

**Titre:** Localisation d'usager dans un environnement interne en utilisant une technique pour contrer la variation temporelle du canal  
Title: [Localisation d'usager dans un environnement interne en utilisant une technique pour contrer la variation temporelle du canal](#)

**Auteur:** Lamia Hamza  
Author: [Lamia Hamza](#)

**Date:** 2008

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Hamza, L. (2008). Localisation d'usager dans un environnement interne en utilisant une technique pour contrer la variation temporelle du canal [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.  
Citation: <https://publications.polymtl.ca/8221/>

## Document en libre accès dans PolyPublie

Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/8221/>  
PolyPublie URL: [https://publications.polymtl.ca/8221/](#)

**Directeurs de recherche:** Chahe Nerguizian  
Advisors: [Chahe Nerguizian](#)

**Programme:** Non spécifié  
Program: [Non spécifié](#)

UNIVERSITÉ DE MONTRÉAL

LOCALISATION D'USAGER DANS UN ENVIRONNEMENT INTERNE EN  
UTILISANT UNE TECHNIQUE POUR CONTRER LA VARIATION  
TEMPORELLE DU CANAL

LAMIA HAMZA

DÉPARTEMENT DE GÉNIE ÉLECTRIQUE  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES  
(GÉNIE ÉLECTRIQUE)

NOVEMBRE 2008



Library and  
Archives Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*

ISBN: 978-0-494-48921-5

*Our file* *Notre référence*

ISBN: 978-0-494-48921-5

#### NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

#### AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

LOCALISATION D'USAGER DANS UN ENVIRONNEMENT INTERNE EN  
UTILISANT UNE TECHNIQUE POUR CONTRER LA VARIATION  
TEMPORELLE DU CANAL

présenté par: HAMZA Lamia

en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées  
a été dûment accepté par le jury d'examen constitué de:

M. CARDINAL Christian, Ph.D., président

M. NERGUIZIAN Chahé, Ph.D., membre et directeur de recherche

M. FRIGON Jean François, Ph.D., membre

À mon père,  
À ma mère,  
À ma sœur Leila,  
À mon frère Salim,  
Merci pour vos précieux encouragements.

## REMERCIEMENTS

Tout d'abord, je tiens à exprimer ma plus sincère reconnaissance à mon directeur de recherche, Dr Chahé Nerguizian, pour son encadrement, sa confiance, sa patience, sa rigueur ainsi que son esprit critique qui m'ont permis d'avancer dans le cadre de ma maîtrise.

Je tiens aussi à remercier tous les professeurs qui ont contribué à ma formation ainsi que tout les techniciens fort sympathiques à qui j'ai souvent eu affaire pour leur disponibilité et leur aide. Merci à Jean Bouchard, Jean-Sébastien Décarie et Maxime Thibault.

Je tient à exprimer ma gratitude à messieurs les membres de jury; Dr Christian Cardinal et Dr Jean-François Frigon, qui ont accepté de juger mon travail de recherche.

Mes sincères remerciements vont à l'ensemble de mes collègues d'études et amis, qui ont contribué à l'élaboration de ce travail par leurs encouragements et leurs conseils. Donc merci à Adam, belkacem, Florian, Hawoye, Laure, Nadia, Olfa, Rachid, Salim, Sergine, Stéphane et Zhao.

Enfin, je remercie tous ceux qui ont contribué à la réussite de ce mémoire, et que tous ceux que j'ai involontairement oubliés trouvent ici l'expression de ma gratitude.

## RÉSUMÉ

Plusieurs recherches proposent des systèmes permettant d'estimer la localisation dans un environnement interne et dans un canal invariant dans le temps. Cependant, peu de ces travaux prennent en compte le cas d'un canal variant dans le temps. Ce mémoire propose un système de localisation permettant de contrer la variation du canal dans le temps dans un environnement interne.

Les puissances reçues (RSS) d'un signal de type radio fréquence, dans un système WLAN, sont mesurées et utilisées avec l'approche des signatures afin de définir l'empreinte du canal. Une base de données est ainsi créée dans un canal invariant dans le temps à  $t_0$  en 3D. L'un des points faibles de l'approche des signatures est sa mauvaise performance lorsque le canal varie dans le temps. Dans ce cas, les mesures sont habituellement reprises pour créer une nouvelle base de données. La méthode de localisation proposée dans ce projet permet une localisation dans un canal variable dans le temps lorsqu'il est difficile ou impossible de reprendre ces mesures.

Ce système permet de générer une base de données à  $t_x$  grâce à des mesures de quelques points de référence et d'une technique d'exploitation de données qui est la méthode d'arbre de décision. Cette méthode utilise des points de référence et l'analyse de régression afin d'effectuer la prédiction temporelle des puissances reçues au mobile au temps  $t_x$  ce qui permet de générer une nouvelle base de données. Ensuite, un réseau de neurones artificielles (ANN) est appliqué aux données pour estimer la localisation.

La génération de la nouvelle base de données avec la méthode d'arbre de décision a engendré une dégradation inférieure à 9% pour une précision de 3 mètres. Trois bases de données sont générées à différents temps ( $t_1$ ,  $t_2$  et  $t_3$ ). Le système propose d'utiliser ces trois bases de données simultanément comme entrées du réseau de

neurones afin d'estimer la position du mobile. Cette méthode donne une précision de 3 mètres pour 61.38% des cas et de 5 mètres pour 78.22% des cas. Ceci entraîne une dégradation moyenne supplémentaire d'environ 7% en 2D.

L'estimation en 3D engendre une dégradation supplémentaires aux deux précédentes de 3.5% à cause de l'ajout de la variable  $z$  qui détermine l'étage où se trouve le mobile. Cette méthode donne une précision de 3 mètres pour 57.06% des cas et de 5 mètres pour 74.81% des cas en 3D. La méthode de localisation proposée dans ce mémoire aboutie à des résultats convenables particulièrement en 2D. Elle est recommandée lorsque les ressources matérielles et temporelles sont insuffisants.

## ABSTRACT

Several researches suggest systems allowing indoor location estimation in no-time-varying channels. However, few of those studies take into account the case of time-varying channels. This thesis propose a location system allowing to counter time-varying channel in an indoor environment. The received signal strengths (RSS), in a WLAN system, are measured and used in a *Fingerprinting technique* in order to define the channel signature. The database is therefore created, in 3D, in no-time-varying channel at  $t_0$ . The *Fingerprinting technique* suffers degradations in location accuracy in time-varying channels. In that case, new measurements are carried out to create a new database. This thesis propose a method allowing the location in time-varying channel when it is difficult or impossible to collect measurements.

The system generates a database with measurements of few reference points and a *Data Mining technique* called the *Model Tree*. The method uses a regression analysis to learn the temporal predictive relationship between the RSS values of the mobile and the reference points and generates a new database at  $t_x$ . Afterwards, an artificial neural network (ANN) is applied to the database for location estimation. The method tree generates a degradation less than 9% within 3 meters accuracy. In order to validate the proposed system, 3 databases are generated at different times ( $t_1$ ,  $t_2$  and  $t_3$ ) and used simultaneously as inputs for the ANN to estimate the position of the mobile. The distance location accuracy of 3 meters is obtained for 61.38% of the cases and of 5 meters for 78.22% of the cases. This causes a supplementary degradation of approximately 7% in 2D.

The 3D estimations adds a 3.5% supplementary degradation to the two previous ones because of the  $z$  variable which defines the mobile floor location (in a building). This approach allows a 3 meters accuracy in 57.06% of the cases and of 5 meters in 74.81% of the cases in 3D. The location system proposed in this thesis achieves acceptable results particularly in 2D. It is recommended when time and material

means are insuffisants or limited.

## TABLE DES MATIÈRES

DÉDICACE . . . . .	iv
REMERCIEMENTS . . . . .	v
RÉSUMÉ . . . . .	vi
ABSTRACT . . . . .	viii
TABLE DES MATIÈRES . . . . .	x
LISTE DES FIGURES . . . . .	xiii
LISTE DES NOTATIONS ET DES SYMBOLES . . . . .	xvi
LISTE DES TABLEAUX . . . . .	xvii
LISTE DES ANNEXES . . . . .	xix
CHAPITRE 1 INTRODUCTION . . . . .	1
1.1 Contribution et méthodologie . . . . .	3
1.2 Structure du mémoire . . . . .	5
CHAPITRE 2 CONCEPTS DE BASE ET TRAVAUX ANTÉRIEURS . . . . .	7
2.1 Concepts de base . . . . .	8
2.1.1 Mécanismes de propagation . . . . .	8
2.1.1.1 La théorie de la réflexion spéculaire et diffuse . . . . .	8
2.1.1.2 La théorie de la diffraction . . . . .	9
2.1.2 Signaux LOS et NLOS . . . . .	9
2.1.3 Phénomène multi-trajets (Multipath) . . . . .	10
2.1.4 Phénomène d'évanouissement . . . . .	11

2.2	Caractéristiques du canal . . . . .	11
2.3	Paramètres de localisation . . . . .	12
2.3.1	Méthode basée sur les mesures d'angle d'arrivée . . . . .	12
2.3.2	Méthode basée sur les mesures de délai d'émission / réception	13
2.3.2.1	Mesure selon le TOA (Time Of Arrival) et le TDOA (Time Difference Of Arrival) . . . . .	14
2.3.2.2	Mesure selon le RTT (Round-Trip Time) . . . . .	15
2.3.3	Méthode basée sur les mesures de la puissance reçue . . . . .	17
2.4	Algorithmes de détection et d'estimation . . . . .	18
2.4.1	L'approche des signatures . . . . .	18
2.4.2	La technique de triangulation . . . . .	19
2.4.3	Méthode des réseaux à neurones artificiels . . . . .	20
2.4.3.1	Phase entraînement . . . . .	21
2.4.3.2	Phase estimation . . . . .	22
2.5	Système WLAN . . . . .	22
2.6	Travaux antérieurs . . . . .	24

CHAPITRE 3	MESURES EN RSS POUR L'APPROCHE DES SIGNA- TURES DANS UN CANAL STATIQUE . . . . .	29
3.1	Description du système de localisation proposé . . . . .	29
3.2	Méthodologie . . . . .	30
3.2.1	Prise de mesures . . . . .	30
3.2.1.1	Paramètres d'erreur . . . . .	30
3.2.1.2	Paramètres de mesure . . . . .	32
3.2.1.3	Équipements de mesure . . . . .	32
3.2.1.3.1	Matériels de mesure . . . . .	32
3.2.1.3.2	Programme de mesure . . . . .	33
3.2.1.4	Protocole de mesures . . . . .	34

3.2.2	Obtention de la base de donnée (Approche des signatures) . . . . .	35
3.2.3	Méthode d'estimation . . . . .	37
3.3	Résultats et analyse . . . . .	38
	 CHAPITRE 4 VARIABILITÉ DU CANAL . . . . .	48
4.1	Description du système de localisation proposé . . . . .	51
4.2	Méthodologie . . . . .	53
4.2.1	Définition des points de référence . . . . .	53
4.2.2	Prise de mesure . . . . .	54
4.2.3	Traitements des données . . . . .	56
4.2.4	Arbre de décision . . . . .	57
4.2.4.1	Principe de fonctionnement . . . . .	57
4.2.4.2	Méthode de la génération de la base de données . . . . .	61
4.2.4.3	Algorithme de génération de la base de données . . . . .	62
4.2.5	Méthode d'estimation . . . . .	65
4.3	Résultats et analyse . . . . .	66
4.3.1	Dégradation selon l'arbre de décision . . . . .	66
4.3.2	Performances du système de localisation proposé en 2D . . . . .	68
4.3.3	Performances du système de localisation proposé en 3D . . . . .	76
	 CHAPITRE 5 CONCLUSION . . . . .	81
	 RÉFÉRENCES . . . . .	85
	 ANNEXES . . . . .	89

## LISTE DES FIGURES

<b>FIGURE 1.1</b>	Diagramme général du travail effectué dans ce mémoire . . . . .	5
<b>FIGURE 2.1</b>	Diagramme fonctionnel d'un système de localisation dans un environnement interne. . . . .	7
<b>FIGURE 2.2</b>	Phénomène de la réflexion spéculaire . . . . .	8
<b>FIGURE 2.3</b>	Phénomène de la réflexion diffuse . . . . .	9
<b>FIGURE 2.4</b>	Phénomène de diffraction . . . . .	10
<b>FIGURE 2.5</b>	Exemple de réception du signal selon une vue directe (LOS) et une vue non directe (NLOS) . . . . .	10
<b>FIGURE 2.6</b>	Exemple de localisation selon la méthode d'angle d'arrivée .	13
<b>FIGURE 2.7</b>	Exemple d'émission/réception d'un signal en TOA . . . . .	14
<b>FIGURE 2.8</b>	Exemple d'émission/réception d'un signal en RTT . . . . .	15
<b>FIGURE 2.9</b>	Schéma du processus de localisation basée sur l'approche des signatures pour un système WLAN, a) phase de collecte de données, b) phase de localisation . . . . .	18
<b>FIGURE 2.10</b>	a) Structure de base d'un neurone, b) Architecture d'un réseau de neurones . . . . .	20
<b>FIGURE 2.11</b>	a) Phase d'entraînement d'un ANN, b) Phase estimation .	21
<b>FIGURE 3.1</b>	Exemple d'une fenêtre Netstumbler . . . . .	34
<b>FIGURE 3.2</b>	Exemple d'un fichier enregistré grâce à Netstumbler . . . . .	35
<b>FIGURE 3.3</b>	Positions exactes et estimées en x, y et z en fonction de 1092 points de mesure de la station mobile obtenues en phase d'apprentissage . . . . .	40
<b>FIGURE 3.4</b>	Erreurs de localisation en termes de x, y, z et de la distance Euclidienne $d$ en 2D et en 3D en fonction de 1092 points de mesure de la station mobile obtenues en phase d'apprentissage	41

FIGURE 3.5	Fonction de distribution cumulative CDF des erreurs de localisation en termes de x, y et de d en 2D en fonction de l'erreur de localisation (en mètres) obtenues en phase d'apprentissage	42
FIGURE 3.6	Positions exactes et estimées en x, y et z en fonction de 350 points de mesure de la station mobile obtenus en phase d'estimation . . . . .	43
FIGURE 3.7	Erreurs de localisation en termes de x, y, z et de la distance Euclidienne d en 2D et en 3D en fonction de 350 points de mesure de la station mobile obtenues en phase d'estimation	44
FIGURE 3.8	Fonction de distribution cumulative CDF des erreurs de localisation en termes de x, y et de d en 2D en fonction de l'erreur de localisation (en mètres) obtenues en phase d'estimation . . . . .	45
FIGURE 4.1	RSS en fonction du temps pour une position donnée dans un canal statique et dynamique . . . . .	49
FIGURE 4.2	Schéma explicatif de la solution de l'article [23] . . . . .	50
FIGURE 4.3	Schéma général de l'étape de la génération d'une base de donnée . . . . .	51
FIGURE 4.4	Schéma général de l'étape de l'estimation de la position . . . . .	52
FIGURE 4.5	Exemple d'un fichier créé par Traitement.cpp . . . . .	56
FIGURE 4.6	Schéma d'un modèle d'arbre de décision . . . . .	58
FIGURE 4.7	Cheminement de décision pour le point de mesure numéro 755 et l'AP24 . . . . .	60
FIGURE 4.8	Schéma de la création du modèle d'arbre de décision . . . . .	61
FIGURE 4.9	Schéma de la création de nouvelles bases de données au temps $t_1$ et $t_2$ . . . . .	62
FIGURE 4.10	Diagramme du programme <i>ArbreDeDecision.m</i> . . . . .	63

FIGURE 4.11	Diagramme donnant les étapes permettant le calcul de la dégradation due à la génération de la base de données avec le modèle d'arbre de décision . . . . .	65
FIGURE 4.12	Exemple de CDF de l'erreur de localisation pour 139 points de mesures en x, y et en d en 2D pour un canal variant dans le temps ( $t_3$ ) . . . . .	73
FIGURE 4.13	Schéma d'utilisation de tous les bases de données créées en 2D à l'entrée de l'ANN . . . . .	74
FIGURE I.1	6 <sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal . . . . .	89
FIGURE I.2	5 <sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal . . . . .	90
FIGURE I.3	3 <sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal . . . . .	91
FIGURE II.1	Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour un canal invariant dans le temps ( $t_1$ ) . . . . .	92
FIGURE II.2	Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour un canal semi-variant dans le temps ( $t_2$ ) . . . . .	93
FIGURE II.3	Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour tout type de canal ( $t_1$ , $t_2$ et $t_3$ ) . . . . .	94
FIGURE II.4	Exemple de CDF de l'erreur de localisation pour 140 points de mesure en x, y, z et en d en 3D pour un canal semi-variant dans le temps ( $t_2$ ) . . . . .	95

## LISTE DES NOTATIONS ET DES SYMBOLES

<b>Sigles</b>	<b>Définitions</b>
ANN	Artificial Neural Network (Réseau de Neurones Artificiel)
AOA	Angle of Arrival (Angle d'Arrivé)
AP	Access Point (Point d'Accès)
CDF	Cumulative Density Function
dB	Décibel
dBm	Décibel par rapport à 1mW
DOA	Direction Of Arrival
GPS	Global Positioning System
IEEE	Institut of Electrical and Electronics Engineers
IEEE802.11	Standard du réseau WLAN
IP	Internet Protocol (Protocole d'Internet ou Protocole de la Couche Réseau)
LOS	Line-Of-Sight (Ligne de Vue Directe)
MAC	Medium Access Control (Méthode d'Accès au Médium)
NEWGRNN	New Generalized Regression Neural Network
NLOS	No-Line-Of-Sight (Ligne de Vue Indirecte)
RF	Radio Frequency (Radio fréquence)
RSS	Receive Signal Strength (Puissance du Signal Reçu)
RTT	Round Trip Time (Temps Aller-Retour de Paquets)
TDOA	Time Difference of Arrival (Délai du Temps d'Arrivé)
TOA	Time of Arrival (Temps d'Arrivé)
WLAN	Wireless Local Area Network (Réseau Local Sans Fil)
WR API	Wireless Research API (Application Programming Interface)
OSI	Open System Interconnection
SSID	Service Set IDentifier

## LISTE DES TABLEAUX

Tableau 3.1	Valeur du CDF en 3D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation . . . . .	47
Tableau 4.1	Coordonnées, en mètres, des 14 points de référence définis pour les simulations en 2D . . . . .	54
Tableau 4.2	Coordonnées, en mètres, des 25 points de référence définis pour les simulations en 3D . . . . .	55
Tableau 4.3	Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation pour la base de données (BD) générée avec l'arbre de décision et mesurée au temps $t_0$ . . . . .	67
Tableau 4.4	Moyenne de la dégradation à 3 mètres et à 5 mètres, pour 15 simulations pour la phase d'estimation de la base de données $t_0$ . . . . .	67
Tableau 4.5	Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps $t_1$ (pour un canal invariant dans le temps) . . . . .	69
Tableau 4.6	Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps $t_2$ (pour un canal semi-variant dans le temps) . . . . .	70
Tableau 4.7	Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps $t_3$ (pour un canal variant dans le temps) . . . . .	71
Tableau 4.8	Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps $t_1$ , $t_2$ et $t_3$ (les 3 bases de données sont définis comme l'entrée du réseau ANN) . . . . .	76

Tableau 4.9	Valeur du CDF en 3D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps $t_2$ (pour un canal semi-variant dans le temps) . . . . .	77
Tableau 4.10	Erreur de localisation en mètres de l'estimation de la position en 3D au temps $t_2$ . . . . .	79
Tableau 4.11	Erreur de localisation en mètres de l'estimation de la position en 3D au temps $t_2$ pour les positions proches de des positions 1 et 10 du tableau 4.10 . . . . .	79

**LISTE DES ANNEXES**

ANNEXE I	PLAN DU PAVILLON LASSONDE . . . . .	89
ANNEXE II	EXEMPLES DE CDFS DE L'ERREUR DE LOCALISATION POUR DIFFÉRENTS TYPES DE VARIABILITÉ DU CANAL DANS LE TEMPS . . . . .	92
ANNEXE III	FICHIERS MATLAB . . . . .	96
	III.1 <i>ANN_GRNN_2008.m</i> . . . . .	96
	III.2 <i>TestUneMesure.m</i> . . . . .	110
ANNEXE IV	FICHIER DE TRAITEMENT DES DONNÉES <i>TRAITEMENT.CPP</i> . . . . .	117

## CHAPITRE 1

### INTRODUCTION

Les systèmes de localisation ont beaucoup évolué depuis leur apparition durant la 2<sup>e</sup> guerre mondiale. Pendant cette période, et pour un usage strictement militaire, des systèmes de radio-navigation maritimes comme *DECCA* et *LORAN* ont été conçus. La résolution de ces derniers était d'une centaine de mètres.

Lorsque la demande militaire s'est tournée vers des systèmes de positionnement plus précis, le système de positionnement mondial *GPS* (*Global Positioning System*) a vu le jour. Ce dernier a été opérationnel en 1995 et utilise la triangulation comme algorithme de localisation. De nos jours, sa précision peut aller jusqu'à l'ordre du centimètre et est utilisé comme principal système de positionnement par satellites mondial.

L'arrivée de cette technologie dans le domaine public et commercial a permis d'offrir des services à faible prix jusqu'alors impossible. Dans les zones urbaines denses, la localisation est difficile mais possible avec une erreur de localisation plus grande. Ses applications dans des environnements externes sont nombreuses, elle est disponible :

- Pour la navigation aérienne et maritime.
- Pour le service d'urgence 911 en cas d'un appel en détresse via un téléphone cellulaire ou pour le service de police pour traquer des voitures volées.
- Dans les voitures pour l'assistance et la navigation routière.
- Dans les téléphones cellulaires pour l'auto localisation dans les villes et/ou pour trouver les magasins et services les plus proches via les pages jaunes.

Comme la localisation et la navigation en extérieure sont bien implantées, la demande s'est tournée vers la localisation dans des milieux internes. En effet, en prenant comme exemple le service d'urgence 911, même s'il est possible de retracer dans la rue une personne faisant un appel d'urgence avec un cellulaire, seul le bâtiment d'où l'appel a été effectué peut être trouvé. Il est pour le moment impossible de connaître dans quel étage et quel bureau la personne se trouve. Dans ce cas très précis, l'ajout au système existant d'un complément qui permettrait la localisation interne impliquerait une amélioration considérable du service d'urgence.

Les ondes des satellites ne traversent presque pas les bâtiments et les sous terrains, c'est pour cela que le *GPS* est souvent inutile pour la localisation en milieu interne. De plus, la triangulation donne de grandes erreurs à cause du phénomène multi-trajets très présent dans ce type de milieu. Beaucoup de recherches sont effectuées afin de contrer ce phénomène et pour développer de nouveaux systèmes de localisation utilisant divers techniques de détections et ressources matérielles à faible coût. Il existe même, dans le domaine commercial, des produits qui permettent la localisation dans certains types d'environnements internes.

En plus des obstacles cités précédemment, la résolution de la localisation en milieu interne doit être plus grande que celle en externe car l'environnement est plus petit. Pour le moment, aucune résolution standard n'a été définie comme résolution acceptable pour la localisation interne. Dans ce projet, la résolution est définie par la probabilité de se trouver dans un rayon de 3 mètres et de 5 mètres.

La détection en milieu interne pourrait être utilisée dans plusieurs applications en dehors du service d'urgence 911. Si cette technologie se développe, elle pourrait être utilisée :

- Dans les hôpitaux afin de localiser des patients, des médecins ou un équipement de santé le plus proche.
- Dans les mines et les métros pour trouver des victimes en cas d'accident.
- Dans les prisons à la recherche de prisonniers qui manquent à l'appel.
- Dans les centres commerciaux et centre publiques fermés pour le positionnement et l'auto navigation ainsi que pour localiser les enfants perdus.
- Dans le domaine de la construction.
- Pour aider des policiers et des pompiers à accomplir leurs missions.

### 1.1 Contribution et méthodologie

Il existe plusieurs travaux faisant état de la recherche sur la localisation dans un environnement interne. Cependant, peu de ces travaux parlent de la localisation dans un environnement où le canal est variant dans le temps. En effet, selon le système de détection et l'algorithme de localisation, l'erreur de localisation dans un canal variant peut être beaucoup plus grande que celle dans un canal statique (invariant) dans le temps. Le but de ce présent mémoire est de contrer la variation du canal en minimisant les mesures effectuées.

Dans le cadre de cette recherche, la localisation est effectuée en prenant des mesures à bande étroite des signaux de type radio fréquences sous le système de réseau sans fil *WLAN*. Ce dernier a été choisi pour son faible coût et sa facilité d'utilisation.

En effet, plusieurs lieux publiques et universités l'emploient pour accéder à un réseau comme l'Internet. À l'aide d'un ordinateur portable muni d'une carte sans fil, il est possible d'obtenir des informations sur le réseau comme la puissance du signal reçu (*RSS*) et l'identifiant de la station de base émettrice.

Le système de localisation proposé se base sur l'approche des signatures (méthode

du fingerprinting) avec des mesures de *RSS* et un algorithme de jumelage à réseaux de neurones artificielles (*ANN*).

La première partie du projet consiste à faire une localisation en trois dimensions (3D) en collectant les données sur 3 étages dans un canal statique. Les données de 1460 points de mesures sur 48 stations de base ont été enregistrées afin de créer une base de données. En injectant cette dernière dans l'*ANN*, une localisation peut être faite. Le diagramme général *a*) de la figure 1.1 donne les étapes effectuées pour l'estimation de la position dans un canal statique.

Les résultats précédents ne sont valides que si le canal est invariant dans le temps. L'utilisation de la base de données précédente dans un canal variant entraîne une forte dégradation de la résolution. Afin de minimiser cette dégradation, des travaux ont proposé de reprendre les mesures à plusieurs moments pour créer N bases de données différentes. Par la suite, il suffit d'utiliser N réseaux de neurones (un *ANN* par base de données) et un algorithme de décision pour effectuer la localisation.

La deuxième partie du projet consiste à effectuer une localisation en évitant de collecter les 1460 points de mesures à chaque variation du canal dans le temps. Pour ce faire, il suffit de prendre des mesures de quelques points de référence à l'instant  $t_x$ . Ensuite un algorithme de prédiction temporelle utilise ces mesures de référence et l'ancienne base de données afin de générer la nouvelle base de données au temps  $t_x$ . Finalement, cette base de données est employée dans l'*ANN* pour la localisation. Le diagramme général *b*) de la figure 1.1 donne les étapes effectuées pour l'estimation de la position dans un canal variant dans le temps.

L'algorithme de prédiction temporelle du système de localisation proposé est un algorithme d'arbre de décision. Ce dernier a permis de générer des bases de données en 2D et en 3D acceptables pour permettre la localisation. De plus, la localisation en 2D dans un canal variant à donner de bon résultats.

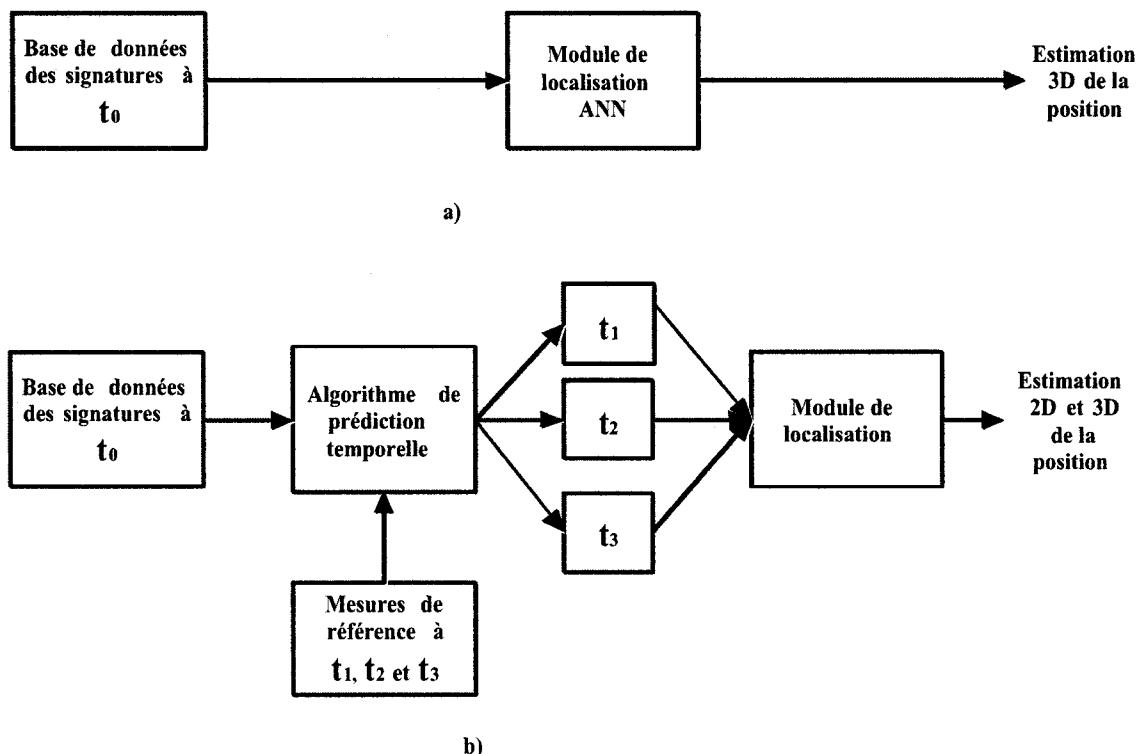


FIGURE 1.1 Diagramme général du travail effectué dans ce mémoire

## 1.2 Structure du mémoire

En plus de ce chapitre, le mémoire est constitué de quatre autres chapitres dont les contenus sont les suivants.

Le chapitre 2 contient les concepts de base ainsi que les travaux antérieurs qui ont permis la définition du projet. Les concepts de base permettent de mieux

comprendre la problématique au niveau de la localisation dans un milieu interne. Dans ce sens, la théorie de la propagation, les paramètres de détection, les algorithmes et méthodes de localisation et la définition du standard et système utilisé sont définis.

Le chapitre 3 contient les mesures en puissance des signaux reçus (*RSS*) dans un canal contenant peu ou pas de variation dans le temps et est divisé en trois parties. La première partie explique en détail la méthode de mesure, les équipements utilisés, la série de collecte de données et le traitement de ces données afin d'obtenir une base de donnée finale. La deuxième partie explique comment utiliser cette base de donnée avec le réseau de neurones afin d'effectuer des tests de localisation. Finalement, la dernière partie donne les résultats de localisation et leur analyse.

Le chapitre 4 propose un système de localisation qui prend en compte la variabilité du canal. Les méthodes de mesure, de traitement des données et de la génération des nouvelles bases de données sont aussi abordées. Finalement, le résultat et l'analyse sont inclus.

Le chapitre 5 conclut en résumant les résultats obtenus, en donnant les limitations du système de localisation et en proposant des travaux de recherches qui pourraient découler de ce mémoire.

## CHAPITRE 2

### CONCEPTS DE BASE ET TRAVAUX ANTERIEURS

Le mécanisme de localisation est généralement constitué de trois parties ; la détection des paramètres de localisation (RSS, AOA, DOA, TOA, TDOA ou RTT), l'utilisation d'un algorithme d'estimation de la position et visualisation et décision des données. Ce mécanisme a été proposé par Pahlavan, Li et *Mäkelä* [13]. Le diagramme suivant, développé à partir de l'article [13], explique le cheminement général de tout système de localisation.

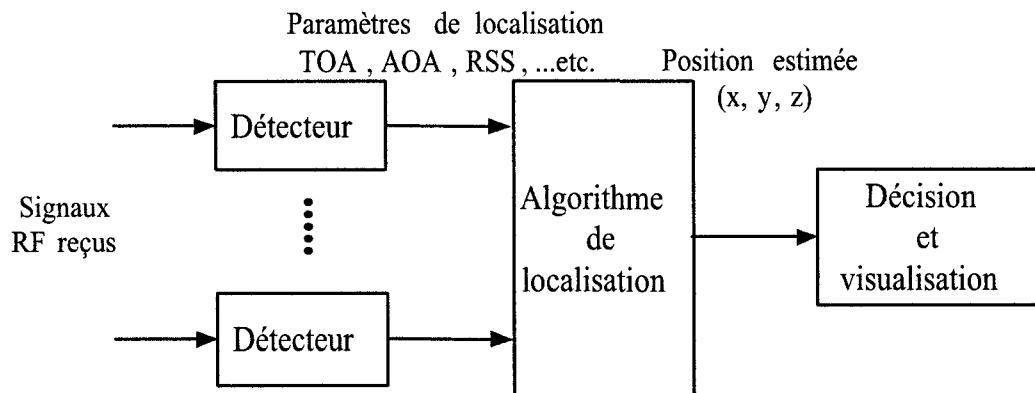


FIGURE 2.1 Diagramme fonctionnel d'un système de localisation dans un environnement interne.

