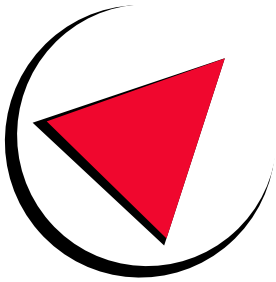
Étude de l'interception et du positionnement de trafic Wi-Fi dans un environnement hétérogène

Matteo Cypriani, Adrien Henriet, Philippe Canalda, François Spies



L I F C

LABORATOIRE D'INFORMATIQUE DE L'UNIVERSITE DE FRANCHE-COMTE

EA 4269

Étude de l'interception et du positionnement de trafic Wi-Fi dans un environnement hétérogène

Matteo Cypriani — Adrien Henriet — Philippe Canalda — François Spies

Rapport de Recherche no RR 2011–05

THÈME 4 – Novembre 2009

Étude de l'interception et du positionnement de trafic Wi-Fi dans un environnement hétérogène

Matteo Cypriani, Adrien Henriet, Philippe Canalda, François Spies

Thème 4
OMNI

Novembre 2009

Résumé : Dans le cadre d'une politique de sécurité dans les réseaux sans fils, il est intéressant de pouvoir identifier la position géographique d'une source de données, pour s'assurer de sa légitimé et de son utilisation du réseau. Dans cette optique, nous proposons ici un aperçu des possibilités de capture de trafic mises en relation avec un système de géolocalisation centralisé. Notre axe d'étude porte en particulier sur le matériel courant, librement accessible au plus grand nombre et à un faible coût.

Mots-clés : Capture de trafic, Réseaux sans fil, Wi-Fi, IEEE 802.11, Sécurité, Détection d'intrusions, Géopositionnement en intérieur, Géolocalisation

Laboratoire d'Informatique de l'Université de Franche-Comté,
Antenne de Montbéliard — UFR STGI,
Pôle universitaire du Pays de Montbéliard,
25200 Montbéliard Cedex (France)
Téléphone : +33 (0)3 81 99 47 43 — Télécopie +33 (0)3 81 99 47 91

Study of Wi-Fi traffic interception and positioning in a heterogeneous environment

Abstract: In the context of a strong security policy for wireless networks, it is interesting to identify the localisation of a data source, to ensure that it is authorised and to control the way it uses the network. In this paper, we propose an overview of the traffic capture possibilities, in relation to a centralised positioning system. We are especially working on low-end materials, widely accessible at a low expense.

Key-words: Traffic capture, Wireless networks, Wi-Fi, IEEE 802.11, Security, Intrusion detection, Indoor positioning

Laboratoire d'Informatique de l'Université de Franche-Comté,
Antenne de Montbéliard — UFR STGI,
Pôle universitaire du Pays de Montbéliard,
25200 Montbéliard Cedex (France)
Téléphone : +33 (0)3 81 99 47 43 — Télécopie +33 (0)3 81 99 47 91

1 Introduction

L'utilisation des communications sans fil a ouvert de nouveaux axes de recherche dans le domaine des réseaux informatiques. Le positionnement des terminaux sans fil en communication entre dans cette catégorie. Cette fonctionnalité ouvre plusieurs perspectives telles que l'informatique ubiquitaire, l'amélioration de la supervision du réseau et un meilleur suivi des utilisateurs. De plus, les réseaux sans fil remettent en cause des postulats bien établis, comme la sécurisation d'un réseau, l'identification des équipements ou l'efficacité du contrôle de congestion.

L'interception d'un trafic filaire devient de plus en plus difficile à réaliser grâce à la généralisation de la commutation des réseaux ; à l'inverse l'interception d'un trafic sans fil devient de plus en plus facile à réaliser même si les procédures à suivre se complexifient. En effet, quelques cartes Wi-Fi actuelles proposent des fonctionnalités étendues qui ne s'inscrivent plus dans la norme IEEE. La qualité de cette interception est variable en fonction des modèles. Ainsi, pour compléter une interception de trafic Wi-Fi, il est souhaitable de pouvoir lui associer une position géographique, pour une bonne gestion de l'infrastructure. L'objectif de cet article est de décrire un type d'application d'interception pouvant combiner le suivi des terminaux dans la couverture d'un réseau sans fil.

L'article est structuré en cinq sections principales. La section suivante décrit la solution utilisée pour effectuer les acquisitions de l'interception et du positionnement. La section 3 concerne l'état de l'art du domaine du positionnement en espace fermé s'appuyant sur la technologie Wi-Fi. Les sections 4 et 5 décrivent l'outil de positionnement OWLPS développé au laboratoire et une expérimentation menée grâce à lui. La dernière section retrace, en quelques périodes, l'historique de l'acquisition des données sur les couches physique et liaison des réseaux sans fil.

2 Analyse de données

2.1 Les données

Pour obtenir un positionnement, il est nécessaire d'avoir un ensemble de données exploitables. Que ce soit les puissances de réception, pour un positionnement de type OWLPS [1], ou le temps précis, comme utilisé par les systèmes GNSS, ces donnés doivent être accessibles. Or, les média de communication sans fil actuels (Wi-Fi, Bluetooth) sont centrés sur les données, et même si le matériel traite ces informations, ces dernières ne sont pas systématiquement remontées au système hôte. L'ensemble des normes de communication sans fil permettent d'assurer un service fiable dans la majorité des situations. Pour les constructeurs, et donc la mise en œuvre de la partie matérielle et logicielle, il existe de nombreuses libertés de réalisation. Certes, pour une communication simple, seules les données sont intéressantes, et les normes sont donc axées sur ces dernières, mais les puces pour les communications sans fil ont accès à de nom-

breuses informations qui pourraient être utiles pour des services annexes, voire pour améliorer l'existant. C'est pourquoi nous nous intéressons aux données complémentaires qui sont traitées par les puces, pour leur bon fonctionnement, mais qui ne font pas partie de la norme. On peut citer la puissance du signal, le bruit, les erreurs, les antennes utilisées, le temps de transmissions, la vitesse, la norme utilisée et les erreurs de communications.

