



Ecole Doctorale
Electronique, Electrotechnique et Automatique



Centre d'Innovation
en Télécommunications & Intégration de services

MEMOIRE DE DEA

DEA Images et Systèmes

Présenté par
Thanh-Hà LÊ

LOCALISATION DES STATIONS MOBILES

DANS LES RESEAUX SANS FIL

Soutenu à l'INSA de Lyon, le 9 Septembre 2004
Devant le jury composé de :

Rapporteur : **M. Hervé Saint Jalmes** Professeur à l'Université Claude Bernard - Lyon 1
Superviseurs : **M. Jean Marie GORCE** Maître de conférence à l'INSA de Lyon
Mlle. Katia RUNSER Doctorante à l'INSA de Lyon

REMERCIEMENTS

Je remercie tout d'abord M. Stéphane UBEDA et tous les membres du laboratoire CITI - INSA de Lyon qui m'ont donné la chance de pouvoir effectuer mon stage de DEA dans de bonnes conditions.

Je tiens à remercier tout particulièrement Monsieur Jean Marie GORCE, mon encadrant de stage, et Mademoiselle Katia RUNSER, ma tutrice, pour leurs précieux conseils et indications, ainsi que pour la disponibilité dont ils ont fait preuve à mon égard tout au long de mon stage.

Je remercie également sincèrement toutes les personnes qui m'ont aidé dans l'avancement de mon stage et dans mes recherches. Elles ont contribué à rendre ce stage agréable et riche en enseignements.

SOMMAIRE

1	INTRODUCTION	4
2	ETUDE BIBLIOGRAPHIQUE	5
2.1	Introduction sur les techniques de localisation.	5
2.2	Principe des techniques de localisation par RF	6
2.3	Méthodes déterministes	7
2.4	Méthodes probabilistes	9
3	MISE EN OEUVRE.....	11
3.1	Réglage des paramètres du moteur de calcul de couverture WILDE.	11
3.2	Structure de la base de données de Simulation et de Mesures	12
4	PRESENTATION ET INTERPRETATION DES RESULTATS.....	14
4.1	Validation du système implanté.....	14
4.2	Localisation avec le moteur WILDE.	17
5	CONCLUSIONS ET PERSPECTIVES	26
5.1	Conclusions.....	26
5.2	Perspectives	27
	BIBLIOGRAPHIE	29
	ANNEXES	30

1 INTRODUCTION

L'utilisation de plus en plus fréquente des dispositifs de calcul portatifs et des réseaux locaux sans fil à haut débit permet utilisateurs de rester connectés tout en se déplaçant à l'intérieur des bâtiments. Ce paradigme a stimulé beaucoup d'intérêt pour le développement d'applications et de services prenant en compte la position physique des utilisateurs mobiles.

Pour cela, il a fallu mettre en place des méthodes de localisation pour permettre à un utilisateur mobile d'interagir avec son environnement physique. Ces interactions regroupent par exemple l'impression d'un document sur l'imprimante la plus proche ou encore le guidage d'un utilisateur à l'intérieur d'un bâtiment. Ces méthodes de localisation peuvent également jouer un rôle prépondérant dans des situations d'urgence où il est nécessaire de localiser rapidement des éléments actifs du réseau.

La localisation d'une station mobile ne peut se faire qu'avec la connaissance a priori de l'environnement radio dans lequel l'opération est effectuée. Cette connaissance passe par l'acquisition d'une carte de couverture radio. Cette acquisition peut se faire soit par mesure radio, soit par calcul de couverture par simulation. Si le modèle de couverture radio est fiable, l'utilisation de celui-ci permet de gagner du temps et rend les systèmes de localisation aisément configurables et portables, contrairement à l'acquisition par mesure.

Le travail de DEA présenté dans ce mémoire a pour objectif final d'estimer l'apport de l'utilisation du simulateur de couverture radio WILDE dans les processus standard de localisation de stations mobiles pour des environnements Indoor (environnements clos).

Le simulateur de couverture WILDE est basé sur le modèle de prédiction de couverture **Multi Resolution Fourier Domain ParFlow** (MR-FDPF) [18][19]. Il a été développé au laboratoire CITI et permet de prédire le niveau de puissance reçu en tout point du plan 2D d'un bâtiment.

Pour valider l'utilité de WILDE pour la localisation indoor, il a fallu concevoir un logiciel qui permette d'implanter les méthodes de localisation. Ce logiciel a permis de comparer les performances des principaux algorithmes de localisation en simulation et en environnement réel.

L'environnement de simulation utilise des cartes de couvertures issues du moteur de calcul de couverture WILDE pour localiser des points simulés à l'aide du même moteur. L'environnement réel utilise un jeu de mesures réelles pour créer la base de données et localise des points de mesure réels.

La technique de localisation RADAR [7], [8], [9] a été retenue pour statuer sur la fiabilité de l'utilisation de cartes de couverture radio créée par Wilde. Pour se faire, les cartes de couverture sont exploitées par RADAR pour localiser des mesures réelles effectuées dans le laboratoire.

Le travail que je vous présente dans ce rapport est effectué au laboratoire CITI - INSA de Lyon, dans le cadre de mon stage de DEA Images et Système (Ecole Doctorale Electronique, Electrotechnique et Automatique). Il a également donné lieu à un travail de projet de fin d'études (PFE) ayant pour objectif la mise en place d'un outil de comparaison des performances d'un système de localisation. Les résultats de ce travail son consultables dans le rapport de PFE [17].

Dans ce mémoire, je présente d'abord mon étude bibliographique sur les techniques de localisation utilisées dans les environnements Indoor (partie 2). La mise en œuvre du projet sera présentée dans la partie 3 et les résultats obtenus seront résumés dans la partie 4. Le mémoire se terminera par des conclusions et des perspectives visant à améliorer les performances du système de localisation utilisant Wilde (partie 5).

2 ETUDE BIBLIOGRAPHIQUE

2.1 Introduction sur les techniques de localisation.

Plusieurs systèmes de détection de position ont été proposés et mis en application. Pour des applications dans des environnements de type « outdoor » (hors des bâtiments), le système GPS, basé sur une constellation de satellites, est utilisé en règle générale. Cependant, les phénomènes de réflexion et de trajets multiples limitent l'utilité du GPS dans les environnements dits « indoor » (à l'intérieur des bâtiments). C'est pourquoi des systèmes de localisation indoor ont été développés. Ils peuvent être classifiés dans trois grandes catégories en fonction du support de transmission utilisé : Infrarouge (IR), Ultrason (US) et Radio-Fréquence (RF).

Le système de localisation « Active Badge » [1] basé sur la **technologie IR** est un des premiers systèmes de localisation proposé pour des environnements indoor. Ce système fournit à chaque personne un badge qui émet périodiquement une identification à l'aide d'un rayon IR. Ce rayon IR est reçu par un ou plusieurs récepteurs dispersés dans un bâtiment. La position du badge est alors déterminée à l'aide de la position du récepteur le plus proche. Tous les systèmes basés sur les rayons IR imposent l'existence d'un chemin direct entre l'émetteur et le récepteur, ce qui n'est pas possible dans tous les cas de figure. De plus, ceci impose le port des badges sur l'extérieur de l'habillement (attaché sur la chemise ou la blouse).

Il existe aussi un certain nombre de systèmes de localisation **ultrasonores** présentés dans la littérature : Constellation system [2], Active Bat [3] et Cricket [4]. Ces systèmes de localisation permettent d'estimer des positions avec précision. Cependant, ils sont sensibles au bruit et nécessitent également la présence d'un chemin direct entre l'émetteur et le récepteur.

Aujourd'hui beaucoup de bâtiments publics et privés sont équipés de la **technologie sans fil IEEE 802.11b**, une technologie RF populaire et peu coûteuse. La plupart des dispositifs 802.11b mesurent la **puissance du signal** des paquets reçus de façon native. Un système de localisation exploitant la seule information du niveau de puissance reçu permet donc d'implanter aisément un service de localisation. C'est pourquoi les systèmes de localisation les plus performants sont basés sur l'analyse du niveau de puissance du signal reçu. Ces systèmes sont limités par la nature complexe du signal radio (les effets des trajets multiples, le bruit et les interférences) mais ont l'avantage de ne pas nécessiter de vue directe entre l'émetteur et le récepteur.

Le principe de fonctionnement des techniques de localisation basées sur la technologie RF sera présenté en détail le paragraphe 2.2. Nous allons également y présenter les deux catégories de technologies existantes : **les techniques déterministes** (paragraphe 2.3) et **les techniques probabilistes** (paragraphe 2.4).

2.2 Principe des techniques de localisation par RF

2.2.1 Fonctionnement Offline/Online

Tous les systèmes de localisation RF présentent deux étapes distinctes :

- la construction de la base de connaissance qui stocke la répartition spatiale des niveaux de puissance enregistrés pour tous les point d'accès présents sur le plan (phase offline ou phase préliminaire).
- la détermination de la position d'un utilisateur mobile en fonction des signaux mesurés (phase online ou phase de localisation).

➤ Phase « offline » – Construction des cartes de couverture radio

Pour estimer la position d'un utilisateur, le système de localisation mesure la puissance des signaux reçus par le mobile en provenance les points d'accès visibles. En espace libre, cette valeur est une fonction logarithmique de la distance entre le point de mesure et le point d'accès. Or, dans des environnements indoor, le canal sans fil est perturbé par la présence de réflexions, de diffractions et de trajets multiples. La puissance du signal n'est plus prédictible aisément. Pour obtenir des cartes de couverture radio réalistes, les systèmes RF construisent une **carte de couverture radio** en mesurant expérimentalement la puissance du signal des points d'accès en un nombre fini de points [6] [7] [10] [11] [12] [15]. La construction de cette base est dénommée « **phase offline** ». La carte de couverture obtenue est stockée dans une base de données qui répertorie les niveaux de signaux reçus de chaque point d'accès en un point du plan. La construction de cartes de couvertures radio peut également se faire à l'aide d'un modèle de prédiction de couverture radio [5] [12] [7],

➤ Phase « online » – Localisation d'un utilisateur mobile

Dans la phase « online », le système mesure la puissance des signaux reçus par les points d'accès visibles. A l'aide de la base de données créée pendant la phase offline, il cherche à localiser le point de la base pour lequel on observe un ensemble de signaux similaire aux signaux mesurés.

Les différentes façons de construire les cartes de couverture radio dans la phase offline et de définir la distance dans l'espace signal entre deux points P1 et P2 permettent de classifier les différentes méthodes de localisations.

2.2.2 Approche « Client Centré » et approche « Serveur Centré »

Concernant l'infrastructure des réseaux sans fils, il existe deux approches fondamentales: l'approche « **client centré** » et l'approche « **serveur centré** ». Pour les modèles de type « Client centré », chaque client reçoit la puissance des signaux des points d'accès et il utilise ces informations pour en déduire sa position actuelle. Dans l'approche « serveur centré », la localisation (phase online) se fait au niveau d'un serveur central. Il détermine quels clients sont en contact avec quels points d'accès et il utilise ces données pour localiser les clients.

Si on compare la performance de ces deux approches, on trouve que chaque approche a son avantage et son inconvénient.