Avant de définir ces étapes, les concepts de base sont abordés afin de mieux comprendre les phénomènes et mécanismes influençant un signal émis dans un environnement interne. La description des caractéristiques de l'environnement étudié et du système WLAN utilisé dans ce projet sont ajoutées. Finalement, les travaux antérieurs au système proposé sont mentionnés.

## 2.1 Concepts de base

### 2.1.1 Mécanismes de propagation

Dans un environnement interne, un signal émis est, dans la majorité des cas, reçu en plusieurs versions (effet trajets multiples) contenant des distorsions et des atténuations. Ces phénomènes, qui s'appliquent dans notre cas au système WLAN, sont causés par les mécanismes de propagation de base qui sont : la réflexion, la diffraction et la dispersion.

#### 2.1.1.1 La théorie de la réflexion spéculaire et diffuse

Un signal émis est réfléchi sur tout obstacle comme les murs, les fenêtres, le sol, le plafond, les bureaux...etc. Comme le montre la figure 2.2, à la rencontre d'un obstacle lisse, une partie du signal est réfléchie à un angle différent. Dans notre cas, à la réception, le signal qui a subit une ou plusieurs réflexions est souvent reçu sous forme de plusieurs versions avec des délais, angles d'arrivés et amplitudes différentes.

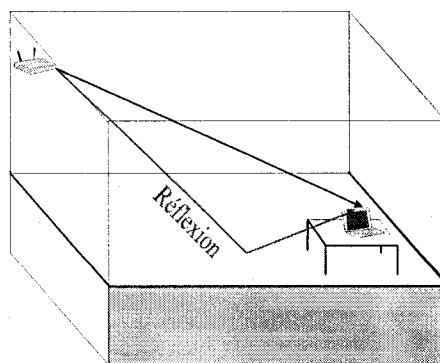


FIGURE 2.2 Phénomène de la réflexion spéculaire

À la rencontre d'un obstacle à surface rugueuse, le signal subit une réflexion diffuse. Dans ce cas, le signal est réfléchi plusieurs fois sous différents angles. La figure 2.3 explique le phénomène.

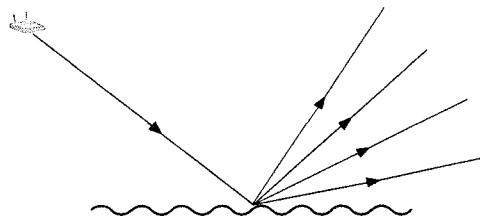


FIGURE 2.3 Phénomène de la réflexion diffuse

#### 2.1.1.2 La théorie de la diffraction

Un signal est diffracté à la rencontre d'un obstacle dont la dimension n'est plus infiniment grande par rapport à la longueur d'onde comme le coin d'une porte ou d'un bureau. La figure 2.4 explique ce phénomène. Dans un environnement interne, où il est rare d'avoir une vue directe entre l'émetteur et le récepteur, la diffraction est un phénomène majeur dans la propagation des signaux.

#### 2.1.2 Signaux LOS et NLOS

Lorsque l'émetteur a une vue directe (LOS) sur le récepteur, le signal dominant reçu représente, habituellement, la version non réfléchie ou dispersée et l'atténuation est faible (dans un environnement interne). Dans le cas où il n'y a pas de vue directe (NLOS), le signal reçu représente la version réfléchie, diffuse ou dispersée du signal émis. Ce dernier, contient des distorsions et le signal dominant est, le plus souvent,

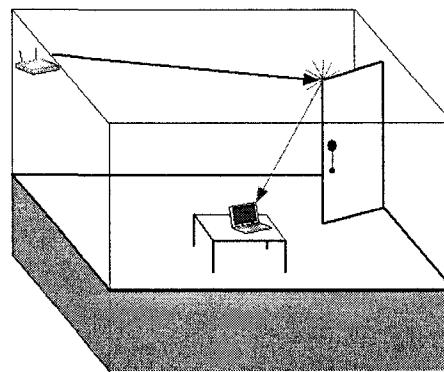


FIGURE 2.4 Phénomène de diffraction

fortement atténué. Dans un bâtiment comme celui étudié dans ce projet, les aires de repos contiennent les deux trajets LOS et NLOS. La figure 2.5 donne un exemple simple des deux trajets.

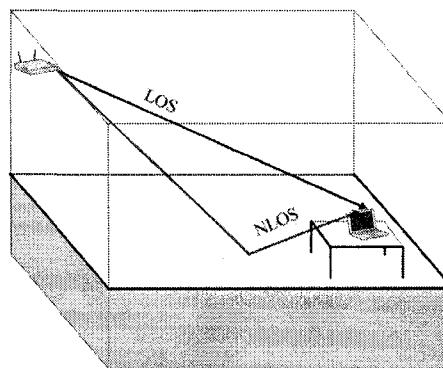


FIGURE 2.5 Exemple de réception du signal selon une vue directe (LOS) et une vue non directe (NLOS)

### 2.1.3 Phénomène multi-trajets (Multipath)

Le phénomène multi-trajets résulte lorsqu'un signal radio se propage par plusieurs chemins avant d'être reçu par une antenne. Dans un environnement externe, ce

phénomène est surtout présent dans une zone urbaine dense. Dans un bâtiment, à cause des différents obstacles, un signal est réfléchi, diffus et dispersé ce qui augmente le nombre de chemin de propagation avant d'arriver au récepteur. C'est pour cette raison que l'utilisation de la triangulation en intérieur donne de mauvaises précisions par rapport à un environnement externe.

#### 2.1.4 Phénomène d'évanouissement

Ce phénomène est décrit par des fluctuations rapides de l'enveloppe du signal dans un court intervalle de temps due aux interférences constructives et destructives. Dans un environnement interne, due au phénomène multi-trajets, l'évanouissement est plus important lors de la caractérisation du canal à petite échelle. Cependant, ce phénomène est négligeable en grande échelle [7].

### 2.2 Caractéristiques du canal

Le modèle de propagation du canal étudié (5<sup>e</sup>, 6<sup>e</sup> et 3<sup>e</sup> étages du pavillon Lassonde) a déjà été établi [1]. À grande échelle, le facteur d'atténuation défini dans la section 2.3.3 est différent selon les étages. Au 3<sup>ème</sup> étage ce facteur est de 2.8. Au 5<sup>ème</sup> étage, il est de 3.1 alors qu'au 6<sup>ème</sup> il est de 4.5. Les trois valeurs du facteur d'atténuation trouvées sont du même ordre de grandeur que celles trouvées dans la littérature [7], [8], [9]. D'après ces références, l'architecture du pavillon Lassonde est plus proche de celle d'un établissement commercial [9] puisque le pavillon comporte des aires ouvertes selon différents étages. D'une façon générale ces valeurs varient en fonction de l'architecture et la nature des obstacles rencontrés par le signal [7]. En effet, le facteur d'atténuation trouvé au sixième étage est un peu supérieure par rapport aux deux autres et s'explique par l'existence d'un grand nombre d'obstacles

entre le point d'accès et une bonne partie des points de mesures. La plus petite valeur du facteur d'atténuation est au 3ème étage et s'explique par le fait que l'architecture de l'espace choisi présente moins d'obstacles par rapport aux autres. La caractérisation à petite échelle varie selon l'emplacement des points d'accès (en LOS ou NLOS) et l'aire étudiée. En général, en NLOS, la distribution la plus proche est celle de Rayleigh alors qu'en LOS c'est la distribution de Rice qui prévaut, ce qui correspond à la définition de ces deux distributions [7], [8], [9].

### 2.3 Paramètres de localisation

Un signal reçu est défini selon 3 paramètres, sa puissance à la réception, le délai entre son émission et sa réception, et son angle d'arrivée. Cette section donne la définition de chacun de ces paramètres.

#### 2.3.1 Méthode basée sur les mesures d'angle d'arrivée

Cette méthode se base sur des mesures d'angles d'arrivées AOA (*Angle Of Arrival*) ou la direction d'arrivée du signal DOA (*Direction Of Arrival*) afin d'effectuer la localisation [12]. En théorie, deux émetteurs suffisent pour localiser le mobile. La figure 2.6 donne un exemple. En pratique, trois émetteurs et plus sont nécessaires afin d'éliminer une partie du phénomène multi-trajets, le bruit et d'augmenter la précision. L'utilisation de plusieurs mesures d'angles d'arrivée simultanément à partir de plusieurs émetteurs permet d'estimer la position à l'aide des relations géométriques. Ainsi une estimation de la localisation de la source est déterminée grâce à l'intersection des lignes définissant les angles d'arrivée obtenus. L'un des avantages de cette méthode est que la synchronisation entre le récepteur et l'émetteur est non nécessaire. En général, les techniques d'estimation se basent sur le

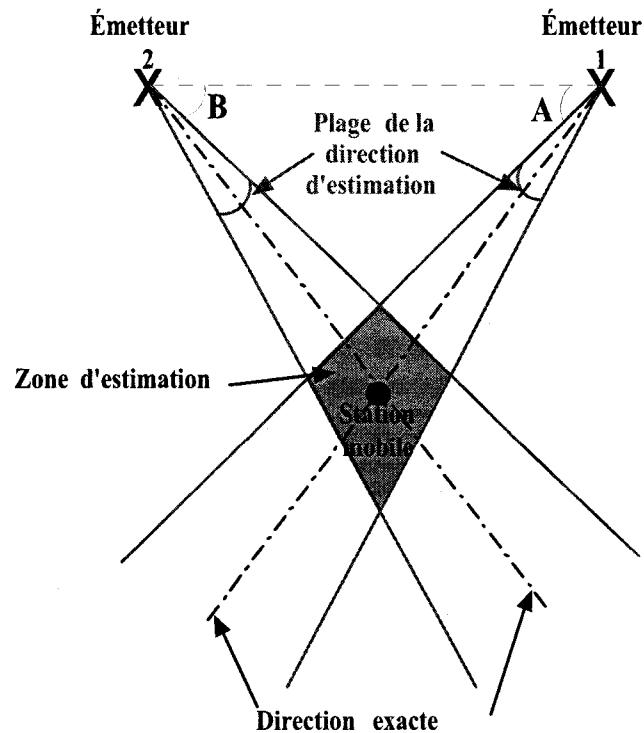


FIGURE 2.6 Exemple de localisation selon la méthode d'angle d'arrivée

signal reçu le plus puissant qui est supposé être en visibilité directe. Cependant, dans un environnement interne il est peu probable d'avoir une vue directe (LOS) comme signal dominant. De plus, en présence d'une vue directe entre les émetteurs et le récepteur, une seule petite erreur de mesure risque de fausser considérablement les résultats. C'est pour cela que cette méthode est rarement utilisée dans un environnement interne ou même dans une zone urbaine dense.

### 2.3.2 Méthode basée sur les mesures de délai d'émission / réception

Pour cette méthode, la localisation est possible avec trois émetteurs et une méthode d'estimation comme la triangulation, quelque soit le type de délai mesuré.

### 2.3.2.1 Mesure selon le TOA (Time Of Arrival) et le TDOA (Time Difference Of Arrival)

Le délai entre la transmission du signal et sa réception diffère selon la distance entre le récepteur et l'émetteur [12]. Dans un milieu externe, l'équation suivante permet de définir la distance entre les deux :

$$d = c \cdot (t_1 - t_0) = c \cdot TOA \quad (2.1)$$

où  $d$  représente la distance,  $c$  la vitesse de la lumière et  $TOA$  le délai de transmission. Dans un système WLAN, due aux limitations des ordinateurs permettant les mesures de délai de transmission, la résolution est de  $1\mu s$ . Avec cette résolution et l'équation 2.1, une localisation minimale de 300m est obtenue ce qui est non acceptable dans des bâtiments.

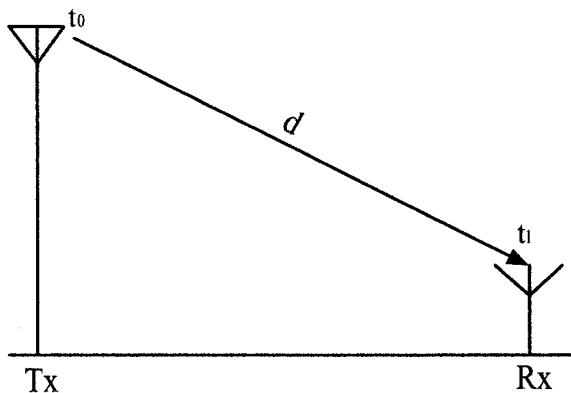


FIGURE 2.7 Exemple d'émission/réception d'un signal en TOA

Dans la figure 2.7, le Tx représente l'émetteur, le Rx le récepteur,  $t_0$  est le temps à l'émission du signal alors que  $t_1$  est le temps à la réception de ce dernier. Le désavantage de cette technique est qu'il faut une bonne synchronisation entre l'émetteur et le récepteur ce qui peut être difficile en pratique. Pour fin de localisation, il

faut au moins trois émetteurs qui permettent d'estimer la position d'un mobile en effectuant une triangulation. La technique du TDOA a le même principe de fonctionnement que la technique du TOA puisqu'elle est basée sur la mesure de distance entre la station de base et le mobile. Son avantage est que la synchronisation entre la station de base et le mobile n'est pas nécessaire. Seule la synchronisation par paires de station de base est utilisée. En effet, cette technique consiste à mesurer la différence du temps d'arrivée d'un signal émis par le mobile par deux stations de base. La précision de localisation est du même ordre que celle du TOA.

### 2.3.2.2 Mesure selon le RTT (Round-Trip Time)

Le RTT est un terme, souvent utilisé en réseaux sans fil (systèmes WLAN), qui représente le délai entre la transmission du signal et la réception d'un accusé de réception (ACK) au même point [6] [15]. La commande *ping* permet de l'obtenir facilement. La figure 2.8 explique le processus.

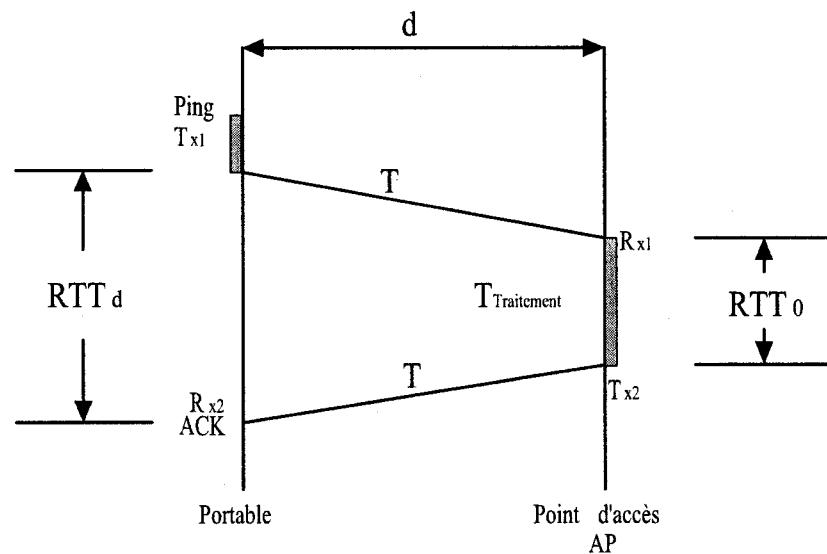


FIGURE 2.8 Exemple d'émission/réception d'un signal en RTT

D'après la figure 2.8, le mobile (ou l'ordinateur portable) joue le rôle de l'émetteur  $T_{x1}$  alors que le point d'accès (AP) joue d'abord le rôle du transmetteur  $R_{x1}$ . Après le traitement du paquet, l'AP envoie ( $T_{x2}$ ) un accusé de réception (ACK) au portable ( $R_{x2}$ ). Les équations suivantes donnent la distance  $d$  recherchée.

$$RTT_d = 2 \cdot T + T_{Traitement} \quad (2.2)$$

et

$$RTT_0 = T_{Traitement} \quad (2.3)$$

où  $RTT_d$  représente le délai à la distance  $d$ ,  $T$  le délai de transmission,  $T_{Traitement}$  le délai de traitement de l'information reçue afin d'envoyer l'ACK et  $RTT_0$  le délai de transmission lorsque la distance  $d$  est nulle (i.e l'émetteur est côte à côte avec le récepteur).

$$\Delta RTT = RTT_d - RTT_0 \quad (2.4)$$

La distance est donc calculée selon l'équation suivante :

$$d = c \cdot \left( \frac{\Delta RTT}{2} \right) \quad (2.5)$$

où  $d$  représente la distance et  $c$  la vitesse de la lumière.

La synchronisation entre l'émetteur et le récepteur est non nécessaire étant donné que le  $RTT_d$  et  $RTT_0$  sont pris au niveau du mobile. De plus le RTT permet d'augmenter la résolution minimale pour le calcul de la distance  $d$ . En effet, pour cette méthode, le temps minimal est de  $0.5\mu s$  au lieu de  $1\mu s$  (en TOA) ce qui diminue la localisation minimale de moitié. En ajoutant un compteur au portable comme indiqué dans l'article [6], la résolution pourrait s'améliorer selon la fréquence

du compteur (dans l'article elle est de 22ns donc une résolution brute de 3.4m au lieu de 150m).

### 2.3.3 Méthode basée sur les mesures de la puissance reçue

Cette technique est basée sur la mesure de la distance à partir de l'atténuation du signal de puissance reçu RSS (*Receive signal strength*) au mobile [12] [7]. En effet, l'atténuation du signal RSS est fonction de la distance séparant l'émetteur et le récepteur. En connaissant le modèle d'atténuation de parcours (*Path-Loss* :  $PL$ ) du canal de propagation, et en prenant en compte l'effet d'ombrage, il est possible de calculer la distance.

$$PL_{dB} = \overline{PL}(d_0)_{dB} + 10n \log\left(\frac{d}{d_0}\right) + X_\sigma \quad (2.6)$$

où  $PL$  indique l'atténuation de parcours à une distance  $d$  entre la station de base et le mobile.  $\overline{PL}(d_0)$  représente la moyenne de l'atténuation du signal à la distance de référence  $d_0$ .  $n$  est le facteur d'atténuation et  $X_\sigma$  représente une variable aléatoire Gaussienne de moyenne nulle et d'un écart type  $\sigma$  difficile à déterminer en pratique.

Après avoir mesuré cette distance, la position du récepteur est estimée en utilisant au moins trois émetteurs fixes et une technique comme la triangulation. Cependant, dans un environnement interne, les limitations de localisation sont liées principalement aux phénomènes multi-trajets et au NLOS. Dans le cadre de ce mémoire, les mesures effectuées sont basées sur les puissances reçues donc tout le processus de mesures, des limitations et avantages sont abordés dans les chapitres suivants.

## 2.4 Algorithmes de détection et d'estimation

### 2.4.1 L'approche des signatures

L'approche des signatures ou la méthode du fingerprinting permet, d'enregistrer l'empreinte du canal et de l'utiliser afin de localiser l'usager. Cette approche est constituée de deux phases ; la collecte de données et l'estimation de la localisation.

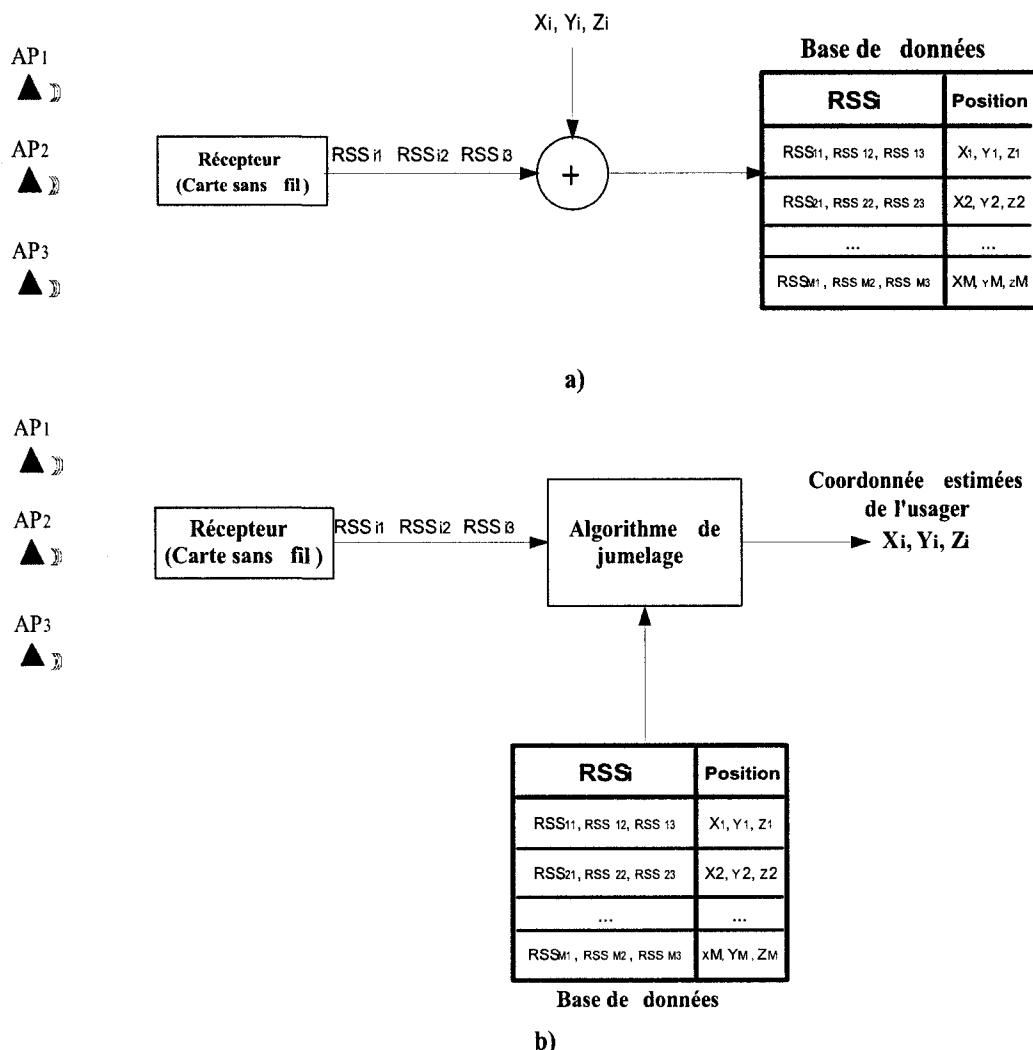


FIGURE 2.9 Schéma du processus de localisation basée sur l'approche des signatures pour un système WLAN, a) phase de collecte de données, b) phase de localisation

La première phase permet de créer une base de données contenant les signatures de l'usager selon les paramètres de détection en fonction de sa position réelle. Lors de la deuxième phase, un algorithme de jumelage est utilisé et dont le rôle est d'identifier la plus proche information d'empreinte (de la base de données) par rapport à celle mesurée et d'en estimer la position de l'usager [11].

Le schéma de la figure 2.9 donne un exemple d'utilisation de cette approche selon des mesures de puissances RSS dans un système WLAN avec trois points d'accès. L'utilisation de cette approche est donnée en détail dans la section 3.2.2.

#### 2.4.2 La technique de triangulation

Cette technique permet de positionner un usager en estimant la distance par rapport à trois points selon l'un des paramètres de localisation définis dans la section 2.3. L'utilisation de la triangulation dans la localisation en extérieur est largement répandue surtout depuis l'arrivée des systèmes GPS dans le marché. Dans ce système, la triangulation est faite en détectant le délai de transmission entre le récepteur et l'émetteur. En théorie, trois satellites suffisent pour une localisation en trois dimensions. Cependant, en pratique un quatrième satellite est nécessaire afin compenser le décalage de l'horloge du récepteur par rapport aux satellites et afin d'augmenter la précision [14]. Dans cette technique, les coordonnées de l'usager sont estimées grâce à l'intersection des trois sphères décrivant le temps écoulé par chacun des signaux émis par le satellite. Ainsi, les coordonnées sont fournies en longitude, latitude et altitude.

Dans un milieu interne, lors de l'existence du LOS entre les émetteurs et le récepteur, la technique de triangulation peut donner des résultats acceptables. Cependant, dans des bâtiments comprenant de nombreux obstacles où l'effet LOS n'est pas présent et avec l'existence du phénomène multi-trajets, cette technique ne donne pas les précisions attendues.

### 2.4.3 Méthode des réseaux à neurones artificiels

Le système de localisation proposé utilise la méthode de neurones artificiel ANN (*Artificial Neural Network*) comme algorithme de jumelage [10] [20] [22]. Cette méthode permet de modéliser le comportement non linéaire entre la position du mobile et les signatures RSS. Le réseau de neurones est composé d'éléments de neurones artificiels simples qui opèrent en parallèles et qui sont inspirés par le système nerveux biologique [17].

La figure 2.10 donne la structure de base d'un neurone et l'architecture d'un réseau de neurones à une couche. L'ANN peut contenir une ou plusieurs couches en série et peut être à circulation de l'information non récurrente (*Feed Forward*), ou récurrente (*Feed Back*).

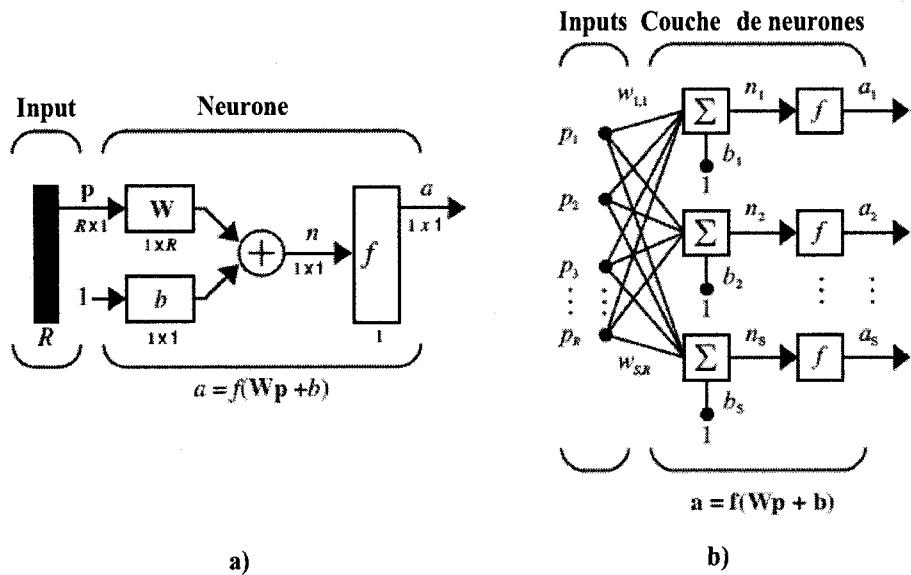


FIGURE 2.10 a) Structure de base d'un neurone, b) Architecture d'un réseau de neurones

$p$  est le vecteur d'entrée à  $R$  éléments,  $W$  représente le vecteur de poids à  $R$  éléments et  $S$  neurones par couche,  $b$  est le biais alors que  $a$  est la sortie estimée.

$f$  est la fonction de transfert du réseau de neurone. Cette fonction peut être de type échelon, linéaire, Log-Sigmoïde, de base radiale,...etc. L'utilisation d'un ANN se fait en deux phases ; la phase d'apprentissage ou d'entraînement et la phase d'estimation ou d'appel.

#### 2.4.3.1 Phase entraînement

Cette phase, dans le cas de l'utilisation de l'approche des signatures, permet de former un ensemble d'empreintes en fonction des positions réelles du mobile par apprentissage. Le schéma a) de la figure 2.11 permet de mieux comprendre cette phase. L'entrée est dans ce cas les puissances reçues RSS par le mobile qui ont

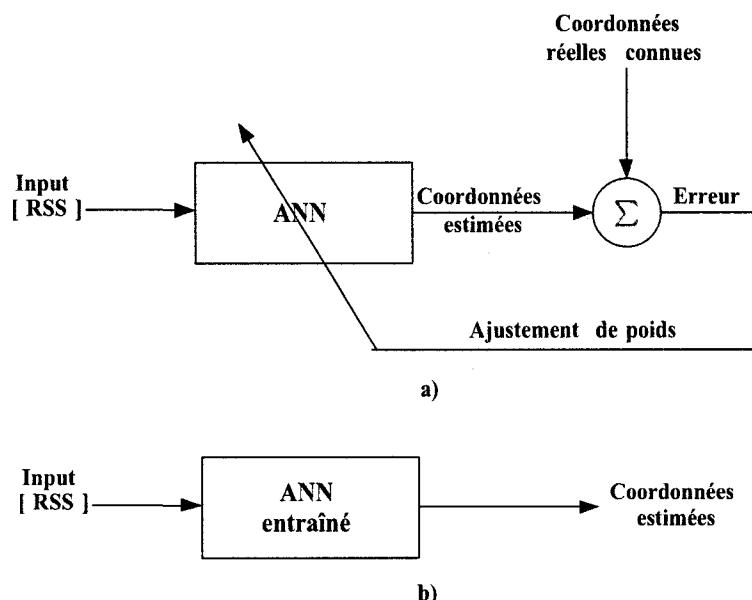


FIGURE 2.11 a) Phase d'entraînement d'un ANN, b) Phase estimation

été émises par les points d'accès par rapport à une position donnée. Le réseau de neurones estime les coordonnées et les compare avec les coordonnées réelles du mobile afin de calculer l'erreur d'estimation. Ensuite, avec l'aide de cette erreur et en se basant sur la minimisation de la fonction de performance, les poids et les

biais sont ajustés. Cette phase est généralement effectuée hors ligne (*Offline*) et, dans le cas de la méthode du fingerprint, permet de créer une base de donnée qui permettra d'estimer la position durant la phase d'appel.

#### 2.4.3.2 Phase estimation

Durant la phase d'appel, les puissances reçues par le mobile et obtenues à partir d'une position spécifique sont appliquées à l'entrée du réseau, entraîné durant la première phase et agissant comme algorithme de jumelage. La sortie donne les coordonnées estimées du mobile. Cette phase est appelée en temps réelle (*Online*) afin d'effectuer une navigation par exemple. Le schéma b) de la figure 2.11 permet de mieux comprendre cette phase.

### 2.5 Système WLAN

Le système de localisation proposé utilise le réseau sans fil WLAN (*Wireless Local Area Network*) comme système de détection. Ce système est employé dans les ordinateurs portables, les assistants personnels PDA (*Personal Digital Assistant*) et des périphériques de bureau comme les imprimantes car il permet un accès facile, rapide et peu coûteux aux serveurs. Cette technologie utilise le standard internationale *IEEE802.11* [16] qui permet la transmission de signaux de type radiofréquence à haut débit (11 Mbit/s en 802.11b et 54 Mbit/s en 802.11a/g). Ce standard est basé sur les couches de base du modèle OSI (*Open System Interconnection*) pour un réseau sans fil utilisant des ondes électromagnétiques. Ces couches sont la couche physique et la couche de liaison de données. Cette dernière est constituée de deux sous couches, la couche de contrôle de la liaison logique LLC (*Logical Link Control*) et la couche de contrôle d'accès au support MAC (*Medium*

*Access Control*). Le standard *IEEE802.11* peut utiliser les bandes de fréquences libre 2.4Ghz (*IEEE802.11b/g*) et 5.8Ghz (*IEEE802.11a*).

Le standard offre le choix entre deux architectures : Ad hoc et celle basée sur une infrastructure. L'architecture Ad hoc permet la communication entre deux stations de manière à ce que chacune d'elle opère de façon autonome afin d'assurer sa connectivité et la connectivité des autres stations. L'Ad hoc est adaptée aux besoins de communications éphémères ou sur des scènes mouvantes nécessitant un déploiement rapide. L'architecture la plus utilisé est celle basée sur une infrastructure, par exemple, pour assurer la liaison entre un point d'accès (comme un routeur) et un ordinateur portable. Ce type d'architecture est adapté pour assurer une connexion dans des lieux clos comme des bureaux et des universités.

