

Calcul de la position exacte d'un appareil à partir de données brutes
par Mahamadou Nouridine Oumarou Mamane
Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de Maître ès sciences (M. Sc.) en informatique
Québec, Canada

© Mahamadou Nouridine Oumarou Mamane, 2022

# RÉSUMÉ

Le *smartphone* est aujourd'hui l'appareil le plus utilisé pour accéder aux services offerts par le numérique. Cette utilisation massive s'explique par nom seulement sa petite taille qui facilite son transport mais surtout du fait des multiples applications qui, par le biais d'innombrables capteurs, offrent des fonctionnalités très utiles. Il faut toutefois noter que l'utilisation de ces applications est à la base d'une génération sans précèdent de données qui va de l'ordre du public au privée (Saint-Louis & Ménélas, 2021). Ces données sont considérées aujourd'hui comme une mine d'or à valoriser. Un usage possible de cela concerne le calcul de la position de l'utilisateur.

Le positionnement de l'utilisateur peut être réalisés en utilisant différents types de technologies comme le GPS, le Bluetooth, le RFID, le réseau cellulaire, le Wifi, etc. Cependant chacune de ces technologies nécessite une infrastructure appropriée, existante ou non, et fournit une précision et une couverture différente. Le GPS est la méthode la plus ancienne et la plus utilisée. Toute fois le GPS nécessite la visibilité directe avec trois satellites pour calculer la position de l'utilisateur et consomme une quantité non négligeable d'énergie. Ces limites du GPS sont à l'origine de la proposition d'autre système de positionnement intérieur et extérieur dont l'utilisation de signaux des requêtes Wifi. Toute fois le mécanisme exact du fonctionnement de ce dernier dans le milieu extérieur n'est connu que par les grandes entreprises et est très peu documenté. En effet le défi de l'utilisation du Wifi dans le positionnement extérieur vient surtout de la difficulté d'accès à des coordonnées globales (latitude et longitude) et la variation de l'intensité du signal RSSI.

Ainsi, dans le cadre de cette recherche, nous avons effectué une revue de littérature sur le positionnement basé sur le Wifi puis étudié et proposé un modèle de positionnement basé sur le Wifi. Dans un premier temps nous avons décrit les données présentes sur les *smartphones*, leurs sources, ainsi que leurs utilisations, puis nous avons détaillé les systèmes de positionnement exploités sur les appareils mobiles. Ensuite, nous avons décrit les systèmes de positionnement basé sur le Wifi et enfin nous avons proposé notre modèle de positionnement basé sur la

combinaison du Wifi avec le GPS. Les résultats nous ont montré l'efficacité du système à pouvoir déterminer la position d'un utilisateur dans des endroits connus et de fournir son historique de déplacements.

# **TABLE DES MATIÈRES**

RÉSUMÉ	i
TABLE DES MATIÈRES	iv
LISTE DES TABLEAUX	v
LISTE DES FIGURES	vi
LISTE DES SIGLES	vii
LISTE DES ABRÉVIATIONS	x
DÉDICACE	<b>x</b> i
REMERCIEMENTS	xi
AVANT-PROPOS	
INTRODUCTION	1
1.1 CONTEXTE	
1.2 PROBLÉMATIQUE	
1.3 CONTRIBUTION	
1.4 ORGANISATION	
CHAPITRE 2 - LES DONNÉES DU SMARTPHONE	
2.1 LE SMARTPHONE	
2.2 LES DONNÉES DU SMARTPHONE	
2.2.1 LES DONNÉES DES APPLICATIONS	
2.2.2 LES DONNÉES DES CAPTEURS	
2.3 UTILISATIONS DES DONNÉES DU SMARTPHONE	
CHAPITRE 3 - POSITIONNEMENT DANS LE MOBILE	
3.1 LE POSITIONNEMENT	
3.1.1 LE POSITIONNEMENT INTÉRIEUR	
3.1.2 LE POSITIONNEMENT EXTÉRIEUR	
3.2 LES CRITÈRES DU POSITIONNEMENT	
3.3 LE POSITIONNEMENT DANS LE MOBILE	
CHAPITRE 4 - POSITIONNEMENT BASÉE SUR LE WIFI	
4.1 LES TECHNIQUES DE POSITIONNEMENT BASÉ SUR LE WIFI	
4.1.1 LE PRINCIPE DE TRILATÉRATION	25
4.1.1.1 LES CONTRAINTES DE LA TRILATÉRATION DANS LE POSITIONNEMENT INTÉRIEUR	28
4.1.1.2 LES CONTRAINTES DE LA TRILATÉRATION DANS LE POSITIONNEMENT EXTÉRIEUR	28
4.1.2 LE PRINCIPE DE L'EMPREINTE DIGITALE	29
4.1.2.1 LES CONTRAINTES DE L'EMPREINTE DIGITALE DANS LE	21

4.1.2.2 LES CONTRAINTES DE L'EMPREINTE DIGITALE DANS LE	
POSITIONNEMENT EXTÉRIEUR	32
4.2 PROBLEMES LIES AUX MÉTHODES BASÉES SUR LA FORCE DU SIGNAL	32
4.3 WIFI ASSOCIÉ AU GPS DANS LE POSITIONNEMENT EXTÉRIEUR	35
4.4 MÉTHODOLOGIE	36
4.4.1 LA PHASE OFFLINE	37
4.4.2 LA PHASE ONLINE	38
CHAPITRE 5 - IMPLÉMENTATION DU MODÈLE PROPOSÉ	40
5.1 ANDROID	40
5.2 OUTILS UTILISÉS	41
5.2.1 Google Maps	41
5.2.2 FIREBASE FIRESTORE	42
5.2.3 KAFKA	43
5.2.4 SPRINGBOOT	43
5.2.5 HEROKU	44
5.2.6 CONFLUENT	47
5.2.7 DBSCAN	49
5.3 ARCHITECTURE DE L'APPLICATION	51
5.3.1 LA COUCHE PRÉSENTATION	52
5.3.2 LA COUCHE MÉTIER	53
5.3.2 LA COUCHE DE DONNÉES	55
5.4 PROCÉDURE DU TEST	57
5.5 RÉSULTAT	57
5.5.1 CAPTURE DE L'APPLICATION	57
5.6 APPRÉCIATION	67
CONCLUSION	70
LISTE DE RÉFÉRENCES	72

# LISTE DES TABLEAUX

TABLEAU 1: TYPES DE SYSTEME DE POSITIONNEMENT	. 19
TABLEAU 2 : METHODE DE POSITIONNEMENT BASE SUR LE WIFI	. 25

# LISTE DES FIGURES

FIGURE 1 : TYPE DE DONNEES GENEREES (DESAI & D'MELLO, 2014)	8
FIGURE 2 : FONCTIONNEMENT DU GPS	17
FIGURE 3: ILLUSTRATION DU RESEAU CELLULAIRE MONTRANT CHAQUE BTS C	
FIGURE 4: ILLUSTRATION DE LA FORCE DU SIGNAL WIFI A PARTIR DE 3 POIN' D'ACCES	
FIGURE 5 : INTERSECTIONS DE 3 CERCLES	27
FIGURE 6: LES DEUX PHASES DE L'APPROCHE DE L'EMPREINTE DIGITALE (LI ET A	
FIGURE 7 : ILLUSTRATION LOS ET NLOS (TABOUBI, 2014)	33
FIGURE 8 : ILLUSTRATION DU PHENOMENE DU TRAJET MULTIPLE (TABOUBI, 2014)	34
FIGURE 9 : CREATION DE L'APPLICATION SUR HEROKU	45
FIGURE 10 : DEPLOIEMENT DE L'APPLICATION SPRINGBOOT SUR HEROKU	47
FIGURE 11 : CREATION DU CLUSTER KAFKA	48
FIGURE 12 : CREATION DU TOPIC KAFKA	48
FIGURE 13 : PROCEDURE D'INTEGRATION DE KAFKA A L'APPLICATION SPRINGBOOT	49
FIGURE 14 : ILLUSTRATION DES CLUSTERS DBSCAN	50
FIGURE 15 : DIAGRAMME DE DEPLOIEMENT DE L'APPLICATION	52
FIGURE 16 : COLLECTION DES POINTS D'ACCES SCANNES	55
FIGURE 17 : COLLECTIONS DE L'HISTORIQUE DE DEPLACEMENT DES UTILISATEURS	56
FIGURE 18 : COLLECTIONS DES CLUSTERS	56
FIGURE 19 : ECRANS D'EMBARQUEMENT DE L'APPLICATION ANDROID	60
FIGURE 20 : RENSEIGNEMENT DU NOM D'UTILISATEUR	61
FIGURE 21 : ECRAN DE SCANNAGE DES POINTS D'ACCES	62
FIGURE 22 : ECRAN DE LOCALISATION D'UTILISATEUR	63
FIGURE 23 : ECRAN DE PARAMETRE	64
FIGURE 24 : MESSAGE D'ACTIVATION DU MODE OFFLINE	65
FIGURE 25 : HISTORIQUE DE DEPLACEMENT DE L'UTILISATEUR	66
FIGURE 26 : VISUALISATION D'UN HISTORIQUE	67

## **LISTE DES SIGLES**

API Application Programming Interface

AOA Angle Of Arrival

BTS Base Transceiver Station

**CAPEX** Capital Expenditures

CID Cell ID

**DBSCAN** Density-Based Spatial Clustering of Applications with Noise

**GIB** Gigabyte

GPS Global Positioning System

**GSM** Global System for Mobile

HTTP Hypertext Transfer Protocol

IEEE Institute of Electrical and Electronics Engineers

iOS iPhone Operating System

JSON JavaScript Object Notation

KNN K-nearest-neighbour Algorithm

**LBS** Location-Based-Services

LOS Line-Of-Sight

LTE Long Term Evolution

MSE Mean Square Error

NLOS Non-Line-Of-Sight

**NN** nearest-neighbour Algorithm

NoSQL Not only SQL

**OPEX** Operational Expenditures

**OS** Operating System

PHP PHP Hypertext Preprocessor

PL Path Loss

**RSS** Received Signal Strength

RSSI Received Signal Strength Indicator

**BSSID** Service Set Identifier

**TDOA** Time Difference Of Arrival

**TOA** Time Of Arrival

RFID Radio Frequency Identification Device

**UWB** Ultra-Wide Band

Wi-Fi Wireless Fidelity

# LISTE DES ABRÉVIATIONS

AMU Appareil Mobile de L'utilisateur

GAFA Google Amazon Facebook Apple

MPts Minimum de points

PAs Points d'Access

PR Point de Reference

**FS** Force du Signal

**Tgps** temps de localisation GPS

**Tscan** temps de balayage Wifi

PA-AMU Distance de séparation entre un point d'accès et l'appareil mobile

de l'utilisateur

# DÉDICACE

Je dédie ce modeste travail :

- À mes très chers parents pour tout l'amour dont vous m'avez entouré, pour tout ce que vous avez fait pour moi. Je ferai de mon mieux pour rester un sujet de fierté à vos yeux avec l'espoir de ne jamais vous décevoir. Que ce modeste travail, soit l'exaucement de vos veux tant formulés et de vos prières quotidiennes.
- À mes très chers sœurs et frères vous occupez une place particulière dans mon cœur. Je vous dédie ce travail en vous souhaitant un avenir radieux, plein de bonheur et de succès.
- À mes formateurs ;
- À ceux qui m'ont soutenu durant ma formation ;
- À ceux qui ont contribué de près ou de loin à l'édition de ce document.