- Le modèle « Client centré » permet d'augmenter la taille du système de localisation car toutes les opérations sont réalisées sur les stations mobiles. Cependant, un système de localisation basée sur un tel modèle doit toujours faire face à la difficulté de la taille mémoire des stations mobiles. Comme la procédure de localisation est effectuée sur la machine cliente, on a besoin d'une taille mémoire relativement grande pour que le temps de localisation ne soit pas trop long. En effet, plus on stocke de données propres à la carte de couverture, plus la localisation est rapide. Pour des stations mobiles réduites (assistants personnels, matériel embarqué,...), la mémoire disponible n'est pas toujours suffisante.
- Le modèle « serveur centré » nécessite un serveur robuste avec une mémoire suffisante. Comme pour la plupart des systèmes centralisés, il faut dimensionner correctement le serveur pour qu'il puisse supporter une augmentation du nombre de requêtes. Son principal inconvénient réside dans la position centrale du serveur. Si une panne survient sur le serveur, le service de localisation n'est plus disponible, ce qui n'est pas le cas pour une architecture répartie.

Dans le tableau ci-après, je regroupe les techniques que j'ai étudiées selon 2 axes :

- Type de techniques de recherche : déterministes/probabilistes,
- Approches Client / Serveur centré.

	Techniques déterministes	Techniques probabilistes
Client centré	- Système CMU-TMI (2001-2002) [5] - Transparent Location Fingerprinting (2002) [6]	- Système Horus (2002-2003) [15] - Robotics-based location sensing (9/2002) [11]
Serveur centré	- Système Radar (2000) [9]	- Système Nibble (2001) [10] - A probabilistic approach (2002) [12]

2.3 Méthodes déterministes

2.3.1 Caractéristiques générales

Les techniques déterministes construisent des cartes de couverture en ne stockant que l'information de **puissance moyenne** des points d'accès pour chaque point de la base. Ainsi, dans la phase de localisation (phase online), on cherchera à localiser le point le plus proche du point mesure (distance euclidienne dans l'espace signal)

Trois systèmes basés sur la méthode déterministe (Radar [9], CMU-TMI [5] et Transparent Location Fingerprinting [9]) utilisent la technique de triangulation pour estimer la position la plus favorable. Parmi ces trois systèmes, le système **Radar** est considéré comme le système de référence pour les techniques déterministes.

Le système **CMU-TMI** [5] (Carnegie Mellon University's Triangulation, Mapping et Interpolation algorithm) est le plus complexe à mettre en œuvre car il existe deux espaces : « *l'espace des signaux* » (signal space) dans lequel la position de l'utilisateur dépend directement des signaux qu'il reçoit et « *l'espace physique* » (physical space) qui représente la position réelle de l'utilisateur. Il faut donc étudier la corrélation entre ces deux espaces. Cette corrélation permet de transformer un point dans l'espace des signaux en un point dans l'espace physique, donc la position réelle du point à localiser.

Le système « **Transparent Location Fingerprinting** » [6], quand à lui, est basé sur la technique « k-nearest-neighbors ». La position estimée est calculée à l'aide de la position des k voisins les plus proches dans l'espace des signaux.

2.3.2 Système Radar

A la suite de mes recherches bibliographiques, j'ai choisi le système Radar [7] [8] [9] comme le représentant des méthodes déterministes. En effet, le système Radar est à la fois le système le plus utilisé et le plus développé de ces méthodes. C'est d'autre part celui qui est cité en référence dans l'ensemble des papiers étudiant ce problème. C'est pourquoi nous avons choisi d'implanter l'algorithme proposé par Radar. Je donne ici ses principales caractéristiques.

Le système Radar (cf. Annexe 2) a été proposé par P.BAHL et al. (Microsoft Research) en 2000. Le processus de localisation comprend phases :

Phase offline

Pendant la phase offline, le système collecte des données pour déduire, dans la phase online, la position de l'utilisateur. La base de données du système Radar est construite de deux façons différentes :

- **Méthode empirique** : Utilisation de mesures expérimentales. En chaque point de mesure, on stocke la puissance moyenne des échantillons de puissance mesurés.
- **Méthode théorique** : Utilisation d'un modèle de propagation de radio.

$$P(d)[dBm] = P(d_0)[dBm] - 10n \log(d/d_0) - nW * WAF \text{ si } nW < C$$

$$C * WAF \text{ si } nW \geq C$$

n : la vitesse à laquelle le "path loss" varie en fonction de la distance

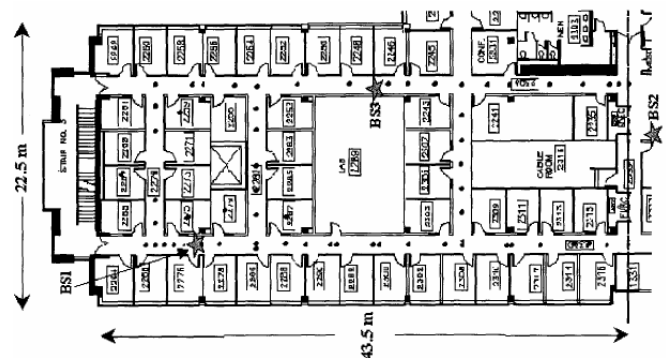
$P(d_0)$: puissance à la distance de référence d_0

d : distance entre l'émetteur et le récepteur

C : nombre maximal d'obstacles jusqu'à lequel le facteur d'atténuation fait une différence

nW : nombre d'obstacles entre le transmetteur et le récepteur

WAF : facteur d'atténuation d'un mur (dépend de disposition de l'étage et le matériau)



Phase online

Soit une position P à laquelle on reçoit la puissance des signaux p_1, p_2, p_3 correspondant aux trois points d'accès AP_1, AP_2, AP_3 placés dans l'environnement de test. Pour ne pas trouver trop facilement le point de test, celui-ci n'est plus pris en compte dans la base de données. Le système Radar cherche dans la phase online le point P' (toujours différent de P) pour que le vecteur de puissance de signal $p' = (p'_1, p'_2, p'_3)$ soit le plus proche du vecteur $p = (p_1, p_2, p_3)$. Il revient à chercher la position P' qui minimise la distance euclidienne dans l'espace signal :

$$d(P, P') = \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - p'_i)^2}$$

(avec N le nombre de points d'accès déployés. Dans l'article, N a été fixé à 3.)

Le banc d'essai expérimental du système (figure 1) est situé au deuxième étage sur trois d'un bâtiment de dimension de 43,5 x 22,5 mètres. Trois stations de base y sont déployées. Dans cet environnement, la moitié des points est localisé avec une précision (erreur médiane) de **2.9 mètres** avec la méthode empirique et **4.3 mètres** avec le modèle de propagation.

Deux autres techniques déterministes sont également présentées afin d'améliorer la performance du système Radar : la technique à voisins multiples dans l'espace signal, **k-NNSS**, et la technique « **Continuous User Tracking** » qui utilise l'information de la position précédente pour estimer la position d'un utilisateur mobile. Dans le cadre de mon stage, je n'ai testé que la technique NNSS.

2.4 Méthodes probabilistes

2.4.1 Caractéristiques générales

Contrairement aux techniques déterministes, les techniques probabilistes utilisent toute l'information obtenue lors d'une mesure en un point donné en stockant l'**histogramme complet des niveaux de puissance** observés au cours de la mesure pour chaque point d'accès. Ainsi, dans la phase de localisation (phase online), on cherche à trouver le point le plus probable à l'aide des histogrammes connaissant les signaux mesurés.

Tous les systèmes basés sur la technique probabiliste calculent la probabilité conjointe p d'avoir les puissances p_1, p_2, \dots, p_N (avec N le nombre de points d'accès) à la location L .

Pour le système « **Robotics – Based Location Sensing** » [11], un groupe de chercheurs de l'Université de Rice (Etats-Unis) a utilisé le modèle HMM (Hidden Markov Modem – HMM) comme un opérateur de mouvement du processus de localisation.

Dans l'article « **A probabilist Approach to WLAN User Location** » [12], Teemu Roos a présenté deux méthodes pour calculer la probabilité en chaque point de test : la méthode « Kernel » dans laquelle on utilise une distribution gaussienne et la méthode « Histogramme » qui prend en compte la distribution locale. L'article a montré que la méthode « Histogramme » donne les meilleurs résultats.

Le système **Nibble** [10], proposé par Paul Castro et al. (UCLA, Los Angeles), utilise le même principe que le système précédent. Cependant, le système Nibble ne cherche qu'à identifier la salle dans laquelle un utilisateur se trouve.

Parmi des systèmes de localisation basés sur la technique probabiliste, le système Horus est apparu comme le système le plus complet avec plusieurs techniques d'optimisation : la technique « Joint Clustering » [14], la technique « Triangular Incremental » [14], la technique « Small Scale Compensation » [13] et la nouvelle technique « Handling Samples Correlation » [15] proposée en mars 2004.

2.4.2 Système Horus

Pour les méthodes probabilistes, j'ai retenu le système Horus [14] [15] car il est connu et les articles relatifs au système Horus sont assez récents (d'octobre 2002 à mars 2004).

Le système Horus (cf. Annexe 3) a été implanté et testé en 2002 - 2003 par un group de chercheurs de l'Université de Maryland, Etats-Unis. La base de données du système est créée dans la phase offline et la détermination de position est réalisée dans la phase online.

Phase offline : Stockage de l'histogramme

Pour construire la base de données du système Horus, 300 échantillons ont été collectés pour chaque position (un échantillon par seconde). Ces informations ont été utilisées pour estimer la distribution de chaque point d'accès pour une position donnée. La figure 2 donne un exemple de l'histogramme de niveau de signal d'un point d'accès pour une position donnée.

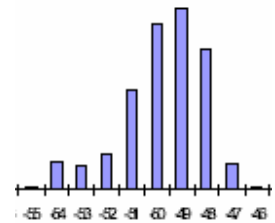


Figure 2

Phase online : Technique Joint Clustering

La phase online du système Horus est réalisée à l'aide de la technique « Joint Clustering » (JC) basée sur l'utilisation de probabilités conjointes et du regroupement des points d'accès visibles. Horus présente aussi la technique IT (Incremental Triangulation) permettant de diminuer le temps de calcul de la phase online. Nous ne traitons pas cette technique dans ce projet.

Le système Horus a été testé pour un étage de dimension 67 x 25,5 mètres (figure 3). La zone est couverte par 12 points d'accès installés aux troisième et quatrième étages d'un bâtiment. La construction de la base de données s'est faite sur une grille de points espacés de 1,7 x 1,7 mètres. Il y a au total 110 positions et chaque position est couverte par 4 points d'accès en moyenne. Dans cet environnement, la moitié des points est localisée avec une précision d'**1mètre** environ (erreur médiane).

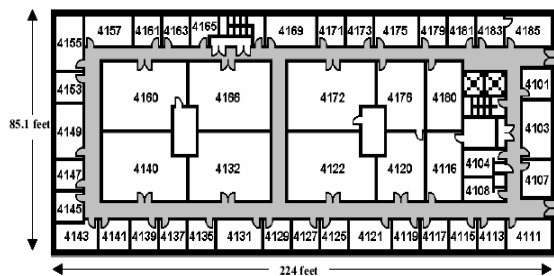


Figure 3: Banc d'essai du système Horus

3 MISE EN OEUVRE

Nous avons choisi d'effectuer la comparaison des systèmes Radar et Horus à l'aide de deux bases de données : une base de données obtenue par la simulation et une base de données créée à l'aide des mesures expérimentales. La partie 3.2 présente la construction des structures de données issues des simulations et des données issues de mesures.

Le paragraphe 3.1 a été introduit pour décrire les paramètres du moteur de couverture WILDE qui doivent être définis pendant la construction de la base de données de simulation.