Pour l'architecture basée sur les points d'accès, les informations sur la communication proviennent des trames émises par les points d'accès [29]. Parmi ces trames, on trouve des trames d'authentification, d'association, d'acquittement, de balise,... etc. Afin d'obtenir l'information voulue pour ce projet, la trame à collecter est la trame de balisage ou *Beacon*. En effet, le point d'accès émet périodiquement ce type de trames (à une intervalle de 100ms) afin de signaler sa présence, et pour émettre des informations comme son identifiant SSID (*Service Set IDentifier*). Les balises sont captées par les clients se trouvant dans la couverture radio du point d'accès émettant. Quand au client, ce dernier scrute en permanence tous les canaux disponibles pour capter ces balises qui lui permettent de déterminer le point d'accès avec lequel il est plus avantageux de s'associer.

Dans le cadre de ce projet, l'utilisation de l'architecture basée sur une infrastructure a été favorisée puisqu'il est nécessaire d'avoir des points d'accès fixes pour effectuer la localisation. Le standard utilisé est le IEEE802.11g qui opère à 2.4GHz à un débit allant jusqu'à 54Mbit/s. Les mesures à prendre pour effectuer la locali-

sation sont de type RSS en dBm pour chaque point d'accès. Cette information est disponible grâce à la carte sans fil du portable qui mesure le RSS qui a été émis par le point d'accès identifié grâce à la trame de balisage.

## 2.6 Travaux antérieurs

Bien que la technique de localisation en extérieur est bien implémentée et utilisée dans un grand nombre de domaines, la localisation en intérieur est moins développée. Cependant, un grand nombre de recherche s'effectuent afin de permettre un positionnement rapide et précis selon le coût et les besoins. Présentement, ces travaux se divisent en deux parties ; des travaux orientés selon la technologie de détection par la nature des signaux émis et reçus (Radio fréquence, Ultrasons, Infrarouge, ...etc.) et des travaux selon la technique de positionnement qui englobe les paramètres (délai, puissance et angle d'arrivée) et méthodes d'estimations (approche des signatures, triangulation, calculs géométriques, ...etc.).

La technologie de détection englobe le type de signal reçu ainsi que l'infrastructure qui permet son émission et sa réception. Ces paramètres influencent grandement la précision de localisation puisque chaque type de signal a une certaine portée et chaque infrastructure a ses propres limitations. Plusieurs systèmes ont été développés et proposés selon le type de signal comme RADAR [21], ACTIVE BAT [27], CRICKET [24], ACTIVE BADGE [26], et LANDMARC [28].

a) RADAR : Conçu par la compagnie *Microsoft* en 2000, il utilise des signaux radio fréquences via le système WLAN et le standard *IEEE802.11*. En utilisant au moins trois points d'accès, une base de données contenant les puissances mesurées aux points d'accès en fonction de la position du mobile est enregistrée ce qui forme l'ensemble des signatures. Ensuite, lorsque le mobile envoie un signal aux APs, les

puissances reçues à leurs niveau sont comparées à la base de donnée. Finalement, un algorithme de jumelage de type K-plus proches voisins (KNN *K-Nearest Neighbors*) est utilisé pour minimiser la distance Euclidienne par rapport à la signature actuelle et, enfin, déduire la position du mobile. Puisque ce système utilise l'approche des signatures, son point faible est qu'il n'est pas robuste par rapport à la variation temporelle du canal [21].

b) ACTIVE BADGE : Ce système est l'un des premiers à avoir été proposé. Il nécessite un badge porté par l'usager et des récepteurs installés au plafond. Il est basé sur l'émission, au niveau du badge, d'un signal infrarouge durant un dixième de secondes à chaque 15s. Les récepteurs sont installés dans chaque pièce et sont reliés entre eux afin de connaître l'emplacement du mobile. Finalement, ce système estime la position de l'usager en déterminant le récepteur le plus proche. Le désavantage de ce système est qu'il nécessite l'existence d'une visibilité directe LOS entre le badge et le récepteur le plus proche, ce qui est rarement le cas dans un bâtiment. De plus, le badge doit être toujours porté sur l'extérieur de l'habillement pour garder la condition du LOS. Mais le plus grand problème de ce système est le type de signal car les signaux IR sont perturbés par la présence de la lumière du jour [26].

c) ACTIVE BAT : Ce système combine les signaux Radio fréquences et Ultrasons. Il nécessite un émetteur Bat au niveau du mobile et des récepteurs installés au plafond. L'émetteur Bat envoi un signal ultrason. Lorsque ce dernier est détecté, un signal RF est émis par les récepteurs vers un ordinateur. Le PC, sert ainsi de centre nerveux au système de récepteurs, il émet vers chacun des récepteurs un signal d'initialisation via la liaison sérielle. Ce signal d'initialisation est émis en même temps que les signaux émis par les récepteurs à destination du Bat à localiser. Puis, le PC interroge chacun des récepteurs composant le réseau, en récupérant les intervalles de temps entre le signal d'initialisation et celui de l'ultrason. La position du Bat est obtenue alors par latération. Son point fort est sa précision qui est de

l'ordre du cm. Son désavantage est que ce système est très coûteux et nécessite beaucoup de ressources pour le déploiement des récepteurs et leurs connexions avec le système de contrôle centralisé (Le contrôleur est le PC) [27].

d) CRICKET : Ce système utilise aussi les signaux RF et les Ultrasons. Son principe est le même que celui du système Active Bat. La différence est que le traitement et le contrôle sont effectués par la station mobile. Donc ce système est dit décentralisé. Les émetteurs placés au plafond envoi les deux types de signaux au même temps, ainsi le mobile reçoit successivement des signaux RF et ultrasons. Ensuite, en effectuant une corrélation de ces deux signaux le TDOA est extrait. Ceci permet d'estimer la distance le séparant de l'émetteur qui a émis ces deux signaux. En utilisant plusieurs émetteurs, on peut déterminer la position du mobile. L'avantage de ce système est son faible coût par rapport au système Active Bat. Son point faible est le risque d'interférences et la nécessité d'une visibilité directe entre l'émetteur et le récepteur [24].

e) EKAHAU : Ekahau est une compagnie qui développe et vend, entre autre, des logiciels et des équipements de localisation et de navigation utilisant le réseau WiFi dans un environnement interne. Le logiciel proposé utilise les plans généraux de l'environnement interne et des mesures RSS pris par la méthode de l'approche des signatures afin d'effectuer la calibration. Ensuite, en attachant un badge ou tag à un mobile ou un usager il est possible de le suivre. En effet, le tag envoi l'information sur ses puissances reçue par rapport aux points d'accès à la station maître afin que cette dernière puisse le localiser et le suivre sur les plans de l'environnement. D'un autre coté, il est possible d'effectuer les même opérations avec ce logiciel afin de localiser un PDA ou un portable puisque tous les signaux envoyés et émis sont de nature radio fréquence et utilisant le protocole IEEE802.11. Le point faible de ce logiciel est qu'il n'est performant que si les points d'accès sont placés selon les conditions du logiciel. De plus, la calibration doit se faire périodiquement ou à

chaque changement des plans de l'environnement puisque le logiciel est basé sur l'approche des signatures.

La technologie de positionnement englobe les paramètres de détection (AOA, TOA/TDOA/RTT, RSS) qui les classifie selon les méthodes d'estimation. C'est à dire, le positionnement selon l'estimation de la distance, de l'angle d'arrivée, le fingerprinting,...etc. Il arrive souvent de proposer des techniques hybrides qui englobe deux paramètres de détection afin d'augmenter la précision. En plus des exemples mentionnés précédemment, d'autres sont donnés ci-dessous.

Dans l'article de A. Hatami et K. Pahlavan [32], une méthode hybride utilisant le TOA et le RSS a été utilisée. L'algorithme de jumelage proposé est constitué de réseaux de neurones. Les motivations qui ont conduit à l'utilisation de cette méthode hybride est le fait que le TOA est plus sensible au phénomène multi-trajets et au NLOS alors que le RSS l'est moins. De plus, le TOA est plus performant lors de l'existence d'un chemin direct dominant (donc d'un LOS) contrairement au RSS. Ces deux particularités permettent, lors de l'utilisation d'une méthode hybride qui utilise l'un ou l'autre paramètre selon les caractéristiques du canal en un point donné, d'augmenter la précision.

L'article de C.H Lim, B. Poh Ng et D. Duan [31] propose d'ajouter des antennes intelligentes dans un portable utilisant un système WLAN afin de déterminer l'angle d'arrivée et la méthode de triangulation pour se positionner. Ce système utilise trois antennes intelligentes à 4 éléments par antenne. Ces antennes, effectuent un balayage de  $-60^\circ$  à  $+60^\circ$  en mesurant le RSS à chaque changement d'angle. Ensuite, l'information sur le RSS par angle de mesure est envoyée vers une station qui effectue le traitement des données permettant de déterminer le AOA. Durant le traitement, une méthode de filtrage moyenne est utilisée. Finalement, avec la triangulation et l'AOA, la position du mobile est estimée. La précision donne un

CDF de 90% à 0.78m ce qui est une excellente précision. Son inconvénient est que l'équipement utilisé est encombrant et difficile à déplacer. De plus, 14 points tests ont été effectués dans un laboratoire informatique.

L'article de J. Ying et L. Ni [3] touche au problème de la variation du canal lors de l'utilisation de l'approche des signatures comme méthode d'estimation. En effet, comme expliqué dans la section 2.9, lorsque le canal est variable les données mesurées pour créer la base de données n'est plus valide. Ying et Ni proposent de mettre à jour la base de données initiale au temps  $t_x$  avec l'aide de quelques points de référence et une technique d'analyse de régression. Après avoir tester plusieurs techniques, les auteurs de l'article ont remarqué que c'est le modèle d'arbre de décision qui donne le meilleur résultat. Cette approche est expliquée en détail dans le chapitre 4 de ce mémoire.

## CHAPITRE 3

### MESURES EN RSS POUR L'APPROCHE DES SIGNATURES DANS UN CANAL STATIQUE

Dans cette partie, le travail consiste à prendre des mesures de puissances de l'environnement étudié au niveau du réseau local sans fil (WLAN) qui utilise le standard IEEE802.11g. Cette puissance est mesurée en dBm à 2.4Ghz au niveau de la carte sans fil d'un ordinateur portable qui joue le rôle du récepteur alors que le point d'accès (AP) représente l'émetteur et ce dernier est généralement un routeur, ou un relais.

Dans ce chapitre, les paramètres de mesure incluant la méthode de prise de mesures, la création de la base de données résultante et l'utilisation de cette dernière pour la localisation sont abordées.

#### 3.1 Description du système de localisation proposé

Le système proposé consiste à utiliser la méthode d'approche des signatures  *Fingerprinting* en mesurant la puissance du signal reçu au niveau de la carte sans fil d'un ordinateur portable à différentes positions.

Cette approche permet de définir une base de données contenant les coordonnées 3D (en mètres) et le RSS de ces coordonnées pour chaque point d'accès. Par la suite, un système à réseau de neurones artificielles  *ANN* permet d'estimer la position d'un mobile selon les données précédentes.

Ce type de système a déjà été testé dans le Pavillon J. Armand Bombardier de l'École Polytechnique [10] pour 431 points de mesure. Une comparaison avec les

résultats obtenus dans ce chapitre sera faite dans la partie d'analyse des résultats. La série de mesures a été faite au Pavillon Lassonde de l'École Polytechnique de Montréal dans un canal statique. Le détail de la méthode utilisée est décrit dans la prochaine section.

## 3.2 Méthodologie

### 3.2.1 Prise de mesures

Cette étape a permis de caractériser le canal et de faire sa topographie suivant les puissances reçues. Les mesures ont été prises avec l'aide d'un étudiant[1] au cours de l'été 2006. Premièrement, une description du signal, des standards, du matériel et des programmes utilisés est donnée. Ensuite, le traitement des mesures prises ainsi que le tri effectué pour l'optimisation des résultats via les algorithmes proposés sont abordés.

#### 3.2.1.1 Paramètres d'erreur

Avant de donner les paramètres des mesures prises, il faut savoir que l'erreur de l'estimation est dépendante de plusieurs paramètres comme : le nombre de points de mesure, la distance entre ces points, la sensibilité de la carte sans fil, le nombre d'AP mesurés dans un même espace, le type d'algorithme utilisé et la variation du canal.

Comme dit précédemment, la méthode utilisée est celle de l'approche des signatures. Cette dernière donne, en général, une estimation plus précise de la

localisation que la méthode de triangulation standard. Par contre, elle est connue pour son manque de robustesse dans un canal variant dans le temps. En effet, la base de données qui définit cette méthode n'est plus valable si le canal varie fortement. La modification de l'architecture, comme ajouter ou éliminer un mur, et le fort achalandage dans le bâtiment fait varier la puissance du signal reçu de manière significative. C'est pour cela que la collecte de données s'est faite dans un canal semi-statique (pas ou peu de mouvements dans le pavillon). De plus, il est peu probable que l'architecture du pavillon soit modifiée dans les prochaines années.

Afin d'effectuer une estimation de la localisation, il faut que le récepteur reçoive le signal de plusieurs points d'accès à la fois car la précision est proportionnelle au nombre de signaux reçus en un même point de mesure. Habituellement, avec la réception de signaux de trois points d'accès ou plus dans un canal, on peut obtenir une bonne estimation de la localisation [10]. De plus, le nombre de APs reçu est dépendant de la sensibilité de la carte sans fil étant donné que plus la sensibilité de la carte est grande, plus le nombre d'APs dans un même point de mesure risque d'être grand.

Le nombre de points de mesure suit de près la distance entre ces points dans une même aire puisque plus le nombre de mesures est grand dans un espace, plus la distance entre les mesures est petite. Il est à noter que dans des mesures à grandes échelles, l'espacement entre deux points de mesure est proportionnelle à l'erreur de localisation.

### 3.2.1.2 Paramètres de mesure

Ce pavillon a un environnement différent des entreprises et bâtiments standards. En effet, c'est un bâtiment à aires ouvertes entre les étages qui contient moins d'obstacles que les immeubles courants. Cette caractéristique permet de collecter plus de signaux dans une même zone (jusqu'à 7 signaux à la fois au lieu des trois nécessaires à une localisation minimale) ce qui diminue l'erreur d'estimation.

Pour ce projet, 1460 points de mesure ont été collectés sur trois étages dans les couloirs, aires de repos, tables et escaliers des 3<sup>e</sup> (figure I.3), 5<sup>e</sup> (figure I.2) et 6<sup>e</sup> (figure I.1) étages du Pavillon Lassonde de l'École Polytechnique de Montréal où 48 points d'accès (APs) ont été détectés. Le Sud est l'orientation considérée pour tous les points de mesure.

La distance entre les mesures a été définie à 1m dans les couloirs et les aires de repos des 5<sup>e</sup> et 6<sup>e</sup> étages alors qu'au 3<sup>e</sup> étage, qui est un environnement constitué de peu d'obstacles au sol, la distance de séparation a été choisie à 1.5m. Le canal était semi-statique (i.e : peu de personnes qui circulent dans le pavillon durant les mesures et pas d'obstacles majeurs dans l'aire de mesure).

### 3.2.1.3 Équipements de mesure

La section suivante donne le matériel et le programme utilisés pour la collecte du niveau de la puissance reçue.

#### 3.2.1.3.1 Matériels de mesure :

L'un des buts de ce projet est d'utiliser les installations déjà existantes et un matériel peu coûteux. C'est pour cela que le réseau local sans fil WLAN a été

choisi. Ce dernier est très répandu et se trouve généralement dans les universités et entreprises ainsi que dans plusieurs lieux publics. Les mesures ont été prises selon le standard IEEE802.11g.

Le signal mesuré représente la puissance reçue en dBm au niveau d'une carte sans fil d'un ordinateur portable. L'émetteur est un point d'accès (AP) qui est, dans ce cas, une station avec une adresse MAC (Medium Access Control) fixe. La matériel utilisé pour ce projet est :

- Un ordinateur portable de marque Dell Inspiron 6400 avec un processeur Core Duo Genuine Intel(R) CPU T2400 @ 1.83GHz.
- Une carte sans fil interne en guise de récepteur de type Intel®PRO/Wireless3945ABG qui a une sensibilité de -93dBm.
- Le réseau sans fil de l'école Polytechnique (réseau Hermes) constitué de points d'accès comme émetteurs.

En général, n'importe quel ordinateur ou portable muni d'une carte sans fil qui utilise le protocol IEEE802.11 peut mesurer la puissance reçue d'un point d'accès du réseau WLAN. Cependant, il faut choisir une carte qui a une grande sensibilité (inférieure à -90dBm) afin de détecter le plus d'AP dans une même zone.

### 3.2.1.3.2 Programme de mesure :

Afin d'obtenir la puissance mesurée, plusieurs logiciels ou applications sont disponibles NDIS, WMI, Netstumbler, WR API, ...etc.). Le *Netstumbler* [25] a été utilisé pour la collecte de données. Ce logiciel donne le niveau de puissance et du bruit et a été configuré afin de collecter la puissance reçue à chaque 0.5 secondes [1].

La figure 3.1 donne un exemple de fenêtre de Netstumbler. La série de mesures a été prise dans un environnement Windows mais aurait pu être prise dans un environnement Linux.

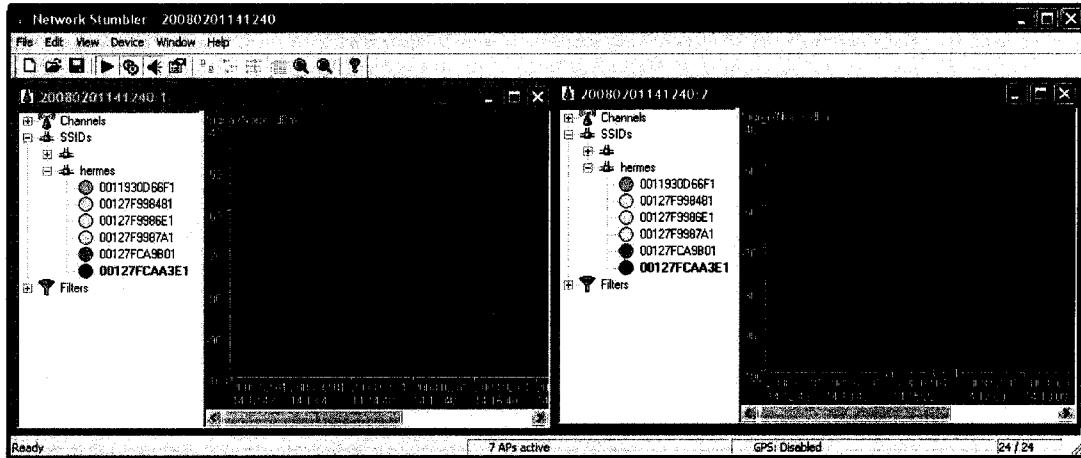


FIGURE 3.1 Exemple d'une fenêtre Netstumbler

### 3.2.1.4 Protocole de mesures

L'ordinateur portable est déposé sur une chaise pivotante afin de faciliter la collecte de données. Durant la campagne de mesures, le portable avait le Sud comme orientation.

Le protocole de la collecte de données utilisé est simple :

- Se positionner sur un point de mesure prédéfini ;
- Lancer la collecte des mesures via le programme Netstumbler pendant 30 secondes ;
- Enregistrer le fichier résultant en incluant les coordonnées cartésiennes du point de mesure ;
- Passer au point de mesure suivant ;

À la fin des mesures, la base de donnée contient 1460 fichiers incluant chacun les coordonnées des points de mesure et environs 60 puissances reçues dans le temps pour chaque AP détecté par la carte. La figure 3.2 donne un exemple de fichier obtenu avec Netstumbler.

Ref_03.txt - Bloc-notes						
Fichier	Édition	Format	Affichage	?		
# \$Creator: NetworkStumbler version 0.4.0						
# \$Format: wi-scan with extensions						
# \$Date: 2008-08-16						
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:42 }	17:29:11 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9c:c0 }	17:29:11 (GMT)	[ 6 55 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:f3 }	17:29:11 (GMT)	[ 8 57 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:43 }	17:29:11 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	SCRIBENS	ad-hoc { 02:04:23:f5:f4:cb }	17:29:11 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9b:el }	17:29:11 (GMT)	[ 9 58 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:53 }	17:29:11 (GMT)	[ 15 64 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:51 }	17:29:11 (GMT)	[ 13 62 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d1 }	17:29:11 (GMT)	[ 34 83 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d2 }	17:29:11 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:40 }	17:29:11 (GMT)	[ 11 60 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:f1 }	17:29:11 (GMT)	[ 14 63 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d2 }	17:29:11 (GMT)	[ 30 99 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:50 }	17:29:11 (GMT)	[ 56 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:44 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d3 }	17:29:12 (GMT)	[ 8 57 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:43 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:40 }	17:29:12 (GMT)	[ 11 60 49 ]
N 0.0000000	E 0.0000000	( )	SCRIBENS	ad-hoc { 02:04:23:f5:f4:cb }	17:29:12 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:53 }	17:29:12 (GMT)	[ 13 62 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9c:c0 }	17:29:12 (GMT)	[ 14 63 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:f1 }	17:29:12 (GMT)	[ 14 63 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:42 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d3 }	17:29:12 (GMT)	[ 50 99 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d2 }	17:29:12 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:53 }	17:29:12 (GMT)	[ 15 64 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:50 }	17:29:12 (GMT)	[ 7 56 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9b:el }	17:29:12 (GMT)	[ 9 58 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:41 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:53 }	17:29:12 (GMT)	[ 15 64 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d3 }	17:29:12 (GMT)	[ 50 99 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:41 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:f1 }	17:29:12 (GMT)	[ 14 63 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9c:c0 }	17:29:12 (GMT)	[ 6 55 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:d2 }	17:29:12 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	SCRIBENS	ad-hoc { 02:04:23:f5:f4:cb }	17:29:12 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:99:87:51 }	17:29:12 (GMT)	[ 13 62 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:43 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:68:40 }	17:29:12 (GMT)	[ 11 60 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:69:42 }	17:29:12 (GMT)	[ 12 61 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9c:c0 }	17:29:13 (GMT)	[ 6 55 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:69 }	17:29:13 (GMT)	[ 33 82 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:12:7f:ca:9b:el }	17:29:13 (GMT)	[ 9 58 49 ]
N 0.0000000	E 0.0000000	( )	BSS	{ 00:11:93:0d:66:f3 }	17:29:13 (GMT)	[ 8 57 49 ]

FIGURE 3.2 Exemple d'un fichier enregistré grâce à Netstumbler

### 3.2.2 Obtention de la base de donnée (Approche des signatures)

Comme dit précédemment, la technique du fingerprint représente la collecte de la base de données à utiliser, dans notre cas, dans la phase d'entraînement du réseau de neurones.

La base de données qui sera utilisée dans l'algorithme d'estimation doit être sous forme de matrice contenant, pour chaque point de mesure, les coordonnées cartésiennes en trois dimensions et la puissance moyenne en dBm de chacun des 48 APs.

Donc on aura la matrice de 51 par 1460 suivante :

$$\mathbf{M}_{(1460,51)} = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1j} & \dots & m_{151} \\ m_{21} & m_{22} & \dots & m_{2j} & \dots & m_{251} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{i1} & m_{i2} & \dots & m_{ij} & \dots & m_{i51} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{14601} & m_{14602} & \dots & m_{1460j} & \dots & m_{146051} \end{bmatrix} \quad (3.1)$$

Avec  $i$  le numéro du point de mesure  $1 \leq i \leq 1460$

et  $j$  le numéro du point d'accès  $1 \leq j \leq 51$

Cette matrice est divisée en deux parties, les trois premières colonnes représentent les coordonnées cartésiennes en 3D (en mètre) des points de mesure alors que la deuxième partie est la puissance moyenne par AP de ces points. La définition de chaque partie de la matrice est donnée ci-après.

Pour les coordonnées des points de mesure :

$$\begin{bmatrix} m_{i1} & m_{i2} & m_{i3} \end{bmatrix} = \begin{bmatrix} X_i & Y_i & Z_i \end{bmatrix} \quad (3.2)$$

Avec  $i$  le numéro du point de mesure et  $1 \leq i \leq 1460$

Pour les RSS par rapport aux points d'accès (AP) :

$$\begin{bmatrix} m_{i4} & m_{i5} & \dots & m_{ij} & \dots & m_{i51} \end{bmatrix} = \begin{bmatrix} AP_{i1} & AP_{i2} & \dots & AP_{ij} & \dots & AP_{i48} \end{bmatrix} \quad (3.3)$$

Avec  $i$  le numéro du point de mesure  $1 \leq i \leq 1460$

et  $j$  le numéro du point d'accès  $1 \leq j \leq 48$

Pour chaque fichier de la base de données et pour chaque AP les 10 premières et dernières mesures sont rejetées alors que les 40 mesures restantes servent à calculer

la puissance moyenne selon l'équation suivante :

$$(P_{moy})_{dBm} = (AP_{ij})_{\substack{1 \leq i \leq 1460 \\ 1 \leq j \leq 48}} = \frac{\sum_{k=11}^{40} P_k}{40} \quad (3.4)$$

Afin d'uniformiser cette base de données, et comme la sensibilité de la carte sans fil est de -93dBm, la valeur de -94dBm est assignée à tous les APs non détecté par la carte pour ce point.

### 3.2.3 Méthode d'estimation

Dans le cadre de ce projet, le fichier *ANN\_GRNN\_2008.m* de l'annexe III.1 appelle et crée le réseau de neurones de type *NEWGRNN* (New Generalized Regression Neural Network) qui est un réseau à regression généralisée non linéaire et non récurrent. Ce réseau est composé de deux couches, la première emploie une fonction de base radiale (*RADBAS*) alors que la deuxième utilise une fonction linéaire (*PURELIN*)[17]. L'entrée de ce réseau est constitué de 48 variables correspondant aux RSS des 48 points d'accès alors que la sortie contient 3 variables représentant les coordonnées 3D du mobile.

Après l'obtention de la base de données, le trois quart (3/4) de cette dernière est utilisée en *Offline* pour la phase d'apprentissage alors que le quart (1/4) restant est employée en *Online* pour estimer la position du mobile. Les données à entraîner et à estimer sont choisies aléatoirement. Finalement, l'erreur d'estimation et la fonction de densité cumulative (*CDF*) sont obtenues. Dans ce cas, l'entrée du réseau pour les deux phases est une matrice contenant le RSS émis par 48 points d'accès selon la position.

### 3.3 Résultats et analyse

Le réseau de neurones de type *NEWGRNN* prend comme argument les RSS et la position du mobile pour l'entraînement du réseau ainsi que la variable *Spread* qui représente la constante de dispersion du réseau. Cette constante permet de faire un compromis entre les propriétés de mémorisation et de généralisation du réseau de neurones. Après plusieurs simulations, en utilisant le programme *ANN\_GRNN\_2008.m*, la valeur de la constante de dispersion donnant le meilleur résultat a été trouvée égale à 4.5.

Lors de la phase d'apprentissage, 75% des points de mesure (75% de 1460 points de mesure) de la base de donnée ont été utilisées alors que le reste des données ont été employées lors de la phase d'estimation. Le programme définit dans l'annexe III.1 permet de choisir aléatoirement les points de mesure utilisés pour la phase d'apprentissage et celle d'estimation. De plus, à chaque simulation, les données choisies pour l'entraînement et l'estimation sont différentes ce qui permet d'obtenir différentes fonctions de distribution cumulative CDF.

Pour l'estimation de l'étage (variable *z*), un seuil a été défini à 0 mètre pour le cinquième étage, à 3.9 mètres pour le sixième étage et à -7.8 mètres pour le troisième étage. Pour cette raison, l'analyse suivant la variable *z* sera faite en comptant le nombre d'erreur suivant l'étage et non par mètres.

À la fin d'une simulation, six figures sont obtenues (trois pour la phase d'apprentissage et trois pour la phase d'estimation). Ces figures donnent, sous forme de courbes, les positions réelles et estimées de l'usager, les erreurs de localisation en

mètres déterminées à partir des différences entre les positions réelles et estimées et le CDF. Ces résultats permettent de donner une mesure de performance du système de localisation pour fin d'analyse. Un exemple de ces derniers est donné aux figures 3.3, 3.4 et 3.5 pour la phase d'apprentissage et 3.6, 3.7 et 3.8 pour la phase d'estimation.

Durant l'entraînement, la figure 3.3 montre que les positions estimées et exactes se superposent presque alors que pour la phase d'appel, de la figure 3.6 une différence entre les positions estimées et exactes est remarquée. Cette différence est normale puisqu'en général, dans un réseau de neurones, la propriété de mémorisation des données est meilleur que celle de la généralisation [20].

Lors de la phase d'apprentissage, selon la figure 3.4, l'erreur de position varie, en général, entre  $-4.27$  et  $9.38$  mètres pour  $x$ ,  $-6.97$  et  $5.13$  mètres pour  $y$  et  $0$  et  $11.69$  mètres pour l'erreur Euclidienne  $d$  en 2D et 3D. En  $z$ , il n'y a eu qu'une seule erreur d'étage. En phase d'estimation, selon la figure 3.7, ces valeurs sont en général, entre  $-4.92$  et  $13.05$  mètres pour  $x$ ,  $-7.38$  et  $10.22$  mètres pour  $y$  et  $0.34$  et  $13.94$  mètres pour l'erreur Euclidienne  $d$  en 2D et 3D. En  $z$ , il n'y a eu qu'une seule erreur d'étage.

Ces données ne prennent pas en compte les valeurs limites des erreur de localisation qui peuvent aller jusqu'à  $40$  mètres et qui représentent les cas extrêmes. Ces cas peuvent être le fruit d'un manque d'informations sur les points de mesure (comme exemple, réception des RSS de 3 APs ou moins) ou d'une erreur de généralisation causée par l'ANN. La distance Euclidienne  $d$  en 2D est sensiblement la même qu'en 3D puisque l'erreur selon l'étage est très faible. Cette dernière (l'erreur selon l'étage) varie entre  $0$  et  $3$  erreurs pour toutes les simulations effectuées dans ce chapitre.