2.2 802.11, les données de la couche MAC

La norme 802.11b/g est très complète, et propose une couche d'abstraction importante, avec de nombreuses options, souvent centrées sur les données. La couche MAC nous permet d'avoir des informations sur les participants aux communications, et sur les types de *trames* utilisées. Lors d'une capture sur le réseau, il est très simple d'extraire l'ensemble de ces données de manière manière rapide et fiable, dans le but d'une exploitation ultérieure.

2.3 *radiotap*, les données pilotes

Avec *radiotap* [2], devenu un standard de fait, il est possible d'accéder à beaucoup d'autres informations. Moyennant une configuration spécifique, il devient ainsi possible d'avoir un nouveau périphérique virtuel qui ajoute des informations à chaque paquet de données, sous la forme d'un en-tête *radiotap*. Ainsi, avec une simple capture et une analyse de ces nouveaux champs, il est possible d'extraire des informations intéressantes.

Actuellement, ce sont ces en-têtes *radiotap* qui sont le plus couramment utilisés et supportés. Il est assez intéressant de noter que la diversité des normes n'a pas facilité un support important de la part des divers constructeurs de matériel. C'est grâce à la restructuration actuelle, qui permet de se focaliser sur une seule extension souple et adaptable, *radiotap*, que l'on a accès de plus en plus facilement à ces données. Malgré tout, les informations retournées dépendent grandement de la puce étudiée.

Les données ainsi exploitables nous permettent de connaître l'antenne de réception, le bruit, l'atténuation du signal, le canal, le type de donnée, la qualité de réception, la vitesse de transmission et le temps.

La figure 1 présente un exemple de champs disponibles, exploitables avec *Wireshark* [3].

2.4 Les pilotes, données supplémentaires

Au vu des possibilités diverses et hétéroclites offertes par les pilotes, un utilitaire permettant d'accéder aux paramètres des cartes Wi-Fi a vu jour sous Linux : *iwpriv* [4]. Cette commande permet d'interroger le pilote de la carte pour obtenir ses spécificités, aussi bien que de configurer de manière fine et spécifique ses paramètres. Avec les pilotes ouverts, comme Madwifi [5], nous avons encore accès à des paramètres de configuration supplémentaires.

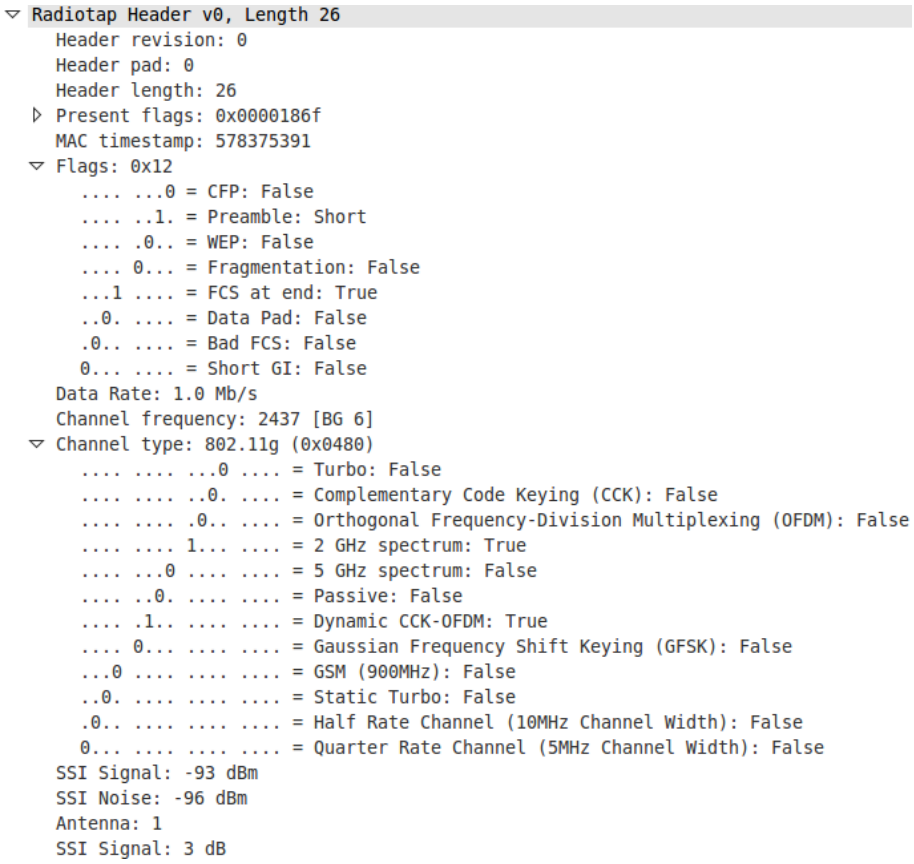


FIGURE 1 – Capture d'écran du logiciel *Wireshark* [3] montrant l'en-tête *radiotap* d'un paquet intercepté.

Ainsi, il est possible d'accéder aux *trames* erronées et de demander au pilote logiciel de remonter ces dernières comme des *trames* valides. Il faudra ensuite, lors du traitement, différencier ces dernières grâce à une vérification du FCS [1]. L'intérêt de l'opération est d'obtenir plus d'information. En effet, même erroné, la majorité du paquet reste souvent correcte et les informations de l'en-tête *radiotap* sont alors exploitables (puissance du signal, canal, etc.), ce qui permet d'améliorer la précision de la localisation, même en cas de fortes interférences. En utilisant la notion de distance et d'historique, déterminer correctement l'adresse MAC, et donc le client émetteur, est donc très probable.

Les pilotes Wi-Fi sous Linux sont actuellement en pleine mutation, et entre ceux requérant un microcode propriétaire (bcm43xxx dans les bornes WRT54GL, ipw2100, ipw2200, ipw3945 centrino, prism54, rt61/73 et ZyDas), ceux n'ayant

1. Le *Frame Check Sequence* permet de vérifier et corriger une trame.

pas l'ensemble des fonctions intégrées (prism54 et Zydas sans WPA, ath5k sans mode Master, b43 très basique), le choix de la puce est primordial. L'utilisation de plusieurs types de cartes peut être intéressante pour les expérimentations, les tests de performances ou le monitoring, mais devient problématique pour le positionnement, cela étant dû aux caractéristiques spécifiques et propres à chaque carte (puissance du signal, qualité du filtrage du bruit, puissance d'émission variable...).

