

# Géolocalisation WiFi et modèles de prédiction de la mobilité dans les réseaux multimédia

Frédéric Lassabe  
Computer Science Laboratory  
of the University of  
Franche-Comté (LIFC)  
Numérica, 1 cour Louis  
Leprince Ringuet  
Montbéliard, France  
lassabe@lifc.univ-  
fcomte.fr

Philippe Canalda  
LIFC  
Numérica, 1 cour Louis  
Leprince Ringuet  
Montbéliard, France  
philippe.canalda  
@lifc.univ-fcomte.fr

Pascal Chatonnay  
LIFC  
Numérica, 1 cour Louis  
Leprince Ringuet  
Montbéliard, France  
pascal.chatonnay  
@lifc.univ-fcomte.fr

## ABSTRACT

Les capacités toujours plus grandes des réseaux sans-fil couplées à la réduction de l'encombrement des terminaux mobiles rendent possibles de nouvelles applications multimédia et mobiles. Parmi les nombreux problèmes soulevés, la continuité des transmissions de contenus multimédia est compromise par la mobilité des terminaux. Dans cet article, nous proposons deux composants majeurs de la gestion de la mobilité. La géolocalisation permet la détection et l'analyse de la mobilité des terminaux. La prédiction de la mobilité permet d'anticiper les déplacements des terminaux et les pertes de connexion. Grâce à la prédiction de la mobilité, la continuité de service peut être assurée par l'intermédiaire de politiques pro-actives. Nos travaux intègrent des expérimentations ainsi qu'une forte orientation vers le transfert technologique.

The increasing wireless network capacities coupled to the decrease of the mobile terminals' weight allow new multimedia and mobile applications. The problems raised are numerous. One in particular, is the multimedia streaming continuity, which is compromised by the mobility of the terminals. In this article, we propose two major components in the field of mobility management. The positioning allows to detect and analyse the mobility of the terminals. Mobility prediction anticipates the terminals movements and the connection losses. Thanks to mobility prediction, the service continuity is ensured through pro-active policies. Our work includes experimentations and has strong implementation possibilities.

## Keywords

Prédiction de la mobilité, multimédia mobile, applications pervasives, modèles de Markov

## 1. INTRODUCTION

La démocratisation des appareils portables, comme les ordinateurs portables et les téléphones mobiles, contribue à la possibilité de développer des applications pervasives. Les applications pervasives s'intègrent à leur environnement et y sont particulièrement sensibles. La connectivité des appareils portables est de plus en plus élevée, avec l'arrivée des réseaux WiFi en particulier et également grâce à la téléphonie 3G, aux réseaux bluetooth et autres réseaux sans-fil. De plus, les appareils portables sont de plus en plus mobiles, leur poids diminuant et leur autonomie augmentant. Actuellement, de plus en plus de données sont transmises entre et en direction des terminaux mobiles. En particulier, les contenus multimédia (voix, vidéo, musique) sont de plus en plus plébiscités par les usagers de terminaux mobiles.

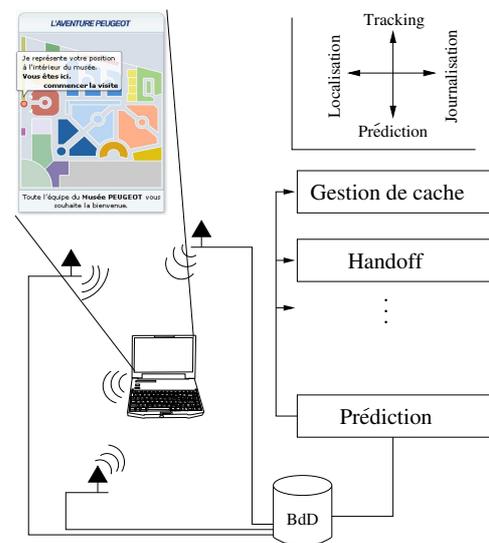


Figure 1: Système de gestion de la mobilité.

Nous nous intéressons beaucoup au WiFi car cette technologie est peu coûteuse, a une portée raisonnable et permet de communiquer à des débits importants. Ces caractéristiques font de WiFi un terrain de développement idéal pour la diffusion de flux multimédia. Cependant, un grand nombre de problèmes se manifestent dans l'utilisation de WiFi dans le domaine des applications multimédia mobiles. En particulier, la continuité des flux de données, assurée dans la technologie GSM, ne l'est pas dans d'autres cas, en partic-

ulier celui du WiFi.

Dans cet articles, nous nous intéressons à deux axes fondamentaux de la gestion de la mobilité : la géolocalisation des terminaux mobiles permet, non seulement de fournir des services adaptés au contexte géographique, mais également de détecter et modéliser la mobilité des terminaux mobiles. Nous nous intéressons uniquement à la géolocalisation en intérieur. La géolocalisation en extérieur est amplement effectuée par le *Global Positioning System* (GPS). Le second axe de travail permis par la géolocalisation est la prédiction de la mobilité. En effet, pour assurer la continuité des services, il faut anticiper les pertes de connexion et appliquer des politiques adaptées pour y remédier. Nous choisissons la prédiction de la mobilité pour anticiper les problèmes de connexion.

La figure 1 décrit l'interopérabilité de notre modèle. Tout d'abord, le terminal mobile est localisé et ses déplacements sont enregistrés. Cette étape de journalisation des déplacements fournit ensuite les données nécessaires à la phase d'apprentissage du modèle de mobilité. Une fois cette phase effectuée, le modèle de mobilité est exploité : on lui fournit des données temps-réel à partir desquelles il déduit, basé sur l'apprentissage, les déplacements futurs.

Dans cet article, nous étudions l'utilisation de la prédiction de la mobilité dans le cadre des réseaux multimédia mobiles. Premièrement, nous traitons la géolocalisation des terminaux mobiles en intérieur, qui est la première phase pour traiter la mobilité. Deuxièmement, nous exposons plusieurs modèles basés sur l'apprentissage afin de prédire les déplacements des terminaux mobiles. La prédiction de la mobilité permet d'anticiper le handoff quand un terminal mobile se déplace d'un point d'accès à un autre. Enfin, nous tirons les conclusions de nos travaux et nous en servons pour déterminer les futurs travaux.

## 2. GÉOLOCALISATION DE TERMINAUX MOBILES

Le premier élément de la gestion de la mobilité est la capacité à déterminer le mouvement d'un terminal mobile. Pour cela, nous nous basons souvent sur sa position, dont les variations permettent la quantification des mouvements. Cette position est déterminée par le processus de géolocalisation, c'est-à-dire la détermination de la localisation géographique du terminal mobile. Dans cette section, nous présentons brièvement les systèmes de géolocalisation de terminaux mobiles, ainsi que nos contributions dans ce domaine.