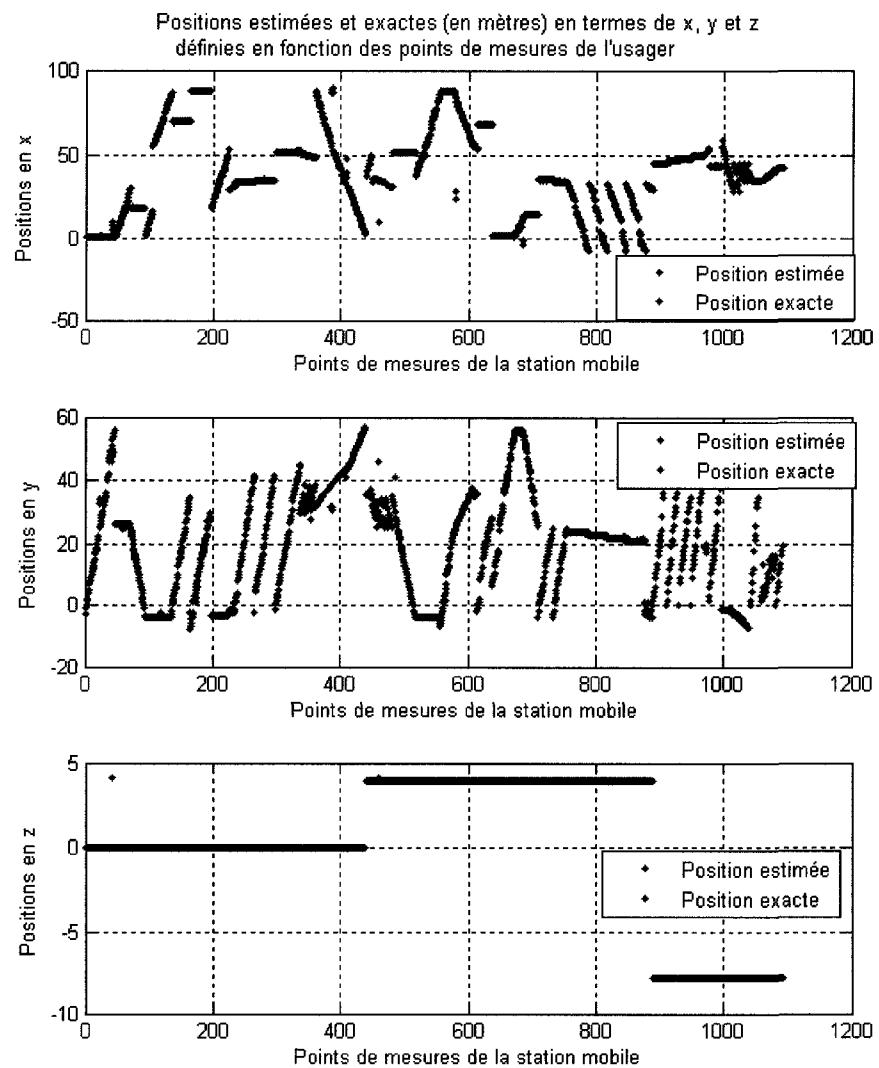


FIGURE 3.3 Positions exactes et estimées en x, y et z en fonction de 1092 points de mesure de la station mobile obtenues en phase d'apprentissage

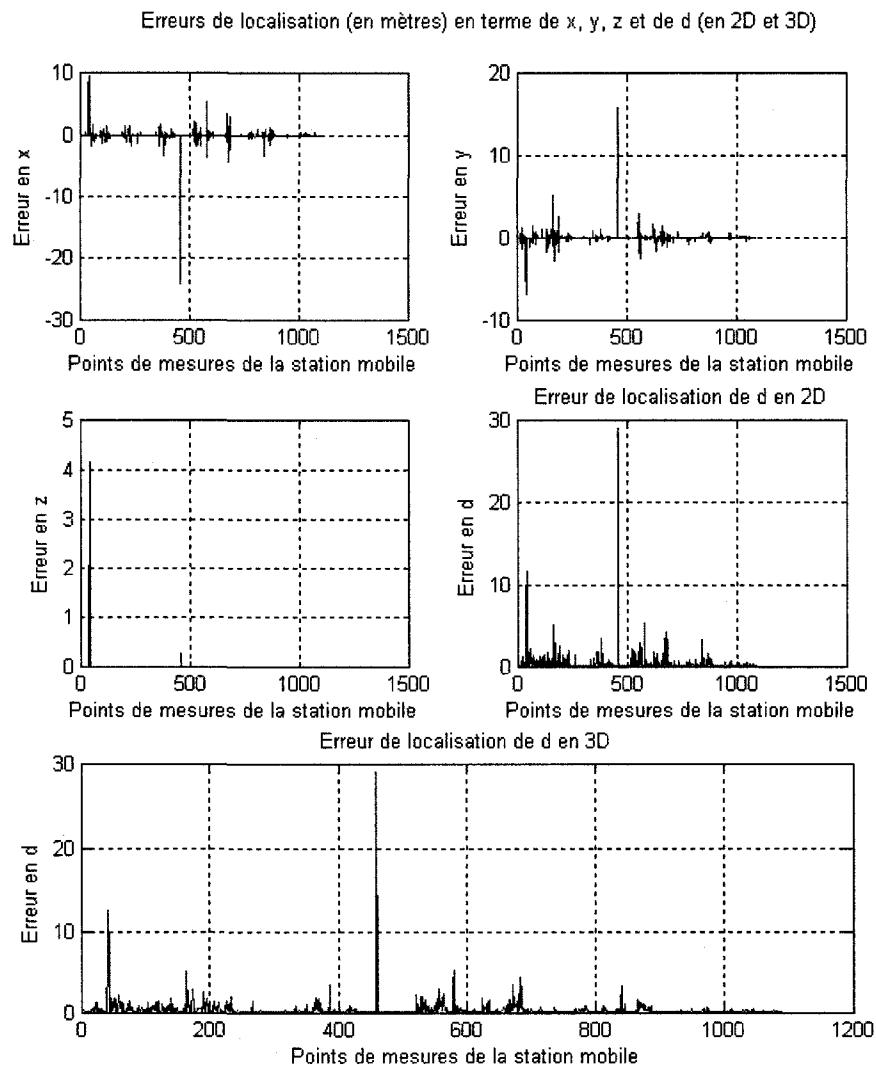


FIGURE 3.4 Erreurs de localisation en termes de x, y, z et de la distance Euclidienne d en 2D et en 3D en fonction de 1092 points de mesure de la station mobile obtenues en phase d'apprentissage

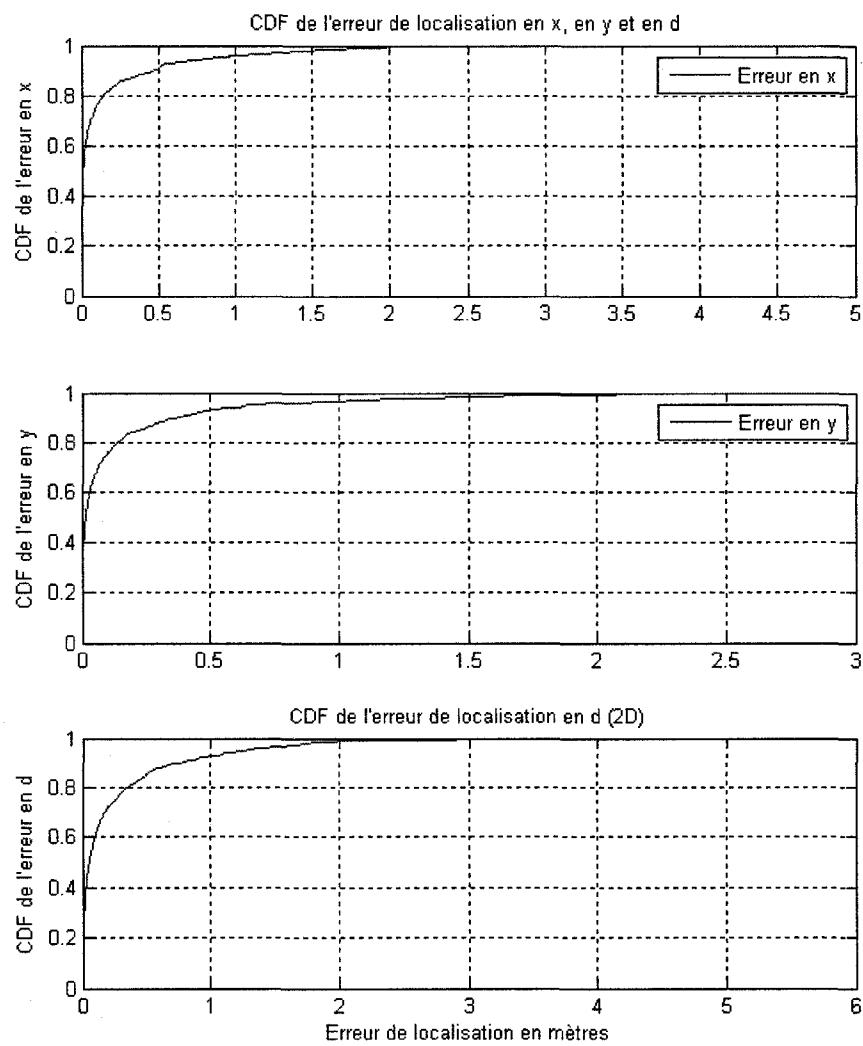


FIGURE 3.5 Fonction de distribution cumulative CDF des erreurs de localisation en termes de x, y et de d en 2D en fonction de l'erreur de localisation (en mètres) obtenues en phase d'apprentissage

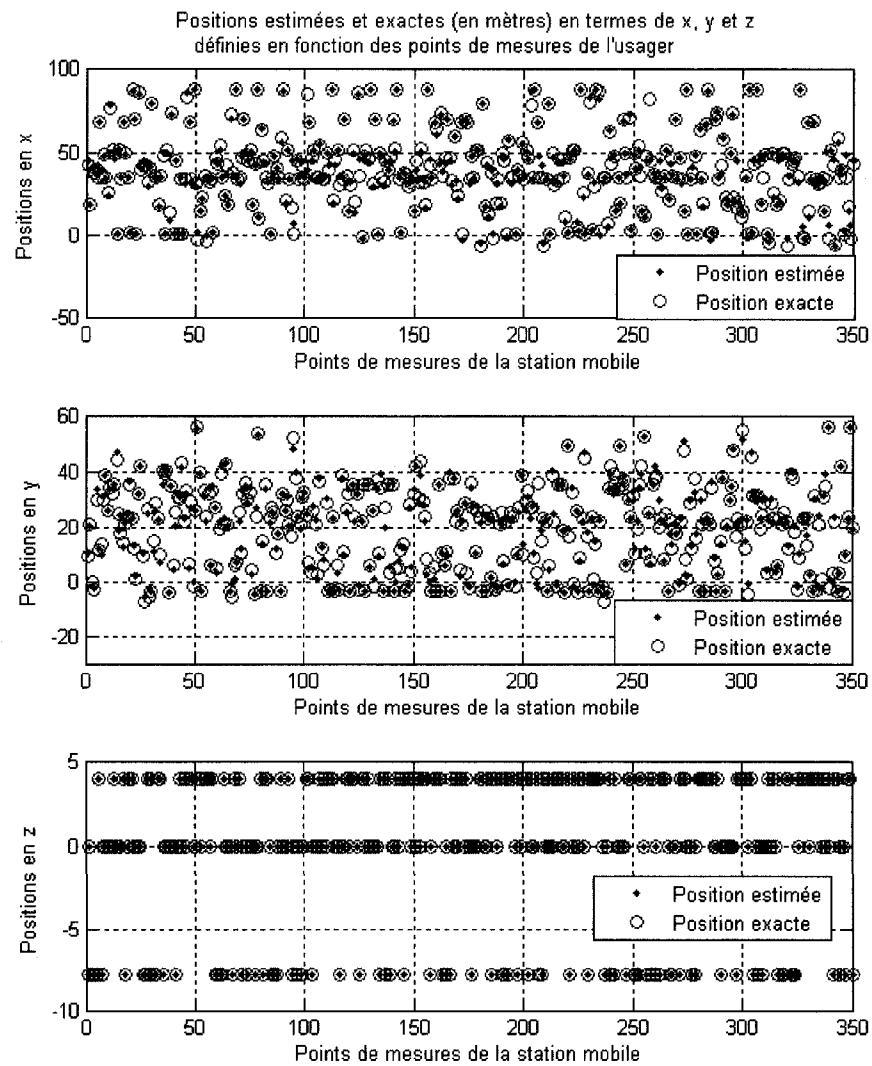


FIGURE 3.6 Positions exactes et estimées en x, y et z en fonction de 350 points de mesure de la station mobile obtenus en phase d'estimation

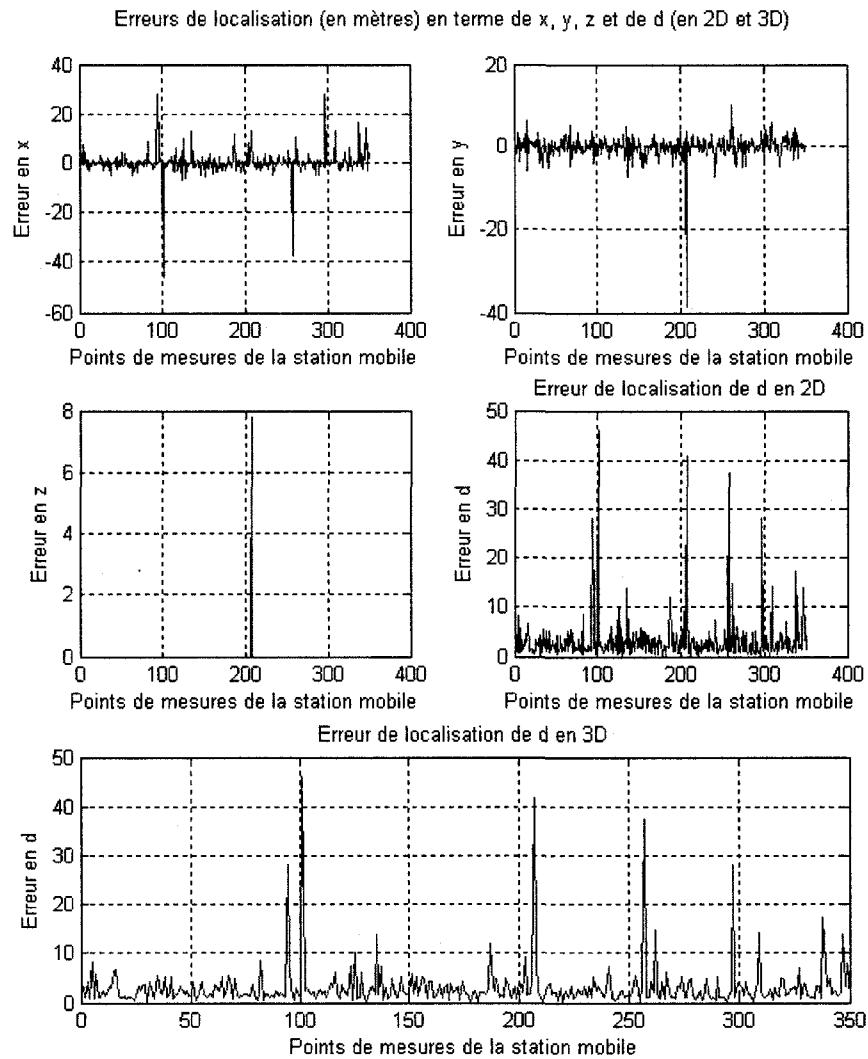


FIGURE 3.7 Erreurs de localisation en termes de x, y, z et de la distance Euclidienne d en 2D et en 3D en fonction de 350 points de mesure de la station mobile obtenues en phase d'estimation

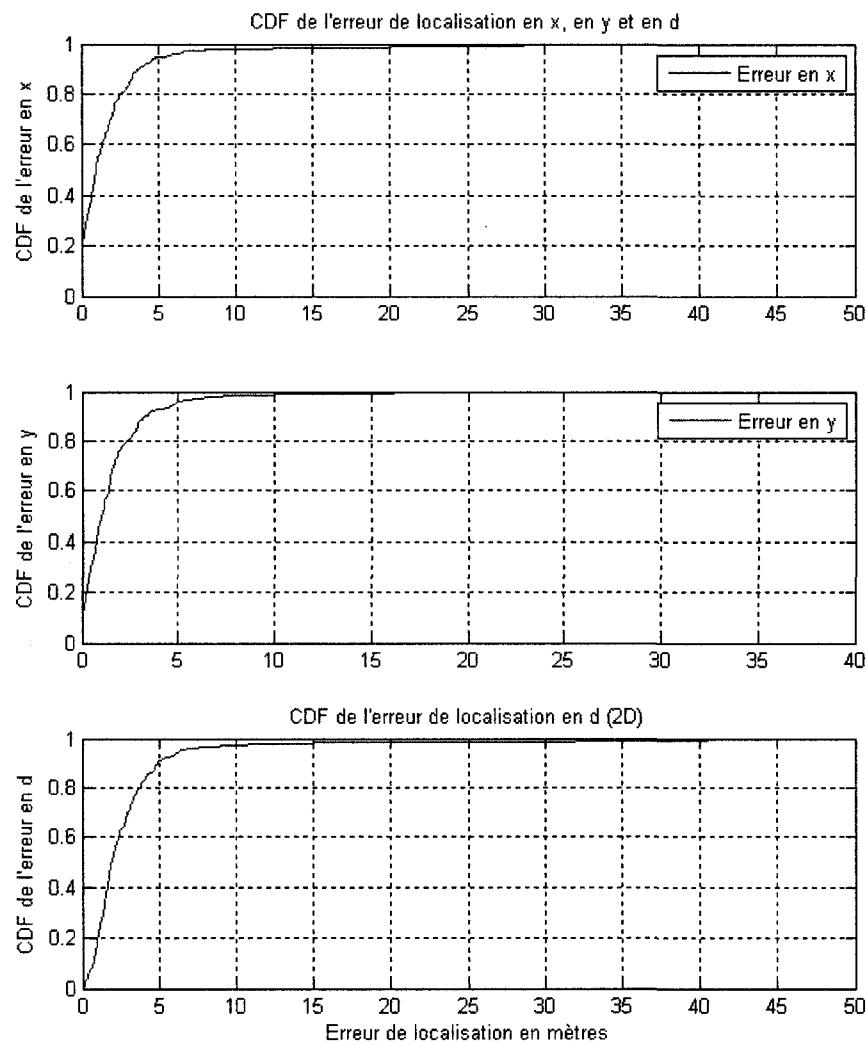


FIGURE 3.8 Fonction de distribution cumulative CDF des erreurs de localisation en termes de x, y et de d en 2D en fonction de l'erreur de localisation (en mètres) obtenues en phase d'estimation

D'un autre coté, les figures 3.5 et 3.8 montrent une précision de localisation (en 2D) de 3 mètres pour 98.81% des cas en phase d'entraînement et pour 68% des cas en phase d'estimation. Pour une précision de 5 mètres, la phase d'entraînement donne une précision proche du 100% alors qu'en phase d'estimation elle est de 88%.

De plus, en z, la précision est de 99.62% à 0.26 mètres en phase d'apprentissage alors qu'en phase d'estimation elle est de 99.71% à 7.8 mètres. En phase d'apprentissage, le 0.26 mètres représente la différence de hauteur entre les tables et les chaises qui n'a pas été définie dans l'ANN. En effet, quelques mesures (aux 6<sup>e</sup> et 5<sup>e</sup> étages) ont été prises sur des tables à la place d'une chaise. En phase d'estimation, et selon la figure 3.7 de l'erreur en z, il n'y a eu qu'une seule erreur d'estimation de l'étage (z estimé est égale à 0 alors que z réel est égale à -7.8 mètres).

Les résultats précédents démontrent que le réseau NEWGRNN permet de faire une bonne généralisation de la localisation en 2D et en 3D afin d'effectuer la reconnaissance des données non apprises par le réseau. De plus, ces résultats sont meilleurs que ceux obtenus dans l'article [10], et qui sont de 78% et 63% à 3 mètres pour les phases d'apprentissages et d'estimation respectivement, car le nombre d'information dans notre cas est plus grande (réception des RSS de 4 APs dans l'article comparé à la réception des RSS des 48 APs dans ce projet). Ce qui confirme que le nombre de point d'accès a un effet sur la précision de la localisation.

Finalement, afin de valider les résultats obtenus, des simulations avec des valeurs pour l'entraînement et pour l'estimation différentes pour chaque simulation ont été effectuées. Par exemple, la série des points de mesure utilisés (pour la phase d'estimation) durant la simulation numéro 1 est différente de celle des autres simulations. Le tableau 3.3 donne le résultat du CDF en 3D à 3 mètres et à 5 mètres

Tableau 3.1 Valeur du CDF en 3D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation

Simulation	à 3 mètres (%)	à 5 mètres (%)
1	77.14	92.86
2	77.86	95
3	80	93.57
4	76.43	94.29
5	75.51	93.57
6	79.29	91.43
7	77.86	94.29
8	70	87.86
9	78.57	91.43
10	79.29	93.57
11	75.71	91.43
12	77.14	93.57
13	76.97	92.14
14	79.29	92.14
15	75.51	95.71
<b>Moyenne</b>	<b>77.1</b>	<b>92.86</b>

pour 15 simulations en phase d'estimation.

Selon le tableau précédent, la précision de 3 mètres est atteinte dans 77.1% des cas. En prenant en compte la règle des 2/3 comme probabilité d'estimation (une précision acceptable selon 67% des cas) on peut dire que la méthode proposée dans ce chapitre (approche des signatures et ANN) donne de bonne performances. Dans ce cas, la base de données créée pour le canal statique et la méthode du réseau de neurones artificielles sont valides. Ces dernières peuvent donc être utilisés dans un canal variable dans le temps pour l'estimation de la position.

## CHAPITRE 4

### VARIABILITÉ DU CANAL

L'un des points faibles de l'approche des signatures est sa faible performance dans un canal non statique dit dynamique. En effet, dans un canal variant dans le temps, la base de données ainsi que le réseau de neurones créé et entraîné pour un instant donné pourraient ne pas refléter le même environnement à un autre instant. Ceci est causé par le changement des puissances reçues à chaque point de la zone étudiée ce qui affecte la base de données précédente et entraîne une réduction considérable de la précision de localisation.

La variation du canal peut être causée par plusieurs facteurs comme l'achalandage humain, le changement des plans initiaux (ajouts de tables, murs, changement de fenêtres, etc.), le changement de l'équipement de mesure (comme utiliser une carte sans fil avec une précision différente) ou la présence de tout objet mobile.

Comme exemple, la figure 4.1 donne la puissance reçue (RSS) en fonction du temps pour une position donnée lorsque le canal est statique et dynamique. Cette figure permet de dire que les fluctuations dans un canal statique sont moins importantes que celles dans un canal dynamique (3dBm de différence entre le minimum et maximum en statique et 12dBm de différence en dynamique). De plus, la moyenne des RSS donne -77.8dBm dans un canal statique et -80.58dBm dans un canal dynamique pour une même position. Dans le cadre de ce projet, la variation temporelle du canal est prise à grande échelle, c'est à dire la variation en terme d'heures ou d'années. Les variations temporelles à petite échelle sont pris en compte pour la création de la base de donnée initiale (chapitre 3) en calculant la moyenne des RSS sur 20 secondes comme expliqué dans la section 3.2.2.

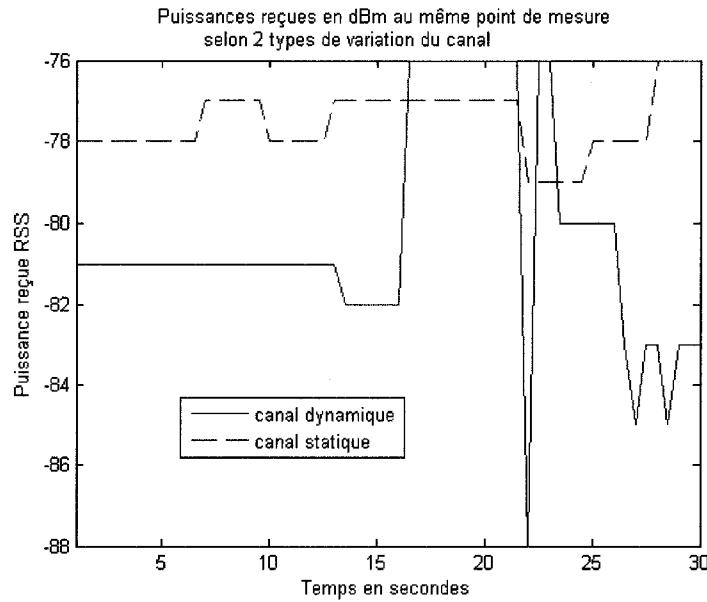


FIGURE 4.1 RSS en fonction du temps pour une position donnée dans un canal statique et dynamique

Dans le cas d'un canal dynamique, il faut normalement reprendre les mesures afin de créer une nouvelle base de données pour ensuite l'utiliser dans le réseau de neurones afin d'effectuer la localisation. Cette méthode demande beaucoup de ressources matérielles et temporelles puisqu'il faut refaire les mesures à chaque modification de l'état du canal avant d'effectuer la localisation. De plus, il faut déterminer l'état présent du canal afin d'utiliser la bonne base de données et ainsi éviter une forte dégradation.

L'article [23] propose de définir plusieurs états significatifs pour utiliser l'approche des signatures à certaines périodes de temps. Une base de données et un réseau de neurones par état sont ainsi créés et entraînés. Le nombre d'ANN est dans ce cas égal au nombre d'états. Finalement, en se basant sur une hiérarchie de réseaux de neurones déjà entraînés et des points de référence pour définir l'état, on peut

utiliser le bon réseau de neurones. Le schéma de la figure 4.2 donne un aperçu de la solution de l'article [23].

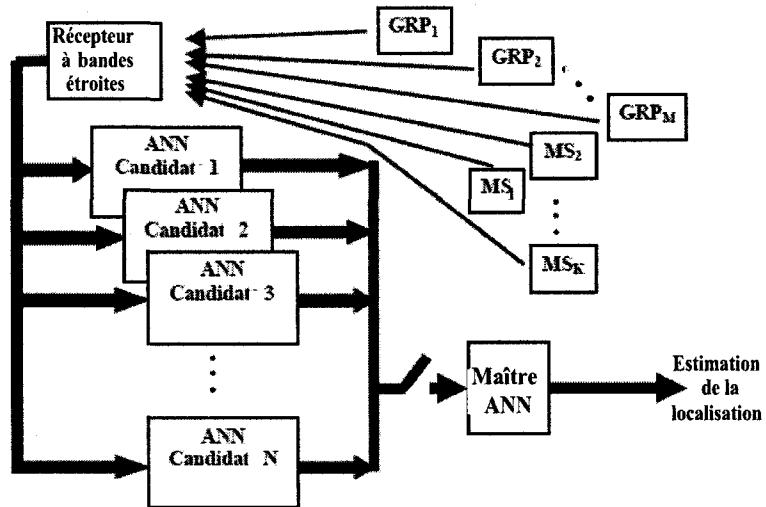


FIGURE 4.2 Schéma explicatif de la solution de l'article [23]

Dans la figure 4.2, les  $GRP_1$  à  $GRP_M$  représentent des émetteurs à coordonnées fixes et connus (points de référence) afin de déterminer l'état du canal.  $M$  est le nombre de points de référence alors que  $MS_1$  à  $MS_K$  sont les stations mobiles à localiser. Le récepteur à bande étroite mesure le RSS du mobile avec le système WLAN.  $ANN \text{ candidat } 1$  à  $ANN \text{ candidat } N$  sont les réseaux de neurones déjà entraînés en offline du premier état jusqu'à l'état  $N$ . Le *Maître ANN* permet de choisir l'état selon les données des points de référence GRPs afin d'obtenir une estimation de la localisation. En utilisant la solution proposée de l'article [23], la fonction de densité cumulative a donné 64% à 5 mètres pour un canal invariant et 57% à 5 mètres pour un canal variant dans le temps.

La solution proposée dans ce chapitre permet de modifier la base de données initiale sans collecter les 1460 points de mesure par état. Cet algorithme permet donc de minimiser les ressources nécessaires pour la localisation en ne prenant que quelques mesures de référence. La section suivante décrit le système proposé.

#### 4.1 Description du système de localisation proposé

Le système proposé permet de s'adapter à la variation du canal en modifiant la base de données initiale grâce à des points de référence et une technique d'exploration de données ou *Data Mining* [3] [4]. Cette technique applique la méthode d'arbres de décision, qui se base sur l'analyse de régression, afin de prédire la relation temporelle entre le RSS par les points de référence et ceux reçus par le mobile à un instant  $t_x$  donné. Le schéma de la figure 4.3 donne les étapes pour générer cette base de données au temps  $t_x$ .

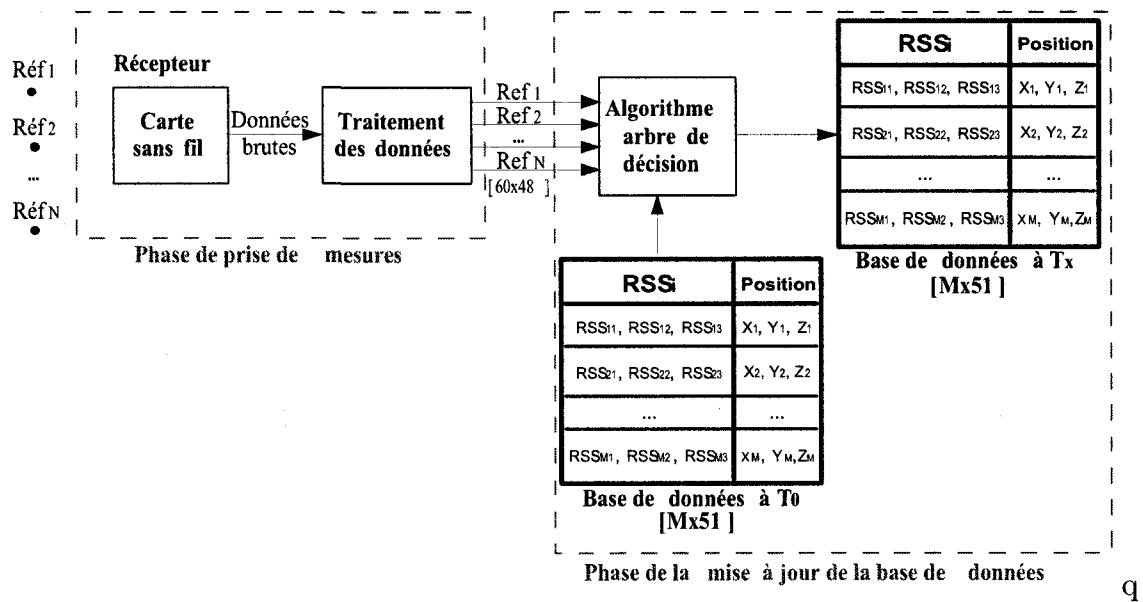


FIGURE 4.3 Schéma général de l'étape de la génération d'une base de donnée

D'après ce schéma, le mobile (récepteur) prend les mesures des  $N$  points de référence au temps  $t_x$  et les traite afin d'obtenir un tableau de données de 60 lignes par 48 colonnes (60 mesures de RSS par 48 points d'accès) pour chaque point de référence. Par la suite, la méthode d'arbre de décision utilise ces données comme référence et la base de données initiale créée au temps  $t_0$  afin de générer la base de donnée au temps  $t_x$ . La base de données contient les coordonnées 3D (X, Y, Z) et le RSS par

rapport à 48 points d'accès de  $M$  points de mesure.

Après avoir créé la base de données à  $t_x$ , ce dernier est utilisé dans les réseaux de neurones ANN pour l'entraînement puis pour l'estimation de la position. Le schéma de la figure 4.4 donne les étapes pour l'estimation des coordonnées.

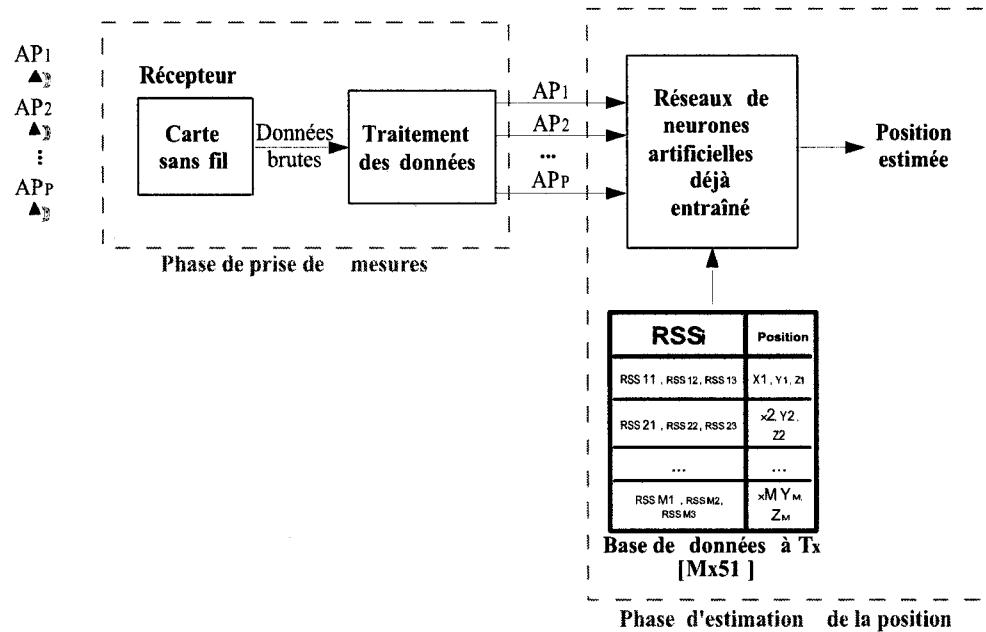


FIGURE 4.4 Schéma général de l'étape de l'estimation de la position

Après l'entraînement de l'ANN et afin d'estimer la position, le mobile mesure le RSS émis par  $P$  points d'accès au temps  $t_x$ . Ces données sont traitées puis envoyées au réseau de neurones déjà entraîné qui utilise la base de données de  $t_x$  pour estimer la position.

Afin de valider le système proposé, une série de mesures a été faite. Des mesures en 2D d'abord pour trois états ou temps différents pour connaître la précision de localisation avec les bases de données créées. Une autre série de mesures en 3D a été réalisée pour un état afin de vérifier l'estimation de la position du mobile selon

l'étage (variable Z). La section suivante donne les détails du travail effectué pour cette partie du projet.