3.1 Réglage des paramètres du moteur de calcul de couverture WILDE.

Le moteur de couverture radio WILDE permet de calculer des cartes de couvertures en deux dimensions pour un environnement donné. En discrétisant l'environnement de test en blocs, ce moteur de couverture permet de propager une source vers tous les points de l'environnement à une résolution de blocs donnée. La figure 4 représente une carte de couverture calculée à la résolution maximale où un pixel représente 10 cm en réalité (pas de discrétisation $dr = 10\text{cm}$). Cette carte représente le laboratoire CITI.

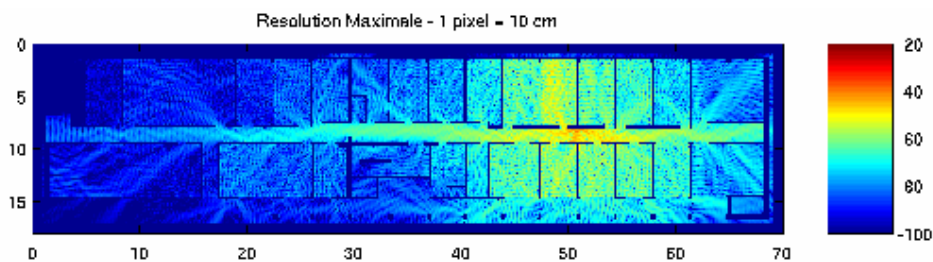


Figure 4: Carte de couverture à la résolution 1x1

Définition de la résolution : La résolution est définie par la taille maximale (largeur * hauteur) des blocs pour lesquels on calcule le niveau moyen de puissance lors de la phase de propagation. Par exemple, à la résolution 4x5, on ne calcule le niveau de puissance que pour les blocs dont la largeur est inférieure ou égale à 4 unités et dont la hauteur est inférieure ou égale à 5 unités. L'unité est définie à l'aide du pas de discrétisation de l'environnement. La figure 5 représente une carte à une résolution de 6 par 6 unités.

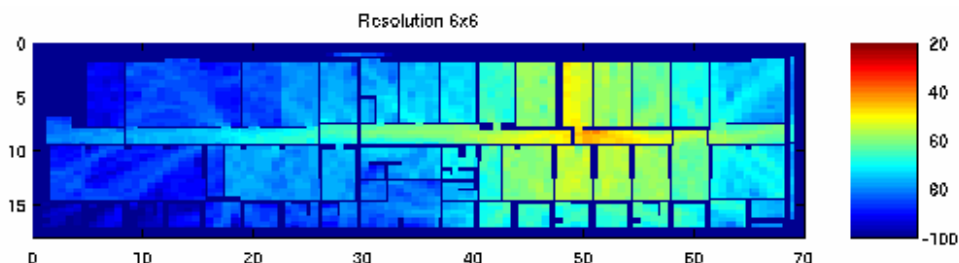


Figure 5 : Carte de couverture à la résolution 6x6

Dans nos calculs, nous avons choisi une résolution de 6x6 unités pour le calcul des puissances moyennes avec un pas de discrétisation de 10 cm. Dans ce cas, nous pouvons simuler un système de localisation qui possède au mieux une précision de 0,6 mètre.

3.2 Structure de la base de données de Simulation et de Mesures

3.2.1 Base de donnée de simulations

La base de données est calculée à partir de cartes de couvertures obtenues après la propagation de N points d'accès placés dans l'environnement du CITI. Dans les tests suivants, nous avons fixé N à 6. Les blocs de taille 6x6 sont appelés « macro bloc », les blocs élémentaires de taille 1x1 sont appelés pixels.

En chaque point de la base de données, il est nécessaire de stocker la puissance moyenne et l'histogramme des signaux provenant des six points d'accès.

Cette structure de base de données permet de simuler à la fois des méthodes déterministes (puissance moyenne du signal) et des méthodes probabilistes (histogramme de distribution du signal). La structure de la base de données est présentée en détail dans la partie 4.4.4.

3.2.2 Base de donnée de mesures

La base de données est obtenue à l'aide de mesures réalisées en environnement réel. 199 points de mesure ont été réalisés au laboratoire CITI et sont représentés sur la figure 6 par des croix noires. Les 6 points d'accès présents lors des mesures sont localisés à l'aide de points rouges.

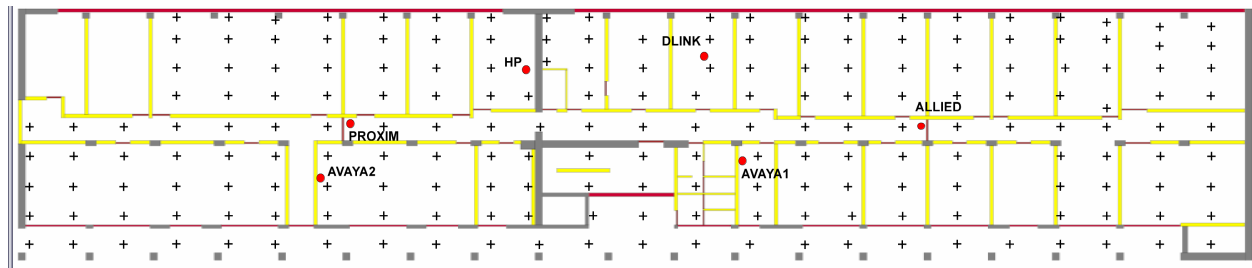


Figure 6: Points de mesures réalisés au laboratoire CITI

La prise de mesures a été effectuée à l'aide de logiciels propres aux cartes d'extension 802.11b utilisées lors des mesures. Nous avons travaillé avec des cartes PCMCIA AVAYA connectées à des PC portables. Le logiciel 'Client Manager' de Lucent a servi de logiciel de mesures.

Pour chaque point de mesure, on a enregistré des mesures de puissance pendant une minute. Pendant toute la durée de la mesure, on réalise des déplacements lents dans toutes les directions sur une zone de 1m² autour du point de test tout en modifiant l'orientation du récepteur. Ceci permet d'obtenir un ensemble représentatif du canal sur une zone d'1 m² environ.

A chaque point de mesure, plusieurs échantillons sont enregistrés. La valeur moyenne des échantillons est stockée dans la base de données. L'histogramme des niveaux de puissance est obtenu à partir de la distribution des échantillons à ce point.

Il est important de noter que l'on ne possède de mesure que pour les points d'accès qui couvrent le point de mesure, c'est-à-dire les points d'accès dont le niveau de puissance est supérieur au seuil de détection des cartes réceptrices. Les points de mesure dans l'environnement du CITI sont répartis tous les 2 mètres environ.

3.2.3 Structure de la base de données

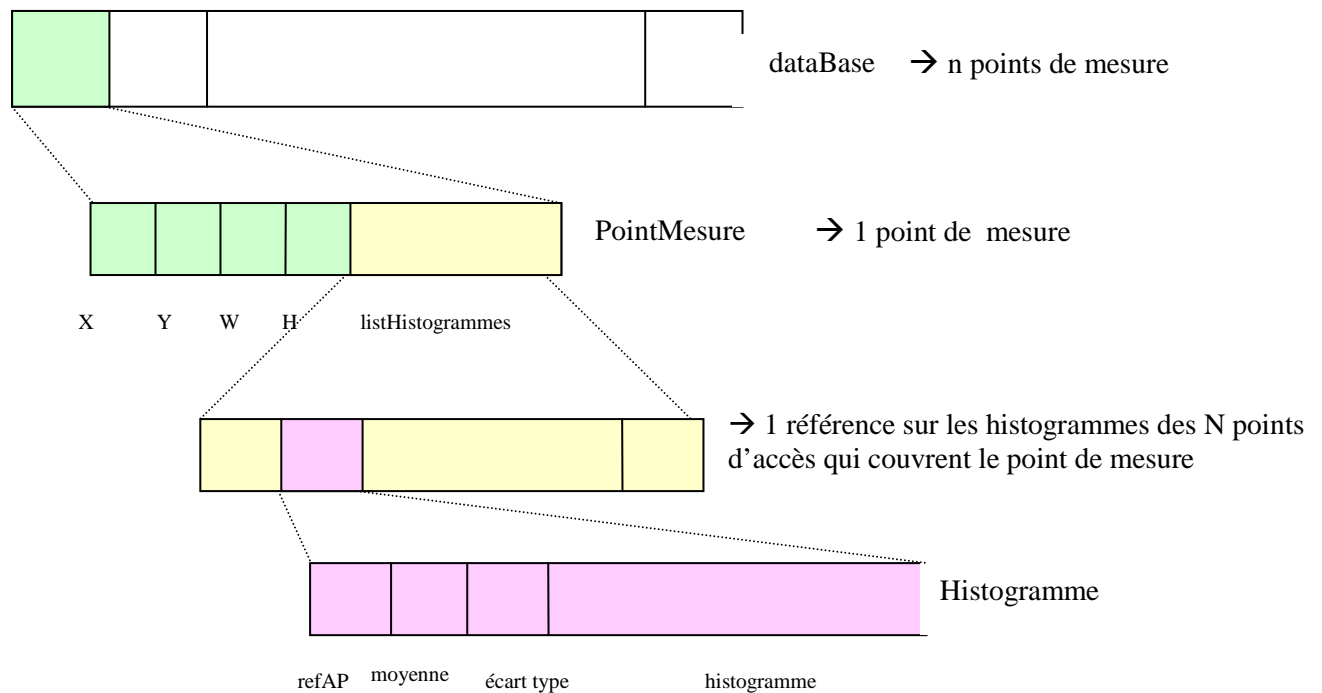


Figure 7: Structure de la base de données

Pour pouvoir exploiter la même base de données pour les méthodes Radar et Horus, il est nécessaire d'avoir une structure commune de données. La description de la structure de données utilisée est donnée dans la figure 7.

L'objet global, la base de données, est un vecteur de n points de mesure (i.e. des objets de classe PointMesure). Une classe PointMesure possède les attributs suivants :

- Position en X et en Y du point de mesure
- Largeur W et hauteur H du bloc de test (seulement pour les simulations)
- Vecteur de N histogrammes (avec N le nombre de points d'accès)

Le vecteur listHistogrammes de la classe PointMesure comprend N histogrammes correspondant à la distribution du signal des N points d'accès. Chaque objet de type Histogramme référence la source AP avec laquelle il a été calculé. Les principaux attributs de cette classe sont donnés par :

- Une référence sur la source AP,
- La puissance moyenne du signal,
- L'écart type (pour la distribution gaussienne),
- L'histogramme $P(AP_i=si)$.

4 PRESENTATION ET INTERPRETATION DES RESULTATS

Nous avons choisi d'implanter les systèmes Radar-NNSS et Horus. Dans un premier temps (cf. partie 4.1), il a fallu valider l'outil de simulation de localisation. Pour cela, on a cherché à localiser des points de mesure réels à l'aide de la base de données construite par les mesures expérimentales. La validation des algorithmes Radar et Horus nous a mené par la suite à comparer leurs performances respectives en terme précision de localisation et de rapidité de traitement. La partie 4.2 présente les résultats les plus important du stage lorsque l'on utilise la base de données de simulation pour localiser les points de mesures réelles.

4.1 Validation du système implanté

Pour comparer la performance des systèmes Radar et Horus, nous avons implanté un système qui permet d'évaluer des techniques de localisation (déterministe et probabiliste). Le but de cette partie est de valider le fonctionnement du système implanté en utilisant la base de données créée à l'aide des mesures expérimentales pour localiser les points de mesure (partie 4.1.2). Le mode d'évaluation utilisé pour cette validation est expliqué dans la partie 4.1.1.