### 2.1 État de l'art

Dans cette section, nous nous concentrons sur les systèmes de géolocalisation en intérieur. En effet, le positionnement en extérieur est aisément obtenu avec le GPS (*Global Positioning System*) [17]. Au contraire, les systèmes de géolocalisation en intérieur sont actuellement étudiés. Dans un premier temps, nous présentons un système de géolocalisation basé sur des capteurs infrarouges qui permet de déduire le plus simple système de géolocalisation de terminaux WiFi. Dans un second temps, nous présentons des systèmes de géolocalisation basés sur la collecte de données de puissance de signal. Dans un troisième temps, nous présentons le principe de la trilatération et plusieurs systèmes de géolocalisation basés sur la trilatération.

Une méthode triviale pour déterminer la position approchée d'un terminal mobile est basée sur des capteurs infrarouges [24]. Comme un capteur infrarouge a une faible portée et n'est capable de trans-

mettre que dans sa ligne de vue, le placement de tels capteurs à des points clés d'un bâtiment permet de connaître ponctuellement la position d'un terminal mobile. Le même principe est utilisé lorsque l'on approxime la position d'un terminal WiFi par celle de son point d'accès.

Dans le projet RADAR [1], une carte des puissances de signal est utilisée pour localiser les terminaux mobiles. Une base de données contenant des points est construite. Chaque point de la base de données est défini par ses coordonnées géographiques et les puissances de signal mesurées pour chaque point d'accès. Par comparaison d'une mesure quelconque avec les points de la base de données, on détermine la position d'un terminal mobile. L'erreur moyenne de RADAR est comprise entre 2 et 3 mètres. D'autres projets sont basés sur la même méthode. Les approches statistiques [26, 25] utilisent des distributions de probabilité des puissances de signal au lieu d'une valeur moyenne.

Une autre méthode de détermination de la position d'un terminal mobile est la trilatération. Elle permet de calculer la position d'un point en connaissant les distances de ce points à d'autres points de coordonnées déjà connues. La première étape de la trilatération consiste à déterminer la distance entre les points connus et le point à localiser. Les techniques de détermination de la distance se basent sur la relation entre la puissance du signal reçu et la distance par rapport à l'émetteur. L'expression *puissance du signal d'un appareil* est utilisée ici au sens de la puissance du signal qui correspond au signal envoyé par cet appareil et mesurée par un autre appareil défini par le contexte.

Deux projets basés sur la trilatération se distinguent. Le premier, SNAP-WPS [23], se base sur des mesures préalables pour déterminer une relation polynômiale entre la distance et la puissance du signal, grâce à une régression polynômiale sur les données mesurées. La seconde méthode, une approche effectuée par l'entreprise Interlink Networks (IN) [11], se base sur une variante de la relation de Friis [2] :

$$\frac{P_R}{P_T} = G_R G_T \left( \frac{\lambda}{4\pi d} \right)^2$$

Dans un espace libre d'obstacle, la puissance du signal reçu  $P_R$  est inversement proportionnelle au carré de la distance  $d$ . Elle est également dépendante de la longueur d'onde  $\lambda$  du signal, de la puissance d'émission  $P_T$  ainsi que des gains d'antennes du récepteur  $G_R$  et de l'émetteur  $G_T$ . De façon à tenir compte d'une absorption plus importante du signal dans un bâtiment, les auteurs de IN proposent d'élever la distance à la puissance 3,5, valeur basée sur des mesures dans de nombreux bâtiments. L'utilisation de la réciproque de cette formule permet de calculer la distance en fonction de la puissance du signal. SNAP-WPS et le système de géolocalisation de IN ont une précision moyenne de 3 mètres avec, toutefois, des points où l'erreur atteint 8 mètres dans le cas d'IN. En s'appuyant sur plusieurs types d'ondes, la détermination de la distance par un calcul basé sur le temps de propagation d'une onde sonore est possible. Cricket Compass [21] est un tel système, dans lequel une onde radio initialise le compteur de temps permettant de calculer le temps de propagation d'une seconde onde, sonore celle-ci, afin de déterminer la distance par rapport à l'émetteur de ces deux signaux.

### 2.2 Analyse

Nous pouvons décomposer les techniques de l'état de l'art en deux familles : les approches discrètes et les approches continues. L'analyse de l'état de l'art est de plus conduite suivant les quatre paramètres suivants :

- coût (temps/argent/calcul/mémoire) ;
- extensibilité ;
- précision ;
- taux de rafraîchissement.

Tous ces paramètres sont critiques dans l'optique du déploiement d'un système de géolocalisation en intérieur.

Les approches discrètes permettent de déterminer la position d'un terminal mobile parmi un ensemble de positions. Dans le système Active Badge [24], les positions possibles sont définies par les coordonnées des capteurs infrarouges. Active Badge est un système rapide à déployer car il ne requiert pas de calculs complexes ou d'importantes masses de données à stocker. Cependant, la précision est limitée par le nombre et la distance séparant les capteurs infrarouges. Active Badge peut atteindre une précision de quelques mètres en déployant un très grand nombre de capteurs, mais ceci se fait au détriment du temps de déploiement et du coût d'achat de tous les capteurs. La même raison limite l'extensibilité de ce système.

Les systèmes basés sur le référencement de points connus [1, 26, 25] ont une bonne précision, d'environ 3 mètres. Ils sont extensibles mais le temps requis pour mettre en œuvre de tels systèmes croît proportionnellement à la taille de la zone de déploiement du fait de l'augmentation du nombre de points de références nécessaires. Le taux de rafraîchissement dépend principalement de l'architecture applicative ainsi que du nombre de clients. Augmenter le nombre de clients réduit la fréquence de rafraîchissement maximale de mise à jour des coordonnées. Par ailleurs, des mesures conduites par le terminal exposent le système à une baisse de la précision à cause de l'hétérogénéité des matériels. Mais la conduite des mesures par l'infrastructure réseau réduit la fréquence de rafraîchissement maximale du positionnement car elle induit des opérations supplémentaires au sein de l'infrastructure. Les coûts en calcul et en mémoire sont élevés mais se limitent au serveur de positionnement. Ainsi, il est peu coûteux de déployer ce type de système de positionnement de terminaux mobiles.

L'inconvénient le plus remarquable des systèmes de géolocalisation basés sur les points de référence est le temps requis pour effectuer les mesures de puissance du signal à chaque point de référence. De plus, ces systèmes ne sont pas réactifs à des changements de topologie. En effet, abattre un mur ou changer la position d'un point d'accès aura pour effet de changer les puissances des signaux transmis aux points de référence, infirmant alors les calculs du système de positionnement.

