (où chaque point représente l'emplacement du transmetteur). Le CSI et RSS change à chaque fois qu'on change la place du récepteur. Cependant, nous pouvons voir sur les figures ci-dessous que les changements produits par RSS sont relativement faibles (il varie entre 29 à 40 dBm), ce qui signifie que RSS est moins sensible aux changements environnementaux par rapport au CSI.

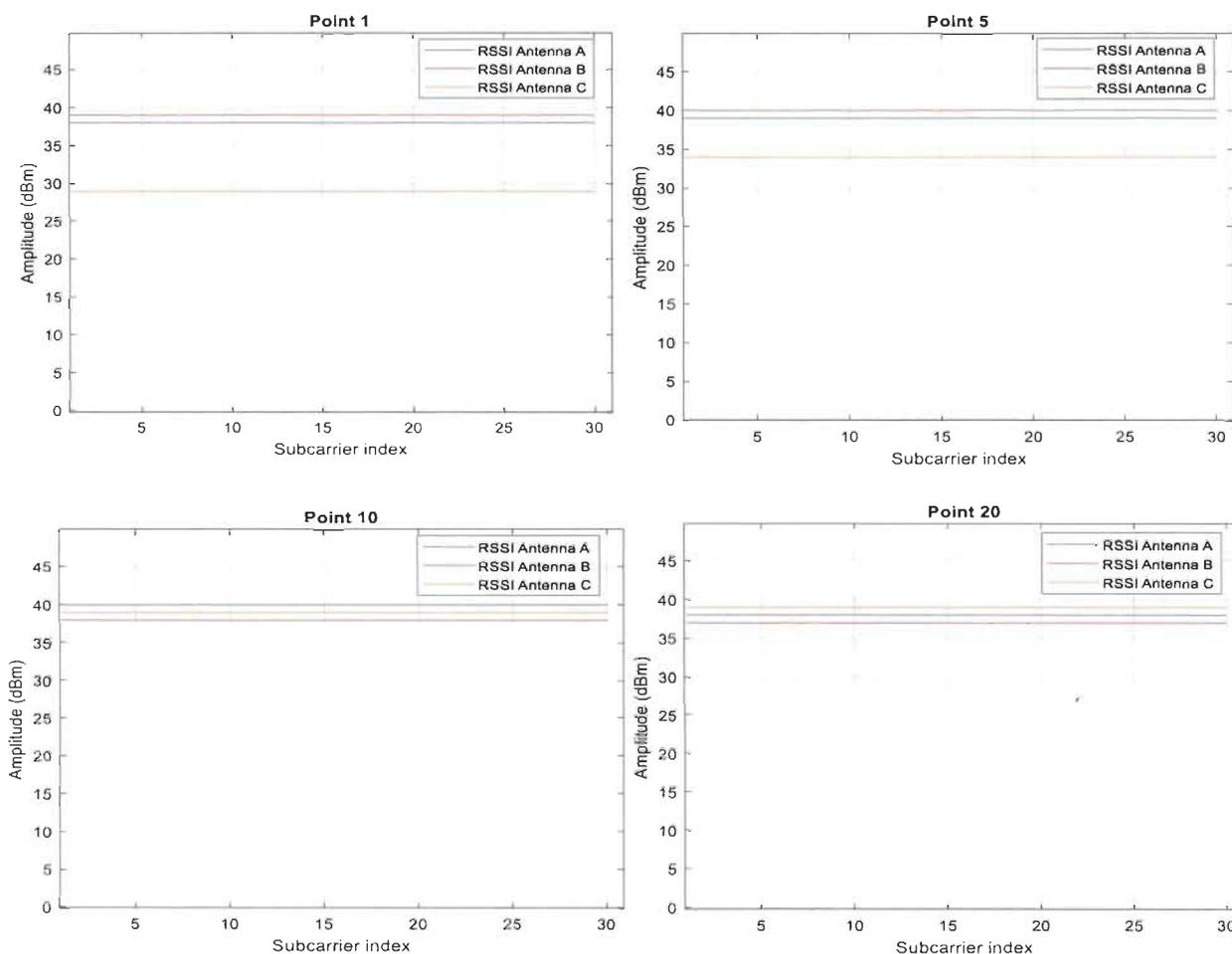


Figure 4-2 Formes d'onde RSS mesurées à différents points de référence.

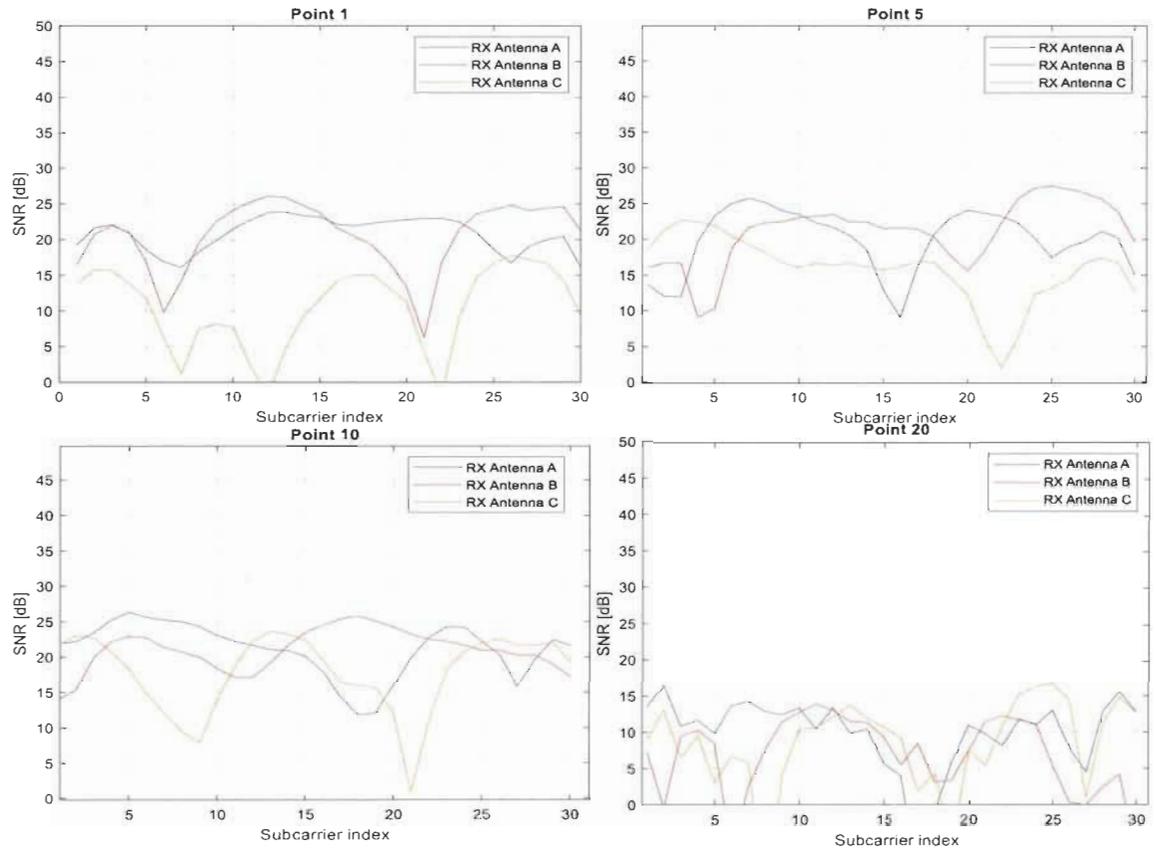


Figure 4-3 Formes d'onde CSI mesurées à différents points de référence.

4.4 Équipements et outils

L'équipement et les outils, y compris le matériel et les logiciels, sélectionnés pour l'enregistrement et la mesure sont présentés comme suit :

4.4.1 Point d'accès

Cette recherche porte sur l'utilisation d'un point d'accès installé dans la zone de test comme transmetteur (avec deux antennes), comme le montre la figure 4-4. La zone de test a une bonne couverture du signal Wi-Fi. Le point d'accès est de type TP-Link Archer C5 V2.0 Routeur Gigabit Wi-Fi Double Bande AC1200 (2.4GHz 300 Mb/s, 5GHz 867Mbps, 5 Ports Gigabit, 2 Ports USB).



Figure 4-4 Photos du modèle du AP

4.4.2 Outil de mesure

L'ordinateur du LSSI a été utilisé comme un récepteur, il est équipé par une carte d'interface Intel NIC5300 (avec trois antennes de réception) comme nous l'avons vu dans la chapitre 3, d'un processeur Intel Xeon(R) CPU E5-2620 v3, une RAM de 32 GB et Nvidia RTX 2080ti. L'architecture du système pour la localisation en intérieur est illustrée de manière conceptuelle sur la figure suivante

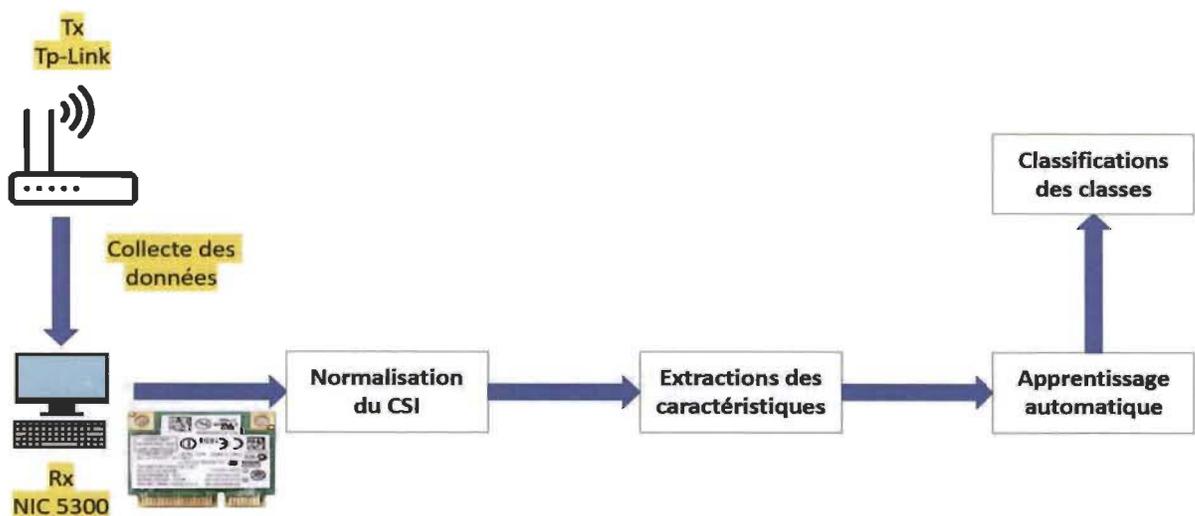


Figure 4-5 Schémas de la collecte des données

4.5 Environnement expérimentale et procédure

Le système de localisation proposé dans notre mémoire collecte des données binaires brutes reçues par le récepteur sans fil et contient diverses informations du processus de transmission, telles que le nombre d'antennes d'émetteur et d'antennes de récepteur, le CSI et le RSS (Figure 4-6).

csi_trace1{1, 1}	
Field ^	Value
timestamp_low	1.4924e+09
bfee_count	2416
Nrx	3
Ntx	2
rss_i_a	38
rss_i_b	40
rss_i_c	32
noise	-78
agc	36
perm	[2,1,3]
rate	2316
csi	<i>2x3x30 complex double</i>

Figure 4-6 Données collectées

Le nombre de CSI dans un paquet est le produit de l'antenne de l'émetteur, de l'antenne du récepteur et du nombre de canaux (sous-porteuses). La carte d'interface réseau Intel 5300 utilisée dans ce travail dispose de 30 canaux. Ainsi pour notre cas, le nombre total de données CSI par paquet est de $2 \times 3 \times 30$.