3 État de l'art des systèmes de positionnement

Depuis près d'une dizaine d'années, les communautés scientifiques et industrielles portent un intérêt croissant à des travaux qui concernent le positionnement en utilisant le Wi-Fi. Bien sûr, l'intérêt de se positionner n'est pas nouveau. Sans remonter aux premiers temps de la navigation et des découvertes de nouveaux territoires, le système de positionnement le plus conventionnel aujourd'hui est celui par satellites dénommé GNSS (GPS, Galiléo, GLONASS...). Dans un tel système, et cela reste vrai pour tous les autres systèmes faisant appel à un réseau sans fil, la précision est influencée par le type d'environnement (urbain, rural...), la densité de celui-ci (immeubles, largeur des rues), la précision de l'horloge du récepteur (besoin d'une horloge atomique pour une précision optimale), les corrections sur les erreurs (ionosphérique, troposphérique...), etc. Différents systèmes d'augmentation ont été créés pour réduire l'impact de ces facteurs sur la position. Seulement, lorsque le service de positionnement doit être rendu dans un milieu intérieur, il est rendu plus ardu (réflexion, réfraction, diffraction). Mais s'il existe des solutions nécessitant une intervention lourde sur l'infrastructure (réseaux de capteurs, détections infra-rouge et sonores...), nous assistons à la pénétration de solutions basées sur les signaux Wi-Fi. Il est désormais possible, et nécessaire, de combiner deux (voire trois) systèmes de positionnement pour garantir l'intégrité et la continuité des données et des services.

Les réalisations basées sur les signaux Wi-Fi ont pour but d'améliorer la précision du positionnement en intérieur, mais pas ou plus seulement, et ensuite de proposer des services contextualisés enrichis. Ces travaux peuvent être classés en trois catégories. La première concerne les travaux basés sur la cartographie des puissances de signal tels que les systèmes RADAR [6], HORUS [7] et Ekahau [8]. Dans ces systèmes la précision est liée à la finesse du maillage de la cartographie. La deuxième catégorie est celle des systèmes fondés sur la multilatération utilisant la puissance du signal, et dont le positionnement repose sur le calcul des distances vers des points d'accès connus, tels que SNAP-WPS [9] et les travaux de Interlink Networks [10]. Dans ces systèmes la précision dépend des modèles de propagation utilisés. Pour affiner le choix de la position, on peut également tenir compte du parcours antérieur du mobile, éventuellement en tenant compte de la topologie du bâtiment pour estimer la distance entre deux points. La troisième catégorie est celle de l'hybridation qui emploie les deux, la cartographie du signal et le calcul de la position. L'inconvénient majeur du

calcul de la position concerne le fait de devoir se baser sur des modèles approximatifs de la propagation des ondes. Plus la topologie est hétérogène, plus le résultat du calcul de la distance est éloigné de la réalité.

Les systèmes appartenant à ces trois catégories nécessitent la mesure du signal, soit sur le client mobile, soit sur les points d'accès. On oppose ainsi le système de géolocalisation supporté par une infrastructure à celui où le mobile, autonome, déduit sa position de ses observations de l'environnement, par exemple les systèmes mécaniques à gyroscopes et accéléromètres.

Dans tout système de géolocalisation supporté par une infrastructure à communication bidirectionnelle, il existe deux solutions de positionnement. La première est que le calcul de la position soit effectué par le mobile. Dans ce cas de figure, les éléments de l'infrastructure émettent des informations, que le mobile écoute et dont il se sert pour déduire sa position. Ces informations peuvent tout simplement être les balises Wi-Fi (*beacons*) émises par de simples points d'accès (AP), auquel cas il est très simple d'ajouter des AP pour affiner le calcul de la position (en prenant toutefois garde à éviter le brouillage mutuel des AP) ; la contrepartie est qu'il est nécessaire que le mobile dispose d'un logiciel dédié et d'une liste à jour des AP environnants avec leurs positions.

La seconde possibilité est que le mobile questionne l'infrastructure quant à sa position, et que ce soient les éléments de cette infrastructure qui effectuent le calcul de la position avant de la transmettre en réponse au mobile. Les avantages de cette solution sont multiples. Tout d'abord, le mobile n'a besoin que d'un petit programme lui permettant de contacter l'infrastructure pour lui demander le calcul de sa position. Mais le principal intérêt réside dans la souplesse dont dispose l'infrastructure pour effectuer le calcul. Puisque les « AP » écoutant les demandes de localisation n'ont pas à émettre eux-même, ils peuvent être entièrement passifs et ainsi éviter toute surcharge du réseau [2] ; on peut donc les multiplier autant qu'on le souhaite sans influer sur la qualité de service du réseau sans fil. Les éléments de l'infrastructure peuvent également communiquer facilement, se coordonner, et on peut imaginer que grâce à ce support le système s'adapte aux évolutions de l'environnement. Cette seconde solution offre également la possibilité de traiter toute émission de la part d'un mobile comme une demande de localisation, afin de trouver la position des mobiles qui n'envoient pas de demande de localisation.

Le tableau 1 dresse un comparatif de plusieurs techniques de géopositionnement, fondées sur le réseau Wi-Fi, actuellement publiées dans les travaux de la communauté. Les colonnes *Cartographie* et *Atténuation* décrivent le cœur du fonctionnement du système expérimental, à savoir la façon dont est employée la puissance du signal : pour réaliser une cartographie, pour évaluer la distance en vue d'une multilatération, ou les deux. Les deux colonnes suivantes, *Historique* et *Topologie*, précisent l'utilisation d'éventuelles données complémentaires permettant d'affiner la position calculée et l'estimation des distances. Enfin, les

2. Il ne s'agit donc pas, dans ce cas, de points d'accès au sens strict du terme, puisqu'étant passifs ils ne fournissent pas d'accès au réseau aux mobiles.