## **REMERCIEMENTS**

Louange à Allah l'Unique, sans associé, le dominateur suprême, le tout puissant, le tout pardon, le tout miséricordieux, seigneur de l'univers pour tous les biens dont il m'a comblé.

Tout d'abord, je tiens à remercier mes parents qui m'ont éduqué, soutenu, encouragé durant toute leur vie.

Ensuite je remercie, Bob Antoine J. Ménélas, Enseignant à l'UQAC pour m'avoir proposé le sujet et accepté de diriger ce travail. Son soutien, et ses compétences m'ont été d'une aide inestimable.

Et enfin, je remercie infiniment tout le corps professoral de l'Université du Québec à Chicoutimi (UQAC).

## **AVANT-PROPOS**

Cette étude entre dans le cadre de la préparation d'un projet de fin d'études pour l'obtention d'un diplôme de maitrise de recherche en informatique. Afin de consolider mes connaissances théoriques par la pratique, l'enseignant Bob Antoine J. Ménélas m'a proposé de travailler sur un système de calcul de la position d'un appareil à partir de données brutes. Le présent document est donc une synthèse de différentes étapes développées dans ce projet.

#### INTRODUCTION

Depuis l'avènement de l'iPhone en 2007, on assiste à une explosion figurante d'achat de cette nouvelle génération de téléphones mobiles. Selon le site web planetoscope, il se vend environ 130 millions de *smartphones* par mois dans le monde soit 1,56 milliard par an malgré la hausse du prix de ces appareils (Planétoscope, 2012). Le chiffre d'affaires de la vente de *smartphones* dans le monde en 2017 est estime à 388,7 milliards d'euros faisant du marché des *smartphones* l'un des plus porteurs dans l'économie mondiale. Cela s'explique par le nombre toujours plus grand d'applications faisant appel à l'intelligence artificielle pour des applications telles que la saisie intuitive, la proposition d'itinéraire, les assistants vocaux, les applications géolocalisées etc., mais aussi par le développement exponentiel d'internet et le faible coût de la connexion.

Comme tout programme informatique les applications utilisent un ensemble de données afin de bien fonctionner, certaines étant fournies par les utilisateurs, d'autres sont générées par l'utilisation des différentes applications comme les réseaux sociaux, les navigateurs web etc., et enfin d'autres sont recueillies par les différents capteurs que comportent les appareils comme l'accéléromètre, le récepteur GPS, le gyroscope, l'analyser Wifi etc. Il en résulte que les *smartphones* contiennent un ensemble de données sur leurs propriétaires issues de sources diverses qui sont de l'ordre du public au privée. Ces données sont devenues une mine d'or et suscite l'attention des chercheurs et entrepreneurs afin d'en tirer un profit.

C'est dans cette thématique que s'inscrit notre travail de recherche qui consiste à déterminer la position d'un utilisateur à partir des données qui sont stockées sur son *smartphone*, en s'appuyant sur les ondes Wifi situés dans l'environnement proche. Cette étude constitue non seulement un moyen de trouver une utilisation de ces données mais

aussi un moyen pour sensibiliser sur les types de données présents sur nos *smartphones* et de réfléchir sur le danger que ces données peuvent représenter.

### 1.1 CONTEXTE

Dans cette période où la digitalisation prévaut dans tous les secteurs, avoir un appareil mobile est devenu la norme. Ainsi, on a en tout temps et partout accès aux multiples possibilités offertes par les différentes applications. Le *smartphone* est aujourd'hui l'outil le plus utilisé en raison de sa facilité d'utilisation et de sa taille qui permet de le transporter partout. Son utilisation principale étant la communication (appels, messageries), mais aujourd'hui à cela s'ajoute les applications participatives comme les réseaux sociaux à partir desquelles on peut non seulement communiquer mais aussi partager du contenu multimédia. On peut aussi considérer les multiples applications qui permettent l'accès à divers services bancaires, administratifs, etc., ainsi que les services de bases offertes par l'appareil à travers les divers capteurs qui y sont intégrés (accéléromètre, gyroscope, thermomètre, boussole, GPS, etc.).

Les applications utilisent les données de ces capteurs pour proposer des fonctionnalités plus pertinentes aux utilisateurs afin de faire face à la concurrence élargie dans un monde de plus en plus ouvert où la qualité de service est au cœur de toutes les entreprises. L'une des données la plus utilisée des services de base est la position géographique. L'idée étant de proposer des fonctionnalités relatives à l'emplacement. On peut citer par exemple la proposition de restaurants ou d'un hôpital proche ainsi que, la proposition d'itinéraire, etc. Cette technique est bien connue sous le sigle LBS, et est déjà utilisée par les géants du web comme Google<sup>1</sup>, Apple<sup>2</sup>, Facebook<sup>3</sup> et Amazon<sup>4</sup>. Ainsi la connaissance de la position de

1 https://www.google.com/

<sup>&</sup>lt;sup>2</sup> https://www.apple.com/

<sup>3</sup> https://www.facebook.com/

<sup>4</sup> https://www.amazon.com/

l'utilisateur apporte une valeur ajoutée aux entreprises et permet la création de nouveaux services, ce qui montre la nécessité indéniable d'un système de localisation plus précis, fiable, accessible, et à faible coût.

# 1.2 PROBLÉMATIQUE

Le *smartphone* est devenu aujourd'hui un compagnon inséparable pour l'Homme et donne accès à une quantité considérable de données allant du public au privé. Afin de tirer profit de ces données, plusieurs travaux de recherche ont été effectués. On peut citer par exemple l'utilisation des données de l'accéléromètre et du gyroscope pour détecter les activités (Lin & Huang, 2017), tracer les déplacements (Sapiezynski et al., 2015), déterminer la quantité de personne présente dans un lieu (Johnson et al., 2019), ou déterminer la position d'un utilisateur (Safei, 2012). Soulignons que la géolocalisation a toujours été un sujet de recherche, elle est généralement réalisée à l'aide du GPS.

La géolocalisation est un procédé qui permet de déterminer la position un objet, un véhicule, ou une personne sur un plan ou une carte à l'aide de ses coordonnées géographique. La méthode la plus connue aujourd'hui pour déterminer la position de l'utilisateur est le système GPS qui est basé sur un système satellitaire. Cependant cette méthode a besoin de la visibilité direct d'au moins trois satellites relativement bien répartis dans l'espace afin de calculer la position (Li et al., 2008). Cela limite l'usage du GPS à la navigation en extérieur. Par ailleurs, soulignons que la précision est fonction du type du récepteur GPS intégré à l'appareil. Dans le cas d'un téléphone portable il est à noter que la précision, des données brutes, est de l'ordre de quelques centaines de mètres. De plus l'usage du GPS tend à décharger plus rapidement la batterie de l'appareil, d'où la nécessité d'une alternative pouvant offrir de meilleurs rendements. Ainsi notre travail vise à répondre à cette problématique en utilisant les signaux des requêtes Wifi pour déterminer la position des utilisateurs. Le Wifi offre plus de portée que la plupart des technologies comme le Bluetooth,

le RFID, etc. Il est utilisable dans les endroits fermés contrairement au GPS et ne nécessite pas l'installation d'une infrastructure supplémentaire. Cependant, l'utilisation du Wifi dans le positionnement extérieur présente un certain défi à savoir son incapacité à fournir des coordonnées globales (latitude et longitude), la variabilité de la mesure de la force du signal RSSI, la non-persistance des données des requêtes du Wifi et surtout l'absence de mécanisme de fonctionnement universel.

C'est dans cette optique que s'inscrit notre travail de recherche qui consiste à proposer un système de positionnement à grande échelle par l'usage des signaux de requête Wifi émis par le téléphone lors de ses tentatives de connexion aux différents points d'accès qui lui sont accessibles.

#### 1.3 CONTRIBUTION

La principale contribution de ce travail est la proposition d'un modèle permettant de déterminer la position de l'utilisateur suivant les points d'accès Wifi accessibles en combinant les différentes méthodes de la littérature. La deuxième contribution est l'implémentation du modèle proposé dans une application Android. La troisième est un apport communautaire par la mise à disposition de notre base de données au service de la communauté pour qu'elle puisse l'utiliser dans les travaux futurs. Enfin l'utilisation du Wifi pour déterminer la position permettrait de réduire la consommation de la batterie des *smartphones* par les requêtes gourmandes du GPS et constitue un moyen d'utilisation des données déjà disponibles sur les *smartphones*.

# 1.4 ORGANISATION

Le reste du document est subdivisé comme suit :

Chapitre deux - les données du smartphone : donne une description des données disponibles sur les *smartphones*, en particulier leurs sources et leurs utilités.

**Chapitre trois - Positionnement dans le mobile :** présente l'état de l'art des méthodes de géolocalisation dans les systèmes mobiles.

Chapitre quatre - Géolocalisation basée sur le Wifi : présente l'état de l'art des méthodes utilisées dans le positionnement basé sur le Wifi, puis expose le modèle proposé.

Chapitre cinq - Implémentation du modèle proposé : décrit comment le modèle proposé est implémenté. Il présente également les choix technologiques, les résultats et enfin il discute les résultats.

**Conclusion et travaux futurs :** une synthèse du travail effectué et des résultats obtenus est faite avant de présenter les perspectives des travaux futurs.

#### **CHAPITRE 2**

# LES DONNÉES DU SMARTPHONE

Dans ce chapitre nous présenterons le *smartphone*, les données qui y sont stockées ainsi que leur usage.