Les systèmes de positionnement continus permettent la détermination précise de la position d'un terminal mobile en s'appuyant sur des modèles analytiques. En particulier, la trilatération permet le calcul de la position d'un terminal mobile dans l'espace ou le plan réels. Le système développé par IN est le plus rapide à mettre en œuvre et le plus extensible. Il est juste nécessaire de lancer le programme chargé des mesures et du calcul de la position sur le client pour que ce dernier puisse déterminer sa position. De cette façon, ce système est totalement dynamique, peut être déployé sur n'importe quel réseau WiFi quelque soit sa taille et est bon marché. La consommation de mémoire et le coût en calcul sont faibles grâce à l'utilisation de la trilatération. De son côté, SNAP-WPS est plus long à installer à cause de la collecte de mesures nécessaires à la régression polynômiale. Par conséquent, bien que

sa précision soit légèrement meilleure que celle du système d'IN, SNAP-WPS est légèrement moins extensible et plus coûteux en terme de temps. Ces deux systèmes de géolocalisation ont des fréquences de rafraîchissement élevées, seulement limitées dans la pratique par la fréquence de scan des périphériques WiFi.

### 2.3 Contributions

Dans [14], nous présentons le FBCM (*Friis-based Calibrated Model*) comme extension au modèle proposé par IN. Dans le FBCM, le coefficient appliqué à la distance dans la relation de Friis est calculé selon le bâtiment de mise en œuvre. Cette étape de calibration est effectuée par un petit ensemble de mesures à différents points du bâtiment. Une fois le modèle calibré, la relation de Friis modifiée est plus adaptée au bâtiment. Le manque de précision obtenu avec l'utilisation du FBCM nous a permis de développer un second modèle de géolocalisation : le *FBCM and RADAR-based Hybrid Model* (FRBHM) [13] qui combine la technique des points de référence avec le FBCM.

Le FRBHM requiert la collecte de données concernant les points de référence. Combiner ce modèle avec le FBCM permet d'utiliser moins de points de référence, ce qui réduit le temps de mise en œuvre de la technique. De plus, le FBCM permet de localiser un terminal mobile dans l'espace ou le plan, sans se limiter aux points de référence. Le principe de la combinaison des deux méthodes est le suivant : les points de référence permettent de restreindre l'espace de recherche au voisinage d'un point, le plus proche en regard des mesures de puissance des signaux reçus. Les points de référence inclus dans cet espace restreint sont utilisés pour calibrer le FBCM et calculer plus finement la position du terminal mobile.

Dans l'application du FBCM, la trilatération est effectuée par un algorithme itératif. En effet, avec l'augmentation du nombre de points d'accès et le fait que les distances calculées par le FBCM sont rarement exactes, le calcul de la position est complexe. L'algorithme parcourt le plan ou l'espace selon un pas prédéfini. Pour chaque point du parcours, on considère les cercles centrés sur les points d'accès et de rayon égal à la distance (obtenue par le FBCM) séparant le point d'accès du terminal mobile. On calcule la distance entre chaque point du parcours et le périmètre de chaque cercle. La plus grande des distances est comparée pour chaque point. Le point qui minimise cette distance est sélectionné comme localisation du terminal mobile. Le FRBHM est applicable à un réseau WiFi en intérieur. Il se base sur du matériel simple et bon marché et requiert moins de calibration que RADAR tout en offrant une précision supérieure.

### 2.4 Expérimentations

Pour juger nos contributions, nous avons mis en œuvre les modèles de IN, SNAP-WPS, une implémentation de RADAR, le FBCM et le FRBHM. Ces tests ont eu lieu dans nos bureaux du centre de développement multimédia à Montbéliard. Tous les tests sont effectués suivant le même procédé. La mesure des puissances reçues en un point est faite. Le point a des coordonnées connues, également transmises au système de géolocalisation. La mesure est transmise au système de géolocalisation qui calcule la position du terminal mobile. Il renvoie pour chaque essai les coordonnées qu'il a déterminées ainsi que l'erreur commise. L'erreur est la distance euclidienne entre la solution du système de géolocalisation et les vraies coordonnées du point de test. Les résultats obtenus sont présentés dans le tableau 1.

Les résultats obtenus tranchent clairement en faveur du FRBHM.

Système	Erreur moyenne (m)	Écart-type (m)
Interlink Networks	29.38	12.17
SNAP-WPS	22.78	14.07
FBCM	15.86	9.34
RADAR	4.32	2.23
FRBHM	1.07	0.32

**Table 1: Précision des modèles de géolocalisation.**

L'analyse des résultats souligne la précision des modèles basés sur les points de référence. En effet, alors que RADAR et le FRBHM exhibent des erreurs inférieures à 5 mètres, le FBCM obtient une erreur moyenne de 15 mètres, alors qu'il est le modèle basé sur la trilatération le plus précis. Son imprécision est un critère important qui nous a poussés à développer le FRBHM. En effet, le FBCM, ainsi que le modèle d'IN et le SNAP-WPS, ne présentent pas une précision suffisante pour être exploités dans le cadre d'applications pervasives. Au contraire, l'erreur moyenne de l'ordre du mètre obtenue avec le FRBHM, permet d'envisager le déploiement d'applications "conscientes" du contexte, comme la diffusion de flux multimédia dans le cadre d'un guide numérique de musée.

### 3. PRÉDICTION DE LA MOBILITÉ

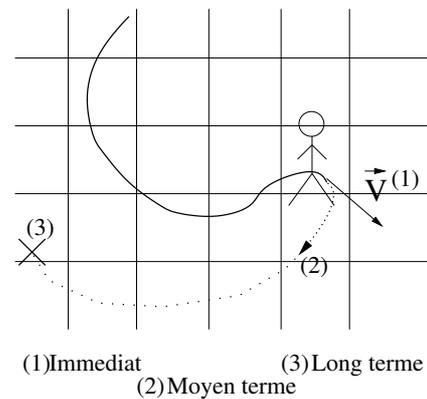
Nous avons choisi la prédiction de la mobilité comme support pour la gestion pro-active du handoff et de la continuité des services dans le cadre des réseaux mobiles. Ayant résolu la géolocalisation qui fournit les données spatio-temporelles nécessaires à l'analyse des mouvements des terminaux, nous avons étudié la prédiction de la mobilité.