## 4.2 Méthodologie

Afin de déterminer les performances de la méthode proposée, trois cas ont été définis pour les tests en deux dimensions et un cas en trois dimensions. L'aire étudiée en deux dimensions représente le cinquième étage alors qu'en trois dimensions se sont les cinquième et sixième étages qui ont été considérés.

### 4.2.1 Définition des points de référence

Les points de référence sont choisis de manière à couvrir toute la zone étudiée. Pour les tests en deux dimensions, 14 points de référence ont été définis et répartis comme suit : 3 au troisième étage, 3 au sixième étage et 8 au cinquième étage. Les points de référence définis au troisième et sixième étages ont été choisis de façon à ce que les puissances reçues fassent partie des puissances émises par les points d'accès se trouvant au cinquième étage (puisque la localisation se fait au niveau du cinquième étage). Les coordonnées de ces 14 points de référence sont définis au tableau 4.1.

Pour les tests en trois dimensions, 25 points de référence ont été définis et répartis comme suit : 14 au cinquième étage et 11 au sixième étage. Les coordonnées de ces 25 points de référence sont définis au tableau 4.2.

Tableau 4.1 Coordonnées, en mètres, des 14 points de référence définis pour les simulations en 2D

Pt de référence	X	Y	Z
1	40.61	42.71	0.26
2	0.75	0	0
3	87.5	8.75	0
4	51.5	8.7	0
5	18.5	18	0
6	0.75	54	0
7	69.75	17.75	0
8	35	0.7	0
9	41.41	36.8	4.16
10	35.24	8.8	3.9
11	26.24	20.8	3.9
12	34.5	36	-7.8
13	45	9	-7.8
14	39	-7.5	-7.8

#### 4.2.2 Prise de mesure

Pour les mesures en deux dimensions, la collecte de données des points de référence a été faite à trois moments différents ; tard le soir, entre 23 heures et minuit, pour obtenir un *canal non variant*, durant un achalandage moyen, entre 14 heures et 15 heures, qui représente un *canal semi-variant* et durant un fort achalandage, de midi à 13 heures, afin d'obtenir un *canal variant*.

Pour les mesures en trois dimensions, la collecte de données a été faite pour le cas d'un *canal semi-variant* (entre 14 heures et 15 heures). Pour la prise de mesure, l'équipement utilisé est le même que celui de la section 3.2.1.3, le protocole de mesure est défini selon la section 3.2.1.4.

À la fin de la prise de mesure, 25 fichiers (un par point de référence) et 42 fichiers (14 fichiers pour chacun des 3 cas de variation de canal) sont obtenus pour le 3D et 2D, respectivement.

Afin de définir le type de canal étudié ou les périodes de prise de mesure ;

Tableau 4.2 Coordonnées, en mètres, des 25 points de référence définis pour les simulations en 3D

Pt de référence	X	Y	Z
1	40.61	42.71	0.26
2	0.75	0	0
3	87.5	8.75	0
4	51.5	8.7	0
5	18.5	9	0
6	0.75	54	0
7	69.75	17.75	0
8	35	0.7	0
9	9.75	26	0
10	41.41	36.8	4.16
11	26.24	20.8	3.9
12	0.98	55.2	3.9
13	13.98	36.7	3.9
14	51.74	0	3.9
15	68.24	9.2	3.9
16	86.74	-3.8	3.9
17	87.74	18.8	3.9
18	1.27	9.6	3.9
19	51.74	18	3.9
20	0.24	21.8	3.9
21	68.75	-3.75	0
22	87.5	30.75	0
23	34	27.7	0
24	51.5	31.7	0
25	26.68	45.71	0

- $t_0$  représente les données obtenues au chapitre 3 pour un canal non variant.
- $t_1$  représente les données obtenues pour le canal non variant 2 ans après les mesures au temps  $t_0$ .
- $t_2$  représente les données obtenues pour le canal semi-variant.
- $t_3$  représente les données obtenues pour le canal variant.

### 4.2.3 Traitement des données

Le programme *Traitement.cpp* est maintenant utilisé avec les fichiers bruts créés par *Netstumbler*. Ce programme permet de trier le fichier brut en ne gardant que les puissances reçues des 48 points d'accès définis dans le chapitre 3. Comme la sensibilité de la carte sans fil est de -93dBm, la valeur de -94dBm est assignée à tous les points d'accès non détectés par la carte pour un instant donné.

Finalement, un fichier contenant 60 puissances reçues pour les 48 points d'accès est créé. La figure 4.5 donne un exemple de fichier généré par le programme Traitement.cpp.

Ref	4005.txt	Bloc notes	1
78	-94	-94	-94
78	-75	-94	-94
78	-75	-93	-65
78	-75	-93	-64
78	-75	-93	-63
78	-75	-93	-62
78	-75	-93	-61
78	-75	-93	-60
78	-75	-93	-59
78	-75	-93	-58
78	-75	-93	-57
78	-75	-93	-56
78	-75	-93	-55
78	-75	-93	-54
78	-75	-93	-53
78	-75	-93	-52
78	-75	-93	-51
78	-75	-93	-50
78	-75	-93	-49
78	-75	-93	-48
78	-75	-93	-47
78	-75	-93	-46
78	-75	-93	-45
78	-75	-93	-44
78	-75	-93	-43
78	-75	-93	-42
78	-75	-93	-41
78	-75	-93	-40
78	-75	-93	-39
78	-75	-93	-38
78	-75	-93	-37
78	-75	-93	-36
78	-75	-93	-35
78	-75	-93	-34
78	-75	-93	-33
78	-75	-93	-32
78	-75	-93	-31
78	-75	-93	-30
78	-75	-93	-29
78	-75	-93	-28
78	-75	-93	-27
78	-75	-93	-26
78	-75	-93	-25
78	-75	-93	-24
78	-75	-93	-23
78	-75	-93	-22
78	-75	-93	-21
78	-75	-93	-20
78	-75	-93	-19
78	-75	-93	-18
78	-75	-93	-17
78	-75	-93	-16
78	-75	-93	-15
78	-75	-93	-14
78	-75	-93	-13
78	-75	-93	-12
78	-75	-93	-11
78	-75	-93	-10
78	-75	-93	-9
78	-75	-93	-8
78	-75	-93	-7
78	-75	-93	-6
78	-75	-93	-5
78	-75	-93	-4
78	-75	-93	-3
78	-75	-93	-2
78	-75	-93	-1
78	-75	-93	0
78	-94	-94	-94
78	-94	-94	-93
78	-94	-94	-92
78	-94	-94	-91
78	-94	-94	-90
78	-94	-94	-89
78	-94	-94	-88
78	-94	-94	-87
78	-94	-94	-86
78	-94	-94	-85
78	-94	-94	-84
78	-94	-94	-83
78	-94	-94	-82
78	-94	-94	-81
78	-94	-94	-80
78	-94	-94	-79
78	-94	-94	-78
78	-94	-94	-77
78	-94	-94	-76
78	-94	-94	-75
78	-94	-94	-74
78	-94	-94	-73
78	-94	-94	-72
78	-94	-94	-71
78	-94	-94	-70
78	-94	-94	-69
78	-94	-94	-68
78	-94	-94	-67
78	-94	-94	-66
78	-94	-94	-65
78	-94	-94	-64
78	-94	-94	-63
78	-94	-94	-62
78	-94	-94	-61
78	-94	-94	-60
78	-94	-94	-59
78	-94	-94	-58
78	-94	-94	-57
78	-94	-94	-56
78	-94	-94	-55
78	-94	-94	-54
78	-94	-94	-53
78	-94	-94	-52
78	-94	-94	-51
78	-94	-94	-50
78	-94	-94	-49
78	-94	-94	-48
78	-94	-94	-47
78	-94	-94	-46
78	-94	-94	-45
78	-94	-94	-44
78	-94	-94	-43
78	-94	-94	-42
78	-94	-94	-41
78	-94	-94	-40
78	-94	-94	-39
78	-94	-94	-38
78	-94	-94	-37
78	-94	-94	-36
78	-94	-94	-35
78	-94	-94	-34
78	-94	-94	-33
78	-94	-94	-32
78	-94	-94	-31
78	-94	-94	-30
78	-94	-94	-29
78	-94	-94	-28
78	-94	-94	-27
78	-94	-94	-26
78	-94	-94	-25
78	-94	-94	-24
78	-94	-94	-23
78	-94	-94	-22
78	-94	-94	-21
78	-94	-94	-20
78	-94	-94	-19
78	-94	-94	-18
78	-94	-94	-17
78	-94	-94	-16
78	-94	-94	-15
78	-94	-94	-14
78	-94	-94	-13
78	-94	-94	-12
78	-94	-94	-11
78	-94	-94	-10
78	-94	-94	-9
78	-94	-94	-8
78	-94	-94	-7
78	-94	-94	-6
78	-94	-94	-5
78	-94	-94	-4
78	-94	-94	-3
78	-94	-94	-2
78	-94	-94	-1
78	-94	-94	0

FIGURE 4.5 Exemple d'un fichier créé par Traitement.cpp

Le programme *Traitement.cpp* est utilisé pour les 67 fichiers bruts (3 \* 14 points de référence en 2D plus 25 points de référence en 3D) des points de référence aux instants  $t_1$ ,  $t_2$  et  $t_3$  et les 1150 fichiers bruts des points de mesure du cinquième et sixième étages. Ces 1150 fichiers bruts ont été créés pour définir la base de données du canal statique à  $t_0$  comme expliqué au chapitre 3. Ce programme est donné dans

l'annexe IV.

#### 4.2.4 Arbre de décision

##### 4.2.4.1 Principe de fonctionnement

La méthode d'arbre de décision permet d'offrir une alternative *non paramétrique* pour l'analyse de données. C'est une méthode d'analyse de *régression non linéaire* qui se base sur la prédiction d'une variable à partir d'un ensemble de données descriptives lorsque la relation entre celles-ci est inconnue. Dans ce cas, il faut faire des hypothèses et des simplifications afin d'obtenir le modèle de prédiction [18]. Généralement ces hypothèses définissent des seuils pour la décision (si ' $A'$  donnée est inférieure au seuil ' $S'$ ', aller à droite sinon aller à gauche). Pour une bonne prédiction (avoir de bonnes hypothèses), il faut avoir un grand nombre d'échantillons pour l'analyse.

Dans le cadre de ce projet, les variables à prédire sont les RSS des points de mesure par rapport aux points d'accès du canal. Comme il n'y a pas de relation connue entre ces derniers et les points de référence définis dans les tableaux 4.1 et 4.2, la méthode d'arbre de décision est utilisée. Dans ce cas, les points de référence mesurés sont les données descriptives. Ces données sont classifiées par l'arbre comme Ref01 à Ref14 pour le 2D et Ref01 à Ref25 pour le 3D.

L'utilisation de cette méthode se fait en deux phases. La phase de la création de l'arbre de décision et la phase d'estimation du RSS. Dans notre cas il faut estimer une puissance reçue par point d'accès et par point de mesure. Afin d'expliquer le

principe de fonctionnement de la méthode, un exemple est donné ci-bas.

Lors de la phase de construction du modèle d'arbre de décision, au lieu de prendre tous les points de référence comme information pour l'analyse, ce dernier ne prend en compte que les points de référence *pertinents* au point de mesure. Donc si ces références sont mal choisies, les bases de données générées seront erronées. La figure 4.6 donne un exemple schématique d'un arbre de décision. Ce dernier représente la prédiction temporelle du point de mesure numéro 755 par rapport au point d'accès AP24 et les points de référence en 3D.

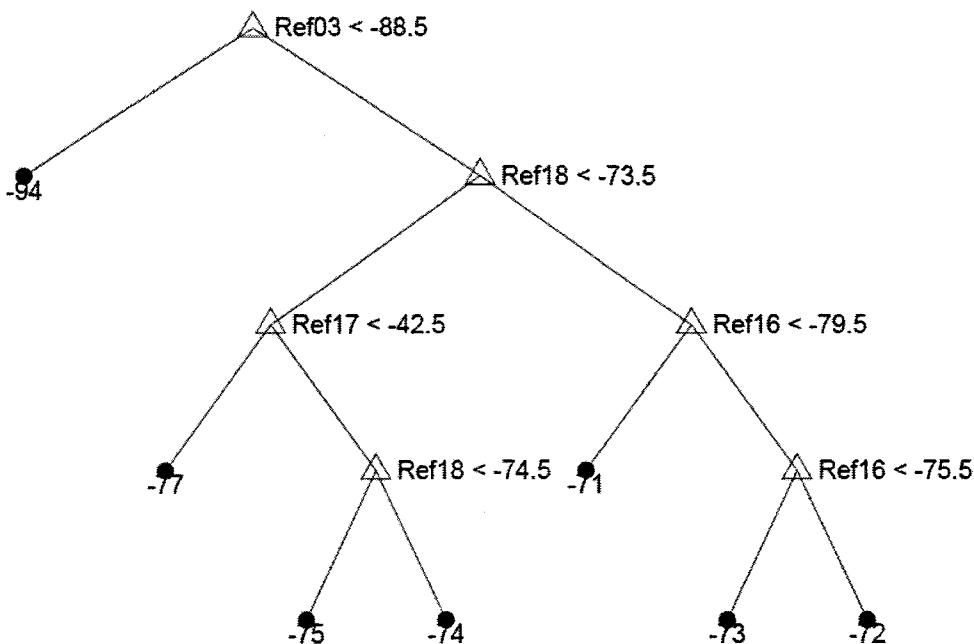


FIGURE 4.6 Schéma d'un modèle d'arbre de décision

Dans l'exemple de la figure 4.6, les RSS mesurées par rapport à l'AP24 pour Ref01, Ref02, Ref04 à Ref15 et Ref19 à Ref25 sont égaux à -94dBm en tout temps (par définition pas de signal reçu à ces points de références par rapport au point d'accès AP24). Par exemple, pour Ref01, on a :

$$Ref01_{i,24} = -94dBm, 1 < i < 60$$

Avec  $i$  qui représente le nombre d'échantillon de RSS selon l'*AP24*.

C'est pour cela que ces points de référence ne sont pas pris en compte lors de la création de l'arbre de décision. D'un autre coté, les signaux envoyés par l'*AP24* sont reçus par les points de référence restants (Ref03, Ref16, Ref17, et Ref18). L'arbre de décision se base donc sur ces derniers pour prédire le RSS. Pour créer l'arbre avec ces données, une analyse est faite selon les 60 échantillons mesurés dans la section 4.2.2. Selon chaque point de référence, l'algorithme choisi un seuil pour chaque échantillon puis teste tous ces seuils sur l'ensemble des 60 échantillons et choisit le seuil optimal (celui qui donne la plus faible erreur d'estimation).

Dans ce mémoire, Cette analyse est faite sur Matlab avec la commande *treefit* qui fait des hypothèses pour l'analyse afin d'obtenir un modèle d'arbre de décision binaire en utilisant l'algorithme du modèle CART (*Classification And Regression Trees*) [19]. Dans l'exemple de la figure 4.6, le premier seuil est  $Ref03 < -88.5$  qui est la racine de l'arbre et est définie comme : *Lorsque le RSS au niveau du point de référence numéro 3 est inférieure à -88.5 dBm aller à droite sinon aller à gauche.*

Lors de la phase d'estimation, l'arbre de décision décide de la valeur du RSS au point de mesure en se référant aux données de références du temps  $t_x$  (pour une seule mesure de RSS). En prenant comme exemple le schéma de la figure 4.6, et en supposant qu'au temps  $t_x$  les valeurs des références utiles sont :

$$Ref03 = -90dBm$$

$$Ref16 = -78dBm$$

$$Ref17 = -45dBm$$

$$Ref18 = -74dBm$$

Le modèle d'arbre de décision décide que le RSS obtenue est égale à  $-71dBm$  au

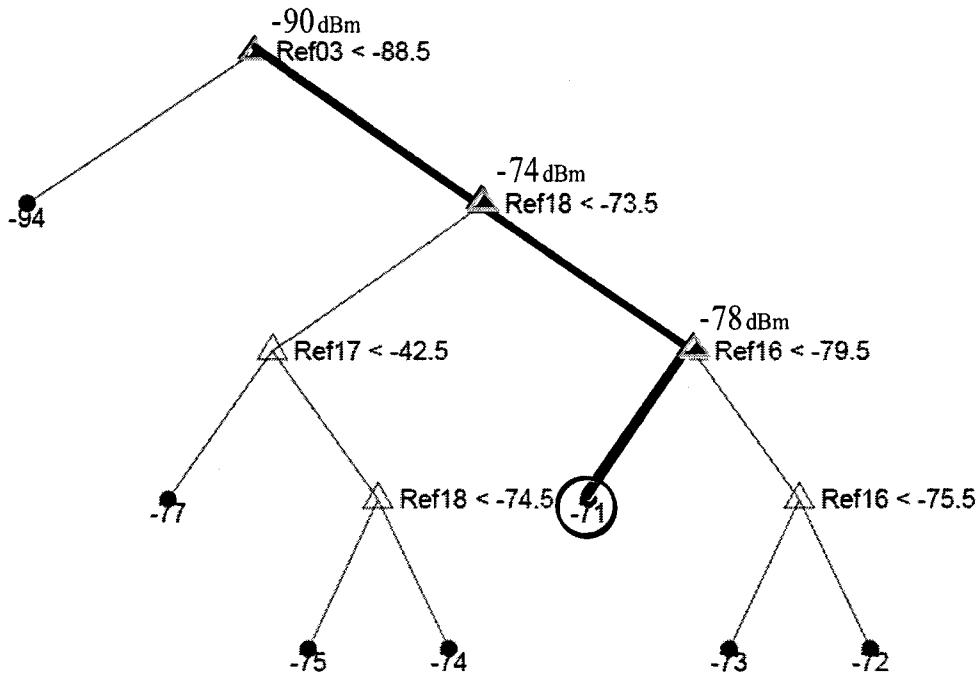


FIGURE 4.7 Cheminement de décision pour le point de mesure numéro 755 et l'AP24

point de mesure numéro 755. Le cheminement de décision est donné sur la figure 4.7. La section suivante donne les étapes permettant de générer une base de données au temps  $t_x$  avec le modèle d'arbre de décision.

Les performances de l'arbre de décision dépendent entièrement du choix des données descriptives permettant sa création. Dans le cas du modèle utilisé dans ce projet, les performances dépendent du choix de l'emplacement et du nombre de point de référence en plus du nombre d'échantillons (60 par point de mesure et par AP) permettant la génération de l'arbre de décision.

#### 4.2.4.2 Méthode de la génération de la base de données

En se basant sur la base de données initiale et sur les points de référence au temps  $t_0$ , le modèle d'arbre de décision est créé. Le schéma de la figure 4.8 permet de mieux visualiser cette partie.

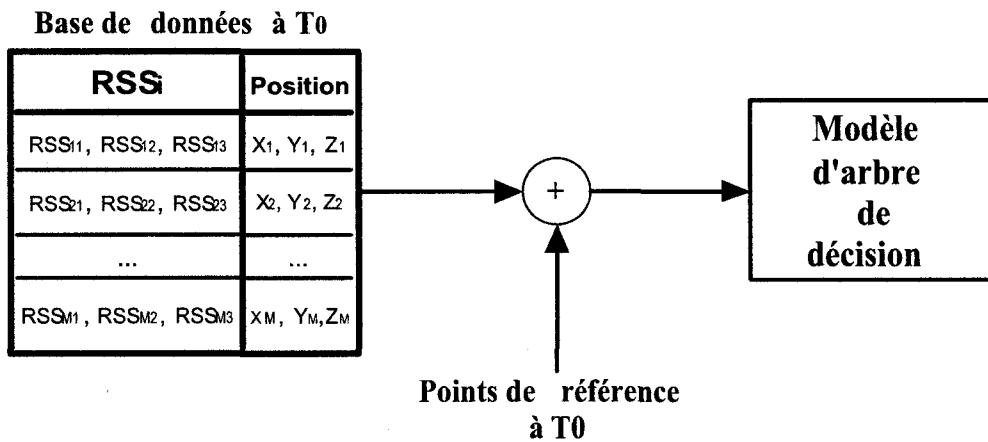


FIGURE 4.8 Schéma de la création du modèle d'arbre de décision

Les points de référence doivent être pré-définis au temps  $t_0$  afin de prédire la relation temporelle entre ces points et les points de mesure de la base de données initiale. Cette étape permet de créer le modèle d'arbre de décision qui sera utilisé afin de définir les nouvelles bases de données au temps  $t_x$ .

En obtenant le modèle de prédiction temporel, la base de données initiale à  $t_0$  et les points de référence mesurés au temps  $t_x$ , une nouvelle base de données peut être créée pour  $t_x$ . Le schéma de la figure 4.9 permet de mieux visualiser cette partie. Les points de référence au temps  $t_1$  et  $t_2$  doivent avoir les mêmes coordonnées que ceux au temps  $t_0$ . Les deux blocs représentant le modèle d'arbre de décision sont identiques puisque la prédiction de la relation temporelle se fait pour le temps

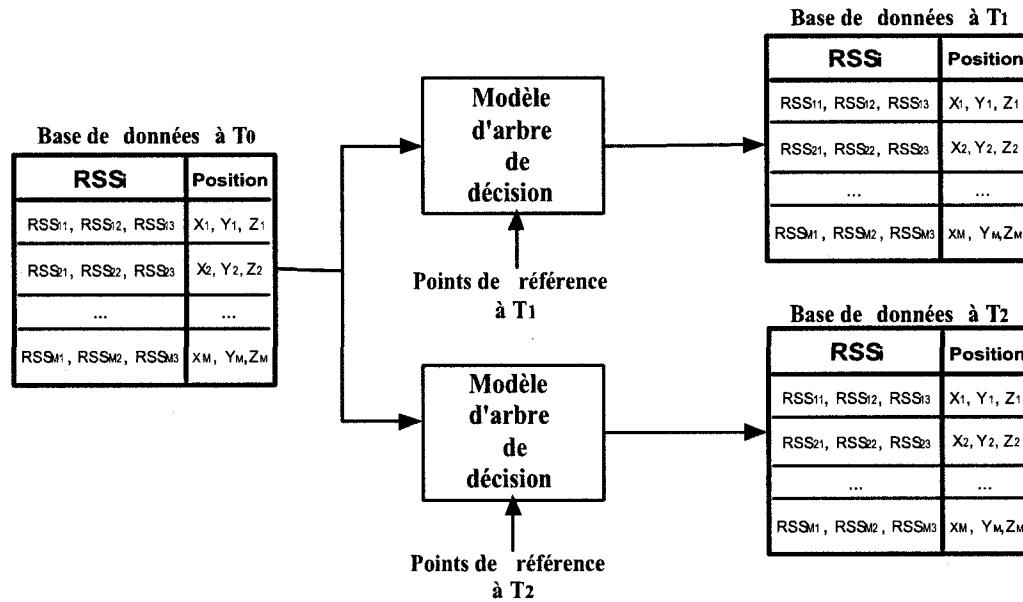


FIGURE 4.9 Schéma de la création de nouvelles bases de données au temps  $t_1$  et  $t_2$

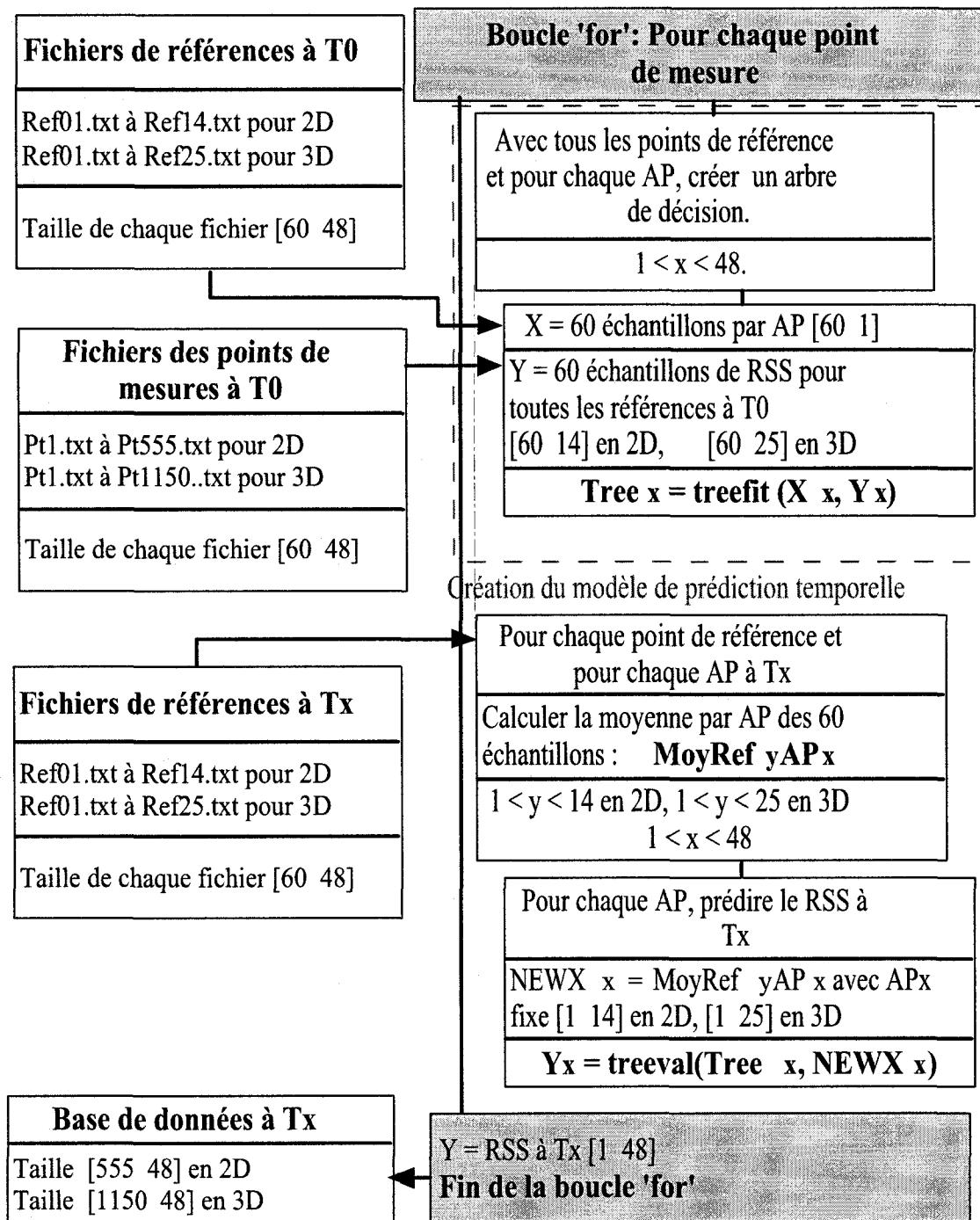
initial  $t_0$ .

#### 4.2.4.3 Algorithme de génération de la base de données

Le programme *ArbreDeDecision.m* a été utilisé pour la création du modèle d'arbre de décision et la génération des bases de données au temps  $t_x$ . Comme ce dernier contient plus de 2500 lignes de codes, seule les étapes les plus importantes sont données en plus de l'algorithme de la figure 4.10.

Ce programme permet de générer un arbre de décision par point de mesure donc 48 arbres de décision sont créés pour un seul point de mesure défini dans la base de données en utilisant la commande suivante :

$$Tree_x = treefit(X_x, Y_x); \quad (4.1)$$

FIGURE 4.10 Diagramme du programme *ArbreDeDecision.m*

où  $Tree_x$  est le modèle d'arbre de décision créé pour un point de mesure et un point d'accès.  $Y_x$  est une matrice de taille [60 1] qui contient les échantillons des puissances reçues RSS  $d'AP_x$  au temps  $t_0$ .  $X_x$  est une matrice de taille [60 14] pour les tests en 2D et [60 25] pour les tests en 3D qui contient les 60 échantillons des RSS d' $AP_x$  de tous les points de référence mesurés au temps  $t_0$ .

Par la suite, la moyenne des 60 échantillons par point d'accès et par point de référence est calculée. Ce qui donne une matrice des moyennes de taille [14 48] pour le 2D et [25 48] pour le 3D. Finalement, en utilisant l'arbre de décision  $Tree_x$  de  $l'AP_x$  créé précédemment et les points de référence au temps  $t_x$ , les RSS au point de mesure et au temps  $t_x$  définis comme  $Y_x$  sont calculés avec :

$$NEWX_x = [MoyRef01AP_x \ MoyRef02AP_x \dots \ MoyRef14AP_x], \text{ pour le 2D}$$

$$NEWX_x = [MoyRef01AP_x \ MoyRef02AP_x \dots \ MoyRef25AP_x], \text{ pour le 3D}$$

Ce qui donne le RSS par rapport au point d'accès  $AP_x$

$$Y_x = treeval(Tree_x, NEWX_x); \quad (4.2)$$

En général, pour générer une base de données, il faut utiliser un nombre d'arbres de décision égale à :

$$Nb \ Arbre \ de \ Decision = Nb \ Points \ de \ Mesure * Nb \ AP \quad (4.3)$$

#### 4.2.5 Méthode d'estimation

Comme dans le système de localisation les bases de données sont générées par un modèle d'arbre de décision au lieu d'être mesurées, une dégradation de l'estimation est attendue. Afin de calculer cette dégradation, la base de données au temps  $t_0$  est générée par la méthode d'arbre de décision et comparée à la base de données mesurée au temps  $t_0$ . Le schéma de la figure 4.11 montre les étapes permettant le calcul de la dégradation. Les résultats de ces calculs sont donnés dans la section suivante.

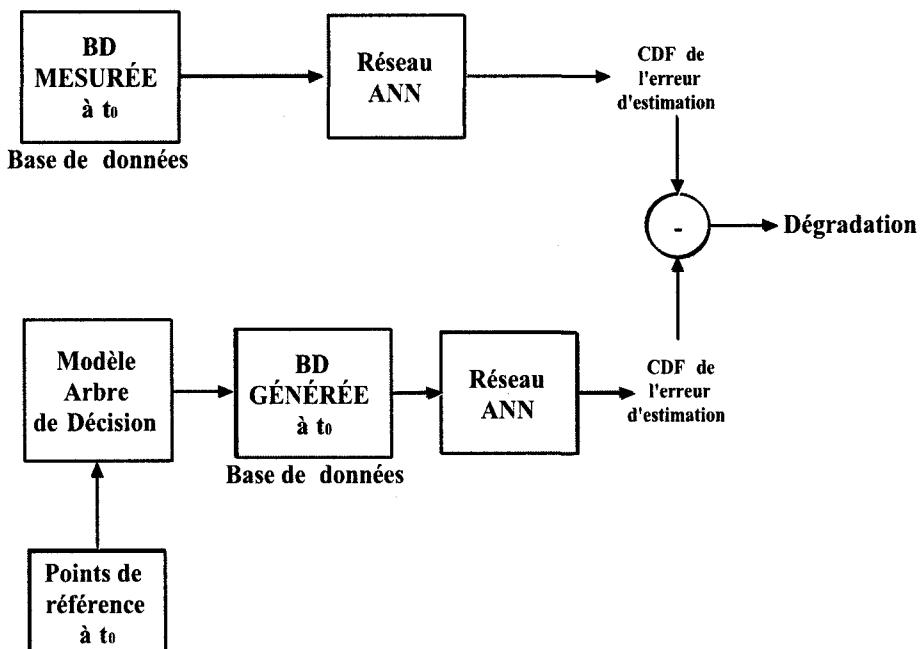


FIGURE 4.11 Diagramme donnant les étapes permettant le calcul de la dégradation due à la génération de la base de données avec le modèle d'arbre de décision

La méthode d'estimation de localisation est la même que celle de la section 3.2.3. Le programme *ANN\_GRNN\_2008.m* a été modifié afin de prendre la base de donnée au temps  $t_x$  souhaité et en estimant la position en 2D et 3D. 75% des points de

mesure ont été utilisés pour l'entraînement. Les 25% des données restantes ont été utilisées pour la phase d'estimation.

### 4.3 Résultats et analyse

#### 4.3.1 Dégradation selon l'arbre de décision

Pour l'analyse des résultats, 4 bases de données en 2D ont été générées pour les temps  $t_0$ ,  $t_1$ ,  $t_2$  et  $t_3$ . Chacune de ces bases de données contient 555 points de mesure donc, selon l'équation 4.3, 26640 arbres de décision ont été générés. En 3D, une base de données contenant 1150 points de mesure (sur 2 étages) a aussi été créée au temps  $t_2$  (55200 arbres de décision générés). Toutes ces données ont été utilisées durant les simulations pour l'estimation de la localisation.