#### 2.1 LE SMARTPHONE

Le smartphone est un téléphone mobile doté d'une forte puissance de calcul et d'un stockage proche de celui de l'ordinateur ainsi que d'une bonne capacité de prise de photos. La première réussite de cette nouvelle génération de téléphones est l'iPhone lancé en 2007 par Apple, regroupant téléphone, iPod, et navigateur internet et a initié la révolution du smartphone en proposant la saisie des données via un écran tactile permettant une facilité d'utilisation par rapport à son prédécesseur. Cependant, certains modèles utilisent toujours des touches physiques tandis que d'autres permettent l'utilisation d'un stylet. Apple a beaucoup contribué au développement du marché des smartphones en apportant chaque année de nouvelles fonctionnalités grâce aux innombrables capteurs intégrés à ces appareils. Les exemples incluent la boussole, la commande vocale, la caméra frontale, le Touch ID, la capture vidéo 3D Touch et 4K, etc. L'inclusion d'un écran de meilleure qualité est la plus grande étape, associée à un processeur plus rapide, qui permet de capturer et de lire du contenu HD. La grande capacité du smartphone en termes de processeur et de RAM lui permet de faire tourner des applications qui auparavant ne pouvaient être utilisées qu'avec un ordinateur, ce qui justifie également sa forte utilisation. Le smartphone est ainsi assimilable à un petit ordinateur de poche. Comme évoqué, l'autre usage fréquent du smartphone concerne la prise de photos, devenue indispensable depuis l'avènement des réseaux sociaux afin de partager son quotidien avec les autres utilisateurs. Ce phénomène a conduit à une amélioration continue de la résolution des caméras, qui représente aujourd'hui un critère majeur de différenciation entre les modèles. Cela explique aussi le remplacement de plus en plus évident des appareils photos numériques par les *smartphones*.

Le *smartphone* fournit, en plus des fonctionnalités du téléphone mobile classique, les fonctionnalités de base suivantes : la messagerie instantanée, le GPS, la boussole, etc. Outre ces applications de base, une variété d'applications peut être téléchargées et installées pour les besoins quotidiens ou le divertissement. Ces applications sont disponibles dans des boutiques en ligne centralisées, également introduites par Apple et spécifiques à chaque type de système d'exploitation, dont les plus connues sont Play store<sup>5</sup> et App store<sup>6</sup> contenant respectivement des applications pour les systèmes Android et IOS. Pour accéder aux contenus disponibles sur internet ou télécharger des applications, une connexion internet est nécessaire, obtenue soit du réseau téléphonique, soit du réseau Wifi. Or, l'utilisation de ces applications est à l'origine de la génération d'une très grande quantité de données dont l'utilité suscite l'intérêt des chercheurs et des entreprises. Cependant, avant toute utilisation, il convient de savoir quelles sont ces données et quelles sont leurs sources.

### 2.2 LES DONNÉES DU SMARTPHONE

Les *smartphones*, en plus d'être utilisés comme appareil de communication, tentent de rivaliser en termes de puissance de calcul avec un ordinateur portable (Desai & D'Mello, 2014). Cette croissance fulgurante de l'utilisation des *smartphones* a conduit à une génération à grande échelle de données. Ces données permettent aujourd'hui d'acquérir de nouvelles connaissances sur divers aspects de l'utilisateur tels que ses habitudes, ses préférences, ainsi que ses déplacements (Desai & D'Mello, 2014). Pour explorer, analyser et extraire des informations utiles à partir de ces données, il est important de connaitre leurs origines, et leurs formes. Selon Desai & D'Mello (2014), ces données collectées par les

<sup>6</sup> https://www.apple.com/ca/fr/app-store/

<sup>&</sup>lt;sup>5</sup> https://play.google.com/store/apps

téléphones intelligent proviennent essentiellement des journaux d'appels, de l'emplacement, de l'utilisation du Web, de l'utilisation des applications, et des données des capteurs. La figure 1 reprend dans une image les données présentent sur le *smartphone*.

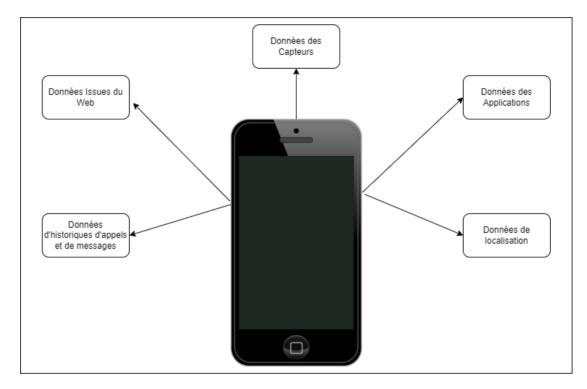


Figure 1 : Type de données générées (Desai & D'Mello, 2014).

# 2.2.1 LES DONNÉES DES APPLICATIONS

Une partie des données stockées sur nos *smartphones* provient des applications que nous utilisons dans notre vie quotidienne pour diverses activités sociales, académiques et professionnelles. Les principales applications gourmandes en stockage sont les réseaux sociaux, les jeux et les navigateurs web car chaque jour plusieurs milliards de personnes dans le monde utilisent des *smartphones* pour passer des appels, tweeter et rechercher des informations sur le web (Fulara, 2018). Ainsi ces navigateurs utilisés pour accéder à Internet génèrent beaucoup de données en enregistrant dans le cache les éléments téléchargés ainsi que l'historique de navigation de l'utilisateur, données à partir desquelles on peut connaître

les centres d'intérêt de ce dernier. En ce qui concerne les réseaux sociaux, il existe tous les types de données, à savoir les textes, les bandes sonores, les vidéos et les images qui sont téléchargées par l'utilisateur ou partagées. Ainsi à chaque seconde, nous créons de nouvelles données grâce aux applications de nos *smartphones*.

### 2.2.2 LES DONNÉES DES CAPTEURS

Si le smartphone propose une panoplie de services comme la température, la localisation, la détection de mouvement, c'est certainement grâce aux nombreux capteurs qui y sont intégrés. En effet, ces capteurs collectent un ensemble de données sur l'environnement, qui sont ensuite transmises par le système de l'appareil aux applications qui y sont abonnées. Cette fonction des capteurs représente ainsi une source majeure de génération de données sur nos smartphones. Chacun de ces capteurs remplit une fonction spécifique et l'utilisation de ces données collectées nécessite souvent l'autorisation de l'utilisateur avant qu'une application puisse y accéder. Ces capteurs sont classés en deux catégories, les capteurs physiques et les capteurs virtuels. Les premiers, les capteurs physiques, sont les plus courants et mesurent directement dans l'environnement une grandeur physique (vitesse, température, etc.). Les plus intégrés de ce type de capteurs sur les appareils d'aujourd'hui sont l'accéléromètre qui détecte la vitesse et le mouvement, le gyroscope qui aide l'accéléromètre à déterminer l'orientation du téléphone, le magnétomètre qui mesure les champs magnétiques pour déterminer l'orientation du téléphone, le capteur de lumière qui mesure l'intensité lumineuse, un système de positionnement GPS global, le capteur sonore (microphone), le capteur d'image (caméra), le capteur tactile (écran), etc. Les capteurs virtuels ne sont pas des dispositifs physiques, bien qu'ils imitent les capteurs physiques. Les capteurs virtuels tirent leurs données d'un ou plusieurs des capteurs physiques. Ce type de capteurs mesurent aussi des données contextuelles depuis des applications ou des services web. Par exemple on peut déterminer la tâche courante d'une personne en requêtant son calendrier personnel, ou savoir quels genres de sites elle visite.

Les téléphones intelligents pourraient avoir dans le futur plus de capteurs que ceux mentionnés ci-dessus. Quel que soit le nombre de capteurs du *smartphone*, ils ont tous une caractéristique commune : ils génèrent d'énormes paquets de données, qui peuvent être utilisés dans de nombreuses applications.

# 2.3 UTILISATIONS DES DONNÉES DU SMARTPHONE

Nos *smartphones* contiennent de nombreuses informations sur nous qui sont en grande partie générées par les applications et les capteurs qui y sont intégrés. En cette période de forte digitalisation dans tous les secteurs et de concurrence accrue, l'une des préoccupations majeures des entreprises est la fidélisation de leurs clients (Marketing-Professionnel, 2012). Pour cela, elles doivent constamment proposer des services qui répondent aux besoins de leurs clients. Pour y parvenir, elles doivent d'abord identifier les besoins potentiels. Ainsi, cette volonté a suscité l'intérêt des entreprises ainsi que des chercheurs à étudier les données stockées sur les *smartphones* pour savoir quelles informations pertinentes sur l'utilisateur ils peuvent en extraire. C'est ainsi que plusieurs études exploratoires ont été menées afin de trouver une utilisation à ses données.

Selon Yao et al. (Yao et al., 2018) la littérature existante a classé l'exploration de données mobiles en trois catégories : l'exploration sensible à la localisation, l'exploration sensible au contexte, et l'exploration sociale. L'exploration sensible à la localisation consiste à récupérer des données relatives à la localisation de l'utilisateur afin de lui proposer des services en fonction de son emplacement. Cette catégorie comprend à la fois la localisation extérieure et intérieure, et les données proviennent généralement du GPS, des réseaux cellulaires, des réseaux Wifi, des connexions Bluetooth, etc. Quant à l'exploration sensible au contexte, elle consiste à acquérir et exploiter des informations sur le contexte de l'utilisateur à partir de données de capteurs du *smartphone* telles que le gyroscope, l'accéléromètre, le magnétomètre etc. Ces informations permettent de déduire les activités en cours de

l'utilisateur telles que la marche, le sommeil, la position assise ou debout, l'état de conduite etc. Et enfin l'exploration sociale, elle consiste à récupérer et exploiter des informations sur l'utilisateur en fonction des activités de celui-ci sur les réseaux sociaux.