#### 3.1 État de l'art

Alors que peu de recherches ont été menées dans le cadre de la prédiction de la mobilité, beaucoup de travaux ont été faits dans le domaine du préchargement de pages web pour améliorer la latence de l'Internet. Ces travaux sont intéressants car ils peuvent être étendus à la prédiction de la mobilité. En fait, les modèles formels tels que les chaînes de Markov (*Markov Model*, MM) peuvent être utilisés pour modéliser de nombreuses données stochastiques, des utilisateurs navigant sur le web aux terminaux mobiles se déplaçant de cellule en cellule. Dans cette section, nous présentons les travaux relatifs à la prédiction. Comme tous les travaux ne sont pas initialement dédiés à la prédiction de la mobilité, nous qualifions leur objectif de *prédiction du comportement*. Les travaux sont classifiés suivant deux grands axes : la prédiction à court terme, qui ne nous concerne pas réellement dans le cadre de la gestion de la mobilité dans les réseaux multimédia mobiles et l'anticipation du handoff. Cependant, de tels travaux s'avèrent intéressants en ce qui concerne le routage dans les réseaux ad hoc. La seconde classe englobe les travaux sur la prédiction à moyen terme, horizon temporel très intéressant dans le cadre de l'anticipation du handoff. Il existe un troisième horizon temporel, non traité ici : le long terme. Il permet d'exploiter les patrons cycliques (par exemple, les individus rentrent chez eux chaque soir). Les trois horizons temporels sont illustrés dans la figure 2.

##### *Prédiction immédiate*

La prédiction de la mobilité peut être obtenue à l'aide des équations de la trajectoire du terminal mobile [15]. Avec les équations de trajectoire, on peut déterminer la position d'un terminal mobile dans un futur proche en calculant son vecteur de vitesse. Seul le court terme peut se prédire ainsi car le terminal mobile est susceptible



**Figure 2: Illustration des horizons temporels de la prédiction de la mobilité.**

de changer de trajectoire à tout instant, ce qui nécessite une mise à jour des équations de trajectoire. Comme le positionnement des terminaux mobiles est rarement exact, diverses méthodes permettent de lisser les données de positionnement pour obtenir une trajectoire plus précise. Ces méthodes sont par exemple les filtres de Kalman [6] et le *Double Exponential Smoothing* [12]. Cependant, ces méthodes requièrent que le positionnement soit précis et suffisamment fréquent pour en déduire l'équation de la trajectoire du terminal mobile. Elles sont particulièrement utiles dans les systèmes de guidage de missiles mais sont moins applicables dans le cas de la gestion de la mobilité des terminaux.

Dans [16], les auteurs présentent un modèle hiérarchique de mobilité. Le modèle local de mobilité utilise le calcul d'équations de trajectoire pour effectuer une prédiction à court terme. Le modèle global de mobilité considère les patrons de mobilité des usagers (*User Mobility Pattern*, UMP). La prédiction est effectuée par la corrélation entre le chemin actuel de l'utilisateur (*User's Actual Path*, UAP) et les UMP stockés. La combinaison des modèles local et global permet de déterminer la cellule de présence du terminal mobile suivante avec précision et corréler les UAP avec les UMP plus facilement.

##### *Court et moyen termes*

Sarukkai [22] propose de modéliser les pérégrinations des usagers du Web par des MM. Il envisage 4 objectifs atteignables par la modélisation des pérégrinations des usagers : la génération d'un parcours sur un site web, la prédiction des accès aux pages web afin de les précharger pour améliorer la latence du WWW, conseiller des liens en accord avec le parcours de l'utilisateur et identifier les "points névralgiques" d'un site web. La seconde application est intéressante dans le cadre de la prédiction de la mobilité. Le MM décrit par Sarukkai est défini par  $(S, A, \lambda)$  où  $S$  est l'espace des états (URL, requête HTTP, action),  $A$  est la matrice probabiliste des transitions et  $\lambda$  est la distribution des états initiaux. L'apprentissage du modèle est basé sur des fichiers de journalisation qui sont analysés pour en extraire les différentes sessions utilisateur. La prédiction est effectuée en choisissant la transition la plus probable issue de la page actuellement visitée. Une précision de 60 à 70% est observée lors des expérimentations sur le MM. Bien que cette précision puisse être qualifiée de bonne, prendre en compte un historique plus long peut améliorer la précision globale d'un modèle prédictif basé sur les MM.

Pirolli et al. [19] étendent la notion de MM pour modéliser des pérégrinations plus longues à l'aide de N-grammes de tailles variées. Un N-gramme est un t-uplet  $(X_1, X_2, \dots, X_N)$  où  $X_i$  est une page web. Les N-grammes extraits des fichiers de journalisation permettent de construire un MM dont les états sont étiquetés avec  $k = N - 1$  pages précédemment visitées par un utilisateur. La  $N^{ième}$  page du N-gramme est utilisée pour l'apprentissage du modèle. Ce type de MM est appelé  $K^{th}$ -order Markov Model (KMM, modèle de Markov d'ordre K). Plus l'ordre du MM est élevé, moins il y a de chances de trouver un état observé parmi les données d'apprentissage. Dans [20], Pitkow et al. établissent un nouveau modèle basé sur le KMM. C'est le All  $K^{th}$ -order Markov Model (AKMM). Ce modèle consiste à construire les KMM d'ordre 1 à K. La prédiction est effectuée par le KMM de plus haut degré pour lequel l'état observé existe dans le modèle. Ainsi, le AKMM pallie les limites du KMM tout en permettant cependant une prédiction suivant le MM d'ordre le plus élevé. Malgré leur efficacité pour effectuer une prédiction, les AKMM consomment beaucoup de mémoire.

Dans [20], J. Pitkow et al. décrivent comment extraire les *longest repeating sequences* (LRS, plus longue séquences répétées) des pérégrinations des utilisateurs. La méthode consiste à identifier dans les pérégrinations les sous-séquences composées d'un ensemble d'entrées consécutives, répétées au moins 2 fois et qui, au moins une fois, ne sont pas incluses dans une LRS plus longue. Extraire les LRS permet de réduire de manière significative la quantité de données avant de construire le AKMM, tout en conservant la même précision lors de la prédiction. En réalité, la précision décroît légèrement mais de façon négligeable eut égard au gain en mémoire.

Dans [9], les auteurs construisent le AKMM complet puis suppriment les états peu utiles. Les états peu utiles sont déterminés au cours de la phase dite de validation. La validation utilise une partie des données d'apprentissage réservées à la validation, donc non utilisées pour l'apprentissage. En vérifiant la précision obtenue individuellement par chaque état, leurs occurrences et la répartition des probabilités des transitions sortantes, il est possible de quantifier la fiabilité et la pertinence de chaque état pour prendre la décision de le garder ou non. Ce modèle est appelé *Selective Markov Model* (SMM, modèle de Markov sélectif). Le SMM est une approche similaire au LRS, agissant cette fois en aval de la création du AKMM. Cela permet de beaucoup réduire la taille du modèle tout en conservant une précision proche de celle du AKMM complet.