Dans notre travail, seule la réponse en amplitude est utilisée pour la localisation en intérieur. Pour plus de simplicité, nous concaténons les 180 ($2 \times 3 \times 30$) données dans un vecteur échantillon CSI et notons :

$$CSI = (|csi_{1,1,1}|, |csi_{1,1,2}|, \dots, |csi_{1,1,30}|, \dots, |csi_{1,2,1}|, \dots, |csi_{2,3,30}|, \dots) \quad (4.1)$$

La figure suivante montre un exemple de 2700 paquets collecté dans un point

```

60 - X = []; % Data
61 - Y = []; % Labels
62
63 % Size of dataset to be defined
64 - NumPacketPerClass = 2700; % Number of packets
65
66 - for c = 1:NumClass
67
68     % Training data collection
69 -     DataLabel = Label(c)*ones(1,NumPacketPerClass); % Data label for the current
70 -     [feature1, label, ~] = getFeatureAndLabel(real(Item1), imag(Item1), DataLabel, c)
71 -     Room1A1 = mat2cell(feature1, size(feature1,1), ones(1, size(feature1,2)))';
72 - end

```

Figure 4-7 Scripts utilisés pour la réorganisation des données

D'où nous obtenons les résultats suivants

180x2700 double												
	1	2	3	4	5	6	7	8	9	10	11	12
1	11.3435	8.5076	-12.4192	-10.1612	8.4695	7.3402	6.8078	6.2405	11.2796	6.7678	10.7155	6
2	5.1046	3.9702	5.6451	3.9516	-9.5988	-7.9049	-8.5098	-6.8078	7.8957	5.6398	-8.4596	-3
3	-7.3733	-1.1344	6.2096	0.5645	19.1975	2.8232	17.0195	-1.1346	-3.9479	-1.1280	-1.6919	-0
4	-15.8809	-2.2687	-18.6288	-2.8225	-1.1293	-1.1293	-1.1346	-1.7020	-16.9195	-3.3839	17.4831	3
5	-10.2092	-5.6718	-9.5967	-7.3386	6.2110	5.0817	9.0771	3.9712	3.9479	5.0758	-6.2037	-2
6	-2.2687	-5.1046	7.3386	1.1290	-10.7280	-5.0817	5.1059	3.9712	-9.0237	-5.0758	-6.2037	-6
7	5.1046	5.6718	6.7741	5.0806	-11.2927	-9.0341	-10.7790	-9.6444	9.5877	8.4597	-11.8434	-6
8	-13.6122	-11.3435	14.6772	14.6772	-10.1634	-10.7280	-7.3751	-7.3751	-13.5356	-8.4597	-12.4074	-9
9	-20.4183	-3.9702	-24.2739	-5.0806	-0.5646	-0.5646	-3.4039	-0.5673	-21.4313	-5.0758	22.5589	5
10	9.0748	0.5672	-9.0322	-0.5645	-24.8439	-3.9524	-22.1254	-3.9712	5.6398	1.1280	-0.5640	0
11	-3.4031	-5.6718	8.4676	2.8225	-12.9866	-5.6463	6.2405	5.6732	-11.8436	-6.7678	-7.8956	-6
12	13.0450	7.3733	13.5482	11.2902	-9.0341	-8.4695	-13.0483	-6.8078	-6.2038	-6.7678	10.1515	3
13	-15.3137	-13.0450	16.9353	14.6772	-11.2927	-11.2927	-7.9425	-7.3751	-14.6635	-9.5877	-11.8434	-7
14	-5.6718	-5.1046	-7.3386	-5.6451	11.2927	11.2927	12.4810	10.2117	-9.0237	-7.3318	13.5353	7
15	9.6420	3.4031	-10.7257	0	-25.9731	-3.9524	-22.6927	-2.8366	6.7678	2.8199	-3.3838	-1
16	22.1199	3.4031	24.8384	4.5161	-0.5646	2.2585	5.1059	2.8366	22.5593	4.5119	-23.6868	-5

Figure 4-8 Matrices de données obtenues après la réorganisation

Pour la classification, il faut tout d'abord réorganiser nos données pour chaque classe.

Les données collectées sont organisées dans la cellule X_input que varie de 1 à Xn, chaque

cellule contient une matrice de 180×1 . Prenons exemple pour la première cellule : la première ligne représente la valeur réelle du CSI au premier récepteur à la première antenne, la deuxième ligne représente la valeur imaginaire du CSI au premier récepteur à la première antenne, la ligne neuf représente la valeur réelle du CSI au deuxième récepteur à la deuxième antenne et la ligne dix représente la valeur imaginaire du CSI au deuxième récepteur à la deuxième antenne. Pour l' Y qui varie de 1 jusqu'à Y_m représente une classe pour exemple de Y_1 à Y_{20} c'est la classe A et ainsi de suite.

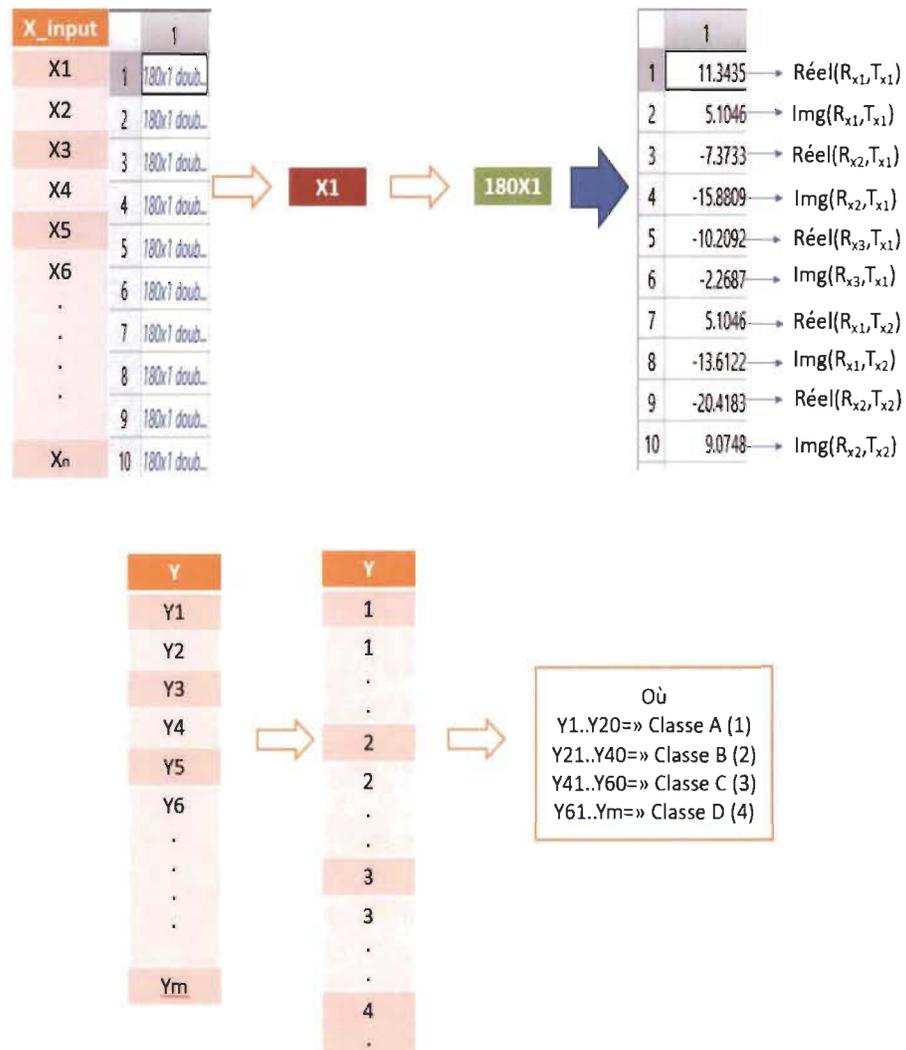


Figure 4-9 Classifications des données pour chaque point de mesure

Pour préparer notre réseau neurone, nous allons séparer ses données en deux catégories qui sont :

- Entraînement (Training)
- Validation

X_Training	X_Validation	Y_Training	Y_Validation
X1	X _{n+1}	Y1	Y _{n+1}
X2	.	Y2	.
.	.	.	.
.	.	.	.
.	.	.	.
X _n	X _m	Y _n	Y _m

Figure 4-10 Schémas de l'organisation de nos données

La figure 4-11 montre une partie du programme où nous allons diviser l'ensemble de données, dans ce cas-là nous utiliserons 80% de ses données pour l'entraînement de notre réseau et 20% pour la validation. Comme nous pouvons voir notre Y représente la classification des classes est un tableau catégoriel et qui est un type de données pour stocker des données avec des valeurs d'un ensemble fini de catégories discrètes.

```

114 % Re-organize the dataset
115 x = X.';
116 Y = Y.';
117 tempX = [];
118 tempY = [];
119 for n = 1:NumPacketPerClass
120     tempX = [tempX;X(n:NumPacketPerClass:end)];
121     tempY = [tempY;Y(n:NumPacketPerClass:end)];
122 end
123
124 % Split the dataset into training set, validation set and testing set
125 TrainSize = 4/5;
126 ValidSize = 1/5;
127
128 NumSample = NumClass*NumPacketPerClass;
129
130 % Training data
131 XTrain = tempX(1:NumSample*TrainSize);
132 YTrain = categorical(tempY(1:NumSample*TrainSize));
133
134 XValid = tempX(NumSample*TrainSize+1:end);
135 YValid = categorical(tempY(NumSample*TrainSize+1:end));

```

Figure 4-11 Scripts pour la réorganisation de nos données d'entraînement et de validation

4.6 Évaluation de différentes méthodes de classification

Dans cette partie, nous mettons en œuvre différentes méthodes ainsi que notre méthode proposée et montrons les performances de chacune. Tout d'abord nous allons proposer deux scénarios différents, dont le premier est celui des mesures effectuées à l'université au LSSI et le second va être celui qui a été appliqué dans une maison où il y avait beaucoup d'obstacles.

4.6.1 Premiers scénarios

Les performances de localisation sont généralement affectées par la complexité de l'environnement intérieur. Afin d'évaluer la robustesse de la méthode proposée, nous avons choisi le LSSI en ajoutant des obstacles pour augmenter la complexité de l'environnement.

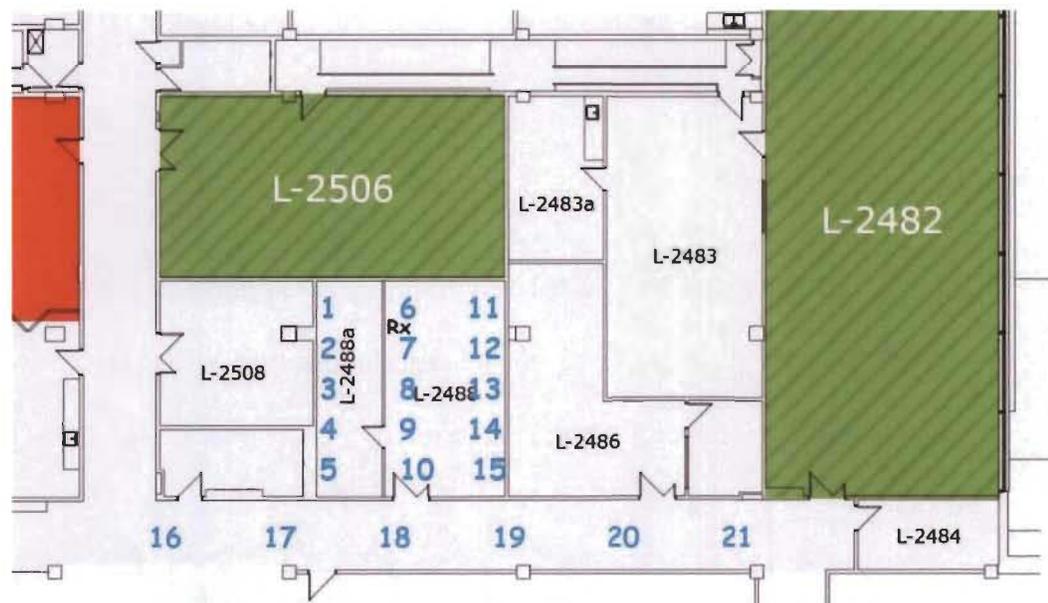


Figure 4-12 Dispositions du pavillon Léon-Provenche au LSSI 2488LP et les placements du point d'accès

Comme le montre la figure 4-12, la pièce est un espace clos de 8,37 mètres de long et 7,44 mètres de large. La zone rectangulaire entre les bureaux de la salle est également divisée

en 3 à 15 points, et nous avons pris aussi 7 autres points dans le corridor en face du LSSI. Le centre de chaque point est le point de référence (ou point de formation), comme indiqué par le cercle bleu sur la figure 4-13.

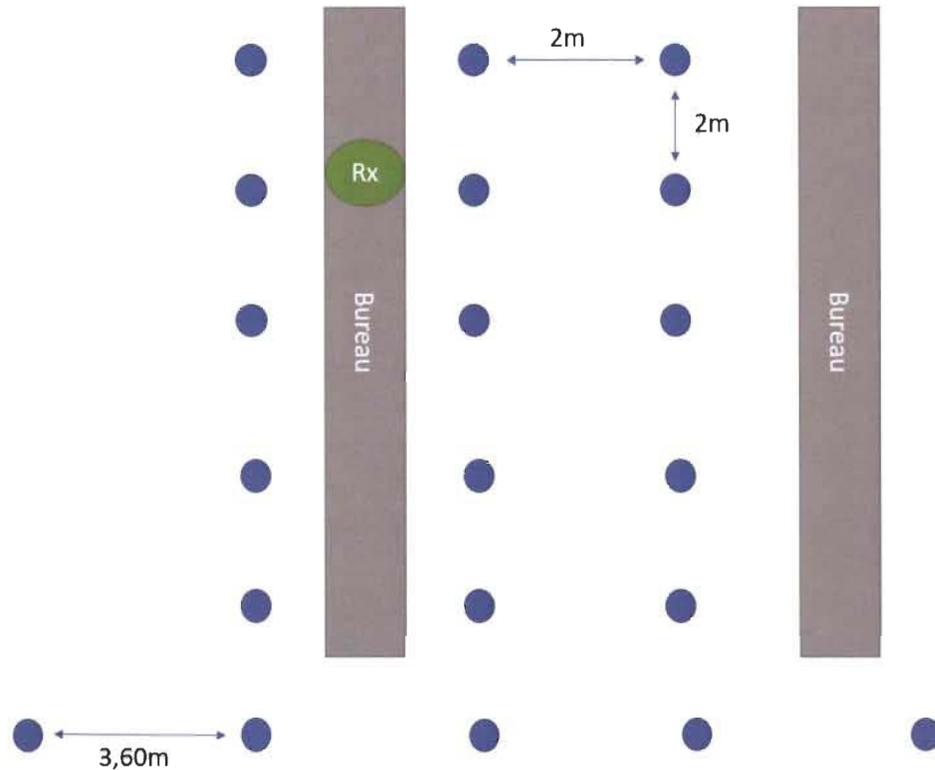


Figure 4-13 Cartes d'environnement expérimental au LSSI (2488LP)

Pour la classification il y a plusieurs techniques de ML et DL, on va débiter par la technique d'un réseau récurrent à mémoire court et long terme (LSTM – Long Short-Term Memory).