En fonction de ces trois types d'exploration de données mobiles, les solutions d'exploration de données mobiles sont utilisées comme méthode de base ou comme technologie d'assistance dans de nombreuses applications repartie en quatre domaines : l'informatique urbaine (Soundararaj et al., 2020), la santé (Francillette et al., 2021), la recherche et publicité géolocalisées (Bassoni & Weygand, 2011), la confidentialité et la sécurité (Yao et al., 2018). Dans l'informatique urbaine, la connexion du smartphone aux tours cellulaires et aux réseaux Wifi disponibles est utilisée pour aider à découvrir les régions fonctionnelles, pour définir les limites de la ville, pour identifier les modes de transport pour les utilisateurs, pour réaliser des enquêtes sur les modes de transport, et surtout pour la surveillance quotidienne du trafic, etc. Dans le domaine de la santé, les données mobiles sont utilisées dans la reconnaissance d'activités telles que la détection des chutes (Menelas & Otis, 2012), l'assistant de rééducation (Bhavnani et al., 2016), etc. Dans la recherche et la publicité basées sur la localisation, l'exploration de mobiles est utilisée dans les services basés sur la localisation tels que les cartes, la navigation, les offres de services et le marketing pour le profilage et la prospection (Bassoni & Weygand, 2011). Dans le domaine de la confidentialité et de la sécurité, ces données mobiles sont souvent utilisées à des fins malveillantes, par exemple, pour sonder la vie privée des utilisateurs ou voler des informations (Yao et al., 2018). Ainsi l'importance des données stockées nécessite un niveau de confidentialité et de sécurité (Androulidakis, 2012).

Dans la suite de notre travail, nous nous concentrerons principalement sur l'exploration sensible à la localisation et plus spécifiquement sur la localisation à grande échelle des utilisateurs de *smartphone*.

#### **CHAPITRE 3**

### POSITIONNEMENT DANS LE MOBILE

Dans le chapitre précédent nous avons décrit les différentes données stockées sur les smartphones ainsi que leurs sources. Nous avons également vu que l'on peut extraire de nombreuses informations à partir de ces données, l'une des plus importantes étant la détermination de la position de l'utilisateur afin de lui proposer des services en fonction de sa localisation. Nous verrons donc dans ce chapitre ce qu'est le positionnement et ensuite nous détaillerons les différents systèmes de positionnement utilisés sur les appareils mobiles.

#### 3.1 LE POSITIONNEMENT

La géolocalisation est une technique permettant de déterminer la position d'un objet ou d'une personne à l'aide de coordonnées géographiques sur un plan ou une carte. Ainsi la détermination des coordonnées est un préalable à la géolocalisation. Cette opération est appelée positionnement. S'il est vrai que positionnement et géolocalisation sont souvent confondus par abus de langage, une distinction claire doit être faite entre ces deux termes. Le positionnement est la phase initiale de la géolocalisation, qui consiste à déterminer les coordonnées globales, à savoir la latitude et la longitude de l'objet à localiser. Cette opération peut être réalisée par un système de positionnement par satellite dans le cas du GPS ou d'une autre technique. Ensuite grâce aux coordonnées, la position de l'objet pourra être visualisée sur une carte comme Google Maps par exemple. Tandis que la localisation est la détermination de coordonnées relatives telles que « urgence hôpital de Chicoutimi salle 201 » (Chan & Sohn, 2015).

De nos jours, la géolocalisation est utilisée à des fins diverses telles que la navigation (maritime, terrestre et aérienne), le suivi des ressources (véhicules et personnes), la proposition d'itinéraire, la proposition de service, etc. Le suivi GPS est le plus ancien et le

plus répandu. Cependant, il présente certaines lacunes puisque le récepteur GPS doit voir au moins trois satellites relativement bien répartis dans le ciel pour calculer la position (Li et al., 2008). Ainsi dans des environnements où le ciel est obstrué comme les immeubles, en forêt et les centres urbains avec de grands immeubles, le positionnement devient difficile, voire impossible. Ainsi plusieurs méthodes ont été proposées pour résoudre le problème de la géolocalisation aussi bien en extérieur qu'en intérieur.

### 3.1.1 LE POSITIONNEMENT INTÉRIEUR

Un système de positionnement intérieur ou système de géolocalisation indoor permet de retrouver la position d'objets ou de personnes dans un espace interne à une structure. Ce type de localisation permet de déterminer la position des visiteurs dans une bibliothèque, une université, un musée ou encore une usine, etc.

Il est important de savoir que toute détermination de position se fait dans un système de référence. Par exemple, dans la localisation extérieure basée sur le GPS, l'emplacement de l'objet est généralement déterminé sur un plan ou une carte en utilisant sa longitude et sa latitude, donc représentées en 2D. Cependant, ce type de référence n'est pas approprié dans les environnements intérieurs qui nécessitent de prendre en compte certaines particularités telles que les sols, les escaliers. Ainsi, dans ces types d'environnements intérieurs à plusieurs niveaux, une représentation appropriée est nécessaire, c'est-à-dire une topographie 3D. Ainsi du fait de la petite taille de l'espace, contrairement à la localisation extérieure, la géolocalisation intérieure doit permettre une localisation fine ou une bonne précision pour pouvoir distinguer facilement les différents espaces et niveau du bâtiment.

La localisation a toujours été un domaine de recherche actif et a conduit à l'émergence de nombreuses solutions de localisation internes qui utilisent différentes architectures et matériels en fonction des besoins et des exigences. Parmi ces technologies, on peut citer :

l'identification par radiofréquence (RFID), les ultrasons, l'ultra large bande, le Bluetooth, le GSM, le réseau local sans fil (Wifi), la localisation par image, le signal TV, etc. (Chan & Sohn, 2015).

### Bluetooth

Le Bluetooth est une norme de télécommunications sans fil qui utilise des ondes radio pour échanger des données sur de courtes distances entre deux appareils électroniques. Cette technologie présente un certain nombre d'avantages tels que le fait que les appareils Bluetooth ne nécessitent pas de ligne de vue directe pour communiquer ; ce qui rend cette approche adaptée dans les environnements contenant des obstacles. De plus, la précision de cette technologie est estimée entre 10 et 2 mètres selon la qualité du signal, et soulignons que l'intensité des ondes Bluetooth est trop faible pour représenter un quelconque danger pour la santé ou des interférences. Toutes ces raisons ont ainsi conduit à l'utilisation massive du Bluetooth en localisation intérieure. Le système de positionnement basé sur le Bluetooth consiste à créer un réseau de balises Bluetooth dans le bâtiment, chacune des balises ayant une zone de couverture bien définie et une position connue. Ainsi, la localisation de l'utilisateur est déduite en fonction des balises qui lui sont accessibles par triangulation ou trilatération.

#### Ultra Wide Band

L'« Ultra Wide Band » (UWB) est une technologie radio à haut débit utilisée pour les communications à courte portée, avec un très faible niveau d'énergie. L'utilisation de cette technologie est largement adoptée en localisation intérieure pour sa bonne précision estimée à 10 cm.

Comme pour le Bluetooth, plusieurs antennes sont placées dans des zones de référence du bâtiment. Ensuite on peut localiser l'objet d'intérêt selon les antennes avec lesquelles il peut communiquer en appliquant soit une méthode de triangulation soit de trilatération.

### RFID

« Radio Frequency Identification » (RFID), est une technologie qui permet de récupérer une information à distance, stockée sur un tag (ou étiquette) afin de pouvoir l'identifier. Pour être détectée par le lecteur, l'étiquette est reliée à une antenne qui émet des ondes. La technologie permet de lire les étiquettes sans même une ligne de vue directe et peut traverser de fines couches de matériaux (peinture, neige, etc.). La RFID a besoin de deux éléments pour fonctionner, un lecteur qui lit les identifiants des étiquettes et l'étiquette, qui n'est rien de plus qu'une puce apposée sur l'objet à identifier. Ainsi cette technologie peut être utilisée pour déterminer la position d'un utilisateur en plaçant des lecteurs d'étiquettes RFID dans des zones de référence puis en déduire la localisation de l'utilisateur en fonction des lecteurs qui détectent l'appareil, avec les positions des lecteurs qui sont déjà connues.

# 3.1.2 LE POSITIONNEMENT EXTÉRIEUR

Le positionnement extérieur permet de retrouver la position d'objets ou de personnes dans un espace étendu (qui ne se limite plus à un seul bâtiment comme un aéroport, un centre commercial, une ville, une région, etc.) voir global. Bien que le système de positionnement global (GPS) ait été largement déployé dans de nombreux domaines, il présente plusieurs inconvénients tels qu'une consommation d'énergie élevée et le fait qu'il nécessite une visibilité directe des satellites. Ces limitations du GPS ont conduit à l'utilisation d'autres systèmes existants comme alternative, y compris l'infrastructure du réseau cellulaire et le réseau Wifi.

#### Le GPS

Le « Global positioning system » (GPS) est une technique permettant de déterminer la position d'un objet, c'est-à-dire ses coordonnées globales à l'échelle planétaire. La

technologie a été initialement développée à des fins militaires par le département américain de la Défense pendant la guerre froide. Cependant, depuis les années 2000 le GPS et ses informations sont désormais accessibles au public et sont embarqués dans de nombreux appareils de notre quotidien comme les *smartphones*, les voitures, les bateaux et même les animaux, etc. Le GPS utilise trois composants pour fonctionner:

- Des satellites en orbite autour de la terre formant une constellation appelée segment spatial. Le nombre de satellites dans la constellation peut varier, mais il faut tout le temps un minimum de 24 satellites pour couvrir toute la planète ainsi que d'autres qui peuvent servir de remplacements en cas de panne. Ces satellites sont organisés en groupes de 4 se déplaçant selon 6 trajectoires différentes formant une orbite d'un rayon d'environ 26 600 km. Leur rôle est d'envoyer en permanence des informations sur leur position.
- Un segment de contrôle composé de cinq stations pour contrôler et assurer le bon fonctionnement des satellites.
- Et un segment d'utilisateurs qui représente n'importe quel appareil avec un récepteur GPS. Leur rôle est de récupérer les informations envoyées par les satellites et de calculer la position.

La position est calculée par une méthode de triangulation plus précisément de trilatération. Ainsi on calcule d'abord la distance entre le récepteur et au moins trois satellites à l'aide des informations envoyées par les satellites contenants, entre autres, la latitude, la longitude, l'altitude, l'heure d'émission. Sachant que les ondes se déplacent à la vitesse de la lumière, la distance se calcule en multipliant le temps parcouru par l'onde entre l'émetteur (satellite) et le récepteur par la vitesse de la lumière. Puis la position est obtenue par trilatération en résolvant un système d'équations à 3 inconnues (sinon généralement 4 si le décalage d'horloge entre le récepteur et les satellites est pris en compte pour améliorer la précision). La figure 2 illustre le fonctionnement du GPS.