Une approche proche de celle des MM est basée sur une *Hypertext Probabilistic Grammar* (HPG, grammaire probabiliste d'hypertexte) [3]. Dans une HPG, les pages sont les symboles non terminaux, les états  $S$  et  $F$  (début et fin d'une séquence de pages) sont les symboles terminaux et les règles de production sont les liens entre les pages. Dans la distribution initiale, les pages qui n'ont encore jamais été les premières dans les sessions d'apprentissage ont cependant une probabilité non nulle dans le modèle.

La première représentation formelle des HPG est l'automate. Dans [4], les auteurs des HPG étendent ce modèle en le représentant avec un MM. Le MM est utilisé pour détecter des patrons. Les patrons sont des sous-séquences dont la probabilité d'apparaître est supérieure à un seuil déterminé en fonction des objectifs. Suivant le seuil, un arbre contenant les patrons est construit. Cette méthode est utile pour déterminer des patrons fréquemment em-

ployés et peut être utilisée comme une base à la prédiction de la mobilité en permettant de ne stocker que les données suffisamment fréquentes pour avec un sens statistique.

Dans [5], les précisions des MM d'ordre 1 et 2 sont comparées afin de déterminer si un état du MM d'ordre 1 doit être cloné pour augmenter la précision du modèle. L'objectif est d'augmenter la précision du MM d'ordre 1 tout en utilisant moins de mémoire qu'un AKMM d'ordre 2. La décision de cloner un état est basée sur la notion d'*état précis*. Un état est dit précis si ses transitions sortantes ont des probabilités proches des transitions équivalentes dans le MM d'ordre 2. L'objectif est de cloner des états afin de virtuellement recréer un historique plus long dans les cas où l'historique conditionne des probabilités différentes.

Dans le cadre de la prédiction de la mobilité, les *Hidden Markov Models* (HMM, modèle de Markov à états cachés) [10] sont également utilisés. Un HMM est un processus doublement stochastique. Un des processus stochastique détermine les probabilités de transition d'un état à l'autre du modèle, comme dans un MM. Le second processus stochastique détermine la distribution de probabilité qu'une observation du système corresponde à un état du système. En effet, dans de nombreux cas, les observations du système ne correspondent pas exactement à l'état réel du système, notamment à cause des imprécisions de mesure. Dans le cas de la mobilité, lorsque la position d'un terminal mobile est déterminée, elle comporte une erreur plus ou moins grande. L'objectif du HMM est alors de pallier cette erreur pour accroître la précision de la prédiction. Le HMM étudié a un bon pouvoir prédictif mais ne prend pas en compte un historique de plus de 1 état. La prédiction à court ou moyen terme nous intéresse beaucoup car elle est bien adaptée aux politiques prédictives de handoff et de gestion des caches multimédia.

## 3.2 Analyse

L'analyse des travaux de l'état de l'art permet de souligner plusieurs points. Premièrement, pourquoi chercher à prendre en compte un historique plus long ? L'exemple qui suit prouve l'utilité de cette approche : soient les états  $A, B, C, D, E, F, G, H$  et les pérégrinations  $ABCD, ABCE$  et  $EBCD$ . Si l'on observe un début de pérégrination de la forme  $ABC$ , quel sera le prochain état ? Les MM d'ordre 1 et 2 calculent une probabilité de 33% d'aller en  $E$  et de 67% d'aller en  $D$ . En considérant les pérégrinations, on remarque cependant qu'un utilisateur ayant suivi le chemin  $ABC$  a 50% de chances d'aller en  $D$  et 50% de chances d'aller en  $E$ . Ces probabilités sont données par le MM d'ordre 3. D'un autre côté, un historique long n'est pas toujours précis et peut mener à des erreurs. Dans tous les cas, plus l'historique est long, plus il est difficile de retrouver un état observé dans le modèle. En effet, augmenter l'historique a pour effet d'augmenter de façon très importante (de l'ordre de  $O(n^k)$ , avec  $n$  le nombre d'états réels et  $k$  la profondeur d'historique) le nombre d'états du MM.

Considérant un AKMM d'ordre 3, l'apprentissage avec les pérégrinations  $ABCD, ABCE, ABCF$  et  $HBCG$  produit les MM d'ordre 1, 2 et 3 présentés dans la figure 3.

## 3.3 Contributions

Nos contributions se basent sur les KMM et AKMM. Nous nommons notre premier modèle le *K-past*. Il consiste en un KMM, construit comme exprimé dans l'état de l'art, pour lequel la prédiction est effectuée suivant un seuil  $\sigma$ . La prédiction est effectuée en considérant l'état actuel du terminal mobile dans le modèle, que nous

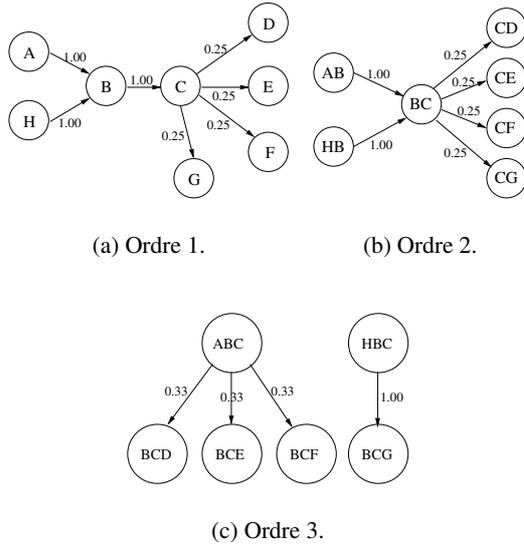


Figure 3: Modèles de Markov d'ordre 1 à 3.

qualifions d'*état courant*. Pour cet état, on étudie les probabilités de toutes les transitions sortantes. Le seuil  $\sigma$  est compris entre 0 et 1. Il s'agit d'une probabilité minimale pour qu'un état soit intégré dans la prédiction. Un état  $E_i$  est inclus dans l'ensemble de prédiction, c'est-à-dire l'ensemble des états pour lesquels une politique sera appliquée, si la probabilité de transition de l'état courant à l'état  $E_i$  est supérieure ou égale à  $\sigma$ .

Le modèle K-past souffre d'inconvénients : les problèmes constatés dans l'analyse de l'état de l'art lui sont totalement applicables. En particulier, un K-past avec  $K$  petit (typiquement égal à 1) ne considère pas avec suffisamment d'importance l'historique des déplacements des utilisateurs. Inversement, quand on augmente  $K$ , il est de plus en plus difficile de trouver un état courant dans le modèle, et les prédictions sont de plus en plus rares. Ainsi, la précision globale est réduite.