L'apprentissage de LSTM est utilisé pour diagnostiquer le comportement du modèle et comme les LSTM sont stochastiques, ce qui signifie que chaque analyse peut donner un tracé de diagnostic différent donc les performances du modèle sont tracées sur plusieurs tests. En examinant ces graphiques, nous décidons des meilleurs hyperparamètres pour obtenir de

meilleures performances. La figure 4.14 montre l'organigramme du fonctionnement de l'algorithme.

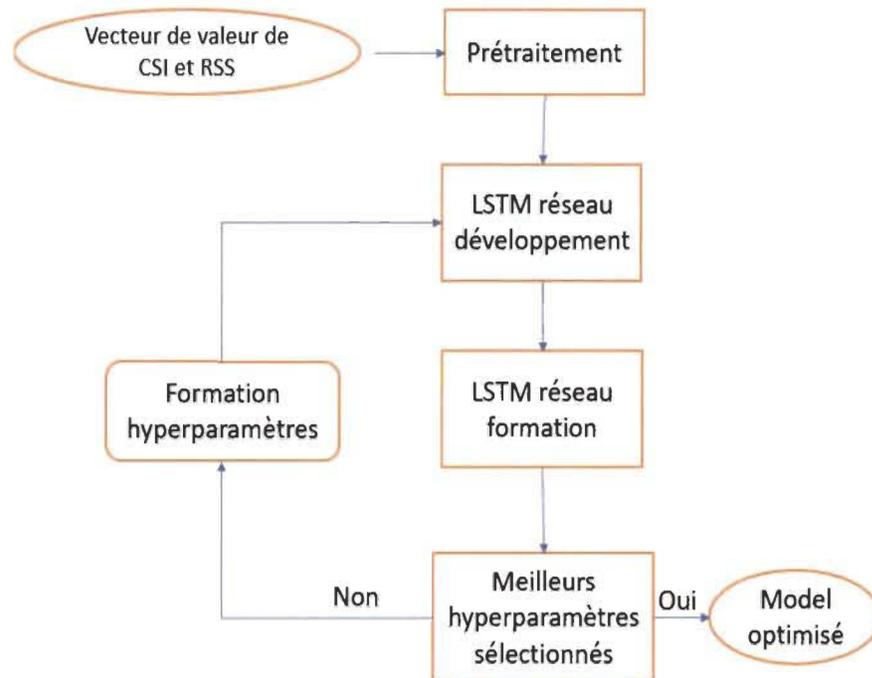


Figure 4-14 Organigrammes du processus de sélection des hyperparamètres

Les meilleurs hyperparamètres du modèle sont obtenus après avoir effectué plusieurs tests comme nous pouvons le voir sur le tableau 4-1

- Une plage de 0,1 à 0,006 est sélectionnée pour trouver le meilleur taux d'apprentissage
- Le nombre d'époques dépend totalement du taux d'apprentissage
- Batch Size détermine le nombre d'échantillons dans chaque itération pour mettre à jour les poids du modèle.
- L'algorithme d'Adam est une belle mise en œuvre de la descente de gradient, car avec cela, un taux d'apprentissage personnalisé est automatiquement déterminé pour chaque paramètre du réseau LSTM.

Tableau 4-1 Ensemble finalisé des meilleurs hyper paramètres pour le scénario 1

Optimizer	Adam
Learning rate	0.007
Input Layers	180
Number of hidden layers	400
Batch Size	2000
Max Epoch	50
L2Regularization	0.0015

Tous les modèles d'apprentissage profond sont stochastiques. Cela signifie que lorsque nous les formons sur un ensemble de données, ils choisissent arbitrairement les poids initiaux et réorganisent également l'ensemble de données de manière aléatoire pour chaque époque d'apprentissage.

Comme nous pouvons voir sur la figure 4-15, notre LSTM s'est bien calibré avec nos données d'où nous avons obtenu une formation de 100% et une validation de 97,90% de précision, nous pouvons implémenter ses résultats et obtenir la matrice de confusion (figure 4-16).

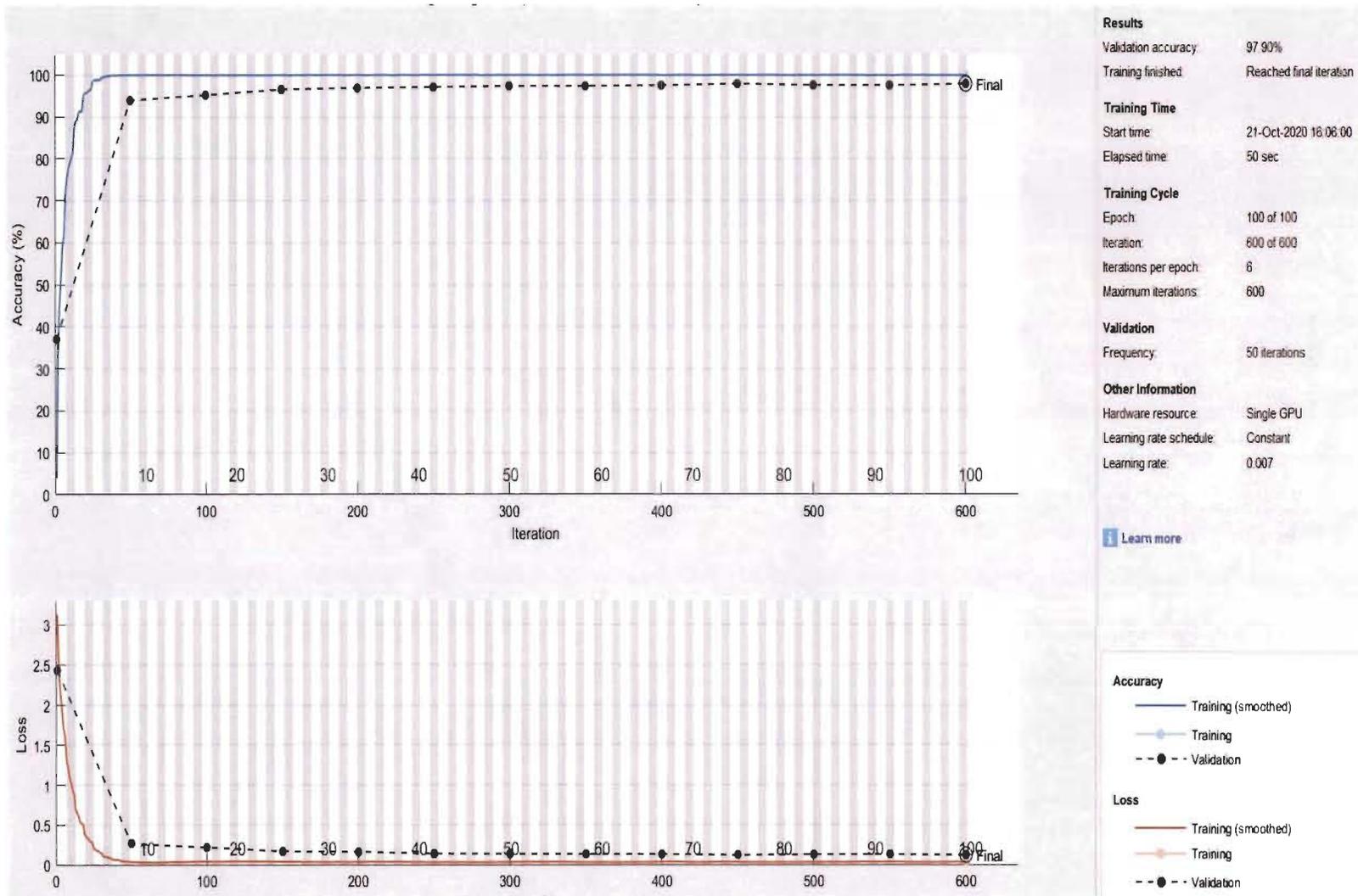
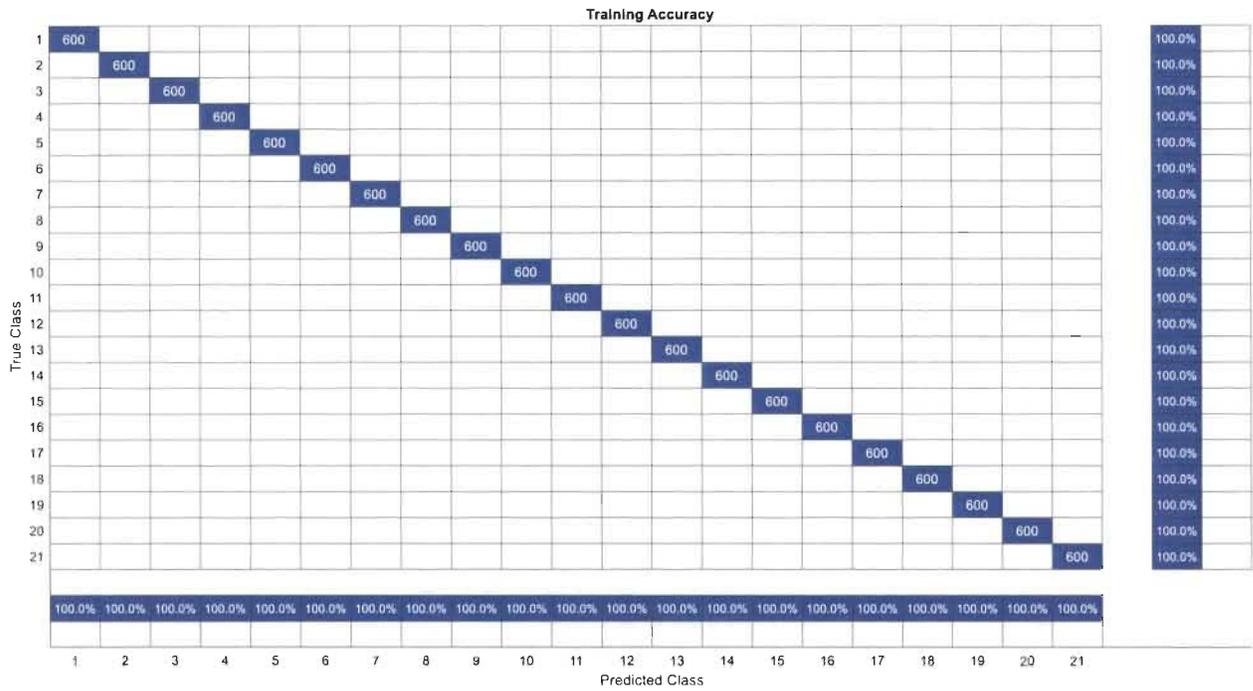


Figure 4-15 Résultats de la simulation au LSSI

(a)



(b)