Technique [a]	**Carto** [b]	**Atnt** [c]	**Histo** [d]	**Topo** [e]	**Centré** [f]	**Déploiement** [g]
RADAR [6]	×				I	moyen / long [h]
RADAR + VL [11]	×		×		I	moyen / long [h]
Interlink Networks [10]		×			I	court [i]
FBCM [12]		×			M ou I	court [i]
FRBHM Basique [13]	×	×			M ou I	moyen [j]
FRBHM Discret [14]	×	×	×	×	M ou I	moyen [k]
FRBHM Continu [14]	×	×	×	×	M ou I	moyen [k]

a. Nom de publication de la technique ou de l'algorithme de géolocalisation.
b. Indique si la technique utilise ou pas une cartographie des puissances de signal.
c. Indique si la technique utilise ou pas un modèle d'atténuation du signal.
d. Indique si l'historique de parcours du mobile est pris en compte pour le calcul des positions suivantes.
e. Indique si la topologie de la zone de mesure est prise en compte pour calculer la distance entre deux points.
f. Indique si le système expérimental, tel que décrit par les auteurs de la technique, effectue les mesures et les calculs sur le mobile (« M ») ou sur l'infrastructure (« I »).
g. Indique si le déploiement du système est aisé et rapide, ou au contraire long et fastidieux.
h. Le temps de déploiement d'un système fondé uniquement sur la cartographie des puissances dépend du maillage de cette cartographie, dont dépendra la précision obtenue ; à cela s'ajoute, comme pour les autres systèmes, le déploiement des AP.
i. Le déploiement consiste seulement à placer les AP et à déterminer leurs coordonnées.
j. Le déploiement consiste en une cartographie des puissances minimaliste (comme pour RADAR dans le cas d'un maillage large) et au placement des AP avec enregistrement de leurs coordonnées (comme pour Interlink Networks et FBCM).
k. La description de la topologie du bâtiment allonge un peu le temps de déploiement par rapport au FRBHM Basique.

TABLE 1 – Comparatif de techniques de géopositionnement basées sur le réseau Wi-Fi.

deux dernières colonnes, *Centré* et *Déploiement*, donnent des informations plus générales sur le système.

La précision du système RADAR [6], qui utilise une cartographie des puissances seule, est dépendante de la finesse du maillage de la cartographie des puissances réalisée lors du déploiement. Selon la précision souhaitée, il est donc possible de consacrer plus ou moins de temps au déploiement : un maillage d'un mètre est très long à réaliser, tandis qu'un maillage de quatre ou cinq mètres (qui correspond à environ un point par pièce dans un environnement de bureaux) est nettement moins fastidieux. Ce système RADAR apporte une première adaptation, selon la précision recherchée, le temps et les moyens dont on dispose, d'une technique dont le temps de déploiement est variable en offrant une précision plus ou moins bonne. La combinaison d'un tel système avec d'autres techniques (ne nécessitant pas de calibration) et d'autres algorithmes (tirant bénéfice d'un contexte : prédiction, topologie) constituerait une contribution remarquable. Le système RADAR a été étendu par des méthodes probabilistes permettant d'accroitre sa précision : Ekahau [8] considère la distribution de la puissance du signal selon une courbe gaussienne ; HORUS [15] utilise une représentation par

histogrammes. Ces méthodes permettent d'obtenir une meilleure précision que l'usage de la moyenne des mesures de calibration de la cartographie.

D'autres techniques de géopositionnement fondées sur des réseaux sans fil (GSM, Wi-Fi) existent ; la façon de déterminer la position est généralement fondée sur un calcul de la distance par atténuation du signal ou par différentiel de temps (*TdoA*), ou sur les cellules du réseau (l'erreur est dans ce cas dépendante de la taille des cellules). Les travaux d'Interlink Networks [10] fonctionnent sur la base de la puissance du signal, en modifiant la formule de Friis [16]. Le principe est le même dans le cas du SNAP-WPS [9], qui établit une relation entre la puissance du signal et la distance entre l'émetteur et le récepteur ; dans le cas de ce système, la relation est obtenue par régression d'ordre 3 sur des données de calibration. D'autres systèmes utilisent des modes de fonctionnement complètement différents : capteurs infrarouges, ultrasons, gyroscopes et accéléromètres [17], etc.

Enfin, des travaux actuels s'emploient à faire collaborer le positionnement Wi-Fi et le positionnement par satellite, afin d'offrir une continuité de positionnement quel que soit l'endroit où se trouve le mobile, à l'intérieur comme à l'extérieur et une amélioration de l'intégrité.

4 Open Wireless Positioning System

Open Wireless Positioning System (OWLPS) [1] est un système de géolocalisation en intérieur, utilisant comme support le réseau sans fil IEEE 802.11 (Wi-Fi). Il met en œuvre plusieurs techniques de positionnement, toutes fondées sur l'analyse de la puissance des signaux Wi-Fi reçus.

4.1 Architecture

OWLPS est un système dit « centré infrastructure », car le calcul de la position des mobiles est effectué par les éléments de l'infrastructure et non par le mobile lui-même (comme c'est le cas pour un système tel que le GPS).

Cette infrastructure est essentiellement composée de points d'accès (AP) écoutant le réseau afin de repérer les informations susceptibles de permettre le positionnement d'un mobile, à savoir les demandes de localisation envoyées par ce dernier. Ces demandes sont ensuite transmises à un serveur de calcul, chargé de traiter l'information de manière à déterminer la position du mobile. Les AP ne communiquant pas entre eux, ils transmettent les demandes sans se soucier d'un quelconque ordre ; un serveur intermédiaire fait donc l'interface entre les AP et le serveur de calcul afin de présenter à ce dernier l'information de manière cohérente.

Sur le plan matériel, les mobiles peuvent être tout type d'appareil doté d'une interface Wi-Fi (ordinateur portable, téléphone cellulaire, PDA communicant, console de jeux portable...). Il en est de même pour les AP, à la différence que leur interface Wi-Fi doit supporter l'en-tête *radiotap*. Enfin, le serveur d'agréga-

tion et le serveur de calcul peuvent être installés sur une machine plus ou moins puissante selon le nombre de mobiles [3] à tracer.

La figure 2 résume les quatre étapes de la résolution de la position d'un mobile.