Pour pallier ces problèmes, nous avons développé le *K-to-1 past* basé sur les AKMM. Il s'agit de la généralisation des k-past avec  $1 \leq k \leq K$ . On commence la prédiction pour le modèle k-past tel que  $k = K$ . Si l'on peut trouver l'état courant dans le modèle, la prédiction est effectuée de la même manière que dans K-past, suivant le seuil  $\sigma$ . K-to-1 past permet de supprimer un des inconvénients de K-past. Cependant, l'historique le plus long n'est pas toujours le meilleur. En voici la démonstration : considérons  $\sigma = 0, 2$ , l'apprentissage de la figure 3 et une nouvelle pégrination  $ABC$ , puis  $G$  (l'état suivant, à prédire). L'algorithme de K-to-1 past cherche d'abord l'état  $ABC$  dans le modèle 3-past. Dans le cas présent, l'état  $ABC$  existe dans le modèle. Il comporte trois transitions sortantes, chacune avec une probabilité de 0,33. Les trois états pointés par les transitions,  $D$ ,  $E$  et  $F$ , sont retournés. Malheureusement, le terminal mobile va en  $G$  donc la prédiction est fautive. De la même façon, la prédiction avec le modèle 2-past retournerait les états  $D$ ,  $E$ ,  $F$  et  $G$ , et la politique de gestion de la mobilité serait un succès.

Pour pallier ce problème, nous avons développé le modèle K-to-

1 past\*. Ce modèle fonctionne sur la base de K-to-1 past. La différence est que l'ensemble de prédiction final est constitué de l'union des ensembles de prédiction des modèles k-past pour  $1 \leq k \leq K$  :

$$S^* = \bigcup_{i=1}^K S_i$$

où  $S^*$  est la solution renvoyée par le modèle K-to-1 past\* et  $S_i$  est la solution renvoyée par le modèle *K-to-1 past*. De cette façon, les états courants sont toujours trouvés au moins dans le modèle 1-past, et pouvoir trouver l'état courant dans un k-past avec  $k \geq 1$  permet de tenir compte de l'historique le plus long possible. De plus, l'union des ensembles de prédiction de chaque k-past permet de pallier les erreurs éventuelles des k-past pour les  $k$  les plus grands.

Les probabilités de transitions des KMM utilisés par le modèle K-to-1 past\* peuvent être étendues à l'utilisation des seuils de préchargement dans une infrastructure de caches répartis [7]. Dans ce cas, la probabilité est confrontée à deux seuils. L'un déclenche un préchargement complet de la séquence multimédia dans le cache. L'autre seuil déclenche un préchargement partiel de la séquence. Ainsi, même si le contenu multimédia est dégradé, il est possible d'assurer la continuité du flux avec plus de certitude tout en réduisant le trafic engendré sur le réseau.

### 3.4 Expérimentations

Dans cette section, nous présentons les expérimentations que nous avons menées sur les modèles K-past, K-to-1 past et K-to-1 past\*. Les tests ont été réalisés sur deux jeux de données réelles, fournies par l'**Institut fur Pervasive Computing**<sup>1</sup>. Le premier jeu de données est l'*Augsburg Indoor Location Tracking Benchmarks*. Les informations complètes sur ces données sont disponibles dans le rapport technique [18]. Le second jeu de données, le *Nokia Context Data*, est basé sur des données GSM incluant les dates et cellules de présence d'un utilisateur. Les déplacements de cet utilisateur ont été enregistrées en plusieurs sessions.

Log file	(learning/validation percent)	1-past
a_fall	(75/25)	0.7841
b_fall	(75/25)	0.8131
c_fall	(75/25)	0.6592
d_fall	(75/25)	0.7127
a_summer	(75/25)	0.5000
b_summer	(75/25)	0.7857
c_summer	(75/25)	0.7375
d_summer	(75/25)	0.3333
Overall	a_fall to c_fall (100/0) d_fall (94/6), *_summer (0/100)	0.6470

Table 2: Précision de la prédiction de la mobilité, MM d'ordre 1, fichiers d'Augsburg.

Dans les Augsburg Indoor Location Tracking, il y a deux séries de fichiers de journalisation. Les \*\_fall ont été constitués en automne et les \*\_summer ont été constitué durant l'été. Les préfixes  $a$ ,  $b$ ,  $c$  et  $d$  identifient les quatre utilisateurs différents. Le tableau 2 présente les résultats des tests du MM d'ordre 1 sur les données des Augsburg Indoor Location Tracking. Les tests consistent en des applications du MM sur les fichiers individuellement considéré et une application du MM sur la totalité des données agrégées. Le test sur les données agrégées est effectué en concaténant les fichiers ensembles. Pour chaque test, l'apprentissage est fait avec les premiers

<sup>1</sup>[http://www.pervasive.jku.at/research/context\\_database/index.php](http://www.pervasive.jku.at/research/context_database/index.php)

75% du fichier de journalisation. Les 25% restant sont utilisés pour mettre à l'épreuve le modèle. On voit dans les résultats que deux fichiers ont une précision inférieure ou égale à 50%. Dans les deux cas, le résultat peut s'expliquer par la petite taille des fichiers concernés. En effet, si le fichier contient peu d'informations, l'apprentissage contiendra également peu d'informations et ne sera pas assez exhaustif pour la prédiction de la mobilité. Globalement, les résultats sont bons car les tests et l'apprentissage sont faits avec les données du même utilisateur. Ceci prouve l'influence du contexte dans le cadre de la prédiction de la mobilité. La précision la plus basse observée pour le test global est normale. En effet, la concaténation des données en un seul fichier conduit à un apprentissage sur les pérégrinations automnales des utilisateurs et leur confrontation avec les données estivales. Comme les utilisateurs peuvent présenter des habitudes différentes selon la saison, la précision est inférieure. Les résultats de K-past, K-to-1-past et K-to-1-past\* sont exposés dans les figures 4 pour les données de Nokia et 5 pour celles d'Augsburg. La précision est le quotient des transitions pour lesquelles une prédiction a été effectuée avec succès par le nombre de transitions pour lesquelles on a tenté de prédire la position future.

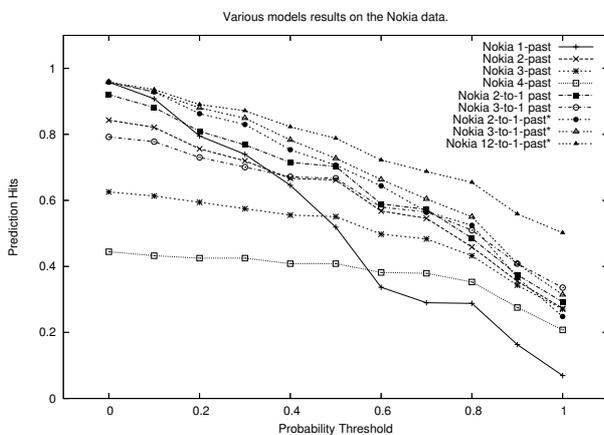


Figure 4: Résultat de la prédiction de la mobilité, données de Nokia.