4.1.1 Mode d'évaluation

Pour tester de façon réaliste les algorithmes implantés dans le système, nous avons fait les choix suivant :

➤ **Phase offline : Construction de la base de donnée**

Afin de prendre en compte la distribution locale de la puissance du signal en chaque point de test, les histogrammes des mesures expérimentales sont construits à l'aide des échantillons collectés durant la mesure. Ces échantillons ont été mesurés en un même point de test pendant une minute. La puissance du point de test est la moyenne de ces échantillons. Ce mode de construction de la base de donnée est identique à celui utilisé par Radar et Horus.

➤ **Phase online : Choix aléatoire des valeurs de l'histogramme**

Dans la phase online, on choisit aléatoirement un échantillon parmi les échantillons appartenant à l'histogramme d'un point de mesure. Cet échantillon est considéré comme un représentant de la puissance du point de test à localiser. Ce choix aléatoire permet de simuler le canal radio réel.

Pour tester un algorithme de localisation, on va lancer la phase online 10 fois pour chaque point de mesure. La base de données comprend 137 points de mesure. Ainsi, on cherchera à localiser $137 \times 10 = 1370$ valeurs de puissance. L'erreur moyenne de localisation est calculée pour ces 1370 valeurs.

4.1.2 Résultat

Description

Chaque point de mesure comprend environ 200 échantillons. On utilise ces échantillons pour construire l'histogramme en chaque point de test pour chaque émetteur. Dans la phase online, pour chaque point P à localiser, on tire aléatoirement un échantillon comme présenté ci avant. On utilise la méthode Radar ou Horus pour chercher P', la position estimée de P, dans la base de données de mesure.

Résultats (cf. tableau de résultats en annexe 1)

Pour évaluer la localisation basée sur les mesures expérimentales, on trace les fonctions de répartition des erreurs de localisation qui présentent le pourcentage des points dont la distance entre le point de test et le point estimé par le système de localisation est inférieur à 0.5m, 1m, 1.5m ,5m (figure 8). Sur ces courbes, on observe les paliers de valeurs tous les 2m (0-2m, 2-4m, 4-5m). Ces paliers correspondent à la résolution des mesures expérimentales (les points de mesures sont espacés de 2m environ)

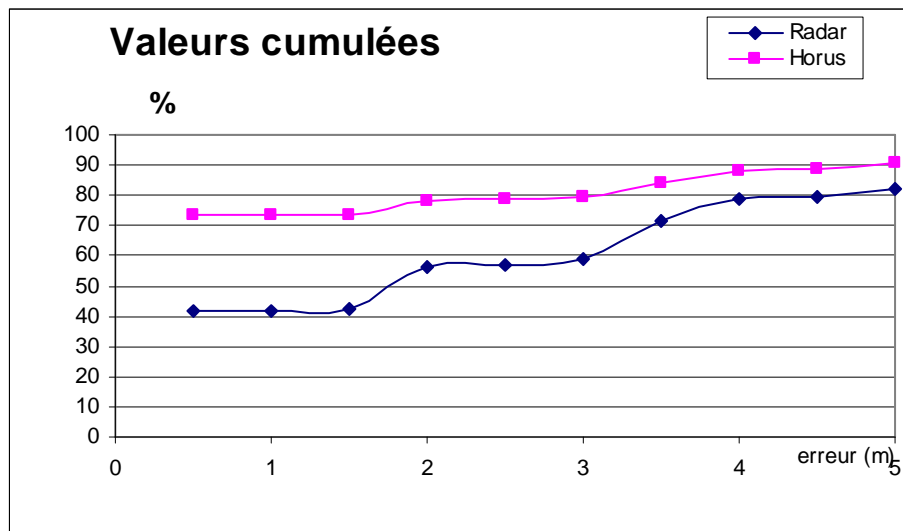


Figure 8: MESURES - Comparaison Radar-NNSS et Horus

La méthode probabiliste donne un meilleur résultat (l'erreur moyenne pour Horus est de **1.39m**, pour Radar elle est de **3.0m**).

Cependant, le tirage aléatoire des échantillons de mesure assure que la valeur des puissances à localiser appartienne toujours à la plage de valeur de l'histogramme local. Or, dans un cas réel, il est possible de recevoir une puissance au point P hors de la plage de l'histogramme local du point P. Dans ce cas, avec Horus, la probabilité conjointe calculée pour tous les N points d'accès est nulle pour P et on estime le point P' en un point différent de P. Ce phénomène crée une erreur de localisation qui n'est prise en compte dans notre mode d'évaluation pour Horus. Il peut considérer que l'erreur moyenne de localisation de Horus est légèrement sous-estimée par notre système.

Les histogrammes ci-dessous permettent de bien observer la distribution des erreurs du système de localisation. Dans ce cas, la méthode probabiliste est très performante car environ 80% des points sont localisés avec une distance d'erreur inférieure à 0.5 mètres.

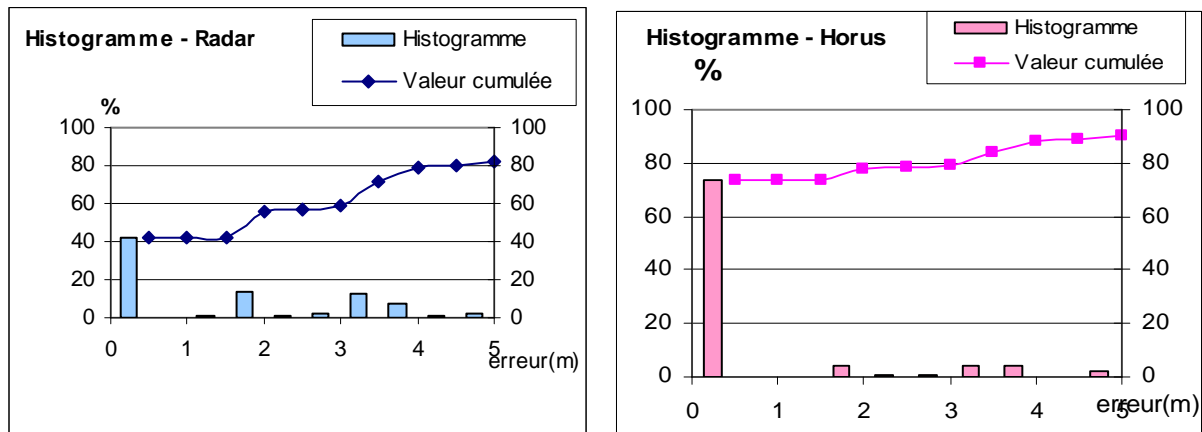


Figure 9 : MESURES - Histogramme des erreurs moyennes

Interprétation

Le tableau 1 présente l’erreur médiane et l’erreur à 75% pour les méthodes Radar et Horus obtenues avec notre système de simulation de localisation. Ces erreurs sont comparées dans ce tableau aux erreurs présentées dans la littérature décrivant les performances de ces systèmes pour un fonctionnement réel [9] [15].

Méthode	Avec nos mesures et notre système		Les systèmes réels	
	Radar - NNSS	Horus	Radar - NNSS	Horus
Erreur à 75% (m)	3.7	1.7	3	1.3
Erreur médiane (m)	1.8	<0.5	2.9	0.5

Tableau 1 : Comparaison des résultats de notre système avec les résultats des systèmes réels

Si on compare les résultats de notre système avec les résultats obtenus par la méthode empirique du système Radar et par la technique Joint Clustering du Horus, on trouve que les erreurs à 75% sont similaires. Par contre, notre erreur médiane est légèrement meilleure que celle trouvée dans la littérature. Ce phénomène peut être expliqué par les raisons suivantes :

Avec le système Radar (cf. annexe 2)

- Nous avons utilisé plus de points d’accès (6 points d’accès) que le système Radar (3 point d’accès) pour construire la base de données. Plus le nombre d’accès augmente, moins il y a d’ambiguïté pour localiser un point. C’est principalement pour cela que notre système présente une meilleure erreur médiane.
- Dans la littérature, le système Radar a été évalué avec un mode de test différent du notre. En effet, le point à localiser P est retiré de la base de donnée dans la phase online. Ainsi, on ne peut jamais arriver à une erreur nulle. Nous considérons que ce mode d’évaluation n’est pas réaliste car le meilleur cas est ignoré. C’est pour cette raison que nous avons choisi de tirer aléatoirement un échantillon de l’histogramme d’un point de mesure.
- Enfin, comme on travaille sur un environnement différent que celui de Radar, on n’obtient donc pas de résultats identiques.

Avec le système Horus (cf. annexe 3)

- Les points de test que nous localisons sont issus des mêmes mesures utilisées dans la construction de la base de données. Ce n'est pas le cas du système Horus qui a été testé par plusieurs personnes et à des moments différents. Le test est ici plus exhaustif que le notre car il prend en compte une variation du canal radio plus importante. Ceci permet de justifier la valeur plus importante de l'erreur médiane présentée dans la littérature.
- Dans nos mesures, nous n'avons utilisé que le calcul de probabilité conjointe sans appliquer la technique de regroupement des points d'accès (clustering). Cette technique permet de diminuer le temps de traitement dans la phase online mais elle a pour conséquence de provoquer plus d'erreurs de localisations. En effet, si lors de la mesure du point à localiser on n'attribue pas cette mesure au bon cluster, on introduit une erreur supplémentaire.

Conclusion

Malgré le fait que les résultats trouvés par notre système soient légèrement différents de ceux obtenus avec les systèmes réels, ils sont du même ordre de grandeur. Nous considérons que le système de test que nous avons implanté est validé.

Il est clair que la technique probabiliste (le système Horus) est plus performante que la technique déterministe (le système Radar). Cependant, la méthode de localisation proposée par Radar est beaucoup plus simple que celle de Horus. De plus, nous considérons qu'avec une erreur moyenne de 3 mètres, les performances de Radar sont très honorables dans un environnement Indoor. Nous avons donc choisi la technique déterministe (Radar) pour tester l'utilité du moteur de prédiction Wilde pour la localisation en temps réel.

4.2 Localisation avec le moteur WILDE.

La base de données la plus réaliste dans la phase offline est obtenue à l'aide de mesures expérimentales en différents points de l'environnement de test. Cependant cette méthode de construction de carte de couverture radio est longue et elle ne permet pas de mettre à jour la base de données lors des modifications de paramètres de l'environnement. Un changement de position d'un point d'accès implique un changement total de la base de données et on devra donc refaire toutes les mesures. De plus, le canal radio n'est pas stationnaire. En effet, il varie en fonction du temps avec l'occupation des lieux qui se modifie au cours de la journée.

De ce fait, la construction d'une base de données par simulation a un grand avantage car la carte de couverture peut être recalculée rapidement et aisément. Il faut juste changer des paramètres de l'environnement et lancer les simulations de couverture dans la phase offline, ce qui ne prend que quelques minutes. Cependant, il est important de noter que la performance du système de localisation dépend fortement de la qualité de la simulation. Si l'écart entre les puissances simulées et les puissances mesurées est grand, il est évident que l'on n'arrivera pas à localiser correctement le point courant.