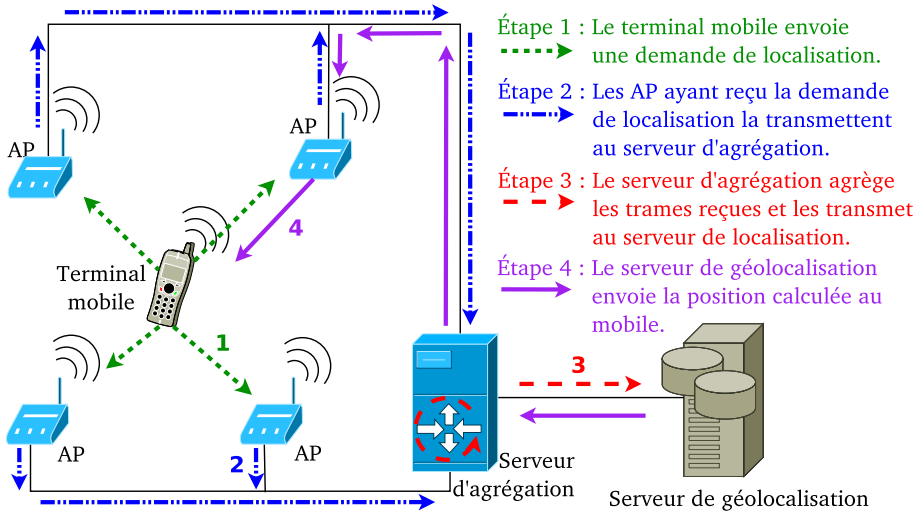


FIGURE 2 – Fonctionnement en quatre étapes du système de géolocalisation centré infrastructure.

4.2 Caractéristiques

OWLPS implante plusieurs techniques de positionnement issues de la littérature (RADAR [6, 11], formule d'Interlink Networks [10]) et de nos propres recherches (FBCM [12], FRBHM [13, 14]). Pour cela, plusieurs caractéristiques ont dû être mises en œuvre : cartographie des puissances du signal, modèles de propagation, prise en compte de la topologie du bâtiment et du parcours du mobile.

La topologie du bâtiment est décrite, lors de la configuration, comme un ensemble de zones homogènes (les pièces) reliées entre elles par des points de passage (les portes et autres ouvertures).

Afin de calibrer le système ou de réaliser une cartographie des puissances du signal, une phase hors ligne est nécessaire, pendant laquelle le mobile envoie, à des positions connues, des requêtes de calibration — qui ne sont autres que des demandes de localisation auxquelles nous ajoutons la position actuelle du mobile. Ces requêtes sont capturées par les AP de la même manière que les demandes de localisation, et les données agrégées sont enregistrées pour êtres utilisées par le serveur de calcul lors de la résolution de la position d'un mobile.

3. Lors de nos expérimentations, une simple machine de bureau assez ancienne a amplement suffi à supporter ces deux services pour un petit nombre de mobiles.

En plus du positionnement des mobiles à la demande, il est possible de considérer chaque paquet émis par ces derniers comme des demandes de localisation. Nous positionnons alors les mobiles de manière systématique, ce qui est intéressant dans le cadre d'une détection d'intrusion.

5 Expérimentations

Nous avons mené une expérimentation visant à positionner un trafic Wi-Fi à l'aide d'OWLPS, au sein d'un bâtiment hétérogène qui héberge notre laboratoire.

5.1 Matériel

Pour les expérimentations réalisées avec le système centré infrastructure, nous avons utilisé cinq mini-PC (processeur Intel Celeron M à 1,50 GHz, 512 Mo de SDRAM, carte Wi-Fi Intel BG2200 dotée d'une antenne de gain 5 dBi ; le système d'exploitation utilisé est Debian GNU/Linux Etch, la version du noyau Linux est 2.6.23.16), un point d'accès Wi-Fi (Linksys WRT54GL), un ordinateur portable (IBM Thinkpad R40, également doté d'une carte Wi-Fi Intel BG2200), et un ordinateur de bureau faisant office de serveur d'agrégation, et exploitant les mesures grâce au serveur de calcul (processeur AMD Athlon 2000+, 1 Go de SDRAM ; le système utilisé est Debian GNU/Linux Lenny).

5.2 Protocole expérimental

Les expérimentations se sont déroulées au rez-de-chaussée et au premier étage de l'aile ouest du bâtiment Numérica, où le LIFC possède ses locaux à Montbéliard. Cette aile mesure 33,50m de long sur 10,30m de large, et comporte deux étages, ainsi qu'un sous-sol. Les dalles de béton et colonnes porteuses ont des épaisseurs variant entre 20cm et 80cm. La plupart des pièces sont des bureaux de 3,60m sur 5m dont les parois extérieures sont entièrement vitrées, alignées du côté ouest et desservies par un couloir faisant toute la longueur du bâtiment côté est ; chaque étage comporte une salle d'eau, des colonnes électriques et d'eau, et deux escaliers. Cet espace est occupé par une trentaine de personnes et est assez passant.

Les cinq mini-PC servant d'AP ont été disposés, deux au rez-de-chaussée (aux deux extrémités du bâtiment), et trois au premier étage (deux disposés dans la longueur, un peu plus resserrés qu'au rez-de-chaussée, et un à l'extérieur, dans une aile perpendiculaire au bâtiment), de manière à former une figure géométrique dans l'espace englobant la majorité des positions potentielles des mobiles.

Nous avons tout d'abord effectué une cartographie des puissances, avec un maillage d'un mètre (une mesure tous les mètres, dans quatre directions correspondant aux quatre points cardinaux), ce qui représente trois à quatre jours de travail (plus de 1200 mesures). Des mesures en mobilité ont ensuite été réalisées,

traçant un déplacement sur les deux étages, en entrant dans plusieurs pièces, et ce à raison d'une mesure par seconde environ, soit un total de 86 points de mesure.

Nous avons pu comparer, pour chaque point de mesure, la position réelle, notée lors du déplacement, à la position calculée par les divers algorithmes, dans des conditions strictement identiques (environnement radio, calibration et cartographie des puissances) puisque le même jeu de mesures est utilisé. Nous avons fait varier le maillage de la cartographie des puissances, d'un mètre à quatre mètres (la distance approximative lorsque l'on ne conserve qu'une seule mesure par pièce, plus une mesure dans le couloir en face de chaque pièce).