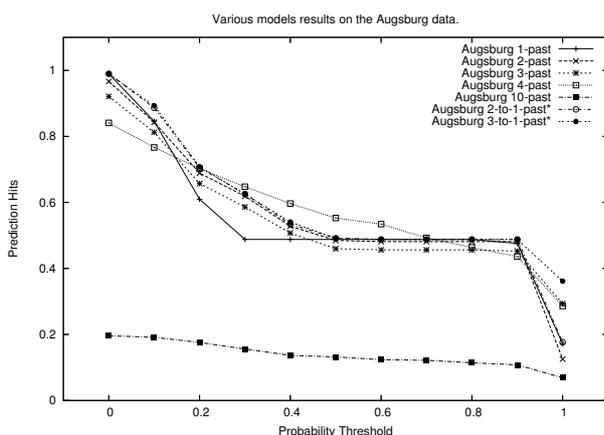


Figure 5: Résultat de la prédiction de la mobilité, données d'Augsburg.

Plusieurs conclusions peuvent être tirées de ces tests. Plus le seuil  $\sigma$  est bas, plus la précision est haute. En effet, diminuer le seuil augmente le nombre d'états sélectionnés pour faire partie de l'ensemble de prédiction, ce qui a pour effet d'augmenter la précision. En dessous d'un seuil, variable selon les données, le modèle 1-past est meilleur que les autres modèles. Comme l'apprentissage de 1-past contient toutes les transitions possibles pour un état donné, celles-ci ont globalement une probabilité inférieure aux transitions des modèles d'ordre supérieur à 1. Ceci explique que 1-past a une précision inférieure pour des seuils élevés. Quand le seuil tend vers 0, sa précision dépasse celles des autres modèles car les transitions sélectionnées sont plus nombreuses.

On remarque également le modèle K-past a une précision dégradée à mesure que  $K$  augmente. En effet, l'augmentation de  $K$  conduit à l'augmentation de la longueur du chemin stocké dans les états du MM. Plus le chemin est long, moins il est probable d'observer le même dans des pérégrinations futures, surtout quand la quantité de données d'apprentissage est faible. Il pourrait y avoir cependant des cas pour lesquels la précision augmenterait en même temps que la profondeur d'historique, mais ces cas ne sont pas généraux.

Augmenter l'ordre du modèle K-to-1 past n'améliore pas toujours la précision. Connaissant les défauts de K-to-1 past (cf. section 3.2), ce n'est guère surprenant. Nous observons que le modèle K-to-1 past a une meilleure précision que le modèle K-past. C'est un résultat attendu car le modèle K-to-1 past pallie la réduction de précision de K-past lors de l'augmentation du degré du modèle. En effet, K-to-1 past sélectionne *au moins* autant d'états que K-past.

Lorsque l'ordre du modèle K-to-1 past\* augmente, la précision augmente. Par exemple, dans le cas de 3-to-1 past\*, on note une amélioration de la précision d'environ 14% comparée à la précision de 3-to-1 past. Cependant, cette augmentation est faite au détriment de la bande passante du réseau car chaque action de handoff a un coût en terme de messages transmis.

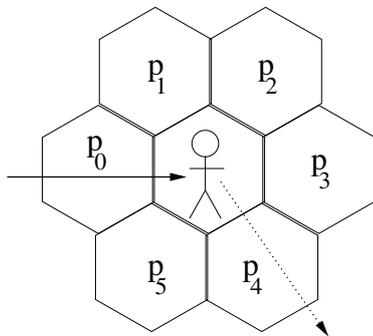
Pour les mêmes raisons que dans le test de K-past, la meilleure précision est atteinte pour un seuil  $\sigma$  égal à 0. On remarque également que, pour  $\sigma = 0$ , la précision du modèle K-to-1 past\* est inchangée quelque soit  $K$ . Comme expliqué auparavant, le MM d'ordre 1 contient toutes les transitions des MM d'ordres supérieurs avec des probabilités faibles. Ainsi, toute transition d'un MM d'ordre supérieur à 1 existe dans le MM d'ordre 1. Lorsque le seuil est égal à 0, le MM d'ordre 1 sélectionne toutes les transitions qui existent dans les MM d'ordres supérieurs à 1, donc sa précision détermine la précision atteinte celle des plus hauts degrés.

#### 4. CONCLUSIONS ET PERSPECTIVES

Dans cet article, nous avons exposé deux éléments fondamentaux de la gestion de la mobilité dans les réseaux sans-fil. Dans un premier temps, nous avons présenté nos travaux passés dans le domaine de la géolocalisation en intérieur des terminaux mobiles. Les résultats que nous avons présentés, en particulier concernant le FRBHM, prouvent la faisabilité de la géolocalisation en intérieur des terminaux mobiles. De plus, pouvoir géolocaliser un terminal mobile au mètre près permet d'envisager la conception d'applications pervasives. En particulier, notre objectif final est d'assurer la continuité des services dans des réseaux mobiles. Cet objectif passe par la modélisation de la mobilité des terminaux, qui s'appuie sur leur détection et localisation ainsi que sur la prédiction de leurs futurs déplacements.

Le point principal de cet article est la modélisation de la mobilité afin de pouvoir prédire les déplacements futurs des terminaux mobiles. Nous nous sommes pour cela appuyés sur les modèles de Markov. Les modèles de Markov permettent de modéliser des systèmes dans lesquels on peut identifier des états, comme la présence d'un terminal mobile dans une cellule d'un réseau sans-fil. Nous avons étudié le comportement des modèles de Markov dans le cas de la prédiction de la mobilité et avons proposé les modèles K-past, K-to-1 past et K-to-1 past\*. Les modèles K-past et K-to-1 past ont tous deux révélé le potentiel des MM dans le cadre de la prédiction de la mobilité. Le modèle K-to-1 past\* quant à lui a permis de valider l'utilisation des MM dans ce cadre.

Modéliser la mobilité et détecter la présence des terminaux mobiles rend possible la diffusion de contenus multimédia à des terminaux mobiles. Nos modèles permettent de communiquer à un système de gestion de caches multimédia [8], les informations nécessaires pour assurer la continuité des flux transmis. Dans la figure 6, l'utilisateur est situé dans la cellule centrale. Il va se diriger vers le sud-est. Les  $p_i$  sont les probabilités de passage de la cellule centrale à chacune de ses voisines. Selon les probabilités, soit rien n'est fait, soit la séquence multimédia est préchargée, totalement ou partiellement.