Afin de connaître l'effet de la génération d'une base de données avec l'arbre de décision par rapport à la base de données mesurée, 15 simulations ont été faites. Le tableau 4.3 donne les résultats pour la base de données mesurée à  $t_0$  et celle générée par l'arbre de décision à 3 et à 5 mètres au temps  $t_0$ . Le tableau 4.4 donne la dégradation de l'estimation provenant de la génération de la base de données avec un arbre de décision.

Tableau 4.3 Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation pour la base de données (BD) générée avec l'arbre de décision et mesurée au temps  $t_0$

Numéro de la simulation	BD mesurée à 3 mètres (%)	BD générée à 3 mètres (%)	BD mesurée à 5 mètres (%)	BD générée à 5 mètres (%)
1	77.14	65	92.86	83.57
2	77.86	70.71	95	85
3	80	73.57	93.57	83.57
4	76.43	69.29	94.29	89.29
5	75.51	71.43	93.57	86.43
6	79.29	70	91.43	84.29
7	77.86	65.71	94.29	82.14
8	70	70	87.86	83.57
9	78.57	68.57	91.43	86.43
10	79.29	67.86	93.57	88.57
11	75.71	67.86	91.43	84.29
12	77.14	70.71	93.57	87
13	76.97	66.43	92.14	80.71
14	79.29	65.71	92.14	83.57
15	75.51	66.43	95.71	86.43
<b>Moyenne</b>	<b>77.1</b>	<b>68.33</b>	<b>92.86</b>	<b>86</b>

Tableau 4.4 Moyenne de la dégradation à 3 mètres et à 5 mètres, pour 15 simulations pour la phase d'estimation de la base de données  $t_0$

Dégradation à 3 mètres	Dégradation à 5 mètres
8.77%	6,86%

La moyenne de l'estimation utilisant la base de données mesurée au temps  $t_0$  est de 77.1% à 3 mètres et 92.86% à 5 mètres. En utilisant la base de données générée au même temps  $t_0$ , cette moyenne est de 68.33% à 3 mètres et 86% à 5 mètres ce qui est supérieur aux précisions trouvés en [10] (63% à 3 mètres, 78% à 5 mètres) donc suffisant pour effectuer une localisation. D'après le tableau 4.4, lors de la génération d'une base de donnée, la dégradation est en moyenne de 8.77% à 3 mètres et de 6.86% à 5 mètres. Ce qui veut dire que dans un canal statique, le

Le système de localisation proposé offre une dégradation acceptable de l'estimation de la position puisqu'elle est inférieure à 9%. Cette dégradation peut venir du choix et du nombre de point de référence comme expliqué dans la section 4.2.4.1.

### 4.3.2 Performances du système de localisation proposé en 2D

Après avoir déterminé les performances du système de localisation proposé dans un canal invariant, les bases de données générées aux temps  $t_1$ ,  $t_2$  et  $t_3$  sont utilisés pour connaître les performances du système quelque soit la variation du canal dans le temps. Comme le programme définissant la méthode d'estimation (ANN) proposée a été validé au chapitre 3, seule les CDFs de la phase d'estimation sont mentionnées dans cette section.

Plusieurs simulations ont été effectuées afin de valider le système de localisation proposé. Les points de mesure choisis aléatoirement pour l'estimation et l'apprentissage ont été enregistrés afin d'être utilisés dans les 3 cas de variation du canal dans le temps. Tout d'abord, des simulations en 2D selon les trois cas de variation du canal dans le temps ( $t_1$ ,  $t_2$  et  $t_3$ ) ont été effectuées. Le résultat de 15 de ces simulations est donné dans les tableaux 4.5, 4.6 et 4.7.

D'après le tableau 4.5, la précision de 3 mètres est obtenue, en moyenne, à 59.76% des cas et celle de 5 mètres est à 79.07%. Lors des simulations, la probabilité d'obtenir une précision de 3 mètres ou moins oscille entre 52.86% et 64.29% selon le choix des données pour la phase d'entraînement et celles de la phase d'estimation. Pour la précision de 5 mètres, la probabilité oscille entre 75.71% et 84.29%. Ces données sont celles pour un canal statique au temps  $t_1$  qui a changé par rapport

Tableau 4.5 Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps  $t_1$  (pour un canal invariant dans le temps)

Simulation	à 3 mètres (%)	à 5 mètres (%)
1	52.86	80
2	60.71	77.86
3	65	84.29
4	60.71	80
5	58.57	77.86
6	64.29	80.71
7	60	75.71
8	55.71	80
9	58.57	77.14
10	62.14	78.57
11	53.57	78
12	67.14	82.86
13	57.14	75.71
14	59.29	78
15	60.71	79.29
<b>Moyenne</b>	<b>59.76</b>	<b>79.07</b>

au temps  $t_0$ . Ceci prouve qu'une base de données mesurée à un temps donné est modifiée dans le temps (en terme d'années).

La dégradation due à la variation dans le temps (en terme d'années) est de 8.57% (68.33%-59.76%) à 3 mètres et de 6.93% (86%-79.07%) à 5 mètres alors que la dégradation du système proposé pour le temps  $t_1$  est de 17.34% (8.77%+8.57%) à 3 mètres et 13.79% (6.86%+6.93%) à 5 mètres.

D'après le tableau 4.6, la précision de 3 mètres est obtenue, en moyenne, à 60.63% des cas et celle de 5 mètres est à 78.02%. Lors des simulations, la probabilité d'obtenir une précision de 3 mètres ou moins oscille entre 55% et 70.71% selon le choix des données pour la phase d'entraînement et celles de la phase d'estimation. Pour la précision de 5 mètres, la probabilité oscille entre 71.43% et 84.57%. Ces données sont celles pour un canal semi-variant dans le temps  $t_2$ . La dégradation due

Tableau 4.6 Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps  $t_2$  (pour un canal semi-variant dans le temps)

Simulation	à 3 mètres (%)	à 5 mètres (%)
1	58.57	81
2	61.43	78.57
3	65	81.83
4	61.43	77.86
5	58	73.57
6	70.71	84.57
7	60.71	75
8	50.71	73.57
9	64.29	78.57
10	70	84.29
11	53.57	77.14
12	62.86	80.71
13	55	71.43
14	55.71	77.86
15	61.43	74.29
<b>Moyenne</b>	<b>60.63</b>	<b>78.02</b>

au système de localisation proposé entre les estimations avec la base de données au temps  $t_0$  générée et celle du canal semi-variant dans le temps ( $t_2$ ) est, en moyenne, de 7.7% (68.33%+60.63%) à 3 mètres et 7.98% (86%+78.02%) à 5 mètres.

D'après le tableau 4.7, la précision de 3 mètres est obtenue, en moyenne, à 63.96% des cas et celle de 5 mètres est à 80.27%. Lors des simulations, la probabilité d'obtenir une précision de 3 mètres ou moins oscille entre 60% et 72.14% selon le choix des données pour la phase d'entraînement et celle de la phase d'estimation. Pour la précision de 5 mètres, la probabilité oscille entre 75.23% et 84.29%. Ces données sont celles pour un canal variant dans le temps  $t_3$ . La dégradation due au système de localisation proposé entre les estimations avec la base de données au temps  $t_0$  générée et celles du canal variant dans le temps ( $t_3$ ) est, en moyenne, de 4.37% (68.33%+63.96%) à 3 mètres et 5.73% (86%+80.27%) à 5 mètres.

Tableau 4.7 Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps  $t_3$  (pour un canal variant dans le temps)

Simulation	à 3 mètres (%)	à 5 mètres (%)
1	62.14	80
2	65.71	81.84
3	72.14	83
4	65	81.43
5	61.02	75.23
6	69.8	84.29
7	61.43	79
8	56.43	76.43
9	66.43	82.86
10	64.29	81.43
11	63.57	82.86
12	65.71	79.87
13	60	78
14	64.29	80
15	61.43	77.86
<b>Moyenne</b>	<b>63.96</b>	<b>80.27</b>

Les trois tableaux précédents (4.5, 4.6 et 4.7) déterminent qu'une variation du canal dans le temps entraîne une dégradation moyenne égale ou inférieure à 8.5% selon l'achalandage humain et matériel. Cette dégradation se rajoute à celle de 9% obtenue lors de la génération d'une base de données avec l'arbre de décision (donc une dégradation totale égale ou inférieure à 17.5%). Dans ce cas, selon l'application nécessitant l'estimation de la localisation, il faut faire un compromis entre l'utilisation des ressources matérielles et temporelles et la précision de la localisation. D'un autre côté, une probabilité d'estimation de 67% des cas est considérée comme une précision acceptable lors d'une localisation. Cette probabilité est atteinte à environs 4 mètres.

La figure 4.12 est un exemple du CDFs pour les simulations en 2D dans un canal

variant (à  $t_3$ ). Ce dernier représente le CDF de la simulation numéro 5 du tableau 4.7. D'autres exemples de CDFs pour chaque type de variation du canal dans le temps ont été ajoutés à l'annexe II pour la phase d'estimation. Ces exemples sont les figures II.1 (simulation numéro 2 du tableau 4.5) et II.2 (simulation numéro 4 du tableau 4.6) qui représentent les CDFs en x, y et en d (distance Euclidienne) de l'erreur de localisation en 2D pour un canal invariant et semi-variant (dans le temps), respectivement.

D'après les figures II.1, II.2 et 4.12, la précision de localisation (en 2D) de 3 mètres est obtenue pour une probabilité entre 60.71% et 61.43% pour les trois cas de type de variation du canal alors qu'à la précision de 5 mètres, cette probabilité est entre 75.23% et 77.86%. Les courbes de ces figures sont "*moins lisses*" en 2D que les courbes de la figure 3.8 puisqu'elles ont été générées avec moins de points de mesure (139 points de mesure pour le 2D contre 350 points de mesures pour la figure 3.8).

L'un des problèmes de la localisation dans un canal variant dans le temps avec le système proposé est de déterminer quelle base de données utiliser dans l'ANN. Lorsque l'utilisateur connaît les caractéristiques du canal au temps  $t_x$  (canal variant fortement, moyennement ou pas du tout), il est facile de déterminer manuellement quelle base de données à utiliser. Cependant, lorsque l'utilisateur ne peut déterminer le type de canal, il faut que le système détermine de lui même la base de données à utiliser. L'article [23] propose d'ajouter un ANN qui permet de choisir la base de données à utiliser pour l'estimation de la localisation. Cette solution rend le système plus complexe. Le système de localisation proposé dans ce mémoire utilise une autre méthode plus simple. Cette dernière, consiste à définir, en même temps, les trois bases de données générées en 2D pour  $t_1$ ,  $t_2$  et  $t_3$  comme entrée du réseau

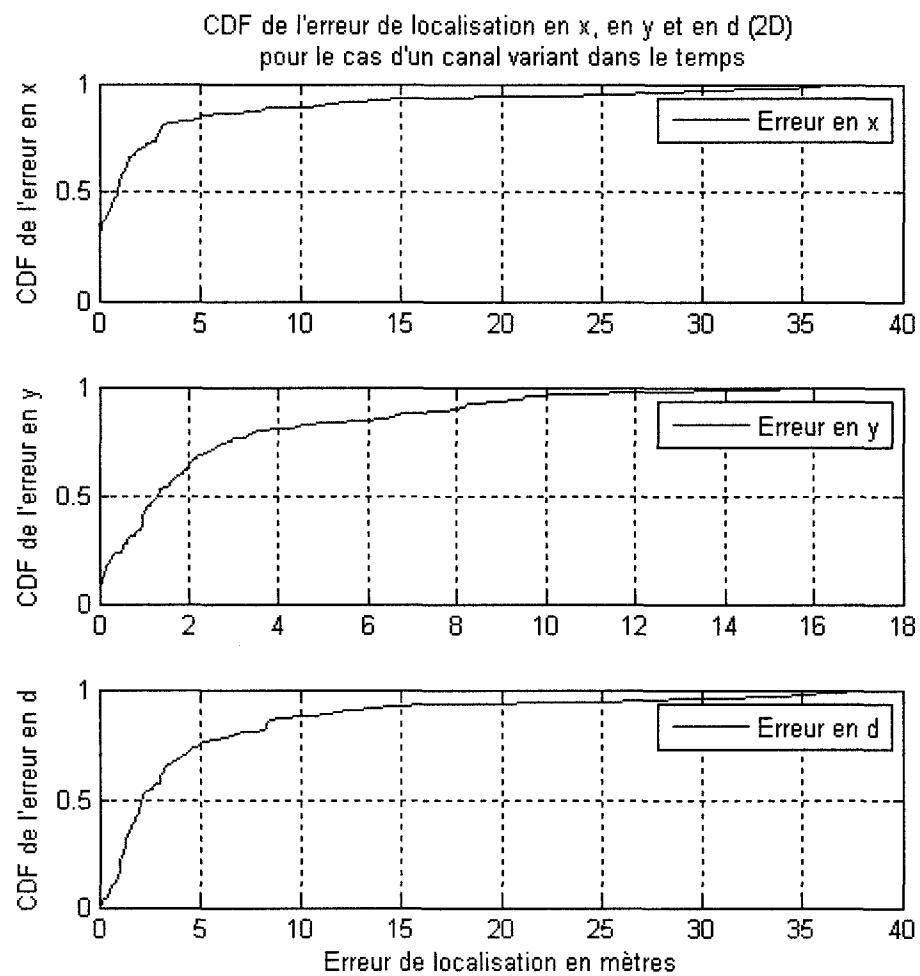


FIGURE 4.12 Exemple de CDF de l'erreur de localisation pour 139 points de mesures en x, y et en d en 2D pour un canal variant dans le temps ( $t_3$ )

de neurones. La figure 4.13 donne le schéma de la méthode.

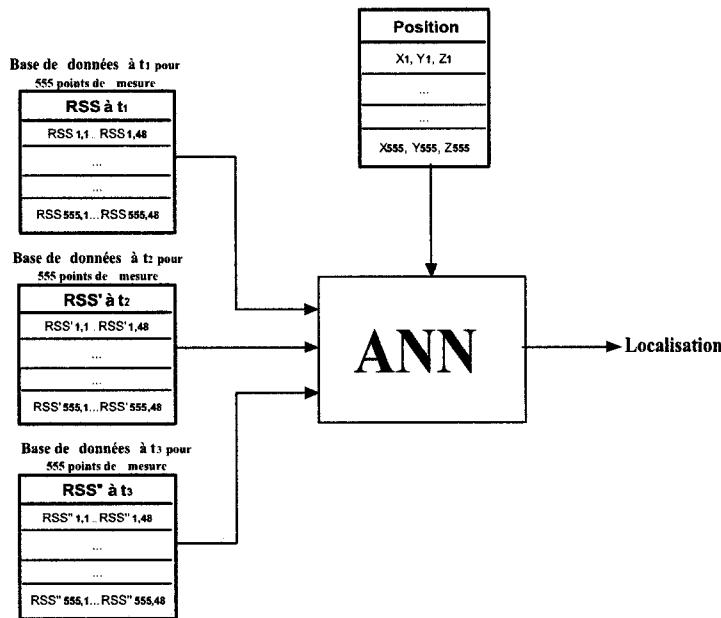


FIGURE 4.13 Schéma d'utilisation de tous les bases de données créées en 2D à l'entrée de l'ANN

Dans ce cas, l'entrée de l'ANN est une matrice de taille [1665 48] (3 \* 555 points de mesure par 48 APs) contenant les RSS des points de mesures pour les trois cas

de variation du canal dans le temps :

$$[\mathbf{RSS}]_{(1665,48)} = \left[ \begin{array}{ccc|c} RSS_{1,1} & \dots & m_{1,48} & t_1 \\ \dots & \dots & \dots & \\ RSS_{555,1} & \dots & RSS_{555,48} & \\ RSS_{556,1} & \dots & RSS_{556,48} & \\ \dots & \dots & \dots & t_2 \\ RSS_{1110,1} & \dots & RSS_{1110,48} & \\ RSS_{1111,1} & \dots & RSS_{1111,48} & \\ \dots & \dots & \dots & t_3 \\ RSS_{1665,1} & \dots & RSS_{1665,48} & \end{array} \right] \quad (4.4)$$

Finalement, afin de valider cette méthode, des simulations ont été effectuées. Le résultat de 15 de ces simulations est donné dans le tableau 4.8. La figure II.3 est un exemple de CDF en  $x$ ,  $y$  en  $d$  de l'erreur de localisation pour la simulation (numéro 7 du tableau 4.8).

D'après le tableau 4.8, la précision de 3 mètres est obtenue, en moyenne, à 61.38% des cas et celle de 5 mètres est à 78.22%. Lors des simulations, la probabilité d'obtenir une précision de 3 mètres ou moins oscille entre 55% et 70% selon le choix des données pour la phase d'entraînement et celles de la phase d'estimation. Pour la précision de 5 mètres, la probabilité oscille entre 75.48% et 83.81%. Ces résultats représentent les valeurs de précisions indépendamment du type de variation du canal dans le temps. La dégradation due au système de localisation proposé entre les estimations avec la base de données au temps  $t_0$  générée et celles du canal quelque soit sa variation dans le temps ( $t_x$ ) est, en moyenne, de 6.95% (68.33%+61.38%) à 3 mètres et 7.78% (86%+78.22%) à 5 mètres.

Tableau 4.8 Valeur du CDF en 2D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps  $t_1$ ,  $t_2$  et  $t_3$  (les 3 bases de données sont définis comme l'entrée du réseau ANN)

Simulation	à 3 mètres (%)	à 5 mètres (%)
1	62.62	80.48
2	60.95	76.9
3	69.29	83.81
4	61.19	77.38
5	60.24	76.43
6	70	81.9
7	60.48	75.48
8	55	77.86
9	61.43	77.14
10	63.81	80.95
11	59.76	77.62
12	62.38	79.52
13	54.52	73.57
14	57.62	76.43
15	61.43	77.86
<b>Moyenne</b>	<b>61.38</b>	<b>78.22</b>

Il est à remarquer que les précisions obtenues sont proches de la moyenne des précisions des 3 types de canal (à 3 mètres la moyenne de 59.76%, 60.63% et 63.96% est égale à 61.45%). En effet, l'ANN estime la position pour les trois cas ( $t_1$ ,  $t_2$  et  $t_3$ ) et calcule la moyenne de l'erreur de l'estimation selon ces trois cas. Ces données (précisions et dégradations) sont du même ordre de grandeur que celles obtenues lors des simulations en  $t_1$ ,  $t_2$  et  $t_3$ . Donc il n'est pas nécessaire de déterminer le type de canal. Il suffit dans ce cas d'utiliser toutes les informations disponibles (toutes les bases de données) afin d'effectuer la localisation.

#### 4.3.3 Performances du système de localisation proposé en 3D

Afin de valider la méthode d'arbre de décision pour la génération d'une base de données en 3D, des simulations (en 3D) selon un canal semi-variant dans le temps

(à  $t_2$ ) ont été effectuées. Le résultat de 15 de ces simulations est donné dans le tableau 4.9. La figure II.4 est un exemple de CDF en x, y, z et en d (2D et 3D) de l'erreur de localisation pour la simulation (numéro 12 du tableau 4.9) en 3D dans un canal semi-variant dans le temps. Pour l'estimation de l'étage (variable z), un seuil a été défini à 0 mètre pour le cinquième étage et à 3.9 mètres pour le sixième étage. Pour cette raison, l'analyse suivant la variable z sera faite en comptant le nombre d'erreur suivant l'étage et non par mètre.

Tableau 4.9 Valeur du CDF en 3D à 3 mètres et à 5 mètres pour 15 simulations pour la phase d'estimation au temps  $t_2$  (pour un canal semi-variant dans le temps)

Simulation	à 3 mètres (%)	à 5 mètres (%)	nombre d'erreur en z
1	62.85	76.39	6
2	55.9	78.47	4
3	54.86	76.04	3
4	53.47	71.53	7
5	58.68	75	4
6	60.42	74.31	5
7	63.54	75.69	8
8	56.25	76.04	6
9	54.86	74.31	2
10	56.6	74.65	4
11	55.9	73.61	3
12	57.29	73.96	8
13	55.56	75.69	4
14	55.21	72.22	6
15	54.51	74.31	6
<b>Moyenne</b>	<b>57.06</b>	<b>74.81</b>	<b>5</b>

D'après le tableau 4.9, la précision de 3 mètres est obtenues, en moyenne, à 57.06% des cas et celle de 5 mètres est à 74.81%. En z, il y a eu, en moyenne 5 erreurs d'étage. Lors des simulations, la probabilité d'obtenir une précision de 3 mètres ou moins oscille entre 53.74% et 63.54% selon le choix des données pour la phase d'entraînement et celles de la phase d'estimation. Pour la précision de 5 mètres, la

probabilité oscille entre 71.57% et 78.47%. En z, l'erreur selon l'étage oscille entre 2 et 8 erreurs sur 140 points de mesures. Ces données sont celles pour un canal semi-variant dans le temps ( $t_2$ ). La dégradation due au système de localisation proposé entre les estimations avec la base de données au temps  $t_0$  générée et celles du canal semi-variant dans le temps ( $t_2$ ) en 3D est, en moyenne, de 11.27% (68.33%-57.06%) à 3 mètres et 11.19% (86%-74.81%) à 5 mètres.

La dégradation due à l'ajout de la variable z (estimation de l'étage) est calculée par rapport à la moyenne de l'estimation dans un canal semi-variant dans le temps ( $t_2$ ) en 2D. Cette dégradation est de 3.57% (60.63%-57.06%) à 3 mètres et de 3.21% (78.02%-74.81%) à 5 mètres. La dégradation due à l'erreur selon z peut être calculée différemment. En effet, cette dernière est donnée par l'équation :

$$\text{Erreur}(\%) = \text{Erreur}(nombre)/\text{Nombre Point De Mesure} = 5/140 = 3.57\%.$$

D'un autre coté, une probabilité d'estimation de 67% des cas est considérée comme une précision acceptable lors d'une localisation. Cette probabilité est atteinte en moyenne à environs 4.5 mètres en 3D.

Comme la dégradation en z est faible (en moyenne 5 erreurs sur 140 positions), l'estimation de la position en 3D est valide. Cependant, puisqu'un seuil a été défini en z, l'estimation de la position en considérant le mouvement d'un mobile entre les étages (dans les escaliers et les ascenseurs) n'a pas été pris en compte. Une plus forte dégradation est attendue s'il est nécessaire de prendre en compte la position du mobile dans les escaliers et les ascenseurs.

Le programme *TestUneMesure.m* de l'annexe III.2 permet d'estimer la position du mobile (une position par simulation) et de calculer l'erreur d'estimation (distance Euclidienne) en se basant sur le système de localisation proposé en 3D. Ce

programme a été ajouté afin de connaître la possibilité de déterminer les raisons des grandes erreurs d'estimation. Dans le système proposé, les erreurs possibles sont dues soit à la base de données générée par l'arbre de décision ou au réseau de neurones artificielles. Pour ce faire, les résultats de quelques simulations avec ce programme sont donnés dans le tableau 4.10 en utilisant la base de données permettant la plus grande erreur d'estimation (donc le canal semi-variant en 3D). Les simulations des positions 2 à 9 représentent des estimations réussies par contre les simulations des positions 1 et 10 sont des estimations à grandes erreurs dues aux limitations du système proposé. L'estimation selon z n'a pas été ajoutée car il n'y a eu aucune erreur d'étages pour ces positions.

Tableau 4.10 Erreur de localisation en mètres de l'estimation de la position en 3D au temps  $t_2$

Test	X réel	Y réel	X estimé	Y estimé	Erreur
1	<b>43.55</b>	<b>42.71</b>	<b>31.86</b>	<b>45.47</b>	<b>12.01</b>
2	34	36	33.99	39.74	3.74
3	23.7	26	20.97	25.87	2.73
4	0.8	27	4.66	27.9	3.96
5	36.93	36.8	40.86	36.35	3.95
6	14	32.7	13.97	34.68	1.98
7	29	-1.8	29.99	-1.57	1.01
8	1	37	0.99	37.91	0.91
9	50.5	32.7	49.96	34.53	1.90
10	<b>87.7</b>	<b>18.8</b>	<b>2.12</b>	<b>0.29</b>	<b>87.55</b>

Tableau 4.11 Erreur de localisation en mètres de l'estimation de la position en 3D au temps  $t_2$  pour les positions proches de des positions 1 et 10 du tableau 4.10

Test	X réel	Y réel	X estimé	Y estimé	Erreur
1'	46.48	41.46	45.38	41.7	1.12
1''	40.61	42.71	43.38	42.1	2.84
10'	87.7	20.8	0.9	0.92	89

Les estimations des positions 2 à 9 du tableau 4.10 donnent des erreurs de localisation entre 0.91 et 3.95 mètres. Ces échantillons représentent bien les estimations

considérées comme réussies du système proposé. Cependant, les simulations des positions numéro 1 et 10 donnent de grandes erreurs d'estimation (12 et 87.55 mètres). Afin de vérifier les raisons de la mauvaise estimation, d'autres tests ont été effectués.

Pour la position numéro 1, en simulant à 3 mètres avant et après la position donnée (voir les positions 1' et 1'' du tableau 4.11), l'erreur est de 1.12 et 2.84 donc la grande erreur trouvée pour la simulation 1 est due au réseau de neurones. Dans ce cas, lors d'une navigation par exemple, il est facile de compenser l'erreur de localisation en fonction des estimations précédentes.

Pour la position numéro 10, en simulant à 2 mètres après la position donnée (position 10' du tableau 4.11), l'erreur d'estimation est de 89 mètres. L'erreur provient donc de la base de donnée générée par l'arbre de décision. Dans ce cas il n'est pas possible d'effectuer une compensation afin de déterminer la bonne position puisque l'erreur se répercute sur les points de mesures proches de la position estimée. Les erreurs dans les bases de données sont souvent dues au manque d'information dans la zone étudiée lors de la création de l'arbre de décision comme expliqué dans la section 4.2.4.1. En effet, en retracant la position numéro 10 (point de mesure numéro 755) dans la base de données, cette dernière ne reçoit que les RSS à faible intensité des 3 points d'accès ( $AP22 = -65dBm$ ,  $AP24 = -75dBm$  et  $AP29 = -82dBm$ ) ce qui n'est pas suffisant pour créer un arbre de décision valide dans cette zone.

## CHAPITRE 5

### CONCLUSION

Dans ce mémoire, un réseau de neurones est appliqué à un système de localisation dans des environnements internes basé sur l'approche des signatures et une technique d'exploitation de données afin de contrer la variation du canal dans le temps. Pour ce faire, des signaux de type radio fréquences d'un système de réseau sans fil WLAN sont utilisés. Ce projet se divise en deux grandes parties, la première consiste à l'estimation de la localisation dans un canal invariant dans le temps (statique) en se basant sur des données empiriques alors que la deuxième partie propose un système permettant d'estimer la localisation dans un canal variable dans le temps en se basant sur les données du canal statique.

Dans la première partie, une base de données contenant des mesures de puissance reçues (RSS en dBm) au niveau du mobile et émis par 48 points d'accès (AP) a été créée. Les mesures ont été effectuées, en 3D, sur 3 étages (3<sup>e</sup>, 5<sup>e</sup> et 6<sup>e</sup> étages) du Pavillon Lassonde de l'École Polytechnique de Montréal. La base de données, définie au temps  $t_0$ , a été utilisée dans un réseau de neurones artificielles (ANN) à régression généralisée (GRNN) afin d'estimer la localisation. Les résultats de l'estimation ont donné des précisions de 3 mètres pour 77.1% des cas et de 5 mètres à 92.86% des cas. Ces grandes valeurs sont dues au grand nombre d'information utilisées à l'entrée de l'ANN (48 APs).

Dans la deuxième partie du projet, afin de contrer la variation du canal dans le temps sans prendre des mesures empiriques, des bases de données à différents temps  $t_x$  ont été générées (en 2D et 3D) grâce à une technique d'exploitation

de données et quelques points de références. Cette technique utilise le modèle d'arbre de décision comme méthode d'analyse de régression afin de faire la prédiction temporelle des RSS au temps  $t_x$ . Des estimations de la localisation ont été effectuées avec toutes les bases de données générées par l'arbre de décision. Les résultats ont permis de démontrer que la génération des données en 2D par la méthode choisie a engendré une dégradation de l'estimation de la position inférieure à 9% par rapport aux estimations avec les données empiriques (de  $t_0$ ).

De plus, dans un canal variant dans le temps et avec le système de localisation proposé, quelque soit le type de variation du canal dans le temps une dégradation inférieure à 7% s'additionne à la dégradation précédente (de 9%). Ce système utilise toutes les bases de données générées comme entrée de l'ANN afin d'estimer la localisation en 2D. Malgré la dégradation, la précision de 3 mètres est atteinte à environs 60% des cas ce qui valide le système proposé en 2D. En 3D, lors de l'estimation de la localisation, une dégradation (dues à la variable z) d'environ 3.5% s'additionne aux précédentes. Dans ce cas la précision de 3 mètres à 60% des cas n'est pas atteinte. Donc selon l'application voulue, il faut faire le choix entre utiliser les ressources matérielles et temporelles pour collecter les données à chaque variation du canal dans le temps ou utiliser le système proposé avec la collecte de quelques mesures de référence au temps  $t_x$  afin d'estimation de la position avec une certaine dégradation de la précision.

Le système de localisation proposé dépend de plusieurs variables influençant ses performances. En effet, le nombre et l'emplacement des points de référence nécessaires pour la génération des bases de données par l'arbre de décision est important puisque c'est cette méthode qui donne la plus forte dégradation de l'estimation. Comme travaux futures, le nombre de points de références (14 points

en 2D et 25 points en 3D) pourrait être augmenté afin d'obtenir une base de données plus proche de la réalité.

De plus, comme le système repose sur la collecte de données initiale (au temps  $t_0$ ), lors de la modification d'un AP (changement de son emplacement ou son élimination), le système perd l'information du RSS émis par cet AP. Comme travaux futures, on peut donc vérifier si l'arbre de décision compense ce manque d'information avec la génération de la nouvelle base de données (puisque celle-ci dépend des mesures de référence actuelles). Cependant, tout ajout de points d'accès (AP) après la création de la base de données (à  $t_0$ ) dans l'environnement étudié n'est pas pertinent pour le système proposé puisqu'il n'accepte pas d'informations additionnelles par rapport aux données initiales.

Le système de localisation proposé, qui est implémenté dans un ordinateur portable, pourrait permettre d'effectuer la navigation et la localisation en temps réel. Dans ce cas, d'autres problèmes se posent. En effet, il faudrait que les points de référence soient des ordinateurs fixes munis d'une carte sans fil acceptant le standard *IEEE802.11*. Ces stations de base joueraient le rôle d'émetteurs/récepteurs en envoyant leurs propre RSS au mobile et en temps réel. Cette contrainte restreint donc l'emplacement et le nombre de références. De plus, pour que la localisation soit en temps réel, il faudrait générer les bases de données à différents temps et entraîner l'ANN en hors ligne. Dans ce cas, le système proposé peut utiliser l'ANN (dont l'entrée représente toutes des bases de données comme expliqué dans la section 4.3.2) afin d'estimer la localisation en temps réel.

De plus, dans le cas d'une localisation à temps réelle, une interface graphique

pourrait être ajoutée au système proposé afin de visualiser le trajet du mobile sur le plan d'un bâtiment [30].

Comme autre possibilité, il serait pertinent de prendre en compte la direction du mobile comme information supplémentaire dans la base de données. Dans le système de localisation proposé, une seule direction du mobile (le Sud) a été prise en compte. En ajoutant cette information, l'estimation de la position pourrait s'améliorer.

D'autres paramètres de détection pourraient être utilisés à la place du RSS comme les mesures de délais de transmission (TOA, TDOA ou RTT) et d'angles d'arrivée (AOA). Cependant, l'implantation de ce type de technologie nécessite l'ajout de modules dans les ordinateurs portables comme un compteur pour le délai de transmission afin d'augmenter la résolution temporelle du portable [6] (qui est pour le moment de  $1\mu$  seconde dans l'environnement Linux) ou un détecteur d'angles pour les mesures d'angles d'arrivées [31].