Dans le paragraphe 4.2.1, je présente le mode d'évaluation utilisé pour tester l'efficacité de Wilde pour la localisation. Je présente ensuite les résultats de localisation des points de mesures à l'aide de la base de données de simulation (partie 4.2.2). Radar est la méthode de localisation utilisée dans ces tests. Afin d'optimiser la localisation, j'ai modifié les données de la base de simulation en appliquant **un seuil de détection** (partie 4.2.3). Ceci a permis de diminuer l'erreur moyenne de localisation. Dans la dernière partie 4.2.4, nous évaluons la qualité de localisation lors d'un déplacement suivant **un trajet continu**.

4.2.1 Mode d'évaluation

Un mode d'évaluation a été défini pour pouvoir utiliser les cartes de couverture de Wilde avec le système de localisation présenté dans la partie 5.1. Dans ce cas, la base de données est construite avec l'ensemble des cartes de couverture et la localisation se fait à partir de mesures.

➤ Phase offline : Construction de la base de donnée

Le moteur de calcul Wilde est utilisé avec une résolution de calcul de 6 par 6 pixels (cf. figure 5) avec un pas de discrétisation de 10 cm par pixel. C'est à cette résolution que les résultats du moteur ont été calibrés.

Des points de la carte de couverture obtenue à cette résolution sont sous échantillonnés à un pas G fixe réglable et sont inclus dans la base de données de simulation. Ainsi, on a sélectionné un ensemble de points qui peuvent être représentés sur une grille de pas G. Les points qui coïncident avec les murs ne sont pas retenus.

➤ Phase online : Choix aléatoire des valeurs de l'histogramme

Les points à localiser sont les points que nous avons mesurés expérimentalement. Cependant, nous ne choisissons que les points qui sont couverts par au moins 4 points d'accès pour assurer la qualité de la localisation. 137 parmi 199 points de mesure sont donc sélectionnés pour évaluer la performance de la base de données de simulation.

Lors de l'évaluation, on lance la phase online 10 fois pour chaque point de mesure. La base de données comprend 137 points de mesure. Ainsi, on cherchera à localiser $137 \times 10 = 1370$ valeurs de puissance. L'erreur moyenne de localisation est calculée pour ces 1370 valeurs.

4.2.2 Résultat.

Description

La base de données utilisée dans cette partie est construite à l'aide du moteur de calcul de couverture Wilde. Le pas de la grille utilisée est de 5 pixels, soit 50 cm.

Trois simulations ont été lancées :

- 1^{ère} simulation : on localise 137 points qui sont couverts par au moins 4 points d'accès
- 2^{ème} simulation : on localise 114 points qui sont couverts par au moins 5 points d'accès
- 3^{ème} simulation : on localise 50 points qui sont couverts par tous les 6 points d'accès.

Résultats

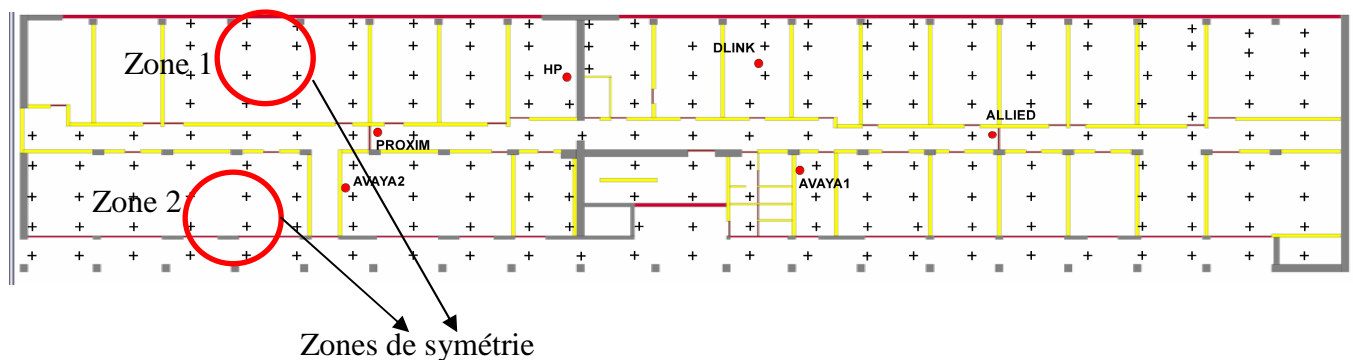
N = nombre d'AP	Nombre de points à localiser	Erreur moyenne
4 ou 5 ou 6	137	6.2 m
5 ou 6	114	5.1 m
6	50	3.5 m

Tableau 2 : Résultats obtenus à l'aide de la base de données de simulation**Interprétations**

- On remarque que la distance d'erreur de localisation devient petite quand le nombre de points d'accès couvrant les points à localiser augmente. On passe d'une erreur moyenne de 6,2 mètres avec au moins 4 émetteurs à une erreur moyenne de 3.5 mètres avec 6 émetteurs. Ce résultat est prévisible car on introduit plus de disparité entre les points de la base.
- Si on considère les points qui sont couverts par tous les six points d'accès, on obtient une erreur moyenne de 3.5 mètres. Ce résultat est acceptable dans les conditions réelles. Cependant l'erreur moyenne de 6.2 m pour 137 points qui sont couverts par au moins 4 points d'accès est encore importante. Si on localise ces mêmes points de tests avec la base de données construite par les mesures expérimentales, l'erreur moyenne est de 3 mètres. Cette erreur de 6.2 m peut s'imputer aux raisons suivantes :

- **Algorithme de localisation – Zones de symétrie**

La disposition des points d'accès peut provoquer des zones de symétrie sur la carte de couverture radio. Si l'on cherche à localiser un point qui est dans la zone 1, on observe deux types de résultats : soit la position estimée de ce point se trouve dans la zone 1 (distance d'erreur faible), soit dans la zone 2 (distance d'erreur grande, environ 10 mètres).

**Figure 10 : Zones de symétrie**

Si on regarde la distribution de puissance des six points d'accès ALLIED, AVAYA1, AVAYA2, DLINK, HP et PROXIM, on observe que les zones 1 et 2 sont proches dans l'espace des signaux. Comme la puissance du signal à une position dépend de plusieurs paramètres (matériaux des murs, nombre de murs traversés...) et phénomènes (chemins multiples, réflexion...), il est donc possible d'avoir des points de la zone 2 qui ont un vecteur de puissance du signal proche de celui d'un point de la zone 1.

Il est intéressant de noter que si l'on souhaite planifier un système radio indoor et que l'on souhaite l'optimiser pour une implantation d'un système de localisation, il est important d'éviter de créer des zones de symétrie dans l'espace des signaux du plan.

○ **Ecart entre les mesures expérimentales et les valeurs de simulation.**

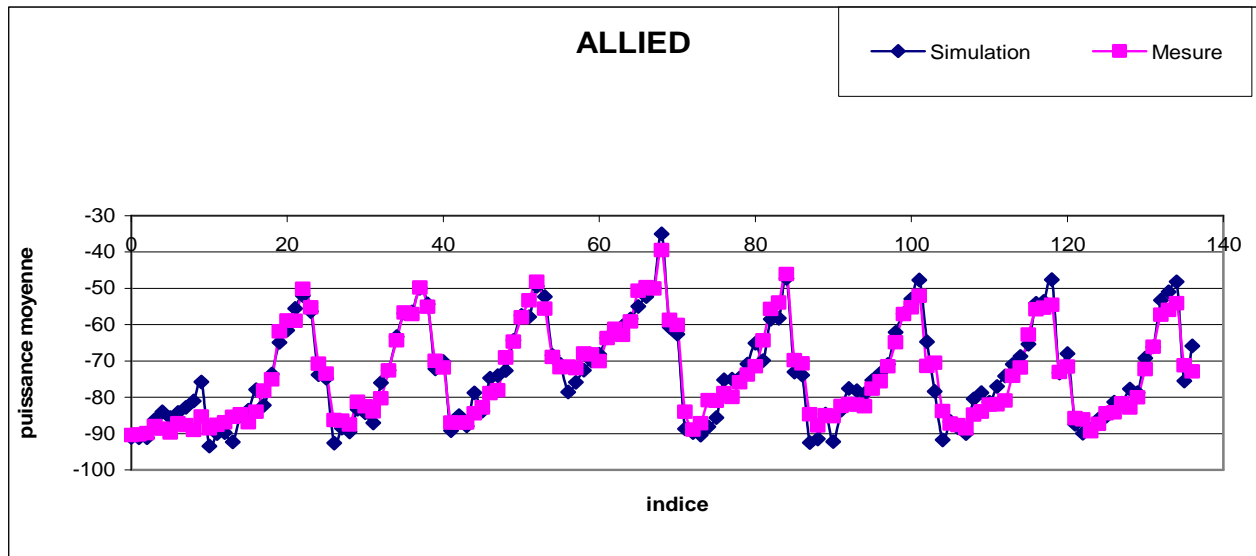


Figure 11 : Puissances mesurées et simulées pour le point d'accès ALLIED.

On constate que si on construit la base de données avec les valeurs mesurées expérimentalement, la distance d'erreur moyenne pour 137 points n'est que de 3 mètres. L'utilisation de la base de données de simulation présente une erreur moyenne de 6,2 mètres. On peut imputer une partie de l'erreur supplémentaire observée avec la base de simulation à l'écart entre les puissances mesurées et les puissances simulées.

En observant la figure 11, qui représente la puissance mesurée et simulée en chacun des 137 points de mesure, on s'aperçoit que ce sont les points en limite de couverture qui sont les moins bien estimés. Plus précisément, ces points sont sous-estimés, c'est-à-dire que la valeur de la puissance simulée est trop faible.

On peut imputer ce phénomène au matériel de mesure utilisé. En effet, la sensibilité du récepteur est fixée. Quand on mesure un point en limite de couverture, la répartition des niveaux de puissance mesurés ne ressemble plus à une distribution gaussienne. En effet, les échantillons de puissance inférieure à la sensibilité du récepteur (-90dBm par exemple) ne sont pas mesurés (figure 12, l'histogramme à droite). Quand on calcule la puissance moyenne du signal en ces points, on aura tendance à surestimer la valeur réelle de la moyenne du champ. C'est pourquoi il est judicieux de ne pas prendre en compte les points de mesure situés en limite de couverture.

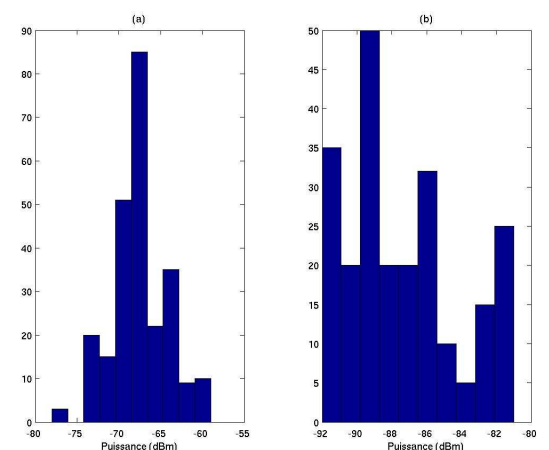


Figure 12 : Histogrammes de points de mesure bien couverts (à droite), en limite de couverture (à gauche).