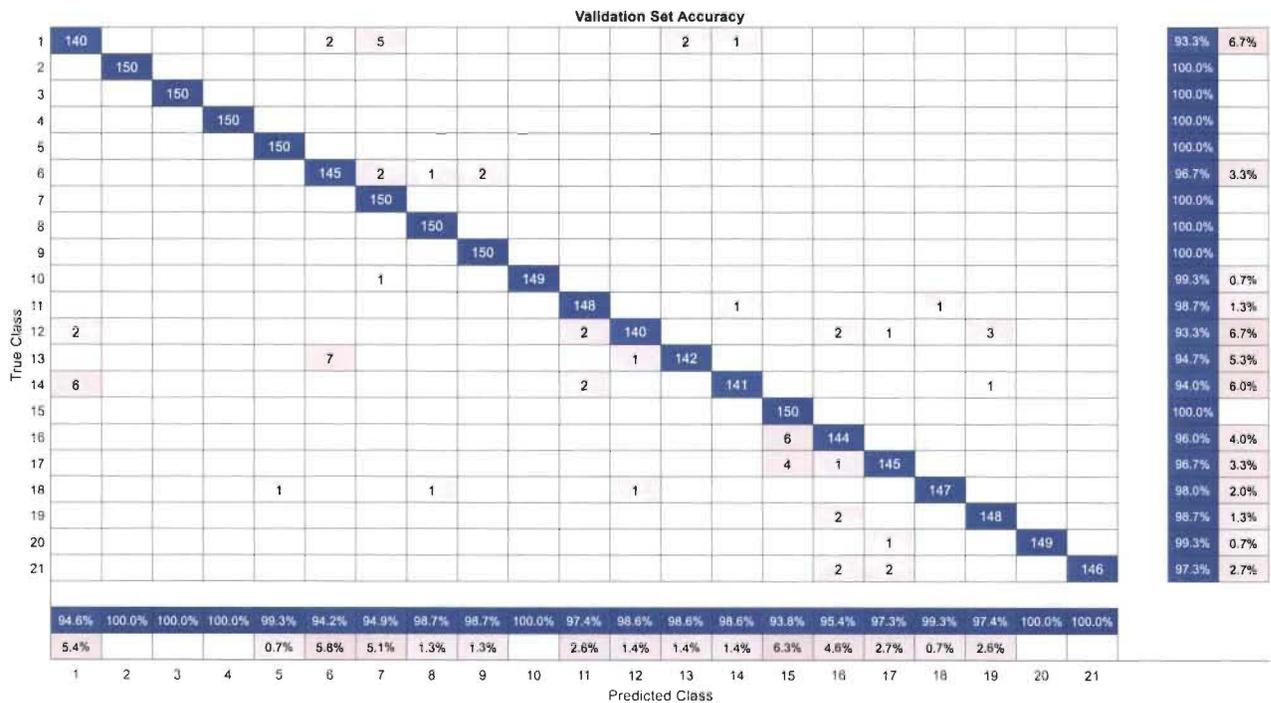
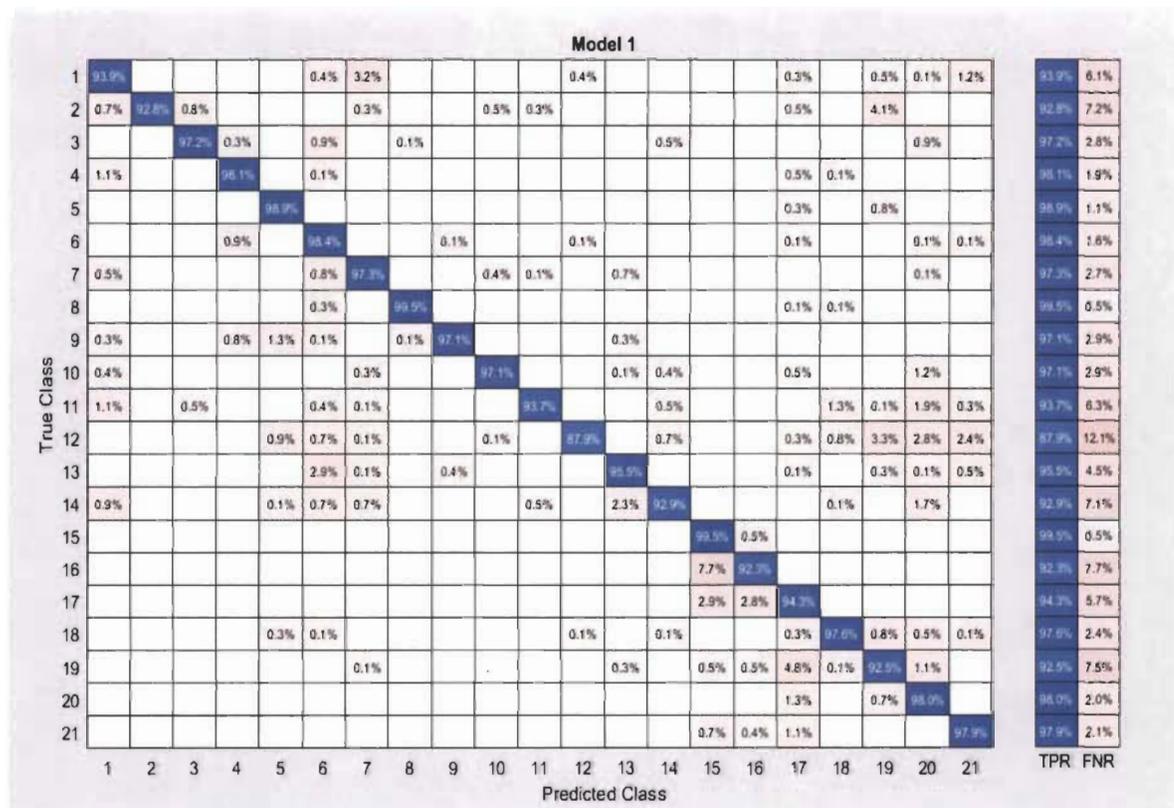
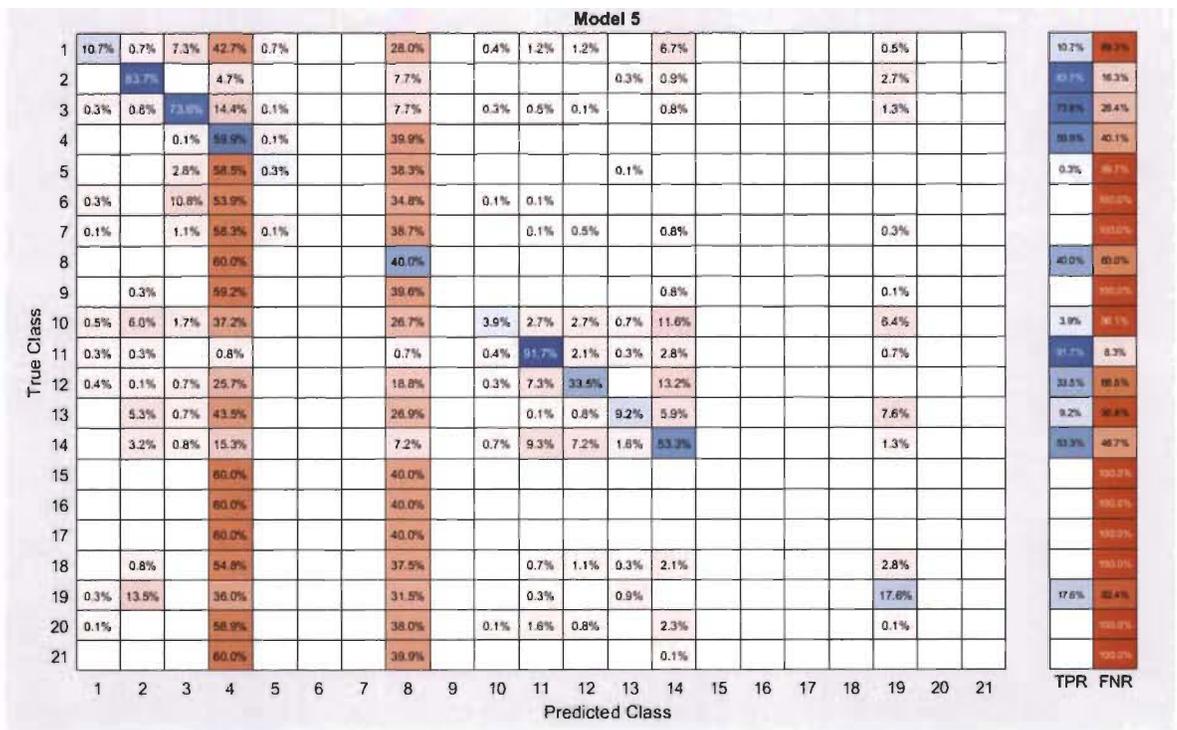
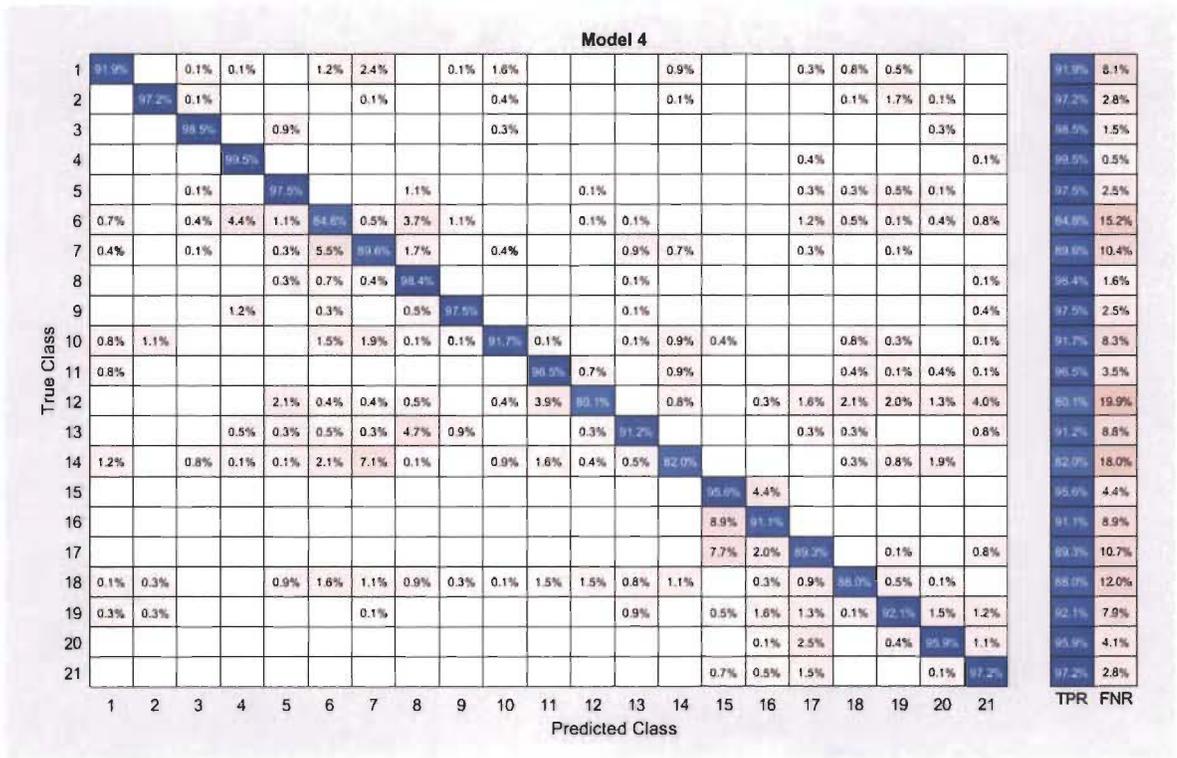


Figure 4-16 a) Matrices de confusion pour l'apprentissage et b) la validation au LSSI

D'après la matrice de confusion lors de l'entraînement de notre LSTM, le réseau s'est bien entraîné avec une précision dans 100% sur tout l'ensemble de données, il s'est entraîné sur 600 points qui représentent 80% de l'ensemble de données. Pour la validation nous pouvons remarquer qu'il arrive à prédire des points à 100% d'où on obtient 150 qui désigne 20% de l'ensemble de données.

Nous avons également mis en œuvre d'autres techniques de ML en utilisant le classificateur de Matlab telle que, l'apprentissage par arbre de décision (Fine Tree), machine à vecteurs de support (SVM), kNN, classificateurs d'ensemble (Ensemble), d'où on va comparer ses résultats par rapport au LSTM. La figure 4-17 présente les matrices de confusion des prédictions sur les données pour les différents modèles de ML. Comme nous pouvons le voir à partir des résultats, la cohérence de toutes les classes n'est pas affectée de la même manière dans le temps et divers profils d'emplacement sont affectés différemment.