5.3 Résultats

Maillage	Propag. [10]	Carto. [6]	Propag. calibré [12]	Carto. + propag. [13]	Carto. + histo. [11]	Carto. + propag. + histo. [14]	Carto. + propag. + histo. [14]
2m (113pts)	11,63	4,48	10,1	4,79	4,52	5,03	5,01
4m (35pts)	11,63	5,03	7	5,94	4,77	5,78	6,07

TABLE 2 – Erreur moyenne (en mètres) du positionnement d'un terminal en mouvement. *(Propag. = modèle de propagation, Carto. = cartographie des puissances du signal, Histo. = historique des positions du mobile.)*

Les principaux résultats obtenus sont présentés dans le tableau 2. Les combinaisons algorithmes et techniques de positionnement offrant les meilleures précisions sont RADAR et RADAR avec Viterbi-like, qui illustrent respectivement la cartographie des puissance du signal seule et combinée avec un historique des positions du mobile. On peut constater que le bénéfice tiré de la mémorisation du parcours antérieur du mobile grâce à un algorithme à la Viterbi n'améliore pas significativement la précision. La série des algorithmes et techniques combinant cartographie et modèle de propagation, et éventuellement un historique des positions du mobile (FRBHM) a une précision légèrement moins bonne, quoique comparable (l'erreur dépasse d'environ un mètre au maximum celle de RADAR) ; les résultats des FRBHM avec historique confirment le fait que l'algorithme à la Viterbi n'apporte pas de précision : dans la plupart des cas, le FRBHM sans historique est plus précis. Enfin, viennent les techniques n'utilisant que l'atténuation du signal (modèle de propagation non calibré ou calibré, respectivement Interlink Networks et FBCM) et la multilatération, qui souffrent d'une erreur beaucoup plus grande ; à noter que l'erreur de la première technique est fixe car elle ne dépend pas du maillage (puisqu'aucune calibration n'est utilisée), tandis que celle du FBCM varie, car il utilise les points de calibration pour modifier la formule d'atténuation du signal employée.

On peut constater que la précision des algorithmes fondés sur une cartographie des puissances ne varie que légèrement en fonction de la finesse du maillage. À noter qu'une étude plus poussée [18] a montré qu'un maillage plus fin (1m)

n'offre pas les meilleurs résultats, et que tous les algorithmes et techniques révèlent leur meilleure précision avec un maillage de deux mètres.

À l'occasion d'évaluations complémentaires [1], nous avons observé le comportement des différentes techniques en faisant varier le nombre d'AP (de 3 à 5). Cela révèle une baisse générale de la précision lorsque l'on diminue le nombre d'AP, sauf pour l'algorithme FBCM.

Ces observations invitent à plusieurs éléments d'analyse. Tout d'abord, pour démontrer la pertinence d'une combinaison (algorithme et technique) telle que la mémorisation de l'historique et son adéquation à un environnement (topologie de bâtiment, exposition à la réflexion, réfraction, absorption en multi-chemin, interférences), il est nécessaire d'introduire un critère de précision topologique, puis de l'appliquer à différents types de bâtiments. Ensuite, il n'existe actuellement pas de modèle d'atténuation satisfaisant en environnement hétérogène et hostile (comme Numérica), alors que dans des espaces clos avec peu ou pas d'obstacles (cloisons, sols, plafonds) des modèles sont efficaces (un rebond sur le sol, rebond sol et plafond). Cela requiert d'évaluer d'autres approches (temps d'arrivée, déphasage). Enfin, la précision n'augmente pas de manière linéaire avec la densité du maillage ; nous pensons que cela peut être imputable à la pérégrination du mobile, les points de calibration, et la prise en compte de la topologie du bâtiment. Toutes ces analyses se doivent d'être validées afin d'aboutir à la compréhension des liens unissant telle approche avec tel environnement.

6 Histoire de l'art de l'acquisition Wi-Fi

6.1 (2000) – La période propriétaire

Initialement, les premiers matériels de communication Wi-Fi étaient composés d'un point d'accès constitué d'un système embarqué propriétaire et d'une carte réseau sans fil, associés à un pilote logiciel et un microcode (*firmware*) embarqué sur la carte. Les accès aux données de la couche physique et de la couche liaison n'étaient pas possibles. En effet, il n'y avait pas d'API pour obtenir des informations de bas niveau. Cependant, à l'aide des logiciels propriétaires livrés par les constructeurs tels que Aironet/Cisco ou Orinoco, l'obtention de statistiques générales concernant la couche physique était possible. Les pilotes de carte Wi-Fi n'existaient pas pour l'environnement UNIX. Il était seulement possible d'imaginer qu'un accès aux données de la couche physique serait disponible assez rapidement du côté des clients. De plus, la plupart des interfaces Wi-Fi en mode client ou station filtrent les paquets en retournant uniquement les paquets de données et en remplaçant l'en-tête 802.11 (Wi-Fi) par un en-tête 802.3 (Ethernet).

6.2 (2002) – La période du mode client libre

Les premiers pilotes libres sont apparus associés aux *chipsets* Prism et Orinoco. Malgré tout, le microcode de la carte Wi-Fi reste très protégé. L'accès

aux données de la couche Wi-Fi est partiellement accessible. Cependant, la plupart des pilotes ne créent que des interfaces réseau de type Ethernet. À partir de ce moment là, les premiers travaux de recherche concernant le positionnement Wi-Fi voient le jour avec une fonction de positionnement intégrée dans les terminaux.

La suite logicielle *wireless-tools*, composée de *iwconfig*, *iwscan*, *iwspy* et *iwpriv*, permet d'obtenir des statistiques générales sur les données de la couche physique en interrogeant le pilote de la carte. *iwscan* et *iwspy* permettent selon le matériel d'obtenir l'atténuation moyenne des émissions Wi-Fi et *iwpriv* configure les paramètres de la carte.