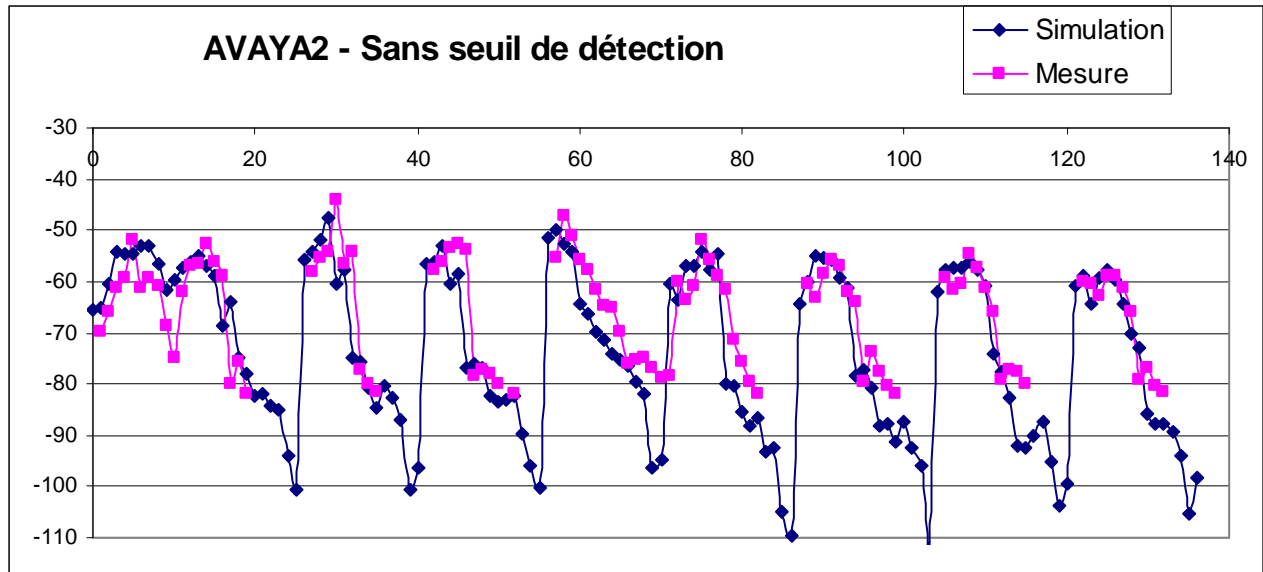


Figure 13 : Puissances mesurées et simulées pour le point d'accès AVAYA2

La figure 13 présente la puissance du signal (simulations et mesures) des 137 points de test. On voit bien que le point d'accès AVAYA2 ne détecte que les signaux dont la puissance est supérieure à -82 dBm. Pour que la simulation soit plus proche de la réalité, on a introduit dans la phase offline un seuil de détection. Ce seuil de détection et les résultats correspondants sont présentés dans le paragraphe suivant.

4.2.3 Seuil de détection

Description

Dans le but d'améliorer la performance de la localisation, un seuil de détection S a été introduit. On a testé des valeurs de seuil pour S allant de -75 dBm à -105 dBm. La valeur $S = -90$ dBm rend le système de localisation le plus performant. Pour une valeur de seuil inférieure à -90 dBm, on garde trop de points erronés et pour une valeur proche de -75 dBm, on dégrade les performances du système car on a trop peu de points pour localiser correctement les signaux. Ce seuil de détection est utilisé pour les simulations et également pour les mesures de la façon suivante :

- Simulations : toutes les puissances inférieures à -90 dBm sont remplacées par -90 dBm
- Mesures : toutes les puissances inférieures à -90 dBm et tous les signaux non détectés sont remplacés par -90 dBm

Résultats

N = nombre d'AP	Nombre de points à localiser	Erreur moyenne
4 ou 5 ou 6	137	5.4 m
5 ou 6	114	4.4 m
6	50	3.5 m

Tableau 3 : Résultats obtenus à l'aide de la base de données avec seuil de détection

Interprétations

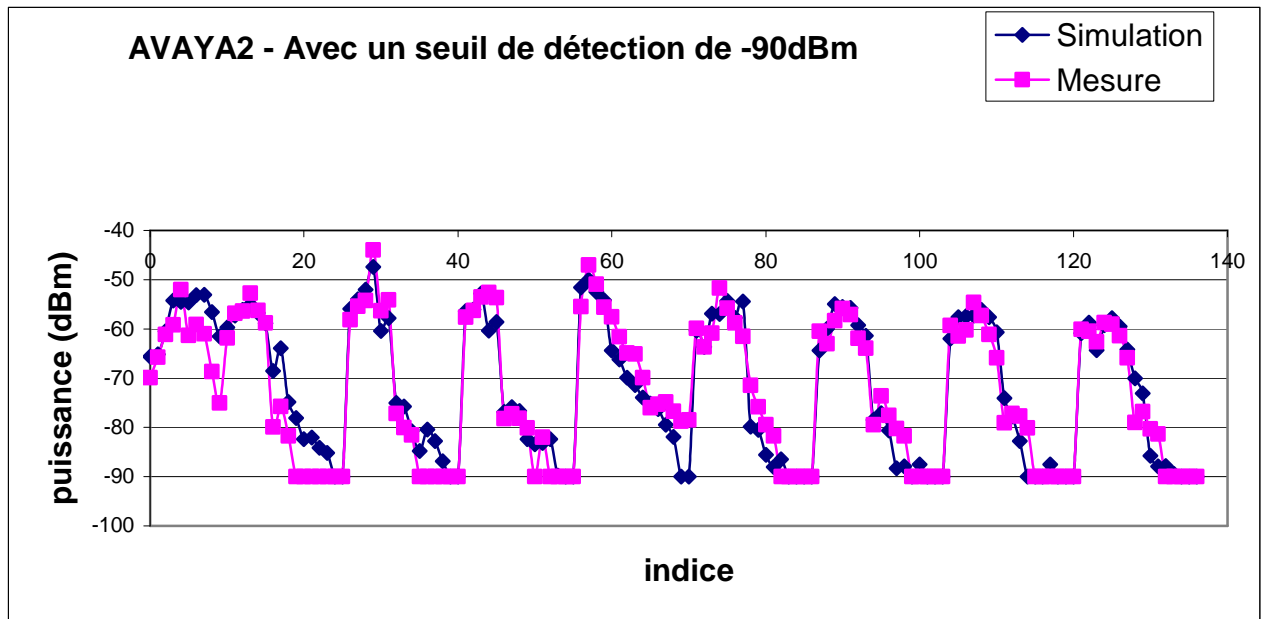


Figure 14 : Puissance des signaux avec $S=-90\text{dBm}$ pour le point d'accès AVAYA2

Avec le seuil de détection à -90dBm , l'erreur moyenne pour 137 points de test a diminué en passant de **6.2 mètres** à **5.4 mètres**. Cependant, l'erreur moyenne pour les 50 points couverts par tous les six points d'accès reste toujours égale à 3.5 mètres. On peut donc conclure que la présence d'un seuil de détection améliore la localisation des points qui ne sont pas très bien couverts par les points d'accès. La qualité de localisation pour les points ayant une bonne couverture (couverts par tous les six points d'accès) n'est pas optimisée.

4.2.4 Localisation des points d'un chemin continu

Description

Le dernier résultat que je voudrais vous montrer dans ce rapport est la variation de l'erreur moyenne de localisation quand on se ballade sur un chemin continu de l'environnement CITI. Le trajet, qui comporte 64 points de test, est représenté sur la figure 15 par la couleur rouge. Le point de départ est B1 et la destination est le point B64.

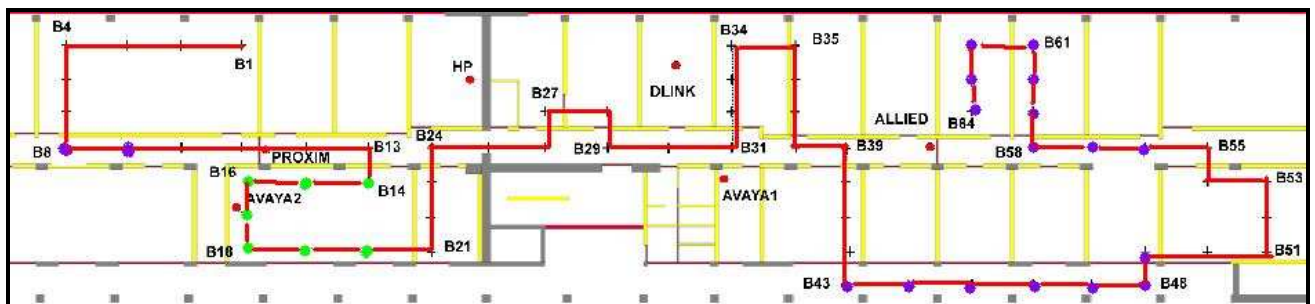


Figure 15 : Localisation d'un trajet continu

Vous trouvez sur le chemin les points violets (B8, B9, B43 → B49, B56 → B64) et les points verts (B14 → B20). Ce sont les points qui ne sont pas bien couverts par les points d'accès. A la position des points violets, on ne reçoit qu'un signal et à la position des points verts, on ne reçoit que trois signaux. Les autres points sont couverts par au moins 4 points d'accès.

Résultats :

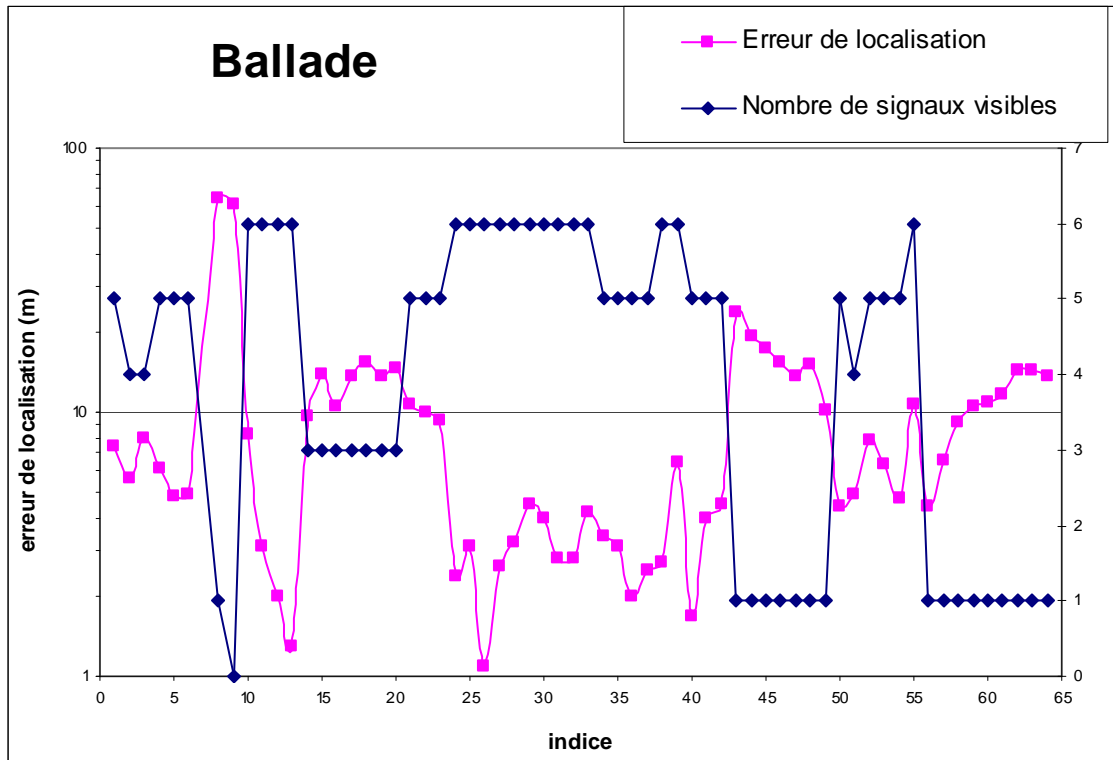


Figure 16 : Erreur de localisation d'un trajet continu

La figure 16 représente l'erreur de localisation à chaque point du trajet. Le nombre de signaux reçus à chaque position est également exprimé dans la figure. Nous constatons qu'en général un point bien couvert (le nombre de signaux visibles est au moins égal à 4) est localisé avec une erreur faible (l'erreur de localisation est inférieure à 5 mètres). Par contre, pour les points mal couverts, cette valeur devient très grande et elle fait augmenter l'erreur moyenne de localisation de l'ensemble de 64 points du trajet de test.