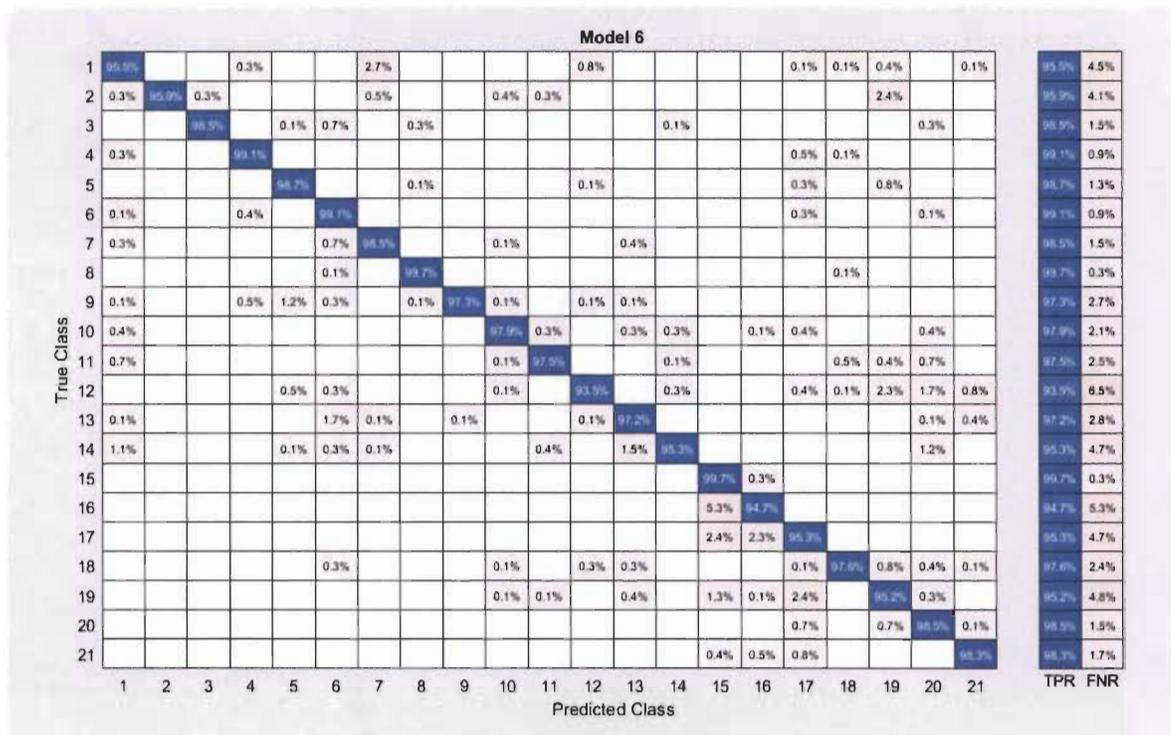


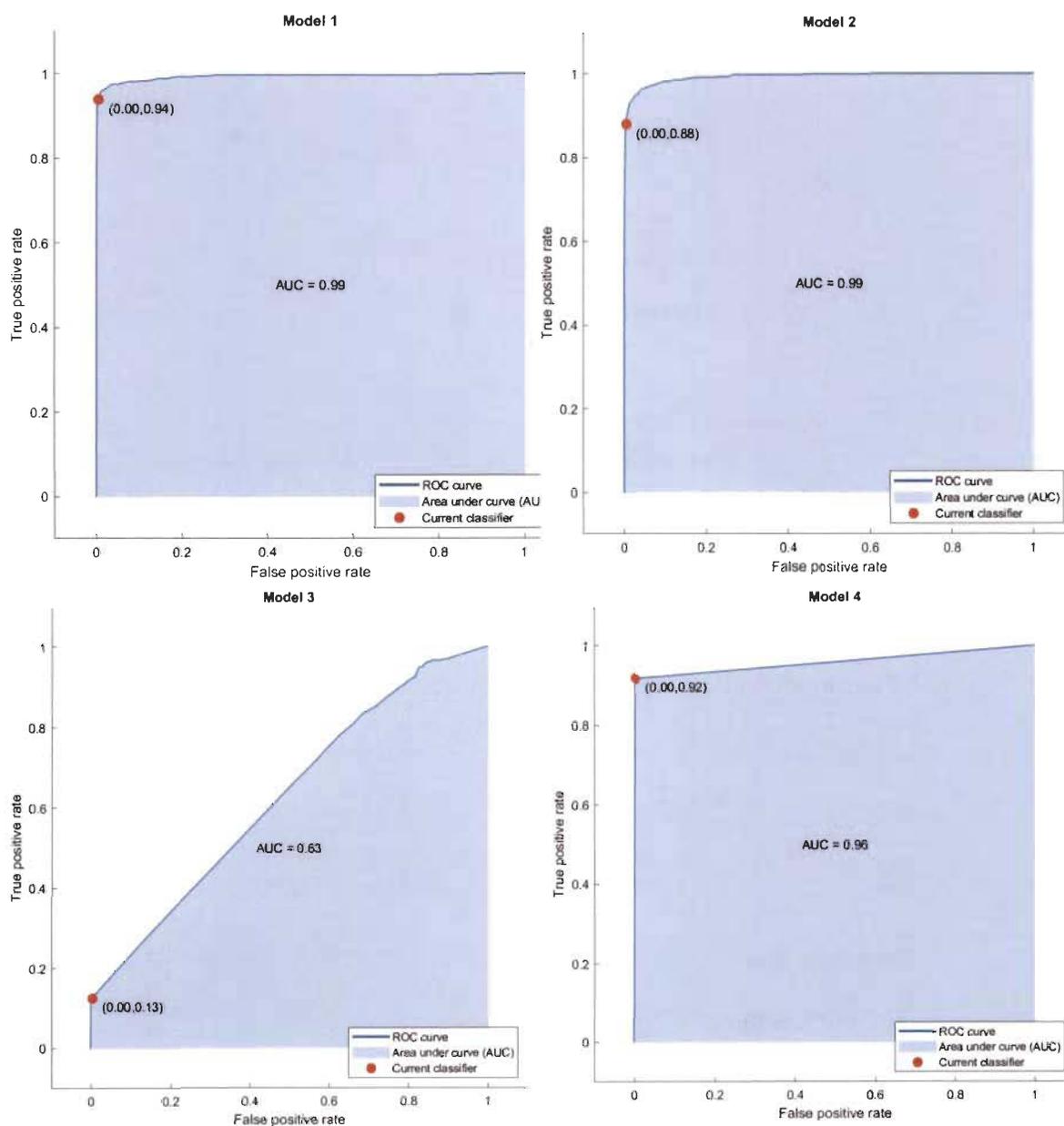
Figure 4-17 Matrices de confusion: (Model 1 : SVM quadratique) (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) au LSSI

Dans l'apprentissage automatique, mesurer les performances est une tâche de base. Par conséquent, en ce qui concerne les problèmes de classification, nous pouvons nous fier à la courbe l'aire sous la courbe et des caractéristiques de fonctionnement du récepteur (AUC-ROC). Lorsqu'il est nécessaire de vérifier ou de visualiser les performances d'un problème de classification multiclassés, nous utilisons la courbe AUC ROC. Il s'agit de l'un des indicateurs d'évaluation les plus importants pour vérifier la performance de tout modèle de classification.

ROC est une courbe de probabilité et AUC représente le degré ou la mesure de la séparabilité. Il indique dans quelle mesure le modèle est capable de distinguer les classes. Plus l'AUC est élevée, mieux le modèle est à prédire les 0 comme 0 et 1s comme 1. Par

analogie, plus l'ASC est élevée, meilleur est le modèle pour prédire la localisation où nous nous trouvons.

La figure 4-18 représente la courbe des caractéristiques de fonctionnement du récepteur (ROC) pour les différents modèles.



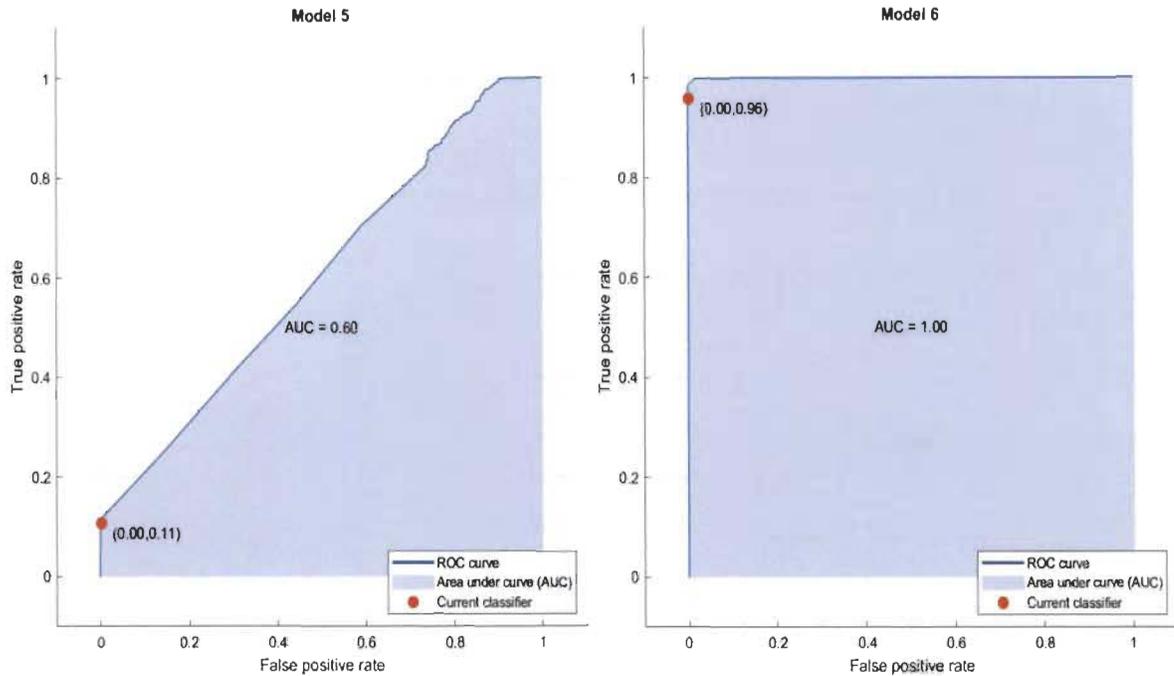


Figure 4-18 ROC et AUC courbes pour (Model 1 : SVM quadratique) , (Model 2 :kNN pondéré) , (Model 3 :Arbre de décision) , (Model 4 :kNN) , (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) au LSSI

4.6.2 Deuxièmes scénarios

Pour le second cas, nous avons fait notre expérience dans une maison où est ce qu'il y aura plein d'obstacle (murs, meubles ...) pour tester la performance de notre setup. Ces mesures sont faites dans 15 points de la maison où chaque point représente une classe. Nous avons choisi quatre points dans la chambre 1, quatre points dans la chambre 2, deux points dans la salle de bain, trois points dans la cuisine et deux points dans le corridor comme le montre dans la figure 4-20.

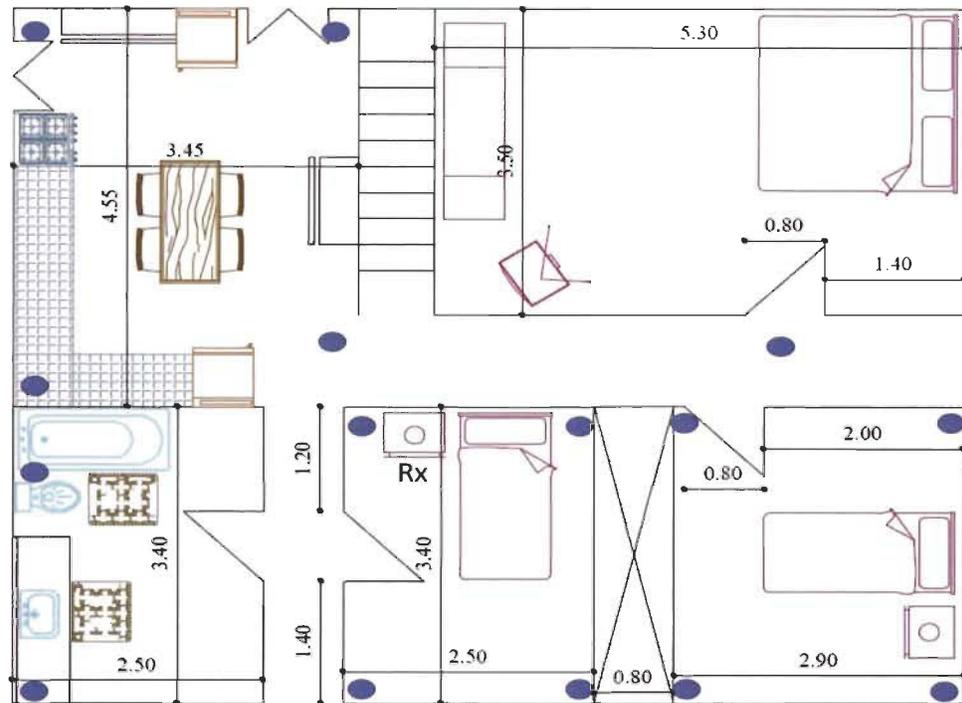


Figure 4-19 Architectures de la maison et les placements du point d'accès

Les meilleurs hyperparamètres du modèle sont obtenus après avoir effectué plusieurs tests comme nous l'avons vu sur l'organigramme de la figure 4-14, d'où nous obtenons le tableau suivant :

Tableau 4-2 Ensemble finalisé des meilleurs hyper paramètres pour le scénario 2

Optimizer	Adam
Learning rate	0.0055
Input Layers	180
Number of hidden layers	400
Batch Size	2000
Max Epoch	50
L2Regularization	0.0004

Les résultats de la simulation du LSTM proposé sont illustrés à la figure 4-20. Nous collectons 2700 mesures CSI pour chaque classe.