Puis, la première solution permettant d'obtenir des informations détaillées par paquet est apparue sous la forme d'une insertion d'en-tête de type *AVS* ou *PRISM* devant l'en-tête de niveau 2. Cela a ouvert une nouvelle voie où les systèmes de positionnement allaient devenir plus réactifs et plus précis. La contre-partie concerne l'impossibilité de faire fonctionner un applicatif simultanément à cause de l'introduction de cet en-tête non prévu. De ce fait, le mode *monitoring*, qui est un mode complètement passif, a été de plus en plus utilisé. C'est ce mode qui permet en général d'intercepter l'intégralité du trafic Wi-Fi tout en conservant les en-têtes d'origine (802.11). Cependant, en fonction du modèle de carte, les informations sont plus ou moins bien renseignées.

6.3 (2003) – La période du mode point d'accès programmable

Les premières cartes Wi-Fi intégrant le mode *master* (appelé aussi *ap*) dans leur microcode apparaissent. Elles sont généralement associées au logiciel *hostap*. Les premiers ordinateurs intégrant une carte Wi-Fi peuvent à présent servir de point d'accès. Toute la puissance du système d'exploitation GNU/Linux permet donc de transformer un ordinateur en point d'accès réseau programmable. La fonction de positionnement peut s'envisager depuis un élément d'infrastructure. Cette manière de positionner élargit le champ d'applications aux terminaux sans logiciel installé, c'est-à-dire qu'il devient possible de positionner tout type de terminal Wi-Fi qu'il soit intégré dans une architecture ou bien en situation tiers. Cependant, le coût de déploiement de chaque point d'accès est assez élevé. Le mode *monitor* permet également de recevoir le trafic, mais c'est un mode passif qui ne permet pas la double fonction : point d'accès et point de mesure.

Une grande avancée a été effectuée lors de l'apparition des en-têtes *radiotap* dans quelques pilotes. Une des premières intégrations a été effectuée dans le système NetBSD en 2001 (FreeBSD en 2003). Elle a été ensuite portée sous GNU/Linux en 2006. Cette bibliothèque fonctionne de manière stable principalement avec le *chipset* Intel des versions Centrino appelé initialement BG2200 et les *chipsets* Atheros.

6.4 (2004) – La période du point d'accès ouvert

La société Linksys, marque grand public de Cisco, met sur le marché en 2003 une borne Wi-Fi nommée WRT54G (*Wireless Router 54Mbps 802.11g*) intégrant un système GNU/Linux embarqué. Cette borne Wi-Fi dispose d'un processeur de type MIPS. En 2004, les premières versions libres d'un système d'exploitation ouvert apparaissent. Les deux distributions les plus répandues sont DD-WRT et OpenWRT. À l'aide d'OpenWRT, il devient possible d'exécuter du code additionnel à partir d'une suite de cross-compilation. Cependant, l'utilisation de l'extension *radiotap* est difficile à mettre en place, car le microcode de la carte Wi-Fi reste protégé. La société Fon diffuse depuis la fin de l'année 2006 une borne Wi-Fi équipée d'un *chipset* Atheros dont le microcode est beaucoup plus ouvert. La bibliothèque *radiotap* s'intègre facilement et fonctionne de manière stable. Les valeurs indiquées dans les champs de l'en-tête *radiotap* sont correctes.

7 Travaux futurs et perspectives

OWLPS est une plate-forme expérimentale permettant l'évaluation de différentes techniques de géolocalisation dans un espace à trois dimensions ; ces techniques sont mises en situation identique, de manière à obtenir des résultats comparables. Les techniques mises en œuvre sont fondées sur une cartographie des puissances, des modèles d'atténuation de la puissance du signal, un historique dynamique des itinéraires des mobiles, et la topologie du bâtiment. Les plus précises offrent une précision d'environ 5 mètres dans un espace intérieur très hétérogène, en requérant une calibration limitée à un peu plus d'un point de référence par pièce, pour une densité de 5 AP pour 600 m^2 sur deux étages.

Un tel système est approprié au calcul de la position de mobiles situés à l'extérieur des bâtiments, avec un couplage éventuel avec d'autres systèmes de positionnement [19, 20], tels que le GPS ou le positionnement par GSM. Il est de plus possible de détecter les intrusions et de calculer la position [21], dans la zone couverte, de mobiles non autorisés, et ce en écoutant tout le trafic réseau, chaque paquet capturé par les AP représentant alors une demande implicite de positionnement (alors que les demandes de localisation émises par les mobiles autorisés, comme dans le cadre de nos expérimentations, sont des demandes explicites).

Le système OWLPS a encore besoin d'un peu de travail afin d'être aisément utilisable, notamment au niveau du serveur de calcul de la position, qui est encore à un stade expérimental. Mais le gros des développements futurs portera sur l'ajout de fonctionnalités.

Le principal axe des recherches futures est l'auto-calibration du système. Il a en effet été observé que l'environnement intérieur des bâtiments, en plus d'être très hétérogène, évolue beaucoup. En fonction de l'heure de la journée par exemple, le nombre de personnes évolue ; des meubles peuvent être déplacés, de l'eau circuler dans les canalisations ou pas... Autant de facteurs influant sur

le comportement du signal. Les éléments de l'infrastructure pourraient donc échanger périodiquement des informations de signalisation, qui serviraient de base pour corriger le paramétrage du système. Ainsi, si la fréquentation de la zone couverte augmente subitement, le signal se trouvera plus atténué, et les distances seront surestimées ; les éléments de l'infrastructure, en échangeant des messages, pourraient se rendre compte du changement, et appliquer une correction sur leurs estimations.

Ce principe pourrait également être appliqué dans le cadre d'une infrastructure mobile. Si les AP sont amenés à se déplacer ou à être déplacés régulièrement, il serait intéressant que l'infrastructure recalcule la position de l'AP déplacé (position absolue calculée grâce à des points fixes connus, ou relative aux autres AP).

Nous avons également en vue la problématique de la continuité des services dépendants du contexte [14, 22], en intérieur comme en extérieur. Pour cela, de nombreux verrous sont à lever. Tout d'abord, la contextualisation des applications nécessite un géopositionnement dans des environnements divers (extérieurs et intérieurs). Ensuite, un service riche contextuel nécessite de la communication multiple (réseaux d'infrastructure mais aussi *ad hoc*, sur différents supports). En outre, il faut proposer des mécanismes permettant de mettre en œuvre une continuité (prédiction de mobilité, systèmes de cache avec pré-chargement, handover, etc.).