Nombre de signaux reçus	Erreur moyenne de localisation
0 → 6	9.7 mètres
3 → 6	5.8 mètres
4 → 6	4.3 mètres

Tableau 4 : Erreur moyenne de localisation d'un trajet continu

En effet, si on prend en compte tous les 63 points du chemin, on obtient une erreur moyenne de 9.7 mètres. Cependant, si on enlève les points violets, qui sont couverts par aucun ou un seul signal, cette erreur diminue à 5.8 mètres. De plus, si on ne considère que les points pour lesquels le nombre de signaux visibles est supérieur ou égal à 4, la localisation est nettement meilleure et on a une erreur moyenne de 4.3 mètres, ce qui est un bon résultat.

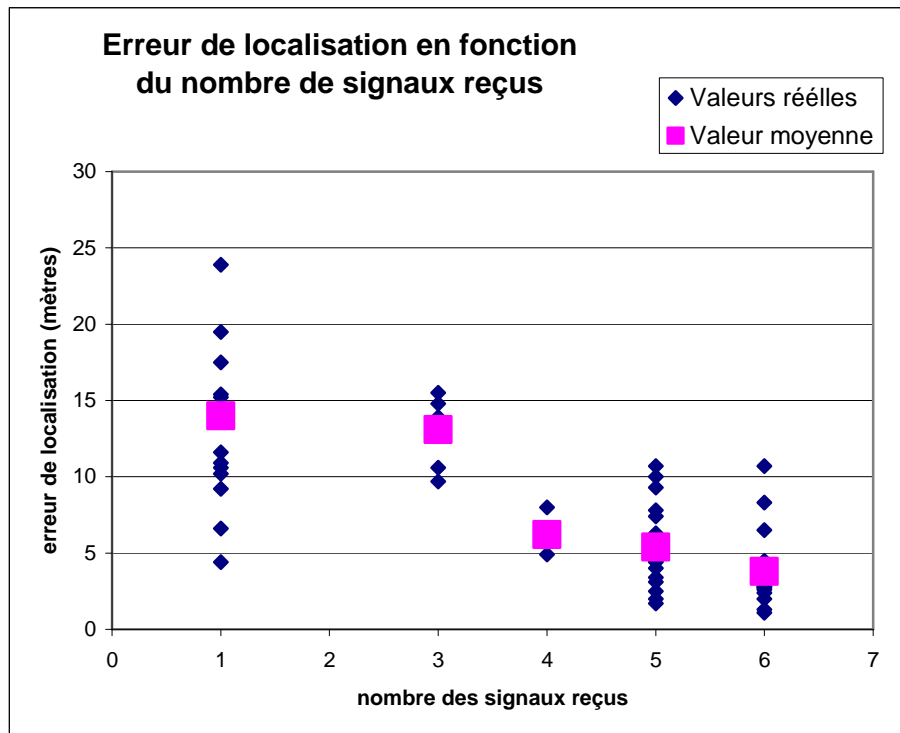


Figure 17 : Erreur de localisation en fonction du nombre de signaux visibles

La figure 17 représente l’erreur de localisation en fonction du nombre de signaux visibles. Afin de visualiser correctement la courbe, les points B8 et B9 présentant une erreur importante ne sont pas représentés. On observe que l’erreur moyenne de localisation diminue nettement quand le nombre des signaux passe de 3 à 4. Ainsi, un environnement est bien couvert si à chaque position, on reçoit au moins quatre signaux provenant des points d’accès.

Interprétations

La majeure partie des points fortement erronés est couverte par moins de 3 points d’accès. Mais on observe que quelques points couverts par plus de 5 points d’accès présentent une erreur de localisation importante. Ces points sont les points (B1, B3, B10, B52, B55) représentés sur la figure 18 par des croix de couleur. Les positions estimées sont représentées sur la figure par de petits rectangles de la même couleur. La position des points d’accès est décrite à l’aide des points noirs.

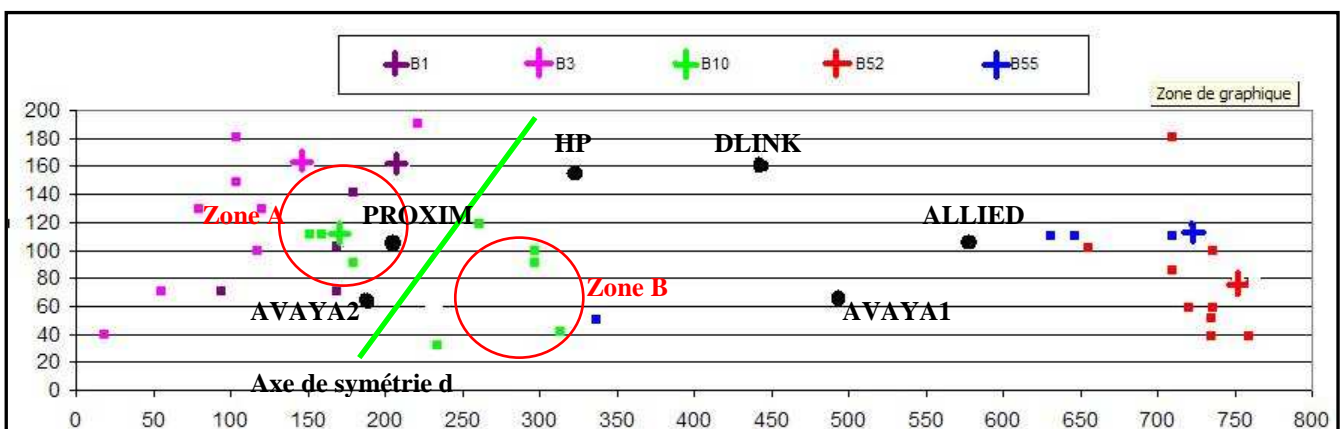


Figure 18 : Quelques points bien couverts mais ils sont mal localisés.

Si les points d'accès sont bien positionnés dans l'espace, une position bien couverte doit être localisée avec une erreur de localisation faible. Cependant, nous constatons que dans ce cas de figure, le point B10 (couleur verte) est couvert par 5 points d'accès mais il est localisé avec une erreur moyenne de 10,7 mètres pour 10 tentatives de localisation (tirage aléatoire d'un échantillon de mesure et localisation de celui-ci). Si on analyse les positions estimées de B10, on trouve des points qui sont symétriques à B10 par rapport à l'axe de symétrie d. On peut expliquer ce phénomène comme suit :

- Le point B10 reçoit 5 signaux venant de 5 points d'accès. Or, les trois signaux les plus forts et donc les plus fiables, sont des signaux d'AVAYA2, de PROXIM et de HP.
- La disposition de ces trois points d'accès a créé dans l'espace de signaux des zones similaires (zone A et zone B). Ces zones sont symétriques par rapport à l'axe de symétrie d. C'est la raison pour laquelle il existe des positions estimées de B10 qui se trouvent dans la zone B au lieu de A.

Encore une fois, le phénomène de zone de symétrie influence sur la qualité de la localisation. Ainsi, la phase de planification des points d'accès joue un rôle non négligeable pour obtenir un bon système de localisation.

Pour les points qui ne possèdent pas de zones de symétrie, l'erreur de localisation peut provenir de l'écart entre les puissances mesurées et les puissances simulées.

5 CONCLUSIONS ET PERSPECTIVES

5.1 Conclusions

L'objectif de mon stage était de montrer la faisabilité de l'utilisation d'un simulateur pour construire de. Après avoir analysé les résultats et étudié les phénomènes apparus, j'ai tiré les conclusions suivantes :

- La localisation des stations mobiles dans les environnements Indoor, basée sur le principe Offline/Online, dépend tout d'abord de la qualité de la base de données. Pour les bases de données construites par les mesures expérimentales, un grand nombre de points de mesures et un temps de mesure suffisant permet d'obtenir une carte de couverture de radio de bon niveau. Les bases de données de simulation, quand à elles, ne sont performantes que si les puissances des signaux obtenues par le moteur de calcul de couverture radio sont proches de la réalité.
- La précision d'un processus de localisation dépend également de la disposition des points d'accès. Il faut que les points d'accès couvrent l'environnement correctement. C'est-à-dire, il faut que chaque zone de l'espace soit couvert par au moins quatre points d'accès. Une autre contrainte de planification est à prendre en compte. Il ne faut pas que la carte de couverture de l'ensemble des signaux présente des zones identiques ou très proches dans l'espace signal. Nous appelons ces zones des zones de symétrie.
- Nous pouvons aussi utiliser la base de données de simulation pour localiser des points de simulation. Bien que les résultats de ces tests n'aient pas de valeurs réelles, ils reflètent quand même les comportements généraux du système de localisation. On peut, par exemple, faire ce type de test pour détecter des zones de symétrie ou pour évaluer et rechercher un bon positionnement des points d'accès.

Les résultats de localisation des mesures réelles avec la base de simulation peuvent être améliorés. Si on ne considère que les points bien couverts de la zone centrale (figure 19 : 94 points sur le laboratoire CITI, de 7 à 18 en abscisse), on a une erreur moyenne de **4,4 mètres**. Ce résultat est meilleur car on se trouve dans une zone du plan où il y a moins de symétries et les puissances des signaux sont élevées.

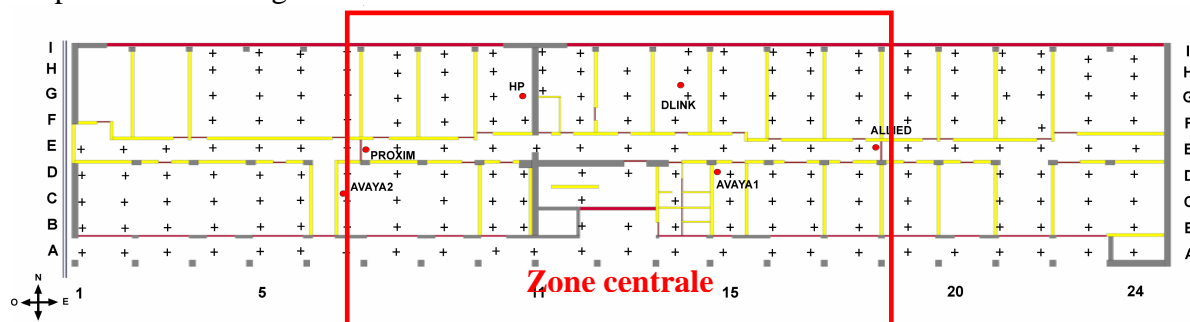


Figure 19 : La zone centrale de l'environnement de test

Dans la partie suivante, je propose quelques idées qui devraient permettre d'optimiser la performance du système de localisation basé sur Wilde.

5.2 Perspectives

Dans la phase d'optimisation de résultats de localisation à la fin du stage, à côté du critère de seuil de détection que j'ai présenté précédemment, nous avons proposé trois autres idées qui pourraient permettre d'améliorer la précision de la localisation.