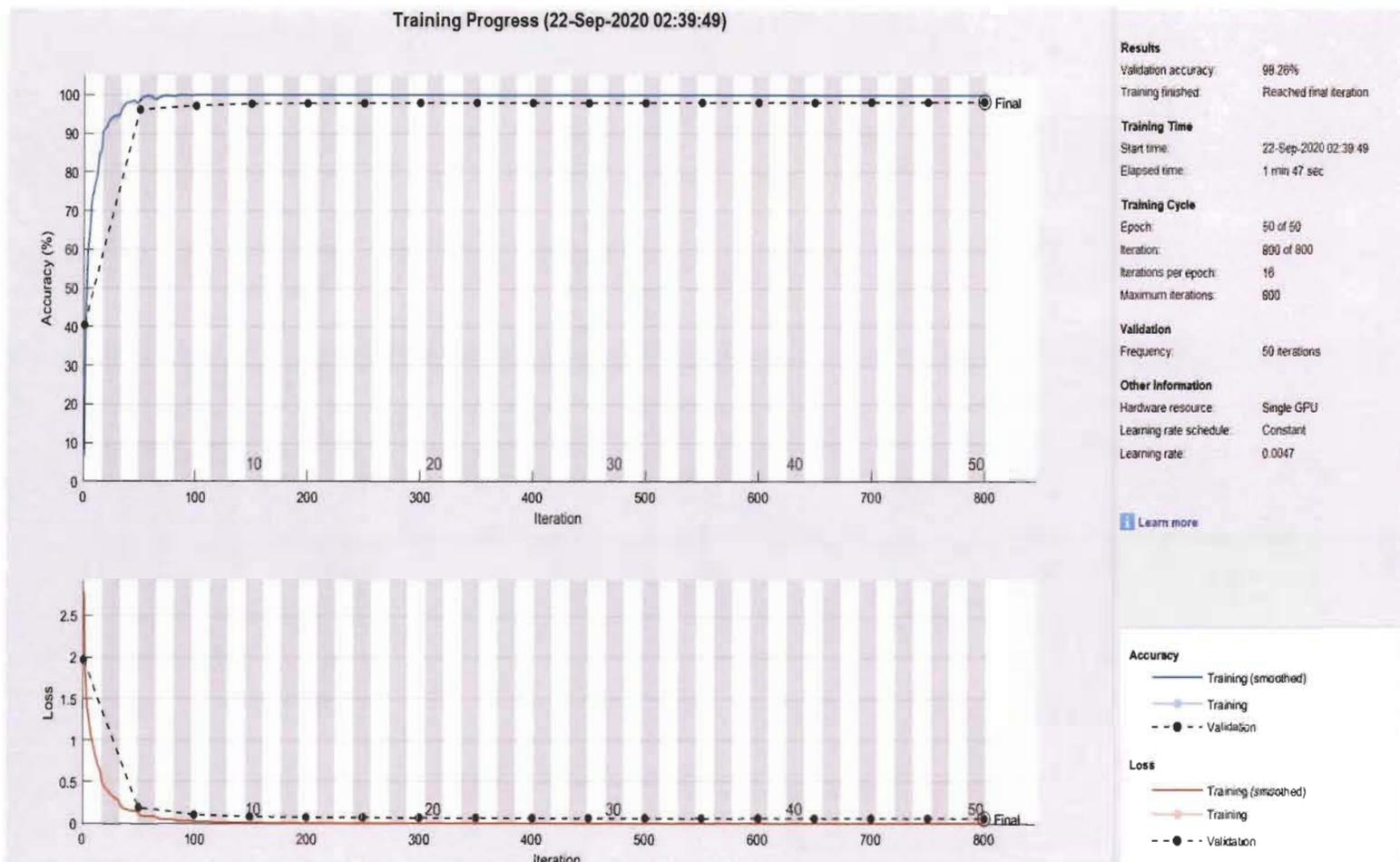


Figure 4-20 Résultats de la simulation à la maison

Model 2

True Class	1	96.4%	0.3%			1.3%	0.3%	0.1%	0.1%		0.0%	0.6%		0.3%	0.5%	0.0%	96.4%	3.6%
	2	0.4%	96.4%		1.3%		0.4%	0.7%	0.6%	0.0%	0.0%	0.1%			0.1%		96.4%	3.6%
	3			99.4%										0.4%	0.2%		99.4%	0.6%
	4		0.1%	0.6%	95.4%		1.2%	1.9%	0.0%	0.1%	0.0%		0.6%		0.1%		95.4%	4.6%
	5	0.0%				95.9%	0.1%	0.1%	0.3%			3.7%					95.9%	4.1%
	6		0.9%		2.0%		96.7%	0.2%				0.1%		0.0%			96.7%	3.3%
	7	0.4%	2.5%	0.1%	8.9%	0.2%	1.1%	75.6%	6.4%	0.3%	0.4%	1.6%	0.4%	0.1%		1.8%	75.8%	24.2%
	8	0.0%						7.2%	91.4%	0.4%	0.6%	0.3%	0.0%				91.4%	8.6%
	9			0.4%			0.4%		0.7%	87.5%	1.2%	8.3%		0.1%	1.4%		87.5%	12.5%
	10	0.2%	0.7%	0.0%	2.6%			1.8%	0.6%	0.1%	93.0%	0.7%			0.3%		93.0%	7.0%
	11	0.3%		0.3%	0.0%	1.3%			2.4%	0.8%	0.7%	93.1%		0.0%	0.9%		93.1%	6.9%
	12	0.3%			2.0%		0.1%	0.0%	0.8%		0.4%	3.2%	93.1%	0.1%			93.1%	6.9%
	13	2.4%	0.0%	0.1%		0.0%		0.1%	0.5%	0.4%	3.6%	1.0%	0.5%	86.4%	2.3%	2.7%	86.4%	13.6%
	14	2.5%	0.0%						0.7%	0.1%	0.3%	1.9%		0.7%	93.0%	0.7%	93.0%	7.0%
	15	0.1%	0.1%		0.5%		0.0%	2.0%	0.1%	0.6%		0.2%	0.0%	0.1%	0.5%	95.6%	95.6%	4.4%
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
		Predicted Class															TPR	FNR

Model 3

True Class	1	6.0%	1.0%	0.0%	0.2%	2.6%	0.9%	3.3%	0.2%	84.8%			0.9%				6.0%	94.0%
	2	0.3%	74.1%		2.3%	0.4%	4.1%	11.4%	0.5%	5.0%	1.3%		0.0%	0.4%		0.1%	74.1%	25.9%
	3	0.4%	0.6%	87.5%	1.2%	0.2%	0.8%	4.2%	0.0%	2.9%			2.0%		0.0%	0.0%	87.5%	12.5%
	4	0.2%	5.1%	0.9%	43.7%	0.2%	9.6%	22.5%	0.8%	12.6%	0.1%		4.1%	0.1%	0.0%	0.1%	43.7%	56.3%
	5	0.0%	6.0%		0.3%	9.9%	0.0%	25.3%	0.1%	58.1%			0.1%	0.1%			9.9%	90.1%
	6	0.0%	5.3%	0.3%	6.2%	1.5%	63.5%	17.9%	0.6%	1.6%	0.5%		2.4%		0.0%	0.1%	63.5%	36.5%
	7	0.2%	8.1%	0.2%	6.9%	4.6%	8.6%	45.7%	2.2%	23.0%			0.3%	0.1%	0.1%		45.7%	54.3%
	8	0.1%	4.3%		4.8%	1.6%	4.4%	26.5%	2.8%	55.1%			0.3%	0.0%			2.8%	97.2%
	9	0.0%	0.2%	0.0%	0.1%	0.2%	0.0%	0.3%	0.0%	98.9%			0.2%				98.9%	1.1%
	10	0.0%	3.2%	0.1%	3.5%	0.3%	3.9%	2.8%	0.4%	50.3%	35.4%			0.2%			35.4%	64.6%
	11		0.1%	0.0%	0.0%	0.3%		0.9%		96.3%			2.3%					100.0%
	12	0.4%	0.9%	0.0%	0.6%	1.4%	0.9%	5.9%	0.2%	31.2%			58.4%				58.4%	41.6%
	13	0.1%	5.1%		0.2%	0.9%	0.1%	4.6%	0.1%	87.6%	0.1%		0.1%	0.9%			0.9%	99.1%
	14	0.7%	0.6%	0.6%	0.9%	0.7%		2.1%	0.0%	94.4%			0.1%		0.0%		0.0%	100.0%
	15		0.2%		0.2%	0.6%	0.2%	2.6%	0.1%	87.3%			7.9%				0.9%	99.1%
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
		Predicted Class															TPR	FNR

Model 4

True Class	1	96.9%	0.1%			1.1%	0.3%	0.1%	0.1%		0.1%	0.4%	0.1%	0.3%	0.5%		96.9%	3.1%	
	2	0.2%	98.0%		0.3%		0.3%	0.8%	0.3%	0.0%	0.0%				0.1%		98.0%	2.0%	
	3			99.6%	0.1%										0.2%	0.0%		99.6%	0.4%
	4		0.0%	0.4%	96.2%		0.6%	1.9%	0.0%		0.1%		0.6%		0.1%			96.2%	3.8%
	5					97.7%		0.1%	0.0%			2.1%		0.0%				97.7%	2.3%
	6		0.6%		1.1%		98.0%	0.3%		0.0%		0.0%						98.0%	2.0%
	7	0.1%	1.7%	0.1%	6.7%	0.0%	0.9%	79.5%	6.5%	0.2%	0.6%	1.0%	0.3%	0.2%		1.9%		79.5%	20.5%
	8							3.8%	95.2%	0.4%	0.5%	0.1%	0.0%					95.2%	4.8%
	9	0.0%		0.0%	0.0%		0.3%		0.6%	91.0%	0.6%	6.4%		0.3%	0.6%			91.0%	9.0%
	10	0.1%	0.2%	0.0%	1.1%			1.2%	0.4%		96.2%	0.4%		0.1%	0.2%			96.2%	3.8%
	11	0.1%		0.3%		0.8%	0.1%	0.1%	0.8%	1.8%	0.4%	94.9%	0.1%	0.1%	0.6%			94.9%	5.1%
	12	0.2%			1.3%		0.0%	0.2%	0.2%		0.3%	2.1%	95.5%	0.1%				95.5%	4.5%
	13	0.5%						0.1%	0.1%	1.0%	1.1%	0.9%	0.5%	93.6%	1.3%	0.7%		93.6%	6.4%
	14	1.0%	0.4%			0.1%		0.0%	0.2%	0.4%	0.3%	1.1%		1.2%	94.9%	0.5%		94.9%	5.1%
	15	0.1%	0.1%		0.1%		0.1%	2.3%	0.1%	0.3%		0.1%		0.1%	0.3%	96.4%		96.4%	3.6%
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		TPR	FNR
		Predicted Class																	

Model 5

True Class	1	28.9%	0.7%	0.6%	0.2%	1.0%	8.8%	1.4%	0.1%	54.6%		2.3%	0.4%		0.9%		28.9%	71.1%	
	2	0.3%	83.6%	0.6%	0.4%		10.0%	1.7%		1.8%	1.6%				0.1%		83.6%	16.4%	
	3		0.0%	98.1%	0.3%		1.2%			0.1%	0.1%		0.2%				98.1%	1.9%	
	4	0.7%	12.6%	5.1%	18.3%	0.0%	39.0%	9.6%	0.3%	5.5%		0.3%	7.6%		1.0%		18.3%	81.7%	
	5	0.6%			0.3%	7.9%	38.9%	7.0%	0.0%	38.6%		8.6%						7.9%	92.1%
	6	0.0%	7.0%		2.9%		81.6%	5.2%	0.0%	0.2%	1.0%		2.0%			0.0%		81.6%	18.4%
	7	0.5%	2.3%	0.8%	3.1%	0.9%	64.5%	12.4%	0.1%	14.1%	0.1%	1.1%	0.0%		0.1%	0.0%		12.4%	87.6%
	8	0.3%	0.4%		2.1%		36.9%	8.3%	0.1%	51.2%		0.2%	0.3%		0.3%			0.1%	99.9%
	9	0.2%	0.1%	0.0%		0.1%	0.6%	0.2%		98.3%		0.2%	0.3%					98.3%	1.7%
	10	0.0%	13.8%	0.4%	1.1%		5.3%	0.4%		50.2%	28.6%	0.0%			0.0%			28.6%	71.4%
	11	0.1%	0.0%	0.0%		0.4%	1.0%	0.1%		96.2%		0.5%	1.6%		0.0%			0.5%	99.5%
	12	0.2%					8.8%	2.1%	0.0%	26.8%		0.7%	61.2%			0.0%		61.2%	38.8%
	13	0.2%	6.6%	0.1%	0.2%		4.6%	1.0%		86.6%	0.1%	0.4%	0.1%						100.0%
	14	1.9%	0.0%	1.3%	0.3%	1.0%	3.0%	0.8%	0.1%	90.3%		0.8%	0.0%		0.6%			0.6%	99.4%
	15	0.6%		0.0%			3.0%	0.9%	0.1%	83.8%		0.5%	9.4%		0.0%	1.8%		1.8%	98.2%
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		TPR	FNR
		Predicted Class																	