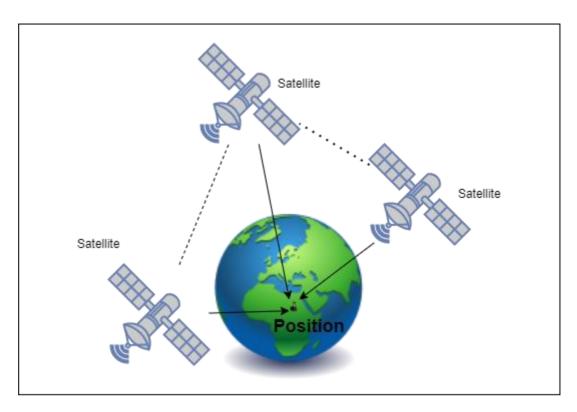


Figure 2: Fonctionnement du GPS.

© Oumarou Nouridine, 2022.

## Le réseau de téléphonie mobile

Le réseau de téléphonie est un réseau permettant la communication simultanée de plusieurs appareils mobiles (y compris des appareils immobiles comme les téléphones fixe) sur une très longue distance. Le réseau a une structure cellulaire. Chacune de ces cellules possède une station de base, ou Base Transceiver Station (BTS), qui émet des ondes radio de même fréquence sur une couverture correspondant à la cellule. Ces signaux sont ensuite reçus par les appareils environnants. Les dispositifs communiquent ainsi avec la station de base en utilisant les canaux de signalisation associés à la BTS de la cellule dans laquelle ils se trouvent et inversement. La figure 3 illustre l'échange d'informations entre un appareil mobile et le BTS de la cellule dans laquelle se situe l'appareil.

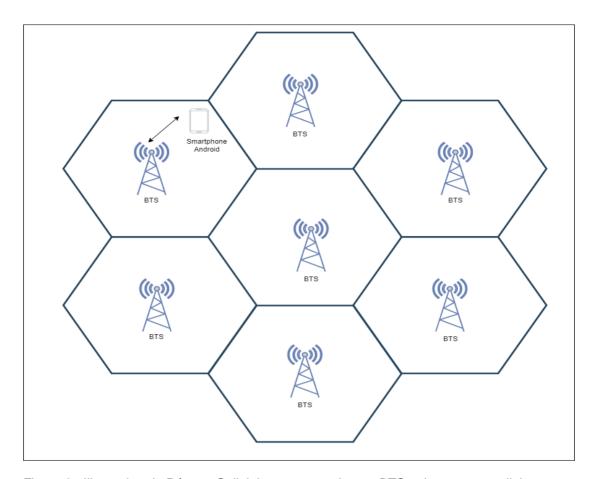


Figure 3 : Illustration du Réseau Cellulaire montrant chaque BTS qui couvre sa cellule. © Oumarou Nouridine, 2022.

Pour permettre une couverture totale de chaque utilisateur, un changement de cellule sans interrompre la communication en cours appelé *handover* est effectué lorsque l'utilisateur se déplace. Comme chaque cellule possède un identifiant appelé CID (CELLID) permettant de l'identifier et une position géographique connue de l'opérateur téléphonique, il est alors possible de faire une géolocalisation basée sur le réseau de téléphonie mobile (tel que GSM, Long Term Evolution ou LTE, etc.) en attribuer la position de la cellule à laquelle l'appareil est connecté à son utilisateur. Cette technique est moins coûteuse car elle utilise une infrastructure déjà existante et ne nécessite aucun matériel supplémentaire. Cependant, la précision peut ne pas être satisfaisante selon le nombre d'antennes BTS car, plus l'antenne est isolée, plus la zone de couverture est grande, ce qui peut détériorer la précision.

### Le Wifi

Le Wifi (souvent orthographié Wi-Fi) acronyme de Wireless Fidelity est un ensemble de règles qui régissent la communication entre des équipements informatiques (ordinateurs, imprimantes, *smartphones*, tablettes, routeurs, etc.) sans connexion filaire. La communication se fait donc par des ondes radio sur les fréquences 2,4 GHz ou 5 GHz. En effet, le nom standard est IEEE 802.11, et a été développé en 1997 par l'Institute of Electrical and Electronics Engineers (IEEE). La première version de la norme était IEEE 802.11a offrant des débits de 1 ou 2 Mbit/s, puis des révisions ont été faites pour améliorer le débit, allant aujourd'hui jusqu'à 100Go/s en théorie avec l'amendement 802.11ay. Grâce aux normes Wifi, il est possible de créer des réseaux locaux sans fil à haut débit (WLAN). Aujourd'hui les réseaux Wifi sont déployés partout (maisons, marchés, aéroports, parcs etc.).

Pour se connecter à ces points d'accès, les appareils envoient en permanence des requêtes pour trouver les points d'accès Wifi environnants. Ainsi comme la plupart de ces points d'accès sont fixes, il est possible de déterminer la position de l'utilisateur en fonction des points d'accès qui lui sont accessibles. Alors même si le Wifi n'est pas déployé à des fins de positionnement, des mesures de l'intensité du signal (SS) transmises en permanence soit par le point d'accès, soit par le *smartphone* permettent d'en déduire la localisation de l'utilisateur.

Tableau 1 : Types de système de positionnement

Systèmes de positionnement Intérieur	Systèmes de positionnement extérieur
Bluetooth, RFID, Ultrasons, Ultra Wide	GPS, Réseau mobile, Wi-Fi
Band, Bluetooth, le GSM, Wifi, localisation	
basée sur l'image, Signal TV	

### 3.2 LES CRITÈRES DU POSITIONNEMENT

Au vu des différents types de localisation et des différents systèmes utilisés dans chacune de ses catégories de positionnement, on comprend que le choix d'un système de localisation repose sur un certain nombre de critères qui doivent être assurés en fonction des besoins et des contraintes. Les principaux paramètres sont : le coût, la précision, le temps de réponse ou latence, l'autonomie et la couverture.

#### Précision

La précision est le degré de concordance entre une série de mesures d'une même quantité, qu'il ne faut pas confondre avec l'accuracy, qui est le degré de concordance entre le résultat expérimental et la valeur. Même si ce n'est pas une mesure qui indique l'exactitude du système, c'est quand même l'attribut le plus important de la plupart des systèmes de localisation, surtout pour la localisation intérieure. Cette caractéristique est régulièrement améliorée et atteint aujourd'hui l'ordre du cm dans le positionnement intérieur

### Coût

Le coût est un critère non négligeable dans le choix de tout système de positionnement et comprend les CAPEX, dépenses d'acquisition du système et les OPEX, dépenses de maintenance. Ainsi, un système basé sur une infrastructure existante réduirait considérablement les coûts.

### ♣ La latence

Un système de positionnement doit avoir un temps de réponse négligeable par rapport à la vitesse de déplacement. Pour ne pas frustrer les utilisateurs, il arrive très souvent que l'algorithme de calcul de position soit optimisé.

#### Couverture

La couverture fait référence à l'étendue géographique sur laquelle les services de positionnement sont accessibles. Cet attribut est particulièrement important dans le

positionnement extérieur et dépend de la technologie utilisée. Par exemple, un système basé sur le Wifi a une couverture limitée à quelques mètres, contrairement au GPS qui couvre l'ensemble du globe. Ainsi, selon le besoin de couverture, ce paramètre peut influencer le choix du système à utiliser.

#### Autonomie

La durée d'autonomie de la batterie fait référence à la durée pendant laquelle un appareil mobile peut fonctionner de manière autonome. Ainsi, un système qui économise la consommation d'énergie de la batterie est très apprécié et important à considérer.

#### 3.3 LE POSITIONNEMENT DANS LE MOBILE

Comme indiqué précédemment, plusieurs technologies peuvent être utilisées pour le positionnement interne et externe nécessitant une infrastructure (existante ou non) et des équipements appropriés, et offrant une précision et une couverture différentes. La géolocalisation dans le mobile fait référence à un système de positionnement dans lequel l'élément à localiser est un *smartphone*. Ainsi un tel système ne devrait se baser que sur des composants disponibles sur un *smartphone* afin d'en limiter le coût tels que le GPS, Bluetooth, Wifi, réseau cellulaire, caméra, etc. Par exemple les systèmes basés sur les ultrasons peuvent avoir des niveaux de précision en millimètres et en centimètres, mais ne sont pas compatibles avec les *smartphones* actuels tout comme le RFID et l'UWB. Ces techniques nécessitent non seulement la mise en place d'une nouvelle infrastructure mais aussi l'intégration d'un nouveau composant dans le *smartphone*. Alors, l'utilisation de systèmes comme le GPS, le Bluetooth, le Wifi, le réseau cellulaire semble mieux adapté. Cependant, le Bluetooth nécessite également le déploiement d'une nouvelle infrastructure sur l'étendue de la zone de couverture. Pour le réseau cellulaire, l'étendue de la couverture

cellulaire peut compromettre l'obtention d'une précision acceptable. Quant au GPS, les problèmes évoqués précédemment ont montré leurs limites dans l'environnement intérieur et les centres urbains. Toutefois, avec l'existence de points d'accès Wifi densément déployés, le positionnement basé sur le Wifi semble très prometteur en termes de précision et de consommation de la batterie. De plus, le fait qu'il ne nécessite aucun apport supplémentaire en termes d'infrastructure et de composant, et qu'il ne nécessite pas non plus de ligne de vue directe pour communiquer en fait un choix pertinent et incontestable (Messaoudi et al., 2020).

#### **CHAPITRE 4**

# **POSITIONNEMENT BASÉE SUR LE WIFI**

Avec l'omniprésence des points d'accès Wifi, le positionnement basé sur le Wifi a suscité l'intérêt des entreprises et des chercheurs et s'avère prometteuse en termes de précision de positionnement, de financement et de consommation d'énergie, justifiant son utilisation à la fois pour le positionnement intérieur et extérieur. Dans ce chapitre nous allons présenter les techniques de positionnement basées sur le Wifi.

#### 4.1 LES TECHNIQUES DE POSITIONNEMENT BASÉ SUR LE WIFI

Il existe plusieurs méthodes pour déterminer l'emplacement des utilisateurs à l'aide de technologies sans fil. La plupart de ces méthodes utilisent la distance entre le mobile et certains points d'accès de coordonnées connus pour calculer la position de l'utilisateur. Ainsi, selon que la distance est utilisée ou non, ces méthodes sont classées en deux catégories. Les premières sont ceux basées sur la distance, tels que l'angle d'arrivée (AOA) des signaux, l'heure d'arrivée (TOA) des signaux, la différence de temps d'arrivée (TDOA) des signaux et l'utilisation d'un modèle de propagation des signaux (Du et al., 2018). La deuxième catégorie sont celles qui ne sont pas basées sur la distance, dont la plus connue est la méthode des empreintes digitales (Du et al., 2018).