- La première idée se rapporte au calcul de distance euclidienne dans la phase online de Radar. Dans la phase online, on sélectionne le point P' qui minimise la distance euclidienne entre le signal mesuré et le signal du point P' de la base. Comme nous avons observé que la mesure de signaux peu puissants était entachée d'une erreur de simulation importante, nous allons **pondérer le calcul de la distance euclidienne** de façon à favoriser les signaux les plus forts.

Les signaux de puissance forte, qui sont plus fiables, seront pondérés par un coefficient important et les signaux faibles, qui sont moins précis, seront multipliés par un coefficient plus faible. Ces coefficients suivent une fonction de puissance du signal décroissante. En fait, j'ai déjà testé ce critère avec des fonctions décroissantes de type linéaire et exponentiel. Cependant, je n'ai pas encore trouvé des meilleurs résultats. L'erreur moyenne pour 137 points de test est toujours proche de 5,5 mètres. Comme dans les tests effectués j'ai déjà pris en compte le seuil de détection, cette pondération peut faire double emploi avec l'intégration du seuil de détection. Il serait intéressant de tester le calcul modifié de la distance euclidienne sans le seuil de détection.

- La deuxième proposition se rapporte encore au choix du point dans la phase online. On présente un **critère de choix dit « géométrique »**. Pour une plage de valeur de puissance du signal d'un point d'accès, on définit une zone géométrique correspondante (un cercle ou une couronne). La position estimée sera l'intersection des zones géométriques des points d'accès les plus forts (figure 20). Comme la réalisation de cette proposition n'est pas évidente, nous n'avons pas eu assez de temps pour la réaliser.

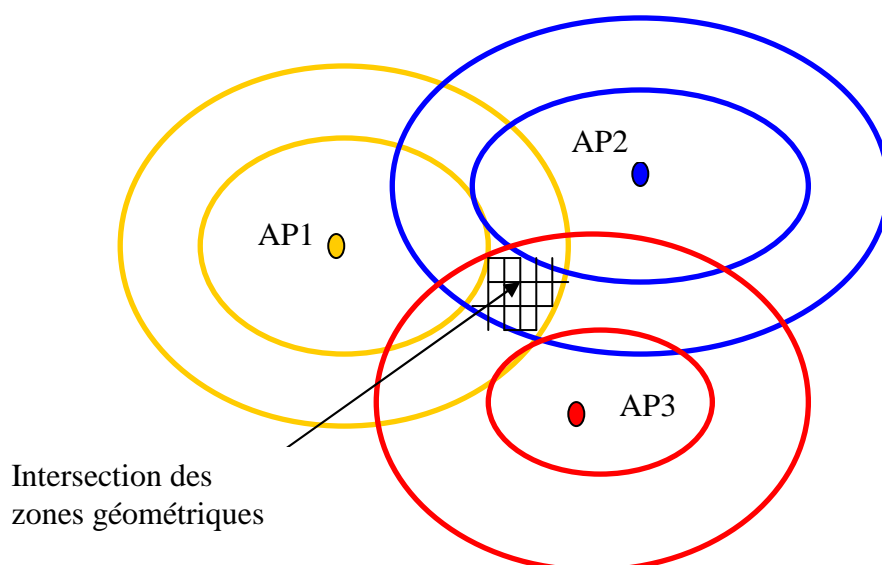


Figure 20 : Deuxième proposition – Critère géométrique

- Jusqu'à présent, toutes les méthodes déterministes de localisation indoor sont basées sur la puissance du signal. Notre troisième idée est de localiser des points en utilisant **les écarts relatifs entre les signaux** de chaque point d'accès en un point donné. On n'exploite pas seulement les valeurs (P_1, P_2, \dots, P_6) mais aussi les écarts ($P_6 - P_5, \dots, P_i - P_j, \dots, P_2 - P_1$). Ainsi, on est tributaire non plus de la valeur réelle des niveaux de champs mais des valeurs relatives d'un point d'accès par rapport à un autre en un même point de test. Ceci devrait rendre la méthode plus robuste aux variations lentes du canal qui peuvent affecter un même point du plan. Nous n'avons pas encore pu tester cette proposition.

En ce moment, si on utilise la base de données issue des simulations Wilde, on obtient une erreur moyenne de **5,4 mètres pour 137 points de test**. Cependant, il faut bien remarquer que dans les tests effectués, les points d'accès ont été répartis de façon arbitraire et on voit donc bien que les positions ne sont pas idéales. De plus, on n'a pas encore implanté la technique « User Tracking », une méthode très efficace pour éliminer les erreurs importantes. Ainsi, je pense vraiment que si l'on applique bien les procédures suivantes, on pourra fortement améliorer la performance de notre système :

- Tout d'abord, pour éviter les zones de symétrie, il faudra optimiser le **positionnement des points d'accès**.
- En suite, on mettra en oeuvre la **technique « User Tracking »** et réalisera une nouvelle expérimentation avec un bon positionnement des points d'accès.
- Enfin, on pourra également appliquer **les procédures de choix des points** en phase online présentées ci avant pour améliorer la qualité du système de localisation.

En conclusion, si on peut faire tendre cette erreur de localisation vers une erreur de 3 mètres, on obtient la même performance qu'avec une base de données créée par des mesures expérimentales. Ainsi, un système de localisation incluant un moteur de calcul de couverture comme Wilde présentera beaucoup d'avantages en terme de temps de calcul et de portabilité du système.

BIBLIOGRAPHIE

- [1] R. WANT, A. HOPPER et al. *The Active badge Location System*. Olivetti Research Ltd, Cambridge, 1992
Disponible sur: < <http://citeseer.nj.nec.com/want92active.html> >
- [2] Constellation system.
Disponible sur: <http://www.isense.com/company/papers/FoxlSig98.pdf>
- [3] A. HARTEK, A. HOPPER et al. *The Anatomy of a context-Aware Application*. AT&T Laboratories Cambridge, 1999.
Disponible sur: < <http://citeseer.nj.nec.com/harter02anatomy.html> >
- [4] N.B. PRIYANTHA, A. CHAKRABORTY, H. BALAKRISNAN. *The Cricket Location-Support System*. MobiCom 2000. Disponible sur: < <http://citeseer.nj.nec.com/priyantha00cricket.html> >
- [5] A. SMAILAGIC, D. KOGAN. *Location sensing and privacy in a context-aware computing environment*. IEEE Wireless Communications, Octobre 2002.
Disponible sur: < <http://citeseer.nj.nec.com/smailagic01location.html> >
- [6] M. BRUNATO, C. K. KALLO. *Transparent Location Fingerprinting for Wireless Services*. MED-HOC-NET 2002, 2002. Disponible sur: < <http://citeseer.nj.nec.com/brunato02transparent.html> >
- [7] P. BAHL, V.N. PADMANABHAN et A. BALACHANDRAN. *A software System for Locating Mobile Users: Design, Evaluation and Lessons*. Rapport technique, Recherche de Microsoft, MSR-TR-2000-12, Avril 2000.
Disponible sur: < <http://citeseer.nj.nec.com/bahl00radar.html> >
- [8] P. BAHL, V.N. PADMANABHAN et A. BALACHANDRAN. *Enhancement to the RADAR User Location and Tracking System*. Rapport Technique. MSR-TR- 2000-12, 2000
Disponible sur: < <http://citeseer.nj.nec.com/bahl00enhancements.html> >
- [9] P. BAHL, V.N. PADMANABHAN et A. BALACHANDRAN. *RADAR : An In-Building RF-Based User Location and Tracking System*. IEEE Infocom 2000, volume 2, p775-784. Mars 2000.
Disponible sur: < <http://citeseer.nj.nec.com/bahl00radar.html> >
- [10] P. CASTRO, P. CHIU et al. *A probabilistic room location service for Wireless Networked Environments*. UCLA, Département Informatique, Los Angeles, 2001.
Disponible sur: < <http://www.fxpal.com/people/chiu/cckm-UBICOMP01.pdf> >
- [11] A. M. Ladd, K. Bekris et al. *Robotics-Based Location Sensing using Wireless Ethernet*. 8ème ACM Mobicom, Atlanta, GA, Septembre 2002.
Disponible sur: < <http://citeseer.nj.nec.com/537210.html> >
- [12] T. ROOS, P., MYLLYMAKI et al. *A probabilistic Approach to WLAN User Location Estimation*. International Journal of Wireless Information Networks, July 2002.
Disponible sur: < <http://netlab18.cis.nctu.edu.tw> >
- [13] M. YOUSSEF, A. AGRAWALA. *Small-Scale Compensation for WLAN Location Determination System*. IEEE WCNC 2003, Mars 2003.
Disponible sur: < <http://citeseer.nj.nec.com/youssef03smallscale.html> >
- [14] M. YOUSSEF, A. AGRAWALA, A. U. SHANKAR, S. H. Noh. *A probabilistic Clustering-Based Indoor Location Determination System*. Rapport technique, UMIACS-TR-2002-30 et CS-TR-444350. University de Maryland, College Park, Mars 2002.
Disponible sur : < http://www.cs.umd.edu/~moustafa/papers/locdet_tr.pdf >
- [15] M. YOUSSEF, A. AGRAWALA, A. U. SHANKAR. *WLAN Location Determination via Clustering and Probability Distributions*. IEEE Percom2003, Mars 2003.
Disponible sur < <http://citeseer.nj.nec.com/youssef03wlan.html> >
- [15] M. YOUSSEF, A. AGRAWALA, A. U. SHANKAR. *Handling Samples Correlation in the Horus System*. Mars 2004
- [17] Thanh Ha LE *Localisation des stations mobiles dans les réseaux sans fils*. Projet de fin d'études, 6/2004, INSA de Lyon.
- [18] J.M. GORCE and S. UBEDA. *Propagation simulation with the ParFlow method : fast computation using a multi-resolution scheme*, IEEE 54th Vehicular Technology Conference Fall 2001, Atlantic City, NJ, USA, Octobre 2001.
- [19] J.M. GORCE, E. Jullo and K. RUNSER. *An adaptative multi-resolution algorithm for 2D simulations of indoor propagation*, Proceedings of the 12th ICAP Conference, Exeter, UK, Avril 2003.

ANNEXES

Annexes 1

Validation du système implanté par les mesures expérimentales

Valeurs cumulées			Histogramme		
Niveau d'erreur	NNSS (%)	Horus (%)	Niveau d'erreur	NNSS (%)	Horus (%)
0 - 0.5m	41.78	73.62	0 - 0.5m	41.78	73.62
0 - 1m	41.78	73.62	0.5 - 1m	0	0
0 - 1.5m	42.37	73.75	1 - 1.5m	0.592	0.132
0 - 2m	56.32	77.96	1.5 - 2m	13.95	4.211
0 - 2.5m	57.24	78.49	2 - 2.5m	0.921	0.526
0 - 3m	59.14	79.41	2.5 - 3m	1.908	0.921
0 - 3.5m	71.78	83.88	3 - 3.5m	12.63	4.474
0 - 4m	79.01	88.36	3.5 - 4m	7.237	4.474
0 - 4.5m	79.67	88.68	4 - 4.5m	0.658	0.329
0 - 5m	82.04	90.46	4.5 - 5m	2.368	1.776
> 5m	17.96	9.539	> 5m	17.96	9.539
Erreur moyenne	3	1.39			

Annexes 2

Systeme Radar

Annexes 3

Systeme Horus