Finalement, afin d'augmenter la précision du système proposé, à la place des mesures à bandes étroites (mesure du RSS), des mesures à bandes larges (mesure de la réponse impulsionnelle du signal) pourrait être prises [20]. Cependant, pour le moment, aucun ordinateur portable ne permet de prendre des mesures de la réponse impulsionnelle d'un signal.

## RÉFÉRENCES

- [1] R. Hachad, "Caractérisation à bande étroite d'un canal de propagation en environnement interne à 2.4Ghz", *Projet de Maîtrise en Ingénierie, École Polytechnique de Montréal*, avril 2007.
- [2] N. Réguigui, "Radiolocalisation en extérieur dans les espaces bidimensionnel et tridimensionnel", *Projet de Maîtrise Ès Sciences Appliquées en Ingénierie, École Polytechnique de Montréal*, novembre 2006.
- [3] Jie Yin, Qiang Yang, Lionel Ni, "Adaptative Temporal Radio Maps for Indoor Location Estimation", *Proceedings. Third IEEE International Conference on Pervasive Computing and Communications*, PerCom 2005, 85-94.
- [4] M.A. Goldberg, H.A. Cho, "Introduction to Regression Analysis", *WIT Press*, 2004.
- [5] Ruiz Angel et Cornieles Ernesto, "Asservissement en niveau et température d'un réservoir d'eau avec un contrôleur neuronal", *EPM/RT, École Polytechnique de Montréal*, 1998.
- [6] M.Ciurana, F.Barcelo, F.Izquierdo, "A ranging system with IEEE802.11 data frames", *IEEE Radio and Wireless Symposium*, Proceeding 2007, p133-136.
- [7] Theodore S. Rappaport, "Wireless Communication Principles and Practice", second edition, *Prentice Hall*, Dec. 2001.
- [8] Theodore S. Rappaport, Clare D. Mcgillem, "UHF Fading in factories", *Journal on selected areas in communications*. vol.7, no.1, Junuary 1989.
- [9] Seong-Cheol Kim, Henry L. Bertoni, and Miklos Stern, "Pulse Propagation Characteristics at 2.4 GHz Inside Buildings", *Transactions on vehicular technology*, vol.45, no.3, August 1996.

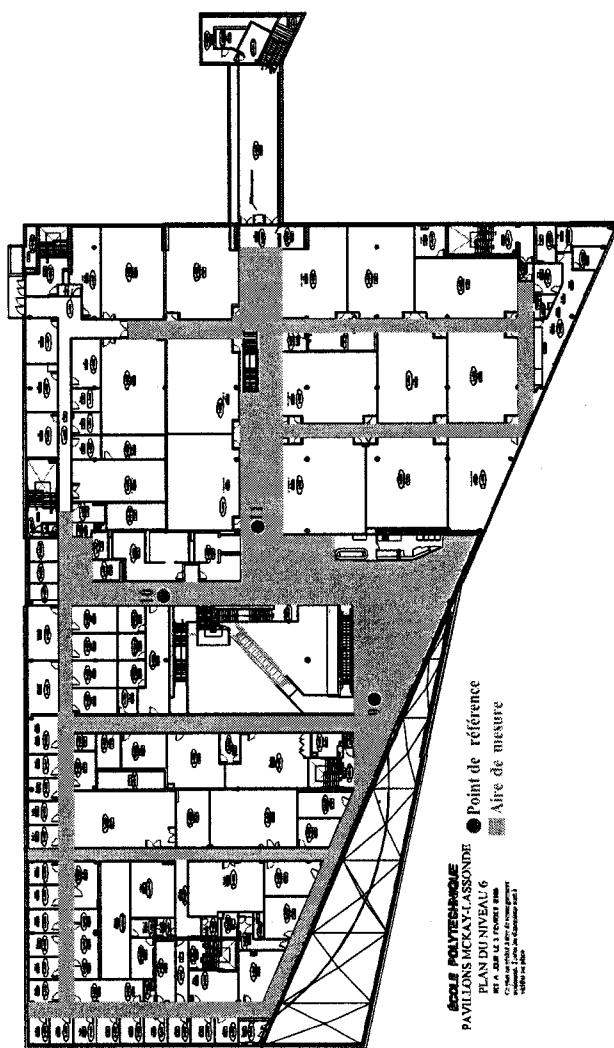
- [10] C. Nerguizian, L. Hamza, V. Nerguizian, M. Saad, "3D Indoor Geolocation With Received Signal Strength Fingerprinting Technique and Neural Networks", *4th WSEAS International Conference on Electronics, Signal Processing and Control*, ESPOCO 2005.
- [11] K. Pahlavan, A.H. Lesvesque, "Wireless information networks", *Wiley Inter-Science Publication*, 1995.
- [12] K. Pahlavan, X. Li, "An Overview of Wirless Idoor Geolocation Techniques and Systems ", *Proceeding of MWCN 2000*, Paris Mai 2000.
- [13] K. Pahlavan, X. Li, J.P. Makela "Indoor Geolocation Science and Technology", *IEEE Commun. Mag.*, vol 40, No 2, pp 112-118, Fevrier 2002.
- [14] E.D. Kaplan, et C.J. Hegarty, "Understanding GPS : Principles and applications", *Artech House*, second edition, 2006.
- [15] A. GuntHer, C. Hoene, "Measuring Round Trip Times to Determine the Distance between WLAN Nodes", *Telecommunication Networks Group, TKN Technical Reports Series*, TKN-04-16, décembre 2004.
- [16] IEEE std 802.11-1997 Information Technology-telecommunications And Information exchange Between Systems-Local And Metropolitan Area Networks specific Requirements-part 11 : Wireless Lan Medium Access Control (MAC) And Physical Layer (PHY) Specifications, 1997.
- [17] H. Demuth, M. Beale, M. Hagan, "Neural Network Toolbox 6, User's Guide", *The MathWorks, Inc*, 1992-2008.
- [18] L. Breiman, et al., "Classification and Regression Trees", *Chapman and Hall*, Boca Raton, 1993.
- [19] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, "Classification and Regression Trees", *Technical report, Wadsworth International*, Monterey, CA, 1984.

- [20] C. Nerguizian, "Radiolocalisation en milieu minier souterrain", *Thèse de doctorat, INRS-SEMT*, Août 2003.
- [21] P. Bahl, V.H. Padmanabhan, "RADAR : an in-building RF-based user location and tracking system", *Proceedings IEEE Infocom 2000*, vol 2, pp 775-784, Tel Aviv, Mars 2000.
- [22] K. Kaemarungsi, "Design of indoor positioning systems based on location fingerprinting technique", *University of Pittsburgh*, 2005.
- [23] C. Nerguizian, C. Despins, S. Affès, G.I. Wassi, D.Grenier, "Neural network and fingerprinting-based geolocation on time-varying channels", *The 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication*, PIMRC 2006.
- [24] N. B. Priyantha, A. Chakraborty, H. Balakrishnan, "The cricket locationsupport system", *Proc. ACM International Conference on Mobile Computing and Networking*, MOBICOM 2000, pp 32-43, Boston, aout 2000.
- [25] Netstumbler, "www.netstumbler.org", décembre 2008.
- [26] R. WANT, A. HOPPER et al, "The Active badge Location System", *Olivertti Research Ltd*, Cambridge, 1992.
- [27] A. Harter, A. Hopper, P. Steggles, A.Ward, P.Webster, "The anatomy of a context-aware application", *Proc. ACM International Conference on Mobile Computing and Networking*, MOBICOM 1999, Seattle, Aout 1999.
- [28] L.M. Ni, Y. Liu, Y.C. Lau, A.P. Patil, "LANDMARC : Indoor Location Sensing Using Active RFID", *Wireless Networks*, vol10, no 6, pp 701-710, novembre 2004.
- [29] J. Geier, "Understanding 802.11 Frame types", *Wi-Fi Panet*, aout 2002.
- [30] Ekahau Inc. , "1851 Alexander Bell Drive, Suite 300, Reston, VA 20191", *Tel : 1-866-4EKAHAU, http://www.ekahau.com*, 2008.

- [31] C.H Lim, B. Poh Ng, D. Duan, "Robust methods for AOA geo-location in a real-time indoor WiFi system", *Journal of Location Based Services*, vol 2, no 2, pp 112-121, juin 2008.
- [32] A. Hatami, K. Pahlavan, "Hybrid TOA-RSS Based Localization Using Neural Network", *IEEE Global Telecommunications Conference, IEEE GLOBECOM 2006*, 2006.

## ANNEXE I

## PLAN DU PAVILLON LASSONDE

FIGURE I.1 6<sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal

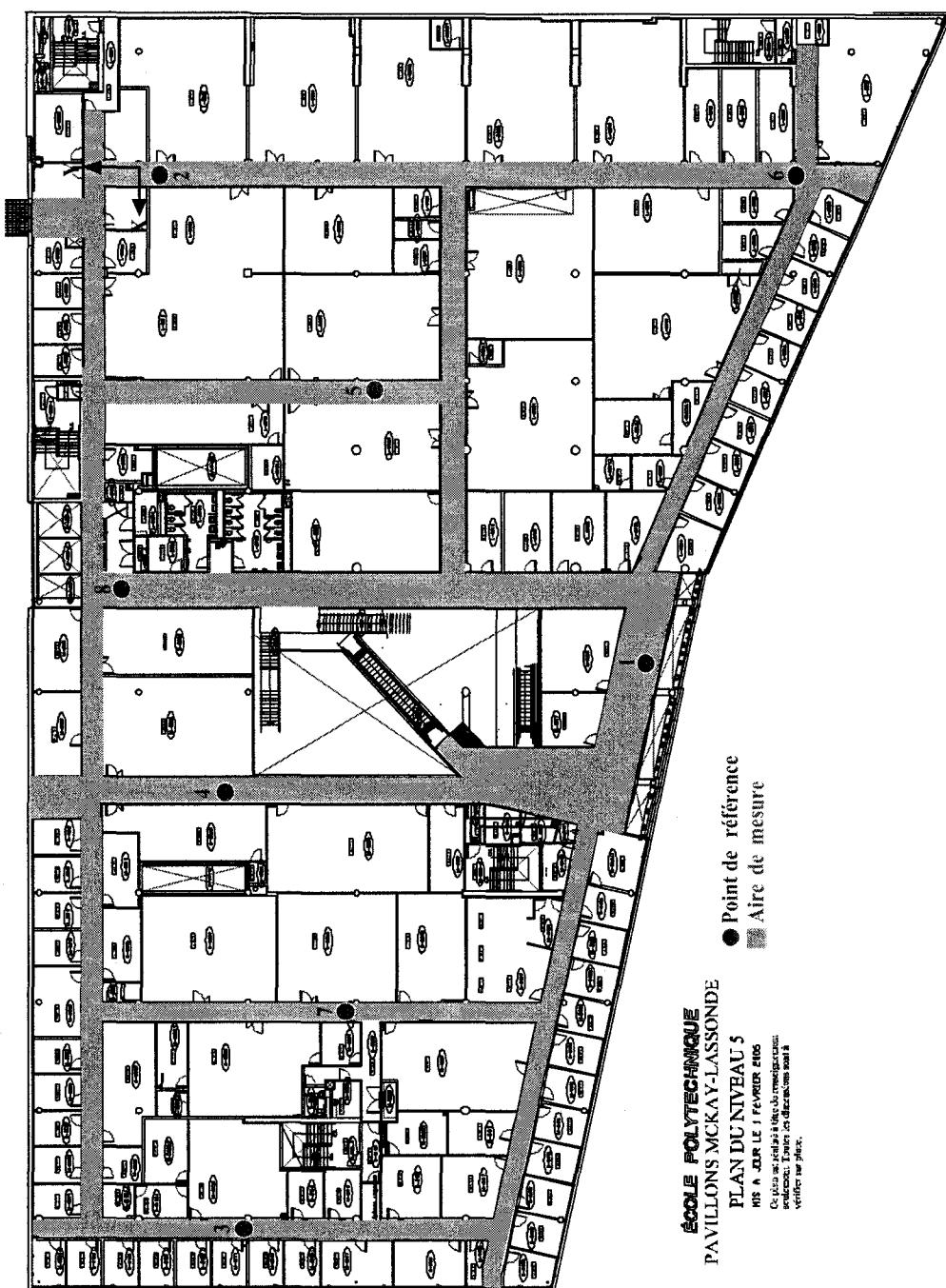


FIGURE I.2 5<sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal

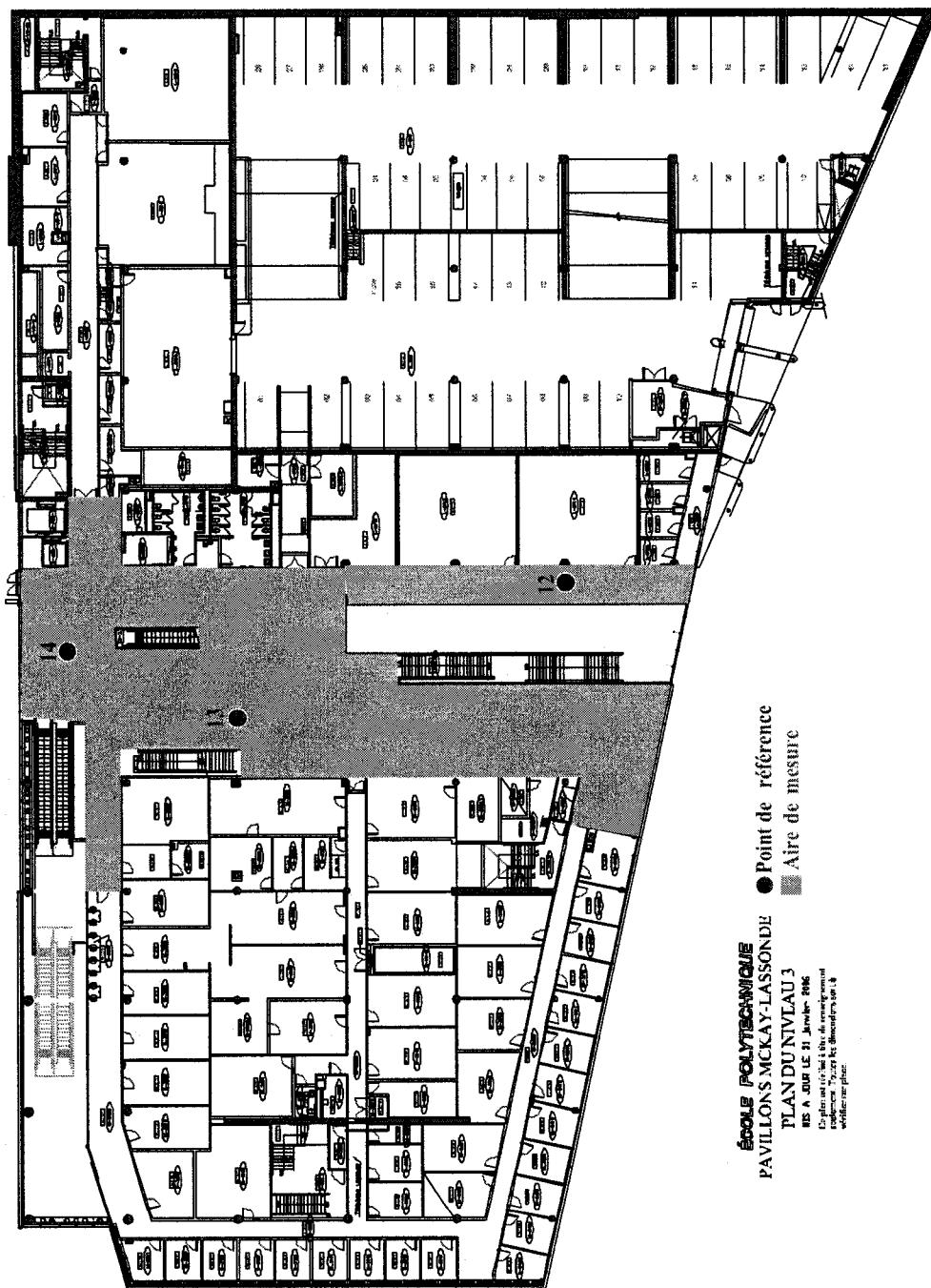


FIGURE I.3 3<sup>e</sup> étage du pavillon Lassonde de l'École Polytechnique de Montréal

## ANNEXE II

**EXEMPLES DE CDFS DE L'ERREUR DE LOCALISATION POUR  
DIFFÉRENTS TYPES DE VARIABILITÉ DU CANAL DANS LE  
TEMPS**

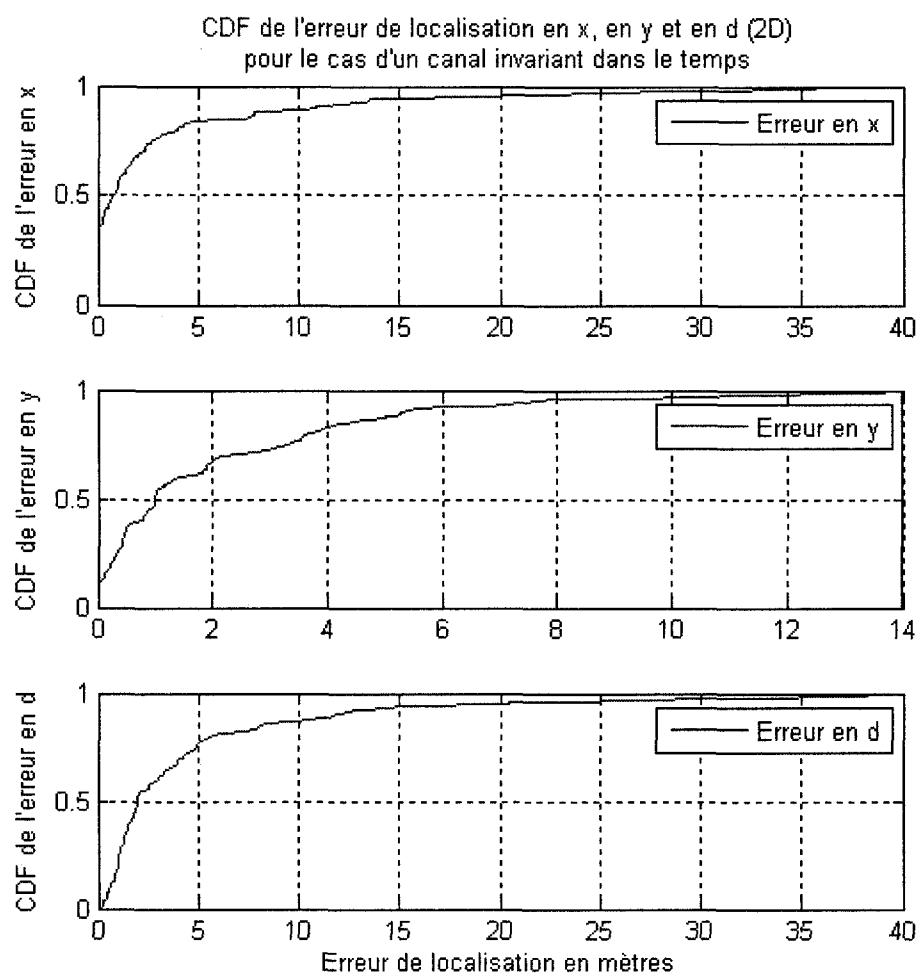


FIGURE II.1 Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour un canal invariant dans le temps ( $t_1$ )

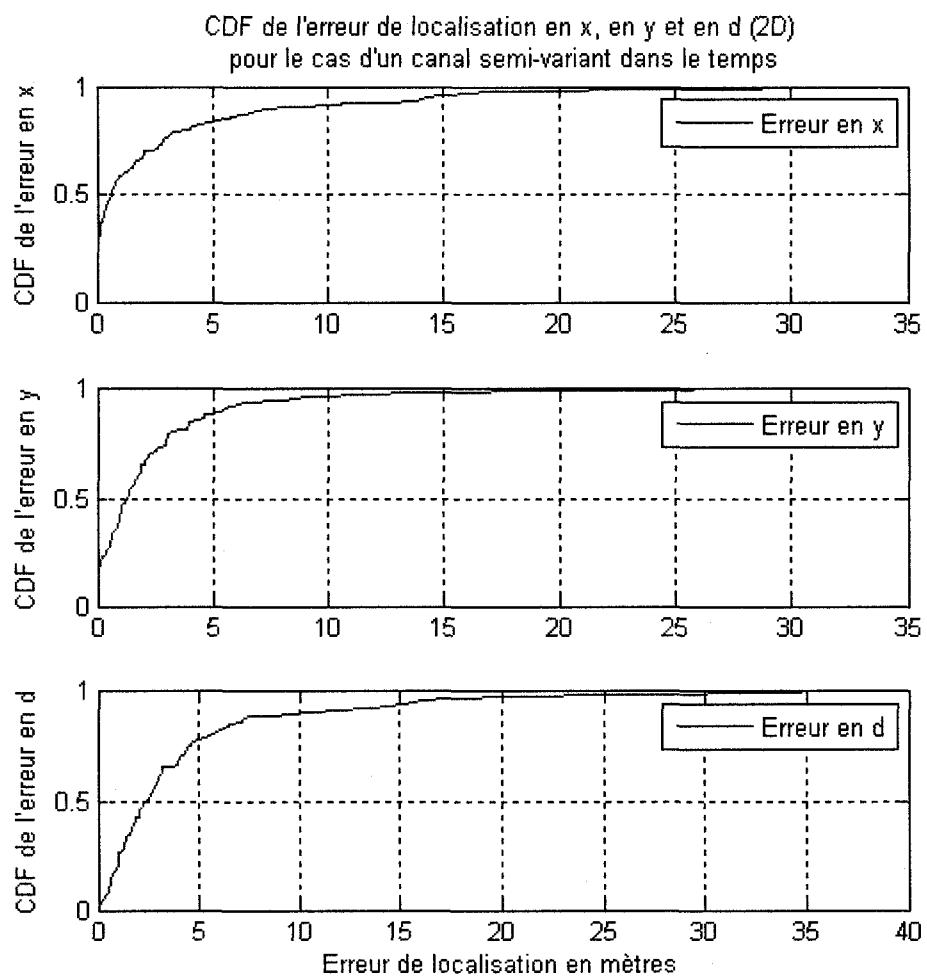


FIGURE II.2 Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour un canal semi-variant dans le temps ( $t_2$ )

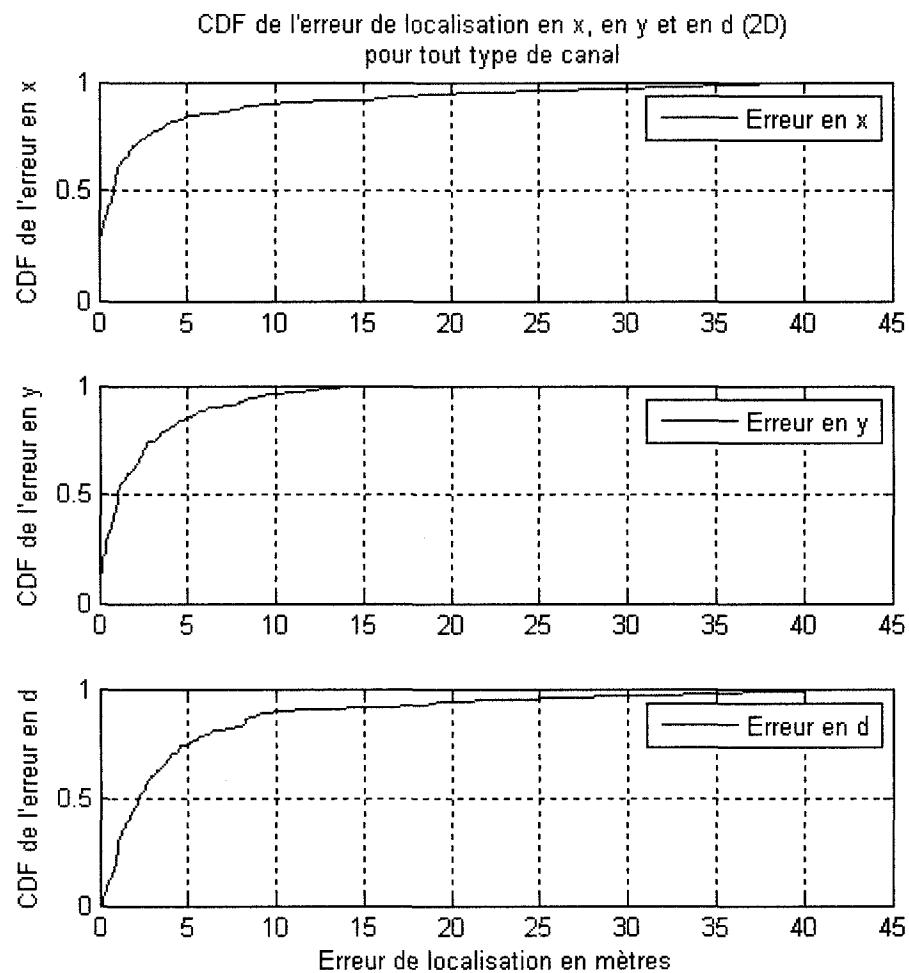


FIGURE II.3 Exemple de CDF de l'erreur de localisation pour 139 points de mesure en x, y et en d en 2D pour tout type de canal ( $t_1$ ,  $t_2$  et  $t_3$ )

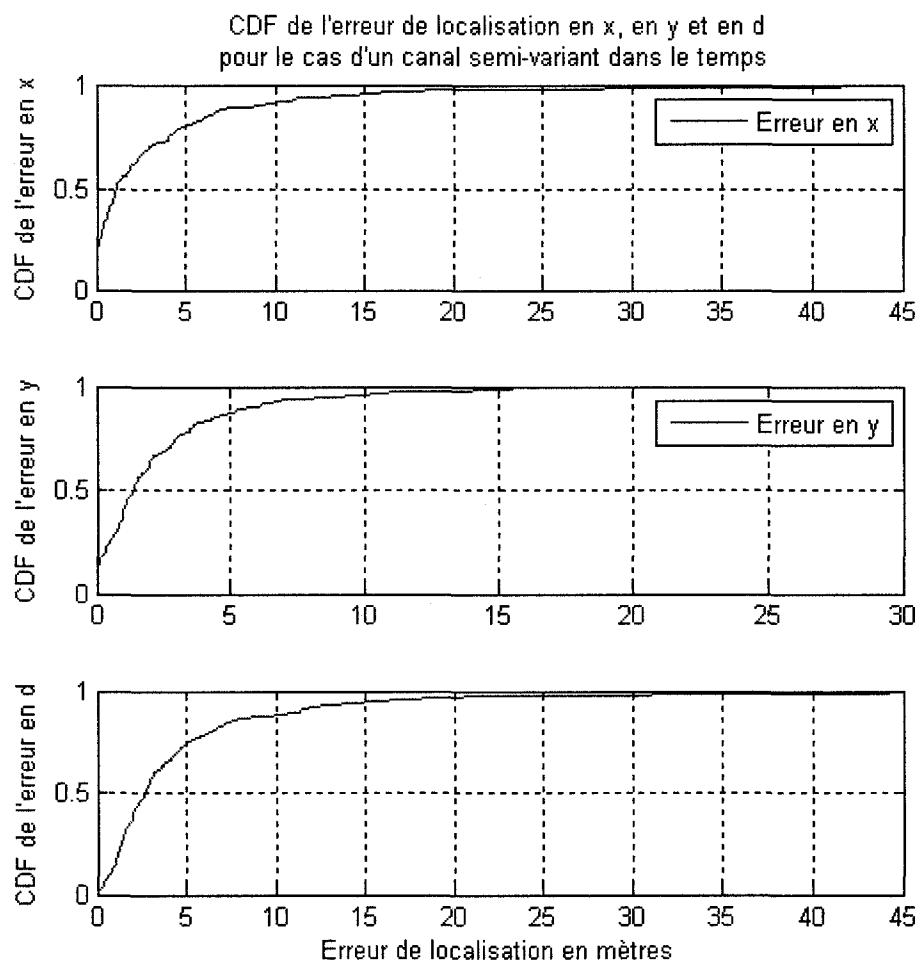


FIGURE II.4 Exemple de CDF de l'erreur de localisation pour 140 points de mesure en x, y, z et en d en 3D pour un canal semi-variant dans le temps ( $t_2$ )

### ANNEXE III

#### FICHIERS MATLAB

##### III.1 ANN\_GRNN\_2008.m

```
%-----%
% Nom du fichier: ANN_GRNN_2008.m
% Description: Fichier permettant d'obtenir le CDF grâce au
% réseau de neurones ANN
% Modifié par: Lamia Hamza
% 2008
%-----%

clear
close all

%-----%
% Lecture du fichier excell contenant la base de donnée t0
A = xlsread('U:\Data\ANN PFE\Resultats_1.xls');
%-----%
% Partie variation du canal
% A = xlsread('U:\Data\ANN PFE\Resultats_DATA_T1.xls');
%-----%

A1 = A(1:length(A),:);
```

X	= A1(:,1); Y	= A1(:,2); Z	= A1(:,3);
P1	= A1(:,4); P2	= A1(:,5); P3	= A1(:,6);
P4	= A1(:,7); P5	= A1(:,8); P6	= A1(:,9);
P7	= A1(:,10); P8	= A1(:,11); P9	= A1(:,12);
P10	= A1(:,13); P11	= A1(:,14); P12	= A1(:,15);
P13	= A1(:,16); P14	= A1(:,17); P15	= A1(:,18);
P16	= A1(:,19); P17	= A1(:,20); P18	= A1(:,21);
P19	= A1(:,22); P20	= A1(:,23); P21	= A1(:,24);
P22	= A1(:,25); P23	= A1(:,26); P24	= A1(:,27);
P25	= A1(:,28); P26	= A1(:,29); P27	= A1(:,30);
P28	= A1(:,31); P29	= A1(:,32); P30	= A1(:,33);
P31	= A1(:,34); P32	= A1(:,35); P33	= A1(:,36);
P34	= A1(:,37); P35	= A1(:,38); P36	= A1(:,39);
P37	= A1(:,40); P38	= A1(:,41); P39	= A1(:,42);
P40	= A1(:,43); P41	= A1(:,44); P42	= A1(:,45);
P43	= A1(:,46); P44	= A1(:,47); P45	= A1(:,48);
P46	= A1(:,49); % P47	= A1(:,50); P48	= A1(:,51);

XX=X'; YY=Y'; ZZ=Z';  
 PP1=P1'; PP2=P2'; PP3=P3'; PP4=P4'; PP5=P5'; PP6=P6'; PP7=P7';  
 PP8=P8'; PP9=P9'; PP10=P10'; PP11=P11'; PP12=P12'; PP13=P13';  
 PP14=P14'; PP15=P15'; PP16=P16'; PP17=P17'; PP18=P18'; PP19=P19';  
 PP20=P20'; PP22=P22'; PP23=P23'; PP24=P24'; PP25=P25'; PP26=P26';  
 PP27=P27'; PP29=P29'; PP30=P30'; PP31=P31'; PP32=P32'; PP33=P33';  
 PP34=P34'; PP36=P36'; PP37=P37'; PP38=P38'; PP40=P40'; PP41=P41';  
 PP42=P42'; PP43=P43'; PP44=P44'; PP45=P45'; PP46=P46'; PP48=P48';  
 % PP21=P21'; PP28=P28'; PP35=P35'; PP39=P39'; PP47=P47';

```
p=[PP1;PP2;PP3;PP4;PP5;PP6;PP7;PP8;PP9;PP10;PP11;PP12;PP13;PP14;  
PP15;PP16;PP17;PP18;PP19;PP20;PP22;PP23;PP24;PP25;PP26;PP27;PP29;  
PP30;PP31;PP32;PP33;PP34;PP36;PP37;PP38;PP40;PP41;PP42;PP43;PP44;  
PP45;PP46;PP48];
```

```
%-----%
```

```
t0=[XX;YY];
```

```
t1=[ZZ];
```

```
tall=[t0; t1];
```

```
[R,Q]=size(p); % Taille des données d'entrée
```

```
% 350 mesures pour test présent 1/4
```

```
%-----%
```

```
% iits=2:4:Q; % Indice pour les données de test (1/4)
```

```
% iitr=[1:4:Q 3:4:Q 4:4:Q]; % Indice pour l'apprentissage
```

```
% %-----%
```

```
% 350 pour test présent aléatoirement
```

```
S = 350;
```

```
%-----%
```

```
% Pour variation du canal sur l'étage 5 avec test aléatoires
```

```
% S = 138;
```

```
%-----%
```

```
Temp=1:1:Q;
```

```
iits=[]; iits_tp=[];
```

```
iitr=[];
j=[]; L=1;

for i=1:S,
    iits(i) = round((Q)*rand);
    iits = [iits iits(i)];
end

iits = iits(1:S);
iits_tp=iits;

while (L<=length(iits_tp))
    P = find(iits_tp==iits_tp(L));
    %ici creer un iits sans redodance
    if length(P)>1
        for k=2:length(P)
            iits_tp = [iits_tp(1:(P(k)-1)) iits_tp((P(k)+1):
length(iits_tp))];
        end
    end
    L=L+1;
end

for i=1:1:length(iits_tp),
    %Maintenant enlever les valeurs iits de iitr (valeurs pour
    % tests non utilisees durant l entrainement)
    j = find(Temp==iits_tp(i));
    if j==1
        Temp = [Temp((j+1):length(Temp))];
```

```

    else
        Temp = [Temp(1:(j-1)) Temp((j+1):length(Temp))];
    end
end
iitr=Temp;
%-----%
% Changement de 1D à 2D (des données en mètres au target).
tz=[];
for i=1:Q,
    if t1(i) == 3.9
        tz = [tz [1; 1]]; % défini le 6ème étage
    elseif t1(i) == min(t1) %-7.8
        tz = [tz [-1 ;-1]]; % défini le 3ème étage
    elseif ((t1(i)>=-0.5)&&(t1(i)<=0.5))
        tz = [tz [-1 ;1]]; % défini le 5ème étage et les tables
    else
        tz = [tz [1 ;-1]]; % défini les tables du 6ème étage
    end
end

t=[t0 ;tz];

%----- ANN -----%
ptr=p(:,iitr);ttr=t(:,iitr); % Données d'entrée pour l'apprentissage
pts=p(:,iits);tts=t(:,iits); % Données d'entrée pour le test

spread=4.5;