L'AOA est basée sur la direction ou l'angle d'arrivée d'au moins deux signaux distincts provenant de deux points d'accès dont la position est connue. Le TOA quant à lui mesure la distance entre le *smartphone* et les points d'accès en calculant le temps de parcours du signal, ce qui nécessite une bonne synchronisation entre l'appareil à localiser et les points d'accès pour avoir une bonne précision. Le TDOA utilise le même principe que le TOA pour calculer la distance, cependant il apporte une réponse à la question de la synchronisation en assurant la synchronisation entre les différents récepteurs (points d'accès). Et enfin, le

modèle de propagation du signal transforme la valeur de l'intensité du signal collecté en distance à l'aide d'une formule de propagation du signal définie par l'équation 1.

$$P(d)[dBm] = P(d_o)[dBm] - 10n\log\left(\frac{d}{d_o}\right) - \begin{cases} nW*WAF & nW < C \\ C*WAF & nW \ge C \end{cases}$$

$$\tag{1}$$

Où n indique le taux auquel l'affaiblissement de trajet augmente avec la distance, P(do) la puissance du signal à la distance de référence do, d la distance de séparation émetteur-récepteur, C le nombre maximal d'obstacles (murs) jusqu'auxquels le facteur d'atténuation fait la différence, nW le nombre d'obstructions (murs) entre l'émetteur et le récepteur, et WAF le facteur d'atténuation du mur.

Dans les méthodes basées sur la distance, la localisation de l'utilisateur correspond à l'intersection de cercles dont les rayons correspondent aux différentes distances calculées entre l'appareil et chacun des points d'accès, cette technique est connue sous le nom de multilatération. Cependant, dans le cas de l'AOA, on parle de multiangulation car on travaille avec des angles, et non de multilatération qui est basée sur des distances. Cependant, par abus de langage, ces deux termes sont souvent utilisés l'un à la place de l'autre. Lorsqu'il s'agit de trois points d'accès, on parle de trilatération ou de triangulation.

On peut donc en déduire que trois méthodes de base permettent de déterminer la localisation des utilisateurs utilisant le Wifi : d'abord l'utilisation de la direction ou de l'angle d'arrivée (AOA) d'au moins deux signaux distincts provenant d'emplacements connus, puis la technique de triangulation est utilisée pour calculer l'emplacement de l'appareil mobile. La seconde, la trilatération qui nécessite au moins trois estimations distinctes de la distance d'un appareil mobile à un récepteur Wifi à partir d'emplacements fixes connus, et la dernière, l'emploi de modèles d'empreintes digitales de localisation (Köbben et al., 2006).

Tableau 2 : Méthode de positionnement basé sur le Wifi

Techniques	Caractéristiques ou Mesure
Triangulation	AOA
Trilatération	TOA, TDOA et RSSI
Empreintes digitaux	RSSI

Cependant, l'AOA nécessite l'utilisation d'antennes spéciales sur les points d'accès pour pouvoir obtenir l'angle d'arrivée et l'obtention du vrai angle est difficile à trouver en raison du problème de trajets multiples (réflexion, absorption, réfraction, etc.) de la propagation du signal causé par les obstacles. Pour cette raison cette méthode n'est pas adaptée au positionnement basé sur le Wifi contrairement aux techniques de trilatération et d'empreintes digitales qui sont basées sur la mesure de la force du signal qui est la seule mesure disponible avec le Wifi.

# 4.1.1 LE PRINCIPE DE TRILATÉRATION

La trilatération est une méthode qui consiste à calculer la position d'un objet connaissant les distances qui le séparent de trois autres éléments de coordonnées connus. Dans le cas du positionnement basé sur le Wifi, ces trois éléments sont des points d'accès Wifi et l'élément à localiser peut-être un *smartphone* ou tout autre appareil disposant d'un adaptateur réseau sans fil.

Le principe de cette approche consiste à tracer trois cercles, dont le rayon R de chacun de ces cercles correspond à la distance D qui sépare un point d'accès de coordonnées connues et l'appareil à localiser. L'intersection des trois cercles est le lieu géométrique où se situe l'objet recherché. Si les trois cercles se coupent en un seul point, on comprend qu'il correspond à la position de l'utilisateur. Les coordonnées de l'utilisateur sont alors calculées soit par la méthode des moindres carrés, soit par toute autre méthode géométrique telle que

l'utilisation de l'équation du cercle (Safei, 2012). Néanmoins dans la très grande majorité des cas, l'intersection n'est pas un point mais une zone, c'est pour cela que d'autres points d'accès sont généralement requis pour affiner la recherche de la position. Il est à noter que les mesures disponibles en Wifi sont la force du signal plutôt que la distance, pour cela il est nécessaire de convertir cette force du signal en distance. Ainsi, l'approche de trilatération consiste en deux étapes. Dans la première étape, l'intensité du signal est convertie en distance à l'aide d'un modèle de propagation du signal, puis dans la deuxième étape, l'emplacement est calculé à l'aide d'une méthode géométrique. (Safei, 2012). La figure 6 illustre la force du signal Wifi de chacun des trois points d'accès.

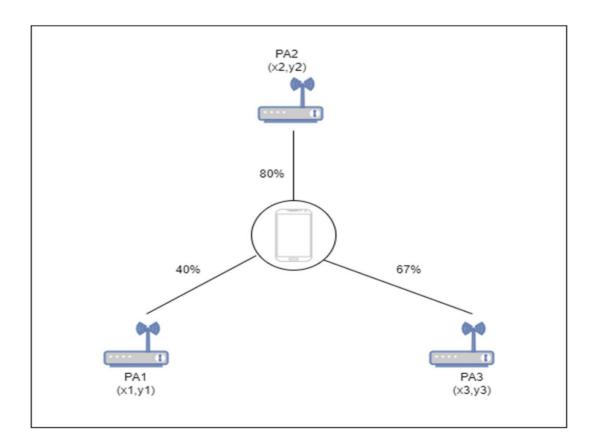


Figure 4 : Illustration de la force du signal Wifi à partir de 3 points d'accès.

© Oumarou Nouridine, 2022.

Pour mieux comprendre, voyons à travers un exemple comment la position d'un utilisateur peut être calculée en utilisant la technique de trilatération. La puissance du signal

représentée en pourcentage par l'application d'analyse Wi-Fi du *smartphone* peut être convertie en distance entre l'utilisateur et chaque point d'accès à l'aide de l'équation 2 (Safei, 2012).

$$d_i = p(1 - m_i) \tag{2}$$

Où *m* est le pourcentage de la puissance du signal, *p* la couverture maximale (soit 100m) de la puissance du signal, et *i* prend les valeurs 1, 2 et 3. Une fois les distances di calculées, les trois cercles de rayons respectives d1, d2 et d3 peuvent être dessinés. L'intersection de ces trois cercles est le lieu géométrique où se situe la position de l'utilisateur. La figure 7 illustre l'intersection des 3 cercles déterminant la position de l'appareil.

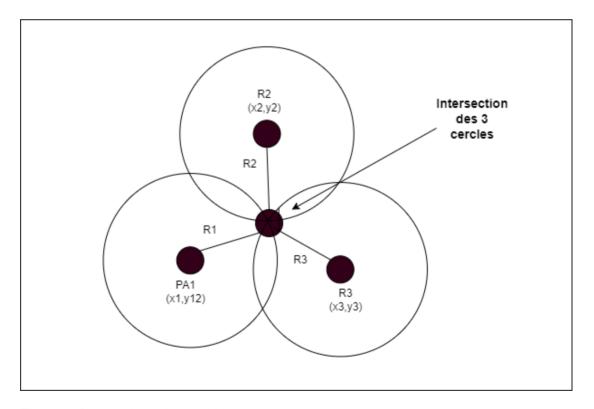


Figure 5: Intersections de 3 cercles.

© Oumarou Nouridine, 2022.

L'équation 3 indique l'équation d'un cercle dans le plan cartésien (Z=0).

$$(X - X_i)^2 + (Y - Y_i)^2 = R_i^2$$
 (3)

Avec  $(X_i, Y_i)$  les coordonnées des points d'accès, (X, Y) celles de l'utilisateur et  $R_i$  la distance séparant le *smartphone* d'un point d'accès. On peut enfin calculer la position en résolvant le système d'équations linéaires constitué des trois équations du cercle. La résolution du système linéaire déterminera les coordonnées X et Y de l'utilisateur. Il est important de souligner que le pourcentage du RSSI est plus important lorsque l'appareil est proche du point d'accès. Cependant, il convient de noter que la présence d'obstacles peut détériorer le signal et faire baisser le pourcentage.

# 4.1.1.1 LES CONTRAINTES DE LA TRILATÉRATION DANS LE POSITIONNEMENT INTÉRIEUR

Pour appliquer la trilatération, les coordonnées de trois points d'accès sont indispensables ainsi que leurs distances au téléphone mobile. Donc pour avoir ces coordonnées une bonne connaissance de la géométrie du bâtiment est nécessaire. Si le détail de l'environnement ne peut être obtenu facilement, une procédure supplémentaire, dite procédure d'apprentissage, est ajoutée. Cette procédure consiste à trouver la relation entre la distance de séparation PA-AMU en collectant des données (force du signal) en certains points de repères de coordonnées connues (Li et al., 2008). Ce qui constitut un travail supplémentaire et requière du temps.

# 4.1.1.2 LES CONTRAINTES DE LA TRILATÉRATION DANS LE POSITIONNEMENT EXTÉRIEUR

Dans le positionnement extérieur, compte tenu de la taille de l'espace de couverture, contrairement au positionnement intérieur où l'on parle d'un seul bâtiment, la détection des points d'accès (principalement des points d'accès dans les infrastructures fixes) et la mesure

de leurs coordonnées est le préalable et le plus grand défi pour appliquer la trilatération, en plus obtenir une carte complète de la zone de couverture est difficile et demande beaucoup de travail.

#### 4.1.2 LE PRINCIPE DE L'EMPREINTE DIGITALE

Le principe de l'empreinte digitale consiste à identifier un lieu en fonction des paramètres (dans le cas du Wifi, la force du signal) d'un ou plusieurs points d'accès présents dans ce lieu, tout comme l'empreinte digitale humaine permet d'identifier une personne (Li et al., 2008). L'idée principale de l'approche par empreintes digitales est de collecter le pourcentage de signaux radio des points d'accès accessibles aux zones d'intérêt, c'est-à-dire de connaître la valeur de l'intensité du signal de certains points d'accès à des points de repères de coordonnées connus pour ensuite les comparer à ceux collectées par le *smartphone* de l'utilisateur. Cette approche compte deux étapes (Li et al., 2005). Une première qui consiste en l'établissement de la base de données d'empreintes de localisation dont chaque enregistrement établit une correspondance entre une coordonnée et des mesures d'intensité de signal d'un certain AP, et la seconde qui consiste à déduire la localisation de l'utilisateur en comparant les pourcentages d'intensité de signal des points d'accès accessibles collectés par l'analyseur Wifi du *smartphone* avec les valeurs qui se trouvent dans la base de données d'empreintes digitales à l'aide d'un algorithme. La figure 7 montre les deux phases de l'approche de l'empreinte digitale.