8 Conclusion

Le système de positionnement intérieur OWLPS, que nous développons, fonctionne actuellement à l'aide de requêtes explicites provenant du terminal. C'est l'infrastructure qui calcule la position des terminaux actifs dans l'environnement. Le positionnement implicite, c'est-à-dire sans l'envoi de requêtes du terminal, est la prochaine étape d'intégration. Pour que les points d'accès Wi-Fi puissent être déployés simplement, il sera nécessaire qu'ils puissent tous détecter automatiquement les variations de puissance, qui peuvent provenir de plusieurs facteurs tels que la présence d'une foule ou des transformations dans la structure du bâtiment. Enfin, concernant l'acquisition qui est actuellement traitée isolément pour chaque points d'accès, une approche de données fusionnées permettra de suivre plus efficacement les utilisateurs dans leurs déplacements.

Remerciements

Ce travail est le fruit d'un projet financé par l'ANR, la Région Franche-Comté, la Communauté d'Agglomération du Pays de Montbéliard, et le Pôle Véhicule du Futur.

Références

[1] Matteo Cypriani, Philippe Canalda, Soumaya Zirari, Frédéric Lassabe et François Spies : Open wireless positioning system. Technical Report RT2008-02, LIFC - Laboratoire d'Informatique de l'Université de Franche Comté, décembre 2008.

[2] radiotap : Site officiel. `http://www.radiotap.org/`.

[3] Wireshark : Site officiel. `http://www.wireshark.org/`.

[4] Wireless Tools for Linux : Site officiel. `http://www.hpl.hp.com/personal/Jean_Tourrilhes/Linux/Tools.html`.

[5] Madwifi : Site officiel. `http://www.madwifi-project.org/`.

[6] Paramvir Bahl et Venkata N. Padmanabhan : RADAR : An in-building RF-based user location and tracking system. *In INFOCOM (2)*, pages 775–784, 2000.

[7] M.A. Youssef, A. Agrawala, A.U. Shankar et S.H. Noh : A probabilistic clustering-based indoor location determination system. UM Computer Science Department ; CS-TR-4350, 2002.

[8] R. Roos, P. Myllymäki, H. Tirri, P. Misikangas et J. Sievänen : A Probabilistic Approach to WLAN User Location Estimation. *International Journal of Wireless Information Networks*, 9(3):155–164, juillet 2002.

[9] Y. Wang, X. Jia et H.K Lee : An indoors wireless positioning system based on wireless local area network infrastructure. *In 6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services*, numéro 54, Melbourne, juillet 2003.

[10] Interlink Networks, Inc. : A practical approach to identifying and tracking unauthorized 802.11 cards and access points. Rapport technique, 2002.

[11] P. Bahl, A. Balachandran et V. Padmanabhan : Enhancements to the radar user location and tracking system. Rapport technique, Microsoft Research, 2000.

[12] F. Lassabe, O. Baala, P. Canalda, P. Chatonnay et F. Spies : A Friis-based calibrated model for WiFi terminals positioning. *In Proceedings of IEEE Int. Symp. on a World of Wireless, Mobile and Multimedia Networks*, pages 382–387, Taormina, Italy, juin 2005.

[13] F. Lassabe, D. Charlet, P. Canalda, P. Chatonnay et F. Spies : Refining WiFi indoor positionning renders pertinent deploying location-based multimedia guide. *In Procs of IEEE 20th Int. Conf. on Advanced Information Networking and Applications*, volume 2, pages 126–130, Vienna, Austria, avril 2006.

[14] Frédéric Lassabe : *Géolocalisation et prédiction dans les réseaux Wi-Fi en intérieur*. Thèse de doctorat, École doctorale SPIM, 2009.

[15] Moustafa A. Youssef, Ashok Agrawala, A. Udaya Shankar et Sam H. Noh : A probabilistic clustering-based indoor location determination system. Tech. Report CS-TR-4350, University of Maryland, mars 2002.

[16] L.V. Blake : *Radar Range-Performance Analysis.* Artech House Radar Library, décembre 1986.

[17] S. Moix, C. Steiner, Q. Ladetto et B. Merminod : Capteurs et analyse de signaux pour la navigation pédestre. *MPG*, pages pp. 512–516, août 2002.

[18] Matteo Cypriani, Frédéric Lassabe, Philippe Canalda, Soumaya Zirari et François Spies : Open wireless positioning system : un système de géopositionnement par wi-fi en intérieur. *In* Alexandre Caminada, éditeur : *JDIR'09, 10èmes Journées Doctorales en Informatique et Réseaux - Session Systèmes de localisation*, pages 73–78, Belfort, France, février 2009.

[19] S. Zirari, P. Canalda et F. Spies : Geometric and signal strength dilution of precision (DoP) Wi-Fi. *Int. Journal of Computer Science Issues*, 3:35–44, août 2009.

[20] S. Zirari, P. Canalda et F. Spies : Modelling and emulation of an extended GDOP for hybrid and combined positioning system. *In ENC-GNSS'09, European Navigation Conference - Global Navigation Satellite Systems*, Naples, Italy, mai 2009. 6 pages, CD-ROM publication.

[21] Matteo Cypriani, Philippe Canalda et François Spies : Problématiques de sécurité dans un système de géolocalisation implicite. *In CFIP'09, Colloque Francophone sur l'Ingénierie des Protocoles*, Strasbourg, France, octobre 2009. 2 pages, Poster session.

[22] Soumaya Zirari, Philippe Canalda et François Spies : Critère de dilution de précision pour un positionnement en intérieur et en extérieur. *In CFIP'09, Colloque Francophone sur l'Ingénierie des Protocoles*, Strasbourg, France, octobre 2009. 2 pages, Poster session.

Laboratoire d'Informatique de l'université de Franche-Comté
UFR Sciences et Techniques, 16, route de Gray - 25030 Besançon Cedex (France)

LIFC - Antenne de Belfort : IUT Belfort-Montbéliard, rue Engel Gros, BP 527 - 90016 Belfort Cedex (France)
LIFC - Antenne de Montbéliard : UFR STGI, Pôle universitaire du Pays de Montbéliard - 25200 Montbéliard Cedex (France)

http://lifc.univ-fcomte.fr