```

```

net=newgrnn(ptr,ttr,spread); % Création d'un réseau grnn

atr=sim(net,ptr); % Phase d'appel avec les données utilisées
                     lors de l'apprentissage

ats=sim(net,pts); % Phase d'appel avec les données utilisées
                     lors du test

%-----
atr1=[];

l=length(iitr);
for i=1:l, % Des donnees target en mètre
  if round(atr(3:4,i)) == [1; 1]
    atr1 = [atr1 3.9]; % défini le 6ème étage
  elseif round(atr(3:4,i)) == [-1 ;-1]
    atr1 = [atr1 -7.8]; % défini le 3ème étage
  elseif round(atr(3:4,i)) == [-1 ;1]
    atr1 = [atr1 t1(1)]; % défini le 5ème étage
  else
    atr1 = [atr1 4.16];
  end
end

atr=[atr(1:2,:); atr1];
%-----

% Courbes de sortie, cas de données pour l'apprentissage
i=1:1:length(iitr);
figure(1)
ttr=tall(:,iitr); % Les cibles pour l'apprentissage

```

```
subplot(3,1,1),
plot(i,atr(1,:),'b.',i,ttr(1,:),'r.'),grid,
ylabel('Positions en x'),
title('Positions estimées et exactes en termes de x et de y
définies en fonction des points de mesures de l''usager')
subplot(3,1,2),
plot(i,atr(2,:),'b.',i,ttr(2,:),'r.'),grid,
xlabel('Points de mesures de la station mobile '),
ylabel('Positions en y'),
title('Positions estimées et exactes en termes de x et de y
définies en fonction des points de mesures de l''usager')
subplot(3,1,3),
plot(i,atr(3,:),'b.',i,ttr(3,:),'r.'),grid,
ylabel('Positions en z'),
title('Positions estimées et exactes en termes de x, de y
et de z définies en fonction des points de mesures de l''usager')

figure(2)
etr(1,:)=atr(1,:)-ttr(1,:);
subplot(2,2,1),
plot(i,etr(1,:)),grid,
ylabel('Erreur en x'),
title('Erreurs de localisation en termes de x, de y et de d
définies en fonction des points de mesures de l''usager'),
etr(2,:)=atr(2,:)-ttr(2,:);
subplot(2,2,2),
plot(i,etr(2,:)),grid,
ylabel('Erreur en y'),
```

```
etr(3,:)=atr(3,:)-ttr(3,:);

subplot(2,2,3),
plot(i,etr(3,:)),grid,
ylabel('Erreur en z'),

% distance Euclidienne entre la position réelle et estimée 3D
edtr3=sqrt(((atr(1,:)-ttr(1,:)).^2)+((atr(2,:)-ttr(2,:)).^2)+
((atr(3,:)-ttr(3,:)).^2));

% % distance Euclidienne entre la position réelle et estimée 2D
edtr=sqrt(((atr(1,:)-ttr(1,:)).^2)+((atr(2,:)-ttr(2,:)).^2));

moyedtr=mean(edtr);
maxedtr=max(edtr);
moyedtr=mean(edtr3);%3D
maxedtr=max(edtr3);%3D
subplot(2,2,4),plot(i,edtr),grid,
xlabel('Points de mesures de la station mobile'),
ylabel('Erreur en d '),

maxextr=max(etr(1,:));
minextr=min(etr(1,:));
maxeytr=max(etr(2,:));
mineytr=min(etr(2,:));
maxeztr=max(etr(3,:));
mineztr=min(etr(3,:));
etra=abs(etr);

figure(3)
[Fextr,extr]=cdfcalc(etra(1,:));
```

```
[Feytr,eytr]=cdfcalc(tra(2,:));  
[Feztr,eztr]=cdfcalc(tra(3,:));  
extr(length(extr)+1)=extr(length(extr));  
eytr(length(eytr)+1)=eytr(length(eytr));  
eztr(length(eztr)+1)=eztr(length(eztr));  
  
[Fetrd,etrd]=cdfcalc(edtr);  
etrd(length(etrd)+1)=etrd(length(etrd));  
  
[Fetrd3,etrd3]=cdfcalc(edtr3);%3D  
etrd3(length(etrd3)+1)=etrd3(length(etrd3));%3D  
  
subplot(3,2,1),  
plot(extr,Fextr),grid  
ylabel('CDF de l''erreur en x'),  
legend('Erreur en x')  
title('CDF de l'erreur de localisation en x,en y,en z et en d')  
subplot(3,2,2),  
plot(eytr,Feytr),grid  
ylabel('CDF de l''erreur en y'),  
legend('Erreur en y')  
subplot(3,2,3),  
plot(eztr,Feztr),grid  
ylabel('CDF de l''erreur en z'),  
legend('Erreur en z')  
subplot(3,2,4),  
plot(etrd,Fetrd),grid  
title('CDFs of Localisation Error in d (2D)')
```

```
xlabel('Localisation error (in meters)')
ylabel('CDF in terms of d'),

subplot(3,2,5:6),
plot(etr3,Fetr3),grid
title('CDFs of Localisation Error in d (3D)')
xlabel('Localisation error (in meters)')
ylabel('CDF in terms of d'),

% Courbes de sortie, cas de données pour le test
ats1=[];
l=length(ats);

for j=1:l,
    if round(ats(3:4,j)) == [1; 1]
        ats1 = [ats1 3.9]; % défini le 5ème étage
    elseif round(ats(3:4,j)) == [-1 ;-1]
        ats1 = [ats1 -7.8]; % défini le 2ème étage
    elseif round(ats(3:4,j)) == [-1 ;1]
        ats1 = [ats1 t1(1)]; % défini le 5ème étage
    else
        ats1 = [ats1 4.16];
    end
end

ats=[ats(1:2,:); ats1];

j=1:1:length(iits);
```

```
figure(4)

tts=tall(:,iits);

subplot(3,1,1),
plot(j,ats(1,:),'b.',j,tts(1,:),'r.'),grid,
ylabel('Positions en x'),
title('Positions estimées et exactes en termes de x et de y
définies en fonction des points de mesures de l''usager')

subplot(3,1,2),
plot(j,ats(2,:),'b.',j,tts(2,:),'r.'),grid,
xlabel('Points de mesures de la station mobile'),
ylabel('Positions en y'),

subplot(3,1,3),
plot(j,ats(3,:),'b.',j,tts(3,:),'r.'),grid,
ylabel('Positions en z'),
title('Positions estimées et exactes en termes de x et de y
définies en fonction des points de mesures de l''usager')

figure(5)

ets(1,:)=ats(1,:)-tts(1,:);

subplot(2,2,1),
plot(j,ets(1,:)),grid,
ylabel('Erreur en x'),
% title('Erreurs de localisation en termes de x, de y et de d
définies en fonction des points de mesures de l''usager'),
ets(2,:)=ats(2,:)-tts(2,:);

subplot(2,2,2),
plot(j,ets(2,:)),grid,
```

```
ylabel('Erreur en y'),
ets(3,:)=ats(3,:)-tts(3,:);
subplot(2,2,3),
plot(j,ets(3,:)),grid,
ylabel('Erreur en z'),
% distance Euclidienne entre la position réelle et estimée 3D
edts3=sqrt(((ats(1,:)-tts(1,:)).^2)+((ats(2,:)-tts(2,:)).^2)+((ats(3,:)-tts(3,:)).^2));
% distance Euclidienne entre la position réelle et estimée 2D
edts=sqrt(((ats(1,:)-tts(1,:)).^2)+((ats(2,:)-tts(2,:)).^2));
moyedts=min(edts);
maxedts=max(edts);
moyedts3=min(edts3);%3D
maxedts3=max(edts3);%3D

subplot(2,2,4),
plot(j,edts),grid,%2D
xlabel('Points de mesures de la station mobile'),
ylabel('Erreur en d'),

maxexts=max(ets(1,:));
minexts=min(ets(1,:));
maxeyts=max(ets(2,:));
mineyts=min(ets(2,:));
maxexts=max(ets(3,:));
minexts=min(ets(3,:));
etsa=abs(ets);
```

```
figure(6)

[Fexts,exts]=cdfcalc(etsa(1,:));
[Feyts,eyts]=cdfcalc(etsa(2,:));
[Fezts,ezts]=cdfcalc(etsa(3,:));
exts(length(exts)+1)=exts(length(exts));
eyts(length(eyts)+1)=eyts(length(eyts));
ezts(length(ezts)+1)=ezts(length(ezts));

[Fetsd,etsd]=cdfcalc(edts);
etsd(length(etsd)+1)=etsd(length(etsd));

[Fetsd3,etsd3]=cdfcalc(edts3);%3D
etsd3(length(etsd3)+1)=etsd3(length(etsd3));%3D

subplot(3,2,1),plot(exts,Fexts),grid
ylabel('CDF de l''erreur en x'),
legend('Erreur en x')
title('CDF de l''erreur de localisation en x, en y et en d')

subplot(3,2,2),plot(eyts,Feyts),grid
ylabel('CDF de l''erreur en y'),
legend('Erreur en y')
subplot(3,2,3),plot(ezts,Fezts),grid
ylabel('CDF de l''erreur en z'),
legend('Erreur en z')
subplot(3,2,4),
plot(etsd,Fetsd,'b'),grid
title('CDFs of Localisation Error in d (2D)')
```

```
 xlabel('Localisation error (in meters)')
 ylabel('CDF in terms of d'),

 subplot(3,2,5:6),
 plot(etsd3,Fetsd3,'b'),grid
 title('CDFs of Localisation Error in d (3D)')
 xlabel('Localisation error (in meters)')
 ylabel('CDF in terms of d'),

 %-----%
 % Écrire dans deux fichiers txt les données avec
 % lesquels on a simulé
 Datas= [iits; ats];
 Datar=[iitr;atr];
 foos = fopen('datas.txt','wt');
 fprintf(foos, '%6.2f %12.8f %12.8f %12.8f\n', Datas);
 foor = fopen('datar.txt','wt');
 fprintf(foor, '%6.2f %12.8f %12.8f %12.8f\n', Datar);
 fclose(foos);
 fclose(foor);
 %-----%
```

III.2 *TestUneMesure.m*

```

%-----%
% Nom du fichier: TestUneMesure.m
% Description: Programme permettant d'obtenir la position
% estimée d'un point de mesure et l'erreur Euclidienne entre
% la position réelle et exacte.
% Créer par : Lamia Hamza
% 2008
%-----%
clear
close all
%-----%
%Charger la base de données à Tx
A =xlsread('U:\Data\ANN PFE\Resultats_DATA_3D_Ref25.xls');
%-----%
% Charger le point test de taille [60 48] à tester
B=load('U:\Data\Point de mesure\3D\Pt327.txt');
%-----%
tts=[34;37.7;-1; 1]; % Les coordonnées réelles du point de test
% pour mesurer l'erreur de position
A1 = A(1:length(A),:);
X = A1(:,1); Y = A1(:,2); Z = A1(:,3);
P1 = A1(:,4); P2 = A1(:,5); P3 = A1(:,6);
P4 = A1(:,7); P5 = A1(:,8); P6 = A1(:,9);
P7 = A1(:,10); P8 = A1(:,11); P9 = A1(:,12);
P10 = A1(:,13); P11 = A1(:,14); P12 = A1(:,15);
P13 = A1(:,16); P14 = A1(:,17); P15 = A1(:,18);

```

P16	= A1(:,19);	P17	= A1(:,20);	P18	= A1(:,21);
P19	= A1(:,22);	P20	= A1(:,23);	P21	= A1(:,24);
P22	= A1(:,25);	P23	= A1(:,26);	P24	= A1(:,27);
P25	= A1(:,28);	P26	= A1(:,29);	P27	= A1(:,30);
P28	= A1(:,31);	P29	= A1(:,32);	P30	= A1(:,33);
P31	= A1(:,34);	P32	= A1(:,35);	P33	= A1(:,36);
P34	= A1(:,37);	P35	= A1(:,38);	P36	= A1(:,39);
P37	= A1(:,40);	P38	= A1(:,41);	P39	= A1(:,42);
P40	= A1(:,43);	P41	= A1(:,44);	P42	= A1(:,45);
P43	= A1(:,46);	P44	= A1(:,47);	P45	= A1(:,48);
P46	= A1(:,49);	P47	= A1(:,50);	P48	= A1(:,51);

XX=X'; YY=Y'; ZZ=Z';

PP1=P1'; PP2=P2'; PP3=P3'; PP4=P4'; PP5=P5'; PP6=P6'; PP7=P7';  
 PP8=P8'; PP9=P9'; PP10=P10'; PP11=P11'; PP12=P12'; PP13=P13';  
 PP14=P14'; PP15=P15'; PP16=P16'; PP17=P17'; PP18=P18'; PP19=P19';  
 PP20=P20'; PP21=P21'; PP22=P22'; PP23=P23'; PP24=P24'; PP25=P25';  
 ; PP26=P26'; PP27=P27'; PP28=P28'; PP29=P29'; PP30=P30';  
 PP31=P31'; PP32=P32'; PP33=P33'; PP34=P34'; PP35=P35'; PP36=P36';  
 PP37=P37'; PP38=P38'; PP39=P39'; PP40=P40'; PP41=P41'; PP42=P42';  
 PP43=P43'; PP44=P44'; PP45=P45'; PP46=P46'; PP47=P47'; PP48=P48';

p=[PP1;PP2;PP3;PP4;PP5;PP6;PP7;PP8;PP9;PP10;PP11;PP12;PP13;  
 PP14;PP15;PP16;PP17;PP18;PP19;PP20;PP22;PP23;PP24;PP25;PP26;  
 PP27;PP29;PP30;PP31;PP32;PP33;PP34;PP36;PP37;PP38;PP40;PP41;  
 PP42;PP43;PP44;PP45;PP46;PP48];  
 %-----%  
 t0=[XX;YY];

```
t1=[ZZ];  
  
tall=[t0; t1];  
  
[R,Q]=size(p); % Taille des données d'entrée  
  
% 350 mesures pour test présent 1/4  
%-----%  
iits=2:4:Q; % Indice pour les données de test (1/4)  
iitr=1:1:Q; % Indice pour les données d'apprentissage  
  
% Changement d'une dimension à 2 dimensions  
% (des données en mètres au target).  
tz=[];  
for i=1:Q,  
    if t1(i) == 3.9  
        tz = [tz [1; 1]]; % défini le 6ème étage  
    elseif t1(i) == min(t1) %-7.8  
        tz = [tz [-1 ;-1]]; % défini le 3ème étage  
    elseif ((t1(i)>=-0.5)&&(t1(i)<=0.5))  
        tz = [tz [-1 ;1]]; % défini le 5ème étage et les tables  
    else  
        tz = [tz [1 ;-1]]; % défini les tables du 6ème étage  
    end  
end  
  
t=[t0 ;tz];
```

```
% Calcul de la moyenne des RSS du point à tester afin de
% créer l'arbre de décision
BAP1 = B(:,1); MoyB1 = mean(BAP1);
BAP2 = B(:,2); MoyB2 = mean(BAP2);
BAP3 = B(:,3); MoyB3 = mean(BAP3);
BAP4 = B(:,4); MoyB4 = mean(BAP4);
BAP5 = B(:,5); MoyB5 = mean(BAP5);
BAP6 = B(:,6); MoyB6 = mean(BAP6);
BAP7 = B(:,7); MoyB7 = mean(BAP7);
BAP8 = B(:,8); MoyB8 = mean(BAP8);
BAP9 = B(:,9); MoyB9 = mean(BAP9);
BAP10 = B(:,10); MoyB10 = mean(BAP10);
BAP11 = B(:,11); MoyB11 = mean(BAP11);
BAP12 = B(:,12); MoyB12 = mean(BAP12);
BAP13 = B(:,13); MoyB13 = mean(BAP13);
BAP14 = B(:,14); MoyB14 = mean(BAP14);
BAP15 = B(:,15); MoyB15 = mean(BAP15);
BAP16 = B(:,16); MoyB16 = mean(BAP16);
BAP17 = B(:,17); MoyB17 = mean(BAP17);
BAP18 = B(:,18); MoyB18 = mean(BAP18);
BAP19 = B(:,19); MoyB19 = mean(BAP19);
BAP20 = B(:,20); MoyB20 = mean(BAP20);
BAP21 = B(:,21); MoyB21 = mean(BAP21);
BAP22 = B(:,22); MoyB22 = mean(BAP22);
BAP23 = B(:,23); MoyB23 = mean(BAP23);
BAP24 = B(:,24); MoyB24 = mean(BAP24);
BAP25 = B(:,25); MoyB25 = mean(BAP25);
BAP26 = B(:,26); MoyB26 = mean(BAP26);
```

```
BAP27 = B(:,27); MoyB27 = mean(BAP27);  
BAP28 = B(:,28); MoyB28 = mean(BAP28);  
BAP29 = B(:,29); MoyB29 = mean(BAP29);  
BAP30 = B(:,30); MoyB30 = mean(BAP30);  
BAP31 = B(:,31); MoyB31 = mean(BAP31);  
BAP32 = B(:,32); MoyB32 = mean(BAP32);  
BAP33 = B(:,33); MoyB33 = mean(BAP33);  
BAP34 = B(:,34); MoyB34 = mean(BAP34);  
BAP35 = B(:,35); MoyB35 = mean(BAP35);  
BAP36 = B(:,36); MoyB36 = mean(BAP36);  
BAP37 = B(:,37); MoyB37 = mean(BAP37);  
BAP38 = B(:,38); MoyB38 = mean(BAP38);  
BAP39 = B(:,39); MoyB39 = mean(BAP39);  
BAP40 = B(:,40); MoyB40 = mean(BAP40);  
BAP41 = B(:,41); MoyB41 = mean(BAP41);  
BAP42 = B(:,42); MoyB42 = mean(BAP42);  
BAP43 = B(:,43); MoyB43 = mean(BAP43);  
BAP44 = B(:,44); MoyB44 = mean(BAP44);  
BAP45 = B(:,45); MoyB45 = mean(BAP45);  
BAP46 = B(:,46); MoyB46 = mean(BAP46);  
BAP47 = B(:,47); MoyB47 = mean(BAP47);  
BAP48 = B(:,48); MoyB48 = mean(BAP48);  
MoyB=[MoyB1 MoyB2 MoyB3 MoyB4 MoyB5 MoyB6 MoyB7 MoyB8 MoyB9 MoyB10  
MoyB11 MoyB12 MoyB13 MoyB14 MoyB15 MoyB16 MoyB17 MoyB18 MoyB19  
MoyB20 MoyB22 MoyB23 MoyB24 MoyB25 MoyB26 MoyB27 MoyB29 MoyB30  
MoyB31 MoyB32 MoyB33 MoyB34 MoyB36 MoyB37 MoyB38 MoyB40 MoyB41  
MoyB42 MoyB43 MoyB44 MoyB45 MoyB46 MoyB48];  
%-----
```

```
% Création et simulation de l'ANN

ptr=p(:,iitr);ttr=t(:,iitr);      % Données d'apprentissage
pts=MoyB';                         % Données d'estimation

spread=4.5;
net=newgrnn(ptr,ttr,spread);       % Création d'un réseau grnn
atr=sim(net,ptr);
ats=sim(net,pts);
%-----
atr1=[];
l=length(iitr);
for i=1:l,
    if round(atr(3:4,i)) == [1; 1]
        atr1 = [atr1 3.9]; % défini le 6ème étage
    elseif round(atr(3:4,i)) == [-1 ;-1]
        atr1 = [atr1 -7.8]; % défini le 3ème étage
    elseif round(atr(3:4,i)) == [-1 ;1]
        atr1 = [atr1 t1(1)]; % défini le 5ème étage
    else
        atr1 = [atr1 4.16];
    end
end

atr=[atr(1:2,:); atr1];

ats1=[];
l=length(ats);
j=1;
```

```
if round(ats(3:4,j)) == [1; 1]
    ats1 = [ats1 3.9]; % défini le 5ème étage
elseif round(ats(3:4,j)) == [-1 ;-1]
    ats1 = [ats1 -7.8]; % défini le 2ème étage
elseif round(ats(3:4,j)) == [-1 ;1]
    ats1 = [ats1 t1(1)]; % défini le 5ème étage
else
    ats1 = [ats1 4.16];
end

Coord_reel=tts
Coord_estime =[ats(1:2,:); ats1]
% Calcul de l'erreur Euclidienne
Erreur_d=sqrt((ats(1)-tts(1))^2 + (ats(2)-tts(2))^2 )
%-----%
```

## ANNEXE IV

FICHIER DE TRAITEMENT DES DONNÉES *TRAITEMENT.CPP*

```
-----  
//-----  
//NOM DU FICHIER: Traitement.cpp  
//DESCRIPTION : Programme permettant d'obtenir à partir du fichier  
//original des mesures de puissances faites par  
//Netstumbler, un fichier contenant les coordonnées du  
//point de mesure et les puissances moyennes reçues  
//des différents points d'accès du réseau Hermès du  
//Pavillon Lassonde de l'Ecole Polytechnique de Montréal et un  
//fichier contenant le nom du point de Mesure, les coordonnées  
//ainsi que les puissances recue sur 60 échantillons consécutifs  
//AUTEURS : Rachid Hachad et Lamia Hamza  
//DATE DE CRÉATION : 06 Juillet 2006  
//DATE DE MODIFICATION: 28 Avril 2008  
//-----  
#include <fstream>  
#include <string>  
#include <iostream>  
#include <iomanip>  
using namespace std;  
void main (void)  
{  
    ifstream Fic_in;  
    ofstream Fic_out, Fic_outt;
```

```
string numero, Fichier, PM;
char ligne[256];
int indice[60];
float signal[60], moy[60], tableau[60][48];
int n=100;
//Initiation des tableaux des moyennes et niveau du signal a zero

for (int l=0;l<61;l++)
indice[l]=1;
for ( l=0;l<61;l++)
moy[l]=0;
for ( l=0;l<61;l++)
signal[l]=0;
for ( l=0;l<61;l++) // Initialiser la matrice a -94dBm
for( int M=0;M<49;M++)
tableau[l][M]=-94;

/*-----*/
/* Saisie des coordonnées du point de mesure */
/*-----*/
float X;
float Y;
float Z;
float stan, K=0;

cout<<"entrez la coordonne en x"<<endl;
cin>>X;
cout<<"entrez la coordonne en y"<<endl;
```

```
cin>>Y;  
cout<<"entrez la coordonne en z"<<endl;  
cin>>Z;  
cout<<"entrez le numre de point de mesure"<<endl;  
cin>>PM;  
  
/*-----*/  
/* Chargement et test de l'ouverture du fichier de mesures */  
/*-----*/  
  
cout<<"entrez le nom du fichier de mesure"<<endl;  
cin>>Fichier;  
  
/*-----*/  
/* Ici le fichier resultat_par_AP.txt est creer. Ce dernier */  
/* contient le numero de point de mesure, les coordonnees 3D */  
/* et les puissances par AP et par temps en dBm */  
/*-----*/  
Fic_outt.open ("test3r.txt",ios::app); // Refxx.txt est a t=t0  
// RefxxAP.txt a t=t1 semi statique  
// Refxxdy.txt est a t=t2 dynamique  
// Refxxst.txt est a t=3 statique  
  
if (Fic_outt.fail())  
cout<<"echec d'ouverture de resultats_par_AP";  
else  
cout<<" ouverture reussie de resultats_par_AP"<<endl;
```

```
//Fic_outt<<"Le nom du fichier est: "<<Fichier<<endl;
//Fic_outt<<"Le numero du point de mesure est: "<<PM<<endl;
//Fic_outt<<"Les coordonnees 3D sont: "<<endl;
//Fic_outt<<X<<setw(8)<<Y<<setw(8)<<Z<<endl;

Fic_in.open(Fichier.c_str());

if (Fic_in.fail())
{
    cout <<"echec d'ouverture du fichier demande";
}
else
{
    cout<<" ouverture réussie du fichier demande"<<endl;

for(int k=0;k<4;k++)
    Fic_in.getline(ligne,256);

/*
 * Extraction des données relatives au réseau Hermès */
/*
do
{
    // Extracte les données non significatives jusqu'à
        // l'arrivée a hermes
    for (int i=1;i<7;i++)
        Fic_in>>numero; // ici numero= hermes ou autre
```

```
if (numero=="hermes") // choix du reseau hermes
{
    for (int i=1;i<5;i++)
        Fic_in>>numero; //ici numero= adresse mac ci bas
        //Fic_outt<<numero<<endl;
        // choix de l'adresse Mac du point d'accès
        if (numero=="00:12:7f:99:77:d1")
            n=0;
        else if (numero=="00:12:7f:99:78:31")
            n=1;
        else if (numero=="00:12:7f:99:78:51")
            n=2;
        else if (numero=="00:12:7f:ca:a7:01")
            n=3;
        else if (numero=="00:12:7f:99:84:81")
            n=4;
        else if (numero=="00:12:7f:99:84:a1")
            n=5;
        else if (numero=="00:12:7f:ca:9d:21")
            n=6;
        else if (numero=="00:12:7f:ca:96:31")
            n=7;
        else if (numero=="00:12:7f:99:83:71")
            n=8;
        else if (numero=="00:12:7f:99:77:e1")
            n=9;
        else if (numero=="00:11:93:0d:67:01")
```

```
n=10;  
else if (numero=="00:12:7f:99:86:e1")  
n=11;  
else if (numero=="00:12:7f:99:87:f1")  
n=12;  
else if (numero=="00:12:7f:ca:99:41")  
n=13;  
else if (numero=="00:11:5c:e5:46:e1")  
n=14;  
else if (numero=="00:12:7f:ca:ab:e1")  
n=15;  
else if (numero=="00:12:7f:99:84:e1")  
n=16;  
else if (numero=="00:12:7f:ca:ab:02")  
n=17;  
else if (numero=="00:12:7f:ca:9e:01")  
n=18;  
else if (numero=="00:11:93:0d:66:b1")  
n=19;  
else if (numero=="00:0e:d7:af:24:00")  
n=20;  
else if (numero=="00:11:93:0d:66:d1")  
n=21;  
else if (numero=="00:11:93:0d:69:41")  
n=22;  
else if (numero=="00:12:7f:ca:9b:e1")  
n=23;  
else if (numero=="00:12:7f:ca:9c:c1")
```

```
n=24;  
else if (numero=="00:12:7f:99:75:41")  
n=25;  
else if (numero=="00:11:93:0d:66:f1")  
n=26;  
else if (numero=="00:12:7f:98:87:a1")  
n=27;  
else if (numero=="00:12:7f:ca:a3:e1")  
n=28;  
else if (numero=="00:12:7f:ca:9c:21")  
n=29;  
else if (numero=="00:12:7f:99:87:51")  
n=30;  
else if (numero=="00:12:7f:ca:9a:41")  
n=31;  
else if (numero=="00:12:7f:ca:9d:11")  
n=32;  
else if (numero=="00:12:7f:ca:a7:71")  
n=33;  
else if (numero=="00:12:7f:99:84:11")  
n=34;  
else if (numero=="00:12:7f:ca:9d:91")  
n=35;  
else if (numero=="00:12:7f:ca:ab:71")  
n=36;  
else if (numero=="00:12:7f:ca:9b:01")  
n=37;  
else if (numero=="00:12:7f:ca:aa:f0")
```

```
n=38;  
else if (numero=="00:12:7f:ca:aa:f1")  
n=39;  
else if (numero=="00:12:7f:ca:aa:31")  
n=40;  
else if (numero=="00:12:7f:99:ab:71")  
n=41;  
else if (numero=="00:11:92:e9:d7:31")  
n=42;  
else if (numero=="00:12:7f:ca:a8:51")  
n=43;  
else if (numero=="00:12:7f:99:84:f1")  
n=44;  
else if (numero=="00:12:7f:ca:a7:91")  
n=45;  
else if (numero=="00:12:7f:ca:98:f1")  
n=46;  
else if (numero=="00:12:7f:99:83:91")  
n=47;  
  
for ( i=1;i<6;i++)  
Fic_in>>numero; // ici numero=SNR=56  
  
Fic_in>>signal[n]; // ici signal[n]= Sig=105  
  
Fic_in.getline(ligne,256);
```

```
if(indice[n]<61) //ca commence avec indice[n]=1
{
/**/ if(signal[n]==-32618) //i.e. on ne capte plus le signal
{
stan=-94;
/**/ tableau[indice[n]][n]=stan;
K++;
//si on ne capte plus apres les 10 premieres mesures
// on met Sig=0 car les 10 prem. mesures sont negligees
// plus tard.
if(indice[n]>10)
signal[n]=0;
indice[n]--;
}
else if(indice[n]>10)//si on capte, on additionne
{
moy[n] = moy[n] + signal[n];
/**/ tableau[indice[n]][n]=signal[n]-149;
K++;
}
else
{
/**/ tableau[indice[n]][n]=signal[n]-149;
K++;
}

indice[n]++;
```

```
}

}

else

Fic_in.getline(ligne,256); // aller a la ligne suivante

}

while(!Fic_in.eof()); // Fermuture du fichier de mesures

}

/*-----*/
/* calcul de la valeur moyenne de la puissance reçue */
/*-----*/
for (l=0;l<52;l++)

{
    if(indice[l] <51 && indice[l]>10)

moy[l]=((moy[l]-((indice[l]-11)*149)+(51-indice[l])* -94))/60;

    else if(indice[l]<10)

moy[l]=-94;

    else

moy[l]=(moy[l]/(indice[l]-11))- 149;

    if(moy[l]==-149)

moy[l]=-94;

}

/*-----*/
/* ouverture et test du fichier qui va recevoir les résultats */
/*-----*/
```

```
Fic_out.open ("resultat.txt",ios::app);

if (Fic_out.fail())
cout<<"echec d'ouverture du fichier Resultat.txt";
else
cout<<" ouverture reussie du fichier Resultat.txt"<<endl;
/*-----*/
/* Ecriture dans le fichier resultat */
/*-----*/
Fic_out<<setw(6)<<X<<setw(8)<<Y<<setw(8)<<Z;

for (l=0;l<52;l++)
Fic_out<<setw(10)<< setiosflags(ios::fixed)<< setprecision(2)<<moy[1];

/**/ // Fic_outt<<endl<<endl<<K<<endl; // Ecrire le signal recu sur
for ( l=0;l<60;l++) // resultat_par_AP.txt
{
for( int M=0;M<48;M++)
    Fic_outt<<setw(5)<<tableau[1][M];
Fic_outt<<endl;
}

Fic_out<<endl;
Fic_in.close ();
Fic_out.close();
Fic_outt.close();
}
```