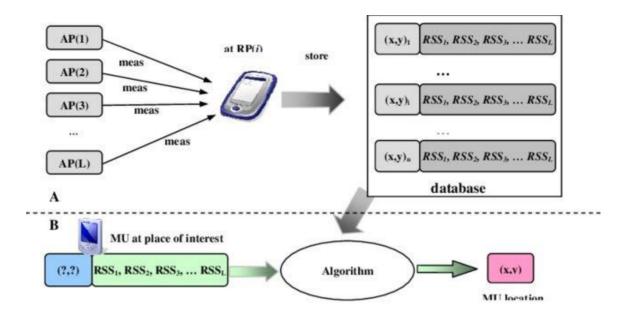


Figure 6 : les deux phases de l'approche de l'empreinte digitale (Li et al., 2008). Image reproduite avec l'autorisation de l'auteur.

Deux types d'algorithmes sont généralement utilisés pour calculer la position en utilisant l'approche des empreintes digitales. La première catégorie et la plus simple est la méthode déterministe qui se base sur un calcul de distance. Pour cela, on peut utiliser, l'algorithme du plus proche voisin (NN nearest neighbour) qui consiste à calculer la différence entre le vecteur d'intensité du signal mesuré par le *smartphone* et chacun des vecteurs de la base de données (Bahl & Padmanabhan, 2000). Le voisin le plus proche est le point de repère avec la distance la plus courte. Par exemple si le vecteur de force du signal mesuré est [s1, s2, ... sn] et l'un des vecteurs de force du signal dans la base de données est [S1, S2, ... Sn], la distance est calculée selon la formule des moindres carrés de l'équation 4.

$$L_q = \left(\sum_{i=1}^n \left| S_i - S_i \right|^q \right)^{\frac{1}{q}} \tag{4}$$

Manhattan (q=1) et la distance euclidienne (q=2) sont les mesures de distance les plus courantes. Cependant, l'augmentation de la valeur de q n'améliore pas la précision de la localisation (Li et al., 2005). On peut aussi parler de kNN (k nearest neighbour) Si k≥2, dans ce cas on choisit les k plus proches voisins (ceux dont la distance euclidienne est la plus courte). Cette approche semble intuitivement meilleure que le NN car on prend en compte tous les voisins proches et ne se contente pas de choisir le plus proche et d'abandonner les autres. La position de l'utilisateur est la moyenne des coordonnées de k points (Li et al., 2005).

La deuxième catégorie est l'approche probabiliste basée sur un modèle qui décrit la distribution de puissance du signal reçu à divers endroits. Cette approche a été proposée afin d'augmenter la précision de positionnement en tenant compte de la variation de la mesure de l'intensité du signal (Roos et al., 2002). Cependant, la force du signal n'est pas gaussienne et varie d'un endroit à l'autre ou au même endroit lorsque l'orientation de l'antenne change (Li et al., 2008). Il faut alors effectuer plusieurs mesures au même endroit et dans toutes les directions. Ce qui augmente le temps de génération de la distribution SS à chaque RP.

Une troisième méthode basée sur l'apprentissage automatique est désormais possible avec la démocratisation des réseaux de neurones. Cette fois, la base de données sera plutôt formée avec un algorithme d'intelligence artificielle comme les machines à vecteurs de support (en anglais support-vector machine, SVM) pour prédire l'emplacement à partir du modèle au lieu d'une méthode déterministe ou probabiliste (Abidin et al., 2021).

# 4.1.2.1 LES CONTRAINTES DE L'EMPREINTE DIGITALE DANS LE POSITIONNEMENT INTÉRIEUR

L'approche par empreintes digitales a été acceptée comme une méthode efficace pour le positionnement intérieur basé sur le Wifi, malgré certains inconvénients, dont le plus

important est le temps de calcul élevé, car il nécessite la distance entre les données collectées par le *smartphone* de l'utilisateur et celles de chaque point de repère. La génération de la base de données et les besoins de maintenance sont également un inconvénient de cette approche car à chaque fois que le détail du bâtiment change (travaux de réhabilitation) il sera nécessaire de régénérer de nouvelles données.

# 4.1.2.2 LES CONTRAINTES DE L'EMPREINTE DIGITALE DANS LE POSITIONNEMENT EXTÉRIEUR

En positionnement extérieur, avec autant de points d'accès, les algorithmes ne peuvent pas toujours trouver le vrai voisin car il est très difficile de retrouver les mêmes séquences et nombre de points d'accès à chaque scan pour pouvoir calculer la distance. De plus, avec la couverture qui est grande pour avoir une bonne précision, il faudra assurer une forte densité de points d'accès et le temps de calcul de la position est d'autant plus élevé qu'il y a plus de points de référence (Li et al., 2005). Enfin, le plus gros problème est la génération de la base, car mesurer les empreintes digitales dans une ville entière, par exemple, est un travail très difficile, si on ne dispose pas d'une infrastructure adaptée, compte tenu de l'étendue de la zone de couverture. Toutes ces raisons rendent cette méthode assez couteuse pour le positionnement externe.

# 4.2 PROBLEMES LIES AUX MÉTHODES BASÉES SUR LA FORCE DU SIGNAL

L'estimation de localisation avec Wifi utilise des techniques basées sur la force du signal (FS) reçu. Ces techniques sont essentiellement divisées en deux catégories. L'un utilise un modèle de propagation du signal et des informations sur la géométrie de la zone pour convertir l'intensité du signal en une mesure de distance. Ensuite, la trilatération est utilisée

pour calculer l'emplacement de l'AMU. L'autre catégorie de positionnement Wifi est l'empreinte digitale qui a reçu plus d'attention récemment (Li et al., 2005).

Chacune des méthodes a ses propres inconvénients. Cependant, il existe également de nombreuses contraintes liées à la mesure de l'intensité du signal qui représente la principale caractéristique sur laquelle reposent toutes ces méthodes. En effet, plusieurs facteurs peuvent influencer la mesure de l'intensité du signal. Les plus importants sont les phénomènes de trajets multiples du signal (en anglais multipath propagation) et d'erreur de propagation sans ligne de visée (en anglais non line of sight, NLOS), tous introduits par des obstacles tels que des bâtiments, des arbres, des interférences, etc., (Li et al., 2008).

La ligne de visée (en anglais line of sight, LOS) est la condition dans laquelle les ondes radio se propagent en ligne droite entre l'émetteur et le récepteur sans aucun obstacle entre eux. Cependant, si un objet obstrue la ligne entre l'émetteur et le récepteur, le terme propagation sans ligne de visée est utilisé comme indique la figure 8.

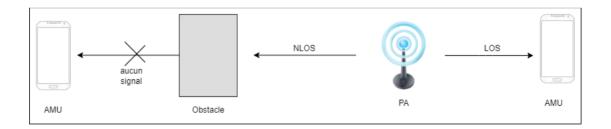


Figure 7 : Illustration LOS ET NLOS (Taboubi, 2014).
Image reproduite avec l'autorisation de l'auteur.

Le phénomène de trajets multiples décrit le phénomène qui se produit lorsque les signaux radio empruntent plusieurs chemins pour atteindre le récepteur en raison des réflexions et des absorptions subies par le signal dues aux obstacles (murs, escaliers, sols, etc.) comme sur la figure 9.

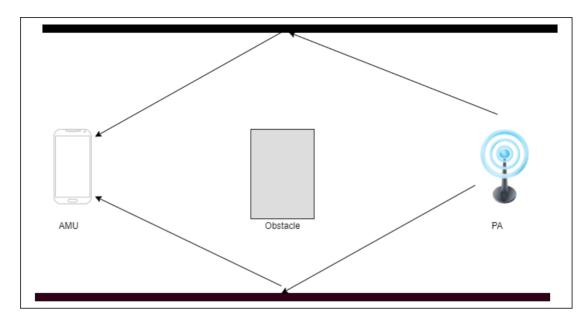


Figure 8 : Illustration du phénomène du trajet multiple (Taboubi, 2014). Image reproduite avec l'autorisation de l'auteur.

L'orientation de l'antenne de l'appareil influence également la mesure de l'intensité du signal. C'est-à-dire que lorsque l'utilisateur change d'orientation, la FS reçu peut changer de manière significative (Li et al., 2008). Ainsi, mesurer la FS dans toutes les directions améliorerait la qualité de la base de données et donc augmenterait la précision du positionnement. Dans ce cas contrairement à la méthode traditionnelle où la FS est mesurée dans une seule direction, cette fois la FS est mesuré à chaque endroit dans les quatre directions pour constituer la base de données afin de prendre en compte la variation du signal selon la direction de l'antenne (Li et al., 2008). D'autre part, cela augmente la taille de la base de données et la charge de calcul.

Un des facteurs influençant également la mesure de la force du signal est le type de *smartphone* (Carreras et al., 2012). Ainsi l'utilisation de plusieurs types de *smartphones* lors de la phase de construction de la base de données permettrait d'en améliorer la précision.

Cependant, cette approche augmente également le coût et le temps de construction de la base ainsi que sa taille.

Ainsi on remarque que le RSSI est un paramètre très variable, donc les solutions basées sur ce paramètre peuvent être imprécises. Ces techniques basées sur le RSSI semblent plus adaptées au positionnement indoor où l'on parvient à avoir une précision inférieure à deux mètres (Roos et al., 2002). Cependant, dans le positionnement extérieur, d'autres défis restent à relever, et ont conduit à associer le Wifi au GPS. Cette nouvelle approche sera discutée dans la section suivante. Cependant, le grand avantage des techniques basées sur RSSI (Wifi) est que nous pouvons utiliser l'infrastructure existante pour déployer un système de positionnement sans dispositifs supplémentaires.