**Figure 6: Prédiction de la mobilité et gestion des caches multimédia.**

L'utilisation des MM dans le domaine de la gestion de la mobilité étant pertinente, des problèmes restent à résoudre. En particulier, la composante temporelle de la mobilité n'est pas encore prise en compte. Nos futurs travaux devront porter sur cet aspect de la mobilité. En effet, les politiques de gestion pro-active se doivent de tenir compte du temps. Nous avons l'intention d'étudier les lois qui régissent le temps des transitions dans les réseaux mobiles afin d'en déduire un modèle statistique adapté à la prédiction de la mobilité. Nous devons également résoudre le problème soulevé par le coût de l'application d'une politique dans plusieurs états potentiellement futurs. Ce coût devra être quantifié et réduit.

À plus long terme, le challenge est la proposition d'une structure applicative de gestion de la mobilité qui poserait non seulement les méthodes de géolocalisation et de prédiction de la mobilité, mais également les protocoles de communication entre ces différents concepts. De plus, nous avons actuellement travaillé sur des réseaux WiFi, et étendre nos travaux à d'autres supports semble également important. Ainsi, les domaines d'applications ciblés seraient étendus à tous les types de terminaux sans-fil.

## 5. ADDITIONAL AUTHORS

Additional authors: Damien Charlet (INRIA Rocquencourt, email: damien.charlet@inria.fr) and François Spiès (LIFC, email: francois.spies@lifc.univ-fcomte.fr).

## 6. REFERENCES

- [1] P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *INFOCOM* (2), pages 775–784, 2000.
- [2] L. Blake. *Radar Range-Performance Analysis*. Artech House Radar Library, December 1986.
- [3] J. Borges and M. Levene. Data mining of user navigation patterns. In *WEBKDD*, pages 92–111, 1999.
- [4] J. Borges and M. Levene. An average linear time algorithm for web usage mining. *International Journal of Information Technology & Decision Making*, 3(2):307–319, June 2004.
- [5] J. Borges and M. Levene. A dynamic clustering-based markov model for web usage mining. *Computer Research Repository*, cs.IR/0406032, May 2004.
- [6] R. Brown and P. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, Inc., 1997.
- [7] D. Charlet, P. Canalda, P. Chatonnay, and F. Spiès. Mobilité dans les caches vidéo : une stratégie adaptative. *TSI (Technique et Science Informatiques)*, 24(7):837–864, 2005.
- [8] D. Charlet, P. Chatonnay, and F. Spiès. Hand-over video cache policy for mobile users. In J. Stefani, I. Demeure, and D. Hagimont, editors, *Proceedings of 6th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 03)*, volume LNCS 2893, pages 176–186, 2003.
- [9] M. Deshpande and G. Karypis. Selective markov models for predicting web page accesses. *ACM Transactions on Internet Technology (TOIT)*, 4(2):163–184, May 2004.
- [10] J. Francois, G. Leduc, and S. Martin. Learning movement patterns in mobile networks : a generic method. In *European Wireless 2004*, pages 128–134, February 2004.
- [11] Interlink Networks, Inc. A practical approach to identifying and tracking unauthorized 802.11 cards and access points. Technical report, 2002.
- [12] J. J. L. Jr. Double exponential smoothing: An alternative to kalman filter-based predictive tracking. In *Proceedings of the workshop on Virtual environments 2003*, pages 199–206. ACM Press, 2003.
- [13] F. Lassabe, P. Canalda, D. Charlet, P. Chatonnay, and F. Spiès. Refining WiFi indoor positioning renders pertinent deploying location-based multimedia guide. In *Procs of IEEE Int. Workshop on Pervasive Computing and Ad Hoc Communications (PCAC06), in conjunction with the IEEE 20th Int. Conf. on Advanced Information Networking and Applications (AINA06)*, pages \*\*\*–\*\*\*, Vienna, Austria, Apr. 2006. 6 pages, to appear.
- [14] F. Lassabe, D. Charlet, P. Canalda, P. Chatonnay, and F. Spiès. Friis and Iterative Trilateration Based WiFi Devices Tracking. In *Proceedings of the 14th Euromicro Conf. on Parallel, Distributed and Network-based Processing (PDP 2006)*, Februar 2006. to be published.

- [15] S. Lee, W. Su, and M. Gerla. Mobility prediction in wireless networks. In *Proceedings of IEEE ICCCN*, pages 22–25, October 2000.
- [16] T. Liu, P. Bahl, and I. Chlamtac. Mobility modeling, location tracking and trajectory prediction in wireless atm networks. *IEEE Journal on Selected Areas in Communications*, 16(6):922–936, August 1998.
- [17] U. A. C. of Engineer, editor. *Engineering and Design - NAVSTAR Global Positioning System Surveying*. Number EM 1110-1-1003. Department of the Army, Washington, DC, July 2003.
- [18] J. Petzold. Augsburg indoor location tracking benchmarks. Tech. Report 2004-9, Institute of Computer Science, University of Augsburg, Germany, February 2004.
- [19] P. Pirolli and J. E. Pitkow. Distributions of surfers' paths through the world wide web: Empirical characterizations. *World Wide Web*, 2(1-2):29–45, 1999.
- [20] J. Pitkow and P. Pirolli. Mining longest repeating subsequences to predict world wide web surfing. In *Proceedings of USITS'99: The 2nd USENIX Symposium on Internet Technologies and Systems*, volume V, pages X–Y. USENIX, USENIX, October 1999.
- [21] N. Priyantha, A. Mui, H. Balakrishnan, and S. Teller. The cricket compass for context-aware mobile applications. In *Proceedings of the 7th Annual ACM International Conference on Mobile Computing and Networking (MobiCom 2001)*, pages 1–14, July 2001.
- [22] R. Sarukkai. Link prediction and path analysis using markov chains. In *Proceedings of the ninth International World Wide Web Conference*, pages XX–YY, May 2000.
- [23] Y. Wang, X. Jia, and H. 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*, number paper 54, Melbourne, July 2003. CD-ROM proc.
- [24] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Transactions on Office Information Systems (TOIS)*, 10(1):91–102, January 1992.
- [25] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao. A wireless lan-based indoor positioning technology. *IBM Technical Journals*, 48(5), 2004.
- [26] M. Youssef, A. Agrawala, A. Shankar, and S. Noh. A probabilistic clustering-based indoor location determination system. Technical report, Maryland Information and Network Dynamics Laboratory, March 2002.