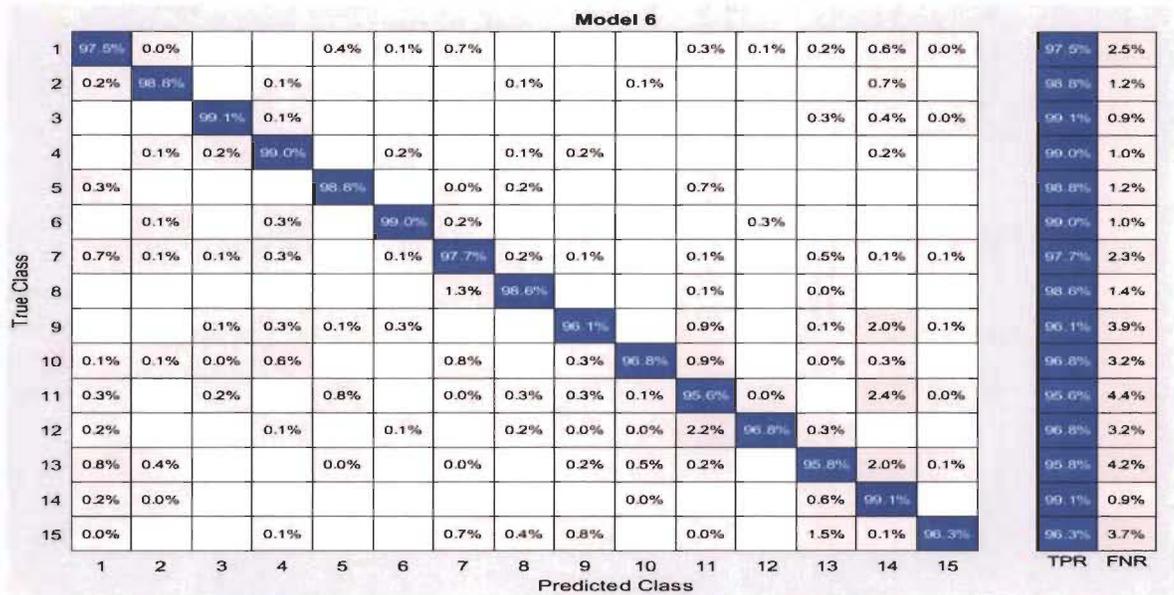
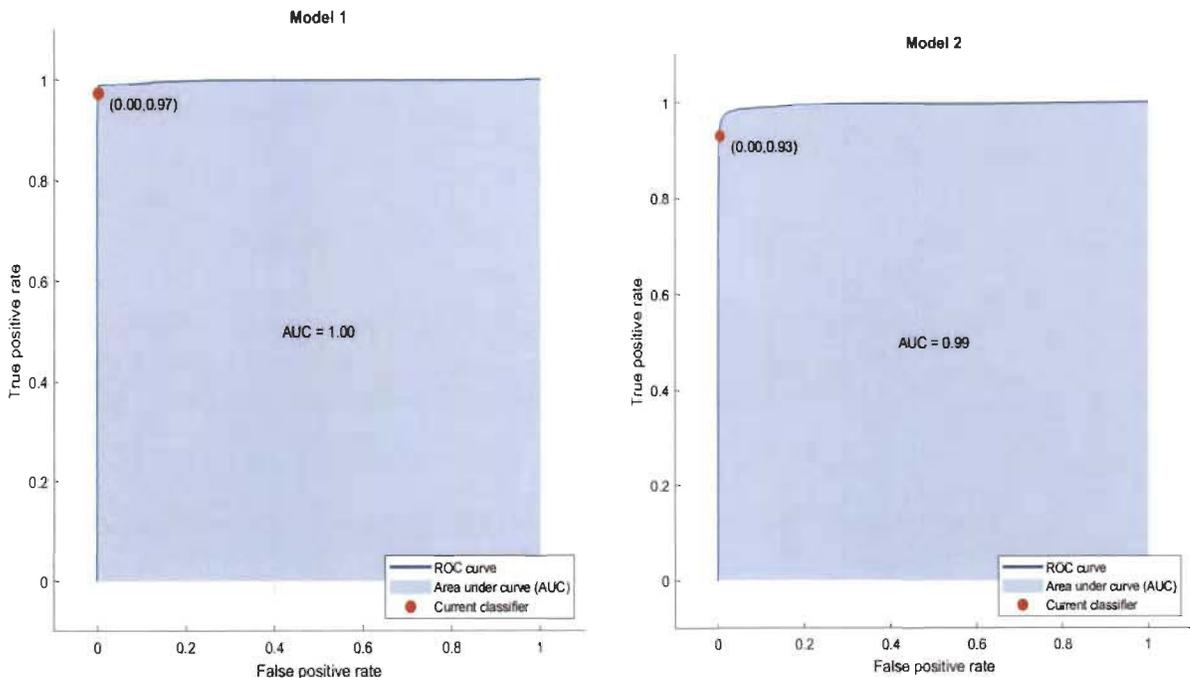


Figure 4-22 Matrices de confusion: (Model 1 : SVM quadratique), (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) à la maison

Le modèle 6 de la figure 4-22 présente une meilleure précision par rapport aux cinq autres modelés, bien qu'avec un temps de calcul plus élevé nécessaire à l'apprentissage.

La figure 4-23 représente la courbe des caractéristiques pour les différents modèles de ML.



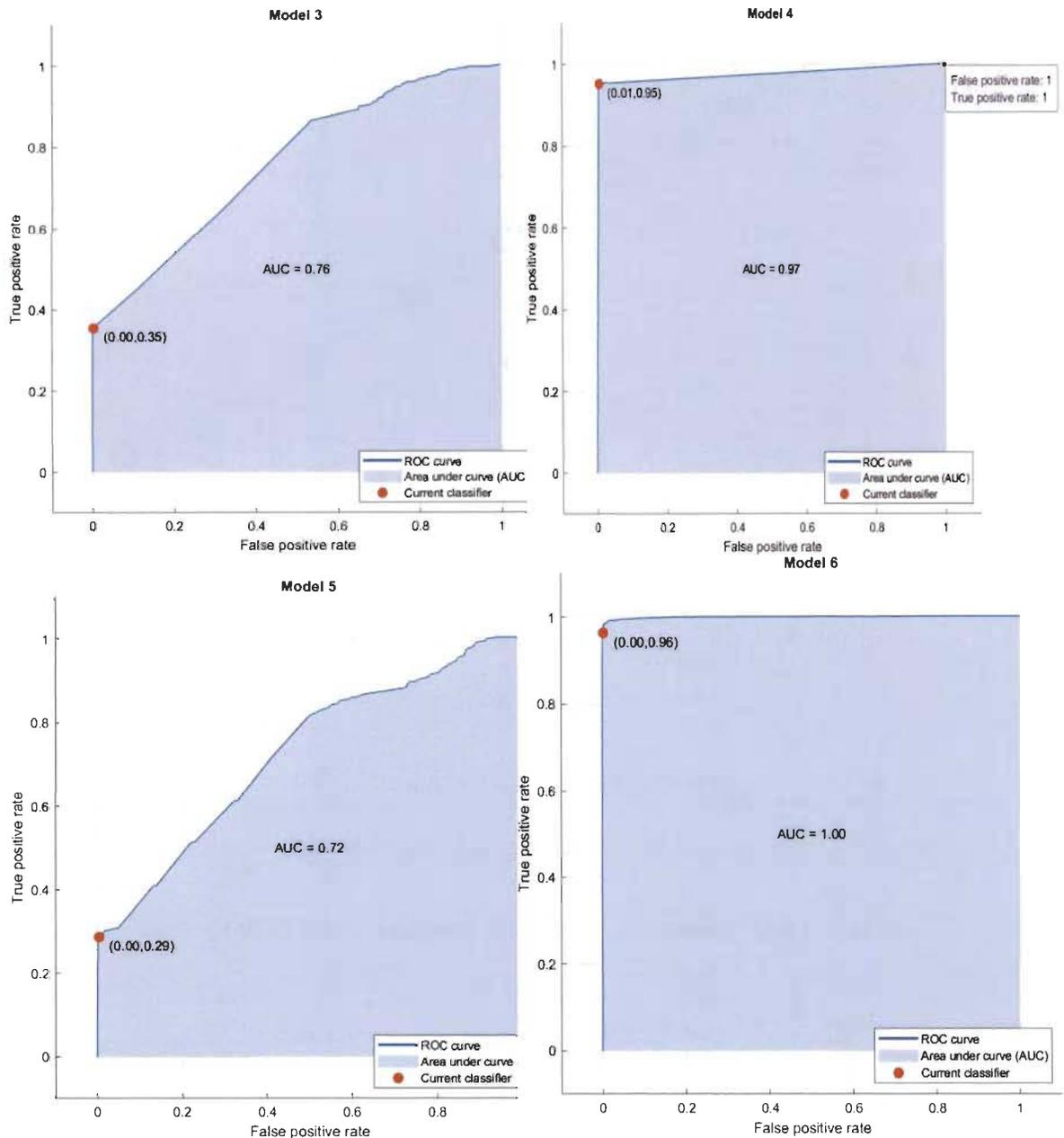


Figure 4-23 ROC et AUC courbes pour (Model 1 : SVM quadratique), (Model 2 :kNN pondéré), (Model 3 :Arbre de décision), (Model 4 :kNN), (Model 5 :Classificateur d'ensembles), (Model 6 : SVM optimisé) à la maison

Comme nous pouvons le voir à partir des résultats, la cohérence de toutes les classes n'est pas affectée de la même manière dans le temps et divers profils d'emplacement sont affectés différemment. Par exemple, il y a des classes qui semblent avoir un comportement plus

stable, tandis que d'autres zones (qui sont physiquement plus proches) ont tendance à être mal classées plus souvent. Par conséquent, une recherche plus approfondie doit être menée pour étudier l'effet des caractéristiques physiques de l'environnement intérieur et du placement de l'appareil sur la précision du classificateur

Comme prévu, les performances de l'algorithme LSTM proposé sont les meilleures parmi d'autres approches choisies. Pour référence, le tableau 4-3 décrit les performances de précision de localisation à l'aide de LSTM et différentes méthodes de ML pour les deux scénarios à l'université au LSSI et à la maison.

Tableau 4-3 Comparaisons de la précision de localisation (%) pour différentes méthodes d'apprentissage

Méthode de classifications	LSTM	SVM quadratique	kNN pondéré	Arbre de décision (Tree)	kNN	Classificateurs d'ensemble (Ensemble)	SVM optimisable
Scénario 1	97.90%	95.8%	87.6%	22.7%	92.6%	22.7%	97.3%
Scénario 2	98.23%	97.5%	92.6%	35.2%	94.9%	34.8%	96.7%

La littérature dans ce domaine est très prometteuse pour obtenir une bonne précision dans les environnements intérieurs. En utilisant notre expérience dans deux scénarios différents, il a été observé qu'en utilisant des techniques d'apprentissage en profondeur telles que LSTM au lieu de méthodes telles que SVM et kNN, une meilleure précision peut être obtenue. Il y a encore des défis dans les travaux futurs, tels que la façon d'utiliser les informations de phase CSI en plus de l'amplitude, comment stabiliser le système dans différents environnements

dynamiques et comment calculer l'erreur de localisation pour une prédiction correcte. Le tableau 4-4 représente la performance de notre LSTM avec d'autres approches à partir de références citées.

Tableau 4-4 Comparaisons de la précision de localisation (%) pour différentes méthodes de mesure basées sur RSS.

Méthodes	Précision
SVM [11]	82.5 %
MLP [11]	82.9 %
CNN [11]	87.4 %
LSTM [15]	90.5 %
HMM [15]	73.3 %
Forêt d'arbres décisionnels [15]	64.6 %
CNN [25]	99.98 %
MLP [25]	99.93 %
Notre LSTM	98.23 %

4.1 Conclusion

Dans ce chapitre, une approche à la localisation en intérieur est proposée et évaluée en utilisant les différents modèles de ML. Les résultats des méthodes de localisation que nous avons proposées et ses données sont spécifiques à notre environnement.

La première expérimentation a tout d'abord commencé par la collecte des données à l'université en choisissant 15 points au LSSI et 6 points dans le corridor puis concevoir notre réseau neurone LSTM en effectuant plusieurs tests pour obtenir les meilleurs hyperparamètres qui nous a permis d'obtenir une précision de 97.90%. Par la suite nous avons exécuté la même tâche mais dans un scénario différent qui est dans ce cas une maison en choisissant 21 points, et même chose notre réseau neurone s'est bien entraîné en donnant une précision de 98.23%. La comparaison des performances a été effectuée entre le LSTM et les autres modèles de ML, les résultats obtenus confirment que le LSTM donne de précision par rapport aux autres modèles que nous avons appliqués. A la fin, un tableau de synthèses des résultats incluant des méthode à partir de différente références pour comparer nos résultats obtenus.

En exploitant les données collectées et en utilisant des techniques d'apprentissage automatique, nous avons montré que ses mesures peuvent être efficacement exploitées pour prédire l'emplacement des personnes à intérieur.

Chapitre 5 - Conclusion

De nos jours, la demande des utilisateurs pour des informations de localisation en temps réel est sans précédent. Pour les environnements extérieurs, l'utilisation des technologies telle que GPS et GSM permettent d'obtenir des résultats de positionnement de haute précision. Cependant, dans un environnement intérieur, la technologie GPS ne peut pas être utilisée efficacement en raison de l'atténuation ou du blocage des signaux GPS. Par conséquent, la technologie de localisation Wi-Fi a été développée pour les environnements intérieurs.

Dans ce travail, nous avons mis en œuvre avec succès un système de localisation en intérieur basé sur Wi-Fi n'utilisant pas CSI au lieu de RSS, mais plutôt la combinaison de RSS et CSI de manière intelligente pour améliorer les performances de la localisation. Nous avons utilisé un routeur de réseau sans fil comme émetteur et une carte d'interface réseau comme récepteur pour collecter les données par paquets. Les performances du système ont été évaluées et comparées sur différente méthode d'apprentissage automatique. Après l'implémentation du système, un certain nombre d'études de cas a été conçu et des expériences ont été menées sur différent scénario et le développement et les exigences des systèmes matériels et logiciels ont été décrits.

Une explication détaillée sur la façon d'entraîner efficacement le LSTM pour la classification et de sélectionner les meilleurs hyper paramètres est présentée. L'évaluation a prouvé avec succès que le LSTM surpasse les autres approches de ML (nous avons atteint des résultats de précision de 98.2% pour LSTM). La collection des ensembles de données est une partie laborieuse du positionnement en intérieur.

Pour améliorer les résultats obtenus dans ce projet, plusieurs aspects peuvent être explorés plus pour améliorer les performances du système de localisation Wi-Fi proposé. La suite des travaux portera sur la résolution du problème de localisation en mesurant la distance des points par rapport à des points de références fixes et nous nous concentrerons aussi sur l'identification et l'extraction des caractéristiques des données brutes qui préservent les profils de localisation même en présence d'effets internes et externes, là où des changements inattendus et indésirables se produisent.

Bibliographies

- [1] E. Asimakopoulou and N. Bessis, "Buildings and crowds: Forming smart cities for more effective disaster management," in *2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2011: IEEE, pp. 229-234.
- [2] H. Borrión, T. Mitchener-Nissen, J. Taylor, and K.-M. Lai, "Countering Bioterrorism: Why smart buildings should have a code of ethics," in *2012 European Intelligence and Security Informatics Conference*, 2012: IEEE, pp. 68-75.
- [3] A. Zelenkauskaitė, N. Bessis, S. Sotiriadis, and E. Asimakopoulou, "Interconnectedness of complex systems of internet of things through social network analysis for disaster management," in *2012 Fourth International Conference on Intelligent Networking and Collaborative Systems*, 2012: IEEE, pp. 503-508.
- [4] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787-2805, 2010.
- [5] K. Su, J. Li, and H. Fu, "Smart city and the applications," in *2011 international conference on electronics, communications and control (ICECC)*, 2011: IEEE, pp. 1028-1031.
- [6] W. Tushar *et al.*, "Internet of Things for Green Building Management: Disruptive Innovations Through Low-Cost Sensor Technology and Artificial Intelligence," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 100-110, 2018, doi: 10.1109/msp.2018.2842096.
- [7] P. Siano, "Demand response and smart grids—A survey," *Renewable and sustainable energy reviews*, vol. 30, pp. 461-478, 2014.
- [8] T. Taleb and A. Kunz, "Machine type communications in 3GPP networks: potential, challenges, and solutions," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 178-184, 2012.
- [9] F. Zafari, A. Gkelias, and K. K. Leung, "A Survey of Indoor Localization Systems and Technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568-2599, 2019, doi: 10.1109/comst.2019.2911558.
- [10] S. He, B. Ji, and S.-H. G. Chan, "Chameleon: Survey-free updating of a fingerprint database for indoor localization," *IEEE Pervasive Computing*, vol. 15, no. 4, pp. 66-75, 2016.
- [11] K. Bregar and M. Mohorčič, "Improving indoor localization using convolutional neural networks on computationally restricted devices," *IEEE Access*, vol. 6, pp. 17429-17441, 2018.
- [12] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581-594, 2016.

- [13] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 27-38.
- [14] V. Bianchi, P. Ciampolini, and I. De Munari, "RSSI-based indoor localization and identification for ZigBee wireless sensor networks in smart homes," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 2, pp. 566-575, 2018.
- [15] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using wifi channel state information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98-104, 2017.
- [16] X. Tang and S. Mandal, "Indoor occupancy awareness and localization using passive electric field sensing," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 4535-4549, 2019.
- [17] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor localization using 802.11 time differences of arrival," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 614-623, 2015.
- [18] P.-H. Tseng, Y.-C. Chan, Y.-J. Lin, D.-B. Lin, N. Wu, and T.-M. Wang, "Ray-tracing-assisted fingerprinting based on channel impulse response measurement for indoor positioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 5, pp. 1032-1045, 2017.
- [19] M. A. Ouameur, M. Caza-Szoka, and D. Massicotte, "Machine Learning Enabled Tools and Methods for Indoor Localization Using Low Power Wireless Network," *Internet of Things*, p. 100300, 2020.
- [20] R. Battiti, N. T. Le, and A. Villani, "Location-aware computing: a neural network model for determining location in wireless LANs," University of Trento, 2002.
- [21] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao, "A wireless LAN-based indoor positioning technology," *IBM Journal of research and development*, vol. 48, no. 5.6, pp. 617-626, 2004.
- [22] A. M. Ladd, K. E. Bekris, A. P. Rudys, D. S. Wallach, and L. E. Kavraki, "On the feasibility of using wireless ethernet for indoor localization," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 3, pp. 555-559, 2004.
- [23] C.-H. Hsieh, J.-Y. Chen, and B.-H. Nien, "Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information," *IEEE Access*, vol. 7, pp. 33256-33267, 2019, doi: 10.1109/access.2019.2903487.
- [24] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067-1080, 2007, doi: 10.1109/tsmcc.2007.905750.
- [25] M. G. Wing, A. Eklund, and L. D. Kellogg, "Consumer-grade global positioning system (GPS) accuracy and reliability," *Journal of forestry*, vol. 103, no. 4, pp. 169-173, 2005.

- [26] A. Bose and C. H. Foh, "A practical path loss model for indoor WiFi positioning enhancement," in *2007 6th International Conference on Information, Communications & Signal Processing*, 2007: IEEE, pp. 1-5.
- [27] J. Barnes, C. Rizos, J. Wang, D. Small, G. Voigt, and N. Gambale, "Locata: the positioning technology of the future," in *Proceedings of the 6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, Australia July, 2003*.
- [28] K. Raja, W. Buchanan, and J. Munoz, "We know where you are [cellular location tracking]," *Communications Engineer*, vol. 2, no. 3, pp. 34-39, 2004.
- [29] J. J. Caffery and G. L. Stuber, "Overview of radiolocation in CDMA cellular systems," *IEEE Communications Magazine*, vol. 36, no. 4, pp. 38-45, 1998.
- [30] V. Otsason, A. Varshavsky, A. LaMarca, and E. De Lara, "Accurate GSM indoor localization," in *International conference on ubiquitous computing*, 2005: Springer, pp. 141-158.
- [31] A. Goetz, S. Zorn, R. Rose, G. Fischer, and R. Weigel, "A time difference of arrival system architecture for GSM mobile phone localization in search and rescue scenarios," in *2011 8th Workshop on Positioning, Navigation and Communication*, 2011: IEEE, pp. 24-27.
- [32] S. Ali, "Indoor geolocation for wireless networks," 2011.
- [33] C. L.-L. M. S. Committee. "802.11n-2009 - IEEE Standard for Information technology." (accessed. https://standards.ieee.org/standard/802_11n-2009.html)
- [34] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, "IEEE 802.11 ah: the WiFi approach for M2M communications," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 144-152, 2014.
- [35] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 60-67, 2016.
- [36] V. Bahl and V. Padmanabhan, "Enhancements to the RADAR user location and tracking system," 2000.
- [37] H.-H. Liu and Y.-N. Yang, "WiFi-based indoor positioning for multi-floor environment," in *TENCON 2011-2011 IEEE Region 10 Conference*, 2011: IEEE, pp. 597-601.
- [38] F. Zafari, I. Papapanagiotou, and K. Christidis, "Microlocation for internet-of-things-equipped smart buildings," *IEEE Internet of Things Journal*, vol. 3, no. 1, pp. 96-112, 2015.
- [39] M. Altini, D. Brunelli, E. Farella, and L. Benini, "Bluetooth indoor localization with multiple neural networks," in *IEEE 5th International Symposium on Wireless Pervasive Computing 2010*, 2010: IEEE, pp. 295-300.
- [40] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen, "Experiments on local positioning with Bluetooth," in *Proceedings ITCC 2003. International*

- Conference on Information Technology: Coding and Computing*, 2003: IEEE, pp. 297-303.
- [41] Apple. "iBeacon." (accessed. <https://developer.apple.com/ibeacon/>)
- [42] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, and Y. F. Hu, "Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and ZigBee standards," *Computer communications*, vol. 30, no. 7, pp. 1655-1695, 2007.
- [43] W. Farjow, A. Chehri, M. Hussein, and X. Fernando, "Support vector machines for indoor sensor localization," in *2011 IEEE Wireless Communications and Networking Conference*, 2011: IEEE, pp. 779-783.
- [44] M. Chiesa, R. Genz, F. Heubler, K. Mingo, and C. Noessel. "RFID." (accessed. <http://people.interactionivrea.it/c.noessel/RFID/research.html>)
- [45] A. A. Honoré, R. W. Thomas, R. K. Martin, and S. H. Kurkowski, "Implementation of collaborative RF localization using a software-defined radio network," in *MILCOM 2009-2009 IEEE Military Communications Conference*, 2009: IEEE, pp. 1-7.
- [46] X. Li, K. Pahlavan, M. Latva-aho, and M. Ylianttila, "Comparison of indoor geolocation methods in DSSS and OFDM wireless LAN systems," in *Vehicular Technology Conference Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conference (Cat. No. 00CH37152)*, 2000, vol. 6: IEEE, pp. 3015-3020.
- [47] B. B. Peterson, C. Kmieciak, R. Hartnett, P. M. Thompson, J. Mendoza, and H. Nguyen, "Spread spectrum indoor geolocation," *Navigation*, vol. 45, no. 2, pp. 97-102, 1998.
- [48] N. S. Correal, S. Kyperountas, Q. Shi, and M. Welborn, "An UWB relative location system," in *IEEE Conference on Ultra Wideband Systems and Technologies, 2003*, 2003: IEEE, pp. 394-397.
- [49] B. G. Yu, G. Lee, H. G. Han, W.-S. Ra, and T. W. Kim, "A time-based angle-of-arrival sensor using CMOS IR-UWB transceivers," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5563-5571, 2016.
- [50] N. Alsindi, X. Li, and K. Pahlavan, "Performance of TOA estimation algorithms in different indoor multipath conditions," in *2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No. 04TH8733)*, 2004, vol. 1: IEEE, pp. 495-500.
- [51] "TOA model." (accessed. <http://www.ciscosky.org/network/wireless/ShenWangLaoZhongDeMoXianDin>)
- [52] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067-1080, 2007.
- [53] M. Aatique, "Evaluation of TDoA techniques for position location in CDMA Systems," Electrical Engineering, Faculty of the Virginia Polytechnic Institute and State University, 1997.
- [54] W. Dargie and C. Poellabauer, *Fundamentals of wireless sensor networks: theory and practice*. John Wiley & Sons, 2010.

- [55] K. Pahlavan, X. Li, and J.-P. Makela, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112-118, 2002.
- [56] C. Drane, M. Macnaughtan, and C. Scott, "Positioning GSM telephones," *IEEE Communications magazine*, vol. 36, no. 4, pp. 46-54, 1998.
- [57] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample ML techniques for parameter estimation and detection in array processing, Ch. 4 in" *Radar Array Processing*" Simon Haykin," ed: Springer-Verlag, 1993.
- [58] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4-24, 1988.
- [59] S. Hara *et al.*, "Propagation characteristics of IEEE 802.15. 4 radio signal and their application for location estimation," in *2005 IEEE 61st Vehicular Technology Conference*, 2005, vol. 1: IEEE, pp. 97-101.
- [60] P. Kumar, L. Reddy, and S. Varma, "Distance measurement and error estimation scheme for RSSI based localization in Wireless Sensor Networks," in *2009 Fifth international conference on wireless communication and sensor networks (WCSN)*, 2009: IEEE, pp. 1-4.
- [61] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, pp. 1-32, 2013.
- [62] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: Passive device-free indoor localization using channel state information," in *2013 IEEE 33rd International Conference on Distributed Computing Systems*, 2013: IEEE, pp. 236-245.
- [63] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004.*, 2004: IEEE, pp. 14-23.
- [64] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piché, "A comparative survey of WLAN location fingerprinting methods," in *2009 6th workshop on positioning, navigation and communication*, 2009: IEEE, pp. 243-251.
- [65] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, 2005, pp. 205-218.
- [66] R. Zhou, "Wireless indoor tracking system (WITS)," *Aktuelle Trends in der Softwareforschung, Tagungsband zum doIT Software-Forschungstag. dpunkt. verlag Heidelberg, Germany. To appear*, pp. 163-177, 2006.
- [67] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *IEE proceedings-communications*, vol. 152, no. 5, pp. 580-586, 2005.
- [68] J. Ma, X. Li, X. Tao, and J. Lu, "Cluster filtered KNN: A WLAN-based indoor positioning scheme," in *2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2008: IEEE, pp. 1-8.

- [69] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53-53, 2011.
- [70] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity Wi-Fi," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1342-1355, 2018.
- [71] Intel. "Intel Ultimate N WiFi Link 5300." (accessed. <https://www.intel.com/content/www/us/en/products/docs/wireless-products/ultimate-n-wifi-link-5300-brief.html>)
- [72] T. Joachims, "Transductive inference for text classification using support vector machines," in *Icml*, 1999, vol. 99, pp. 200-209.
- [73] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [74] L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classification and regression trees. Wadsworth Int," *Group*, vol. 37, no. 15, pp. 237-251, 1984.