# 4.3 WIFI ASSOCIÉ AU GPS DANS LE POSITIONNEMENT EXTÉRIEUR

Le positionnement d'un objet ou d'une personne détermine sa localisation dans un référentiel. Dans un système de référence comme la planète, l'emplacement est déterminé par des coordonnées globales, à savoir la longitude et la latitude. Par conséquent, tout système de positionnement absolu à l'échelle terrestre doit être en mesure de fournir ces informations. Cependant, le Wifi ne peut fournir qu'une mesure de la force du signal. Même si l'utilisation d'un modèle de propagation de signal peut convertir la force du signal en distance PA-AMU, cette distance est très imprécise en raison des problèmes de propagation du signal radio mentionnés ci-dessus et n'est pas adaptée au positionnement extérieur à grande échelle qui, lui a besoin de coordonnées globales.

Ainsi, dans le système de positionnement extérieur basé sur le Wifi, outre les difficultés qui ont été discutées, il y a un autre problème à considérer, à savoir la coordination des points d'accès. Une façon d'acquérir ces informations serait de les demander directement aux propriétaires de ces points d'accès. Cependant, la majorité des points d'accès sont

déployés par des entreprises de télécommunications dans le cadre de leur infrastructure fixe, qui ne donneront pas ces détails gratuitement ou ne veulent pas du tout que ces informations soient publiques. Les autres points d'accès sont établis par des utilisateurs privés pour un usage personnel. Cependant, les propriétaires de *hotspots* privés, le plus souvent, ne connaissent pas ces coordonnées (Li et al., 2008).

Alors compte tenu de la difficulté d'obtenir les coordonnées des points d'accès auprès des propriétaires et de l'incapacité de la technologie Wifi à les fournir, il est évident que les seules données Wifi ne peuvent pas être utilisées dans le positionnement extérieur à grande échelle. Il faut donc d'autres méthodes de positionnement préalable pour obtenir les coordonnées globales des PAs. À l'extérieur, la probabilité de recevoir un signal en visibilité directe d'au moins un satellite GPS est assez élevé. Il est alors possible d'intégrer Wifi et GPS pour estimer la position de l'AMU. Ainsi, comment intégrer Wifi et GPS est un problème intéressant à étudier.

## 4.4 MÉTHODOLOGIE

L'utilisation du Wifi comme moyen de localisation se justifie par la maturation des technologies Wifi et les capacités croissantes du matériel de nos appareils intelligents ainsi que le nombre croissant de points d'accès dans les villes où des études montrent que dans certaines grandes ville 92% de toutes les analyses Wifi détectent au moins un point d'accès et 33% détectent plus de 10 points d'accès et dans les zones densément peuplées, une moyenne de 25 points d'accès sont visibles dans chaque analyse. Cette utilisation s'explique aussi par la faible entropie de la mobilité humaine (Sapiezynski et al., 2015).

Cependant, comme mentionné précédemment, le positionnement absolu nécessite la connaissance de la longitude et de la latitude de l'objet, des informations qui ne peuvent pas être obtenues à partir du Wifi, donc une combinaison de ce dernier est faite avec le GPS

pour améliorer la localisation extérieure. De grandes entreprises telles que Google, Apple, Microsoft ou Skyhook combinent des points d'accès Wifi avec des données GPS pour améliorer le positionnement des utilisateurs de *smartphones*. Bien que largement utilisé, l'utilité et les mécanismes exacts sont largement inconnus, seules des études étroites et non systématiques sont rapportées dans la littérature (Sapiezynski et al., 2015). Par conséquent, on ne sait généralement pas comment les réseaux Wifi peuvent être utilisés pour détecter l'emplacement à l'échelle de la société, cette connaissance appartient aux grandes entreprises.

Ainsi notre objectif est de proposer un système de positionnement à l'échelle de la société basé sur le Wifi en utilisant les méthodes des études systématiques disponibles sur le sujet. Ces études utilisent généralement une approche en deux étapes. La première phase consiste en la coordination des points d'accès en cartographiant les points d'accès avec les coordonnées GPS des lieux où se trouvent ces points d'accès. L'association des points d'accès et des positions géographiques nous servira de base de données, cette étape s'appelle la phase offline. Puis la deuxième phase dite phase online consiste à déduire le positionnement de l'utilisateur à partir de la base de données établie dans la phase offline (Sapiezynski et al., 2015).

#### 4.4.1 LA PHASE OFFLINE

L'utilisation du Wifi est partie du constat que le dispositif Wifi des *smartphones* envoie en permanence des requêtes dites probes afin de trouver des points d'accès Wifi à proximité, ce qui implique que l'utilisateur du *smartphone* se trouve dans la zone de couverture du Wifi (ou Wifi) qui lui est accessible. D'où la possibilité de déduire la position de l'utilisateur de ces points d'accès Wifi à condition de connaître leurs coordonnées.

Les *smartphones* étant désormais tous équipés d'un GPS, pour attribuer une position à un point d'accès nous allons effectuer une estimation de localisation GPS et un scan Wifi en continu sur des intervalles de temps réguliers. Si le balayage se produit quelques secondes après la position GPS estimée, nous pouvons attribuer cette position aux points d'accès qui ont été détectés (Sapiezynski et al., 2015). Par exemple, si la localisation GPS est produite à l'instant  $T_{gps}$  et le scan Wifi à l'instant  $T_{scan}$  le décalage horaire  $diff = |T_{gps} - T_{scan}|$  si diff est inférieur ou égal à epsilon (dans notre cas 5 secondes), alors ces points d'accès scannés sont considérés comme étant à cet emplacement. En effet, la position est attribuée uniquement aux points d'accès avec un RSS élevé, c'est-à-dire inférieur à -60% afin d'augmenter la précision. Ce pourcentage peut encore être augmenté afin d'avoir une meilleure précision. Ces données sont ensuite enregistrées dans une base de données pour être utilisées ultérieurement pour calculer le positionnement de l'utilisateur (Sapiezynski et al., 2015).

# 4.4.2 LA PHASE ONLINE

Une fois la base de données construite, on passe à la phase en ligne, qui est la phase durant laquelle la position de l'utilisateur est calculée. Pour calculer la position de l'utilisateur, on peut penser à utiliser l'une des méthodes vues précédemment, à savoir la trilatération ou la prise d'empreintes digitales. Cependant, ces deux techniques sont peu adaptées au positionnement externe à grande échelle. Par exemple, l'approche par empreintes digitales ne nécessite pas de coordonnées absolues et se base sur la mesure du RSSI qui est un paramètre très variable qui peut fausser le calcul de positionnement. De plus, mesurer les empreintes digitales dans une ville entière est un travail très exhaustif. Sans compter la difficulté à trouver le bon voisin et le temps de calcul de position élevé dû au très grand nombre de points de repères. Pour la trilatération, nous avons besoin des coordonnées d'au moins trois points d'accès ainsi que de la distance du mobile par rapport à ceux-ci. Cependant, le calcul de cette distance peut être faussé en raison de problèmes liés à la

propagation du signal. De plus les coordonnées des points d'accès ne sont pas les coordonnées réelles de ces points d'accès mais plutôt des coordonnées approximatives et étant donné qu'une même position est attribuée à plusieurs points d'accès il y a un risque d'avoir trois AP confondus ce qui montre le manque de fiabilité du calcul de position.

Nous avons donc utilisé une méthode assez simpliste qui consiste à attribuer la position du PA avec la valeur RSSI élevée à l'utilisateur sachant que la position de ce PA est connue à l'avance. Car malgré la variation de la mesure du RSSI, nous sommes sûrs d'une chose, à savoir que si la valeur du RSSI du Wifi est élevée alors l'utilisateur en est proche (Safei, 2012). De plus, comme on est dans un système de positionnement extérieur à grande échelle, on sera toujours dans une précision acceptable qui est la couverture maximale d'un point d'accès Wifi, soit 100m correspondant à celle du système actuel, à savoir le GPS.

#### **CHAPITRE 5**

# IMPLÉMENTATION DU MODÈLE PROPOSÉ

Dans les chapitres précédents, nous avons montré les différentes méthodes de positionnement basées sur le Wifi, utilisé à l'intérieur et à l'extérieur. Nous avons également étudié leurs caractéristiques, spécificités et contraintes, puis nous avons proposé un modèle applicable à grande l'échelle. Dans ce chapitre, nous décrirons comment ce modèle a été mis en œuvre. Tout d'abord, nous commencerons par présenter Android qui est la plateforme sur laquelle le modèle a été implémenté. Le choix est porté sur Android en raison de sa forte utilisation sur le marché des *smartphones*. De plus, sa forte communauté de développeurs et sa bonne documentation facilitent la recherche de solution. Toutes ces raisons expliquent l'utilisation d'Android dans le domaine académique. Ensuite nous présenterons les outils de développement utilisés. La fin du chapitre sera consacrée à la présentation des résultats et aux discussions.

#### **5.1 ANDROID**

Android est un système d'exploitation pour les périphériques mobiles (téléphones, tablettes, montres etc.). Il est fondé sur un noyau linux, et racheté par Google en 2007. Au début le système a été conçu pour les caméras numériques par la société Android Inc., en 2003 avant d'être réadapté pour les *smartphones* en 2004. Cependant depuis le rachat de la société Android Inc., par Google en 2005 le système est supporté en dehors des mobiles et tablettes par divers appareils électroniques comme les télévisions (Android TV), les automobiles (Android Auto), les montres intelligents (Wear OS), les satellites etc. En effet, Android est gratuit et open source (sous licence apache), ce qui permet aux constructeurs qui intègrent le système de pouvoir le modifier et se distinguer de la concurrence. Cette forte adoption du système Android par divers appareils a permis à Android d'être le système d'exploitation mobile le plus utilisé dans le monde, devant le système iOS d'Apple, avec plus

de 80 % de parts de marché dans les *smartphones* pour l'ensemble de ses versions et adaptations (IDC, 2021). Le langage java est le langage de développement utilisé dans notre projet pour sa forte communauté et étant le langage de base du système Android. Dans le guide du développeur, Android est défini comme étant une pile de logiciels, c'est-à-dire un ensemble de logiciels destinés à fournir une solution clé en main pour les appareils mobiles *smartphones* et tablettes tactiles. Ainsi, Android met à la disposition un kit de développement logiciel et de nombreuses librairies facilitant le développement.

## **5.2 OUTILS UTILISÉS**

Dans la réalisation du projet, nous avons dû utiliser différents outils tels que des cadriciels (en anglais frameworks), des algorithmes et des API afin de faciliter le développement et de gagner du temps.

#### 5.2.1 Google Maps

Google Maps<sup>7</sup> est un service mondial de cartographie en ligne créé par Google. Il a été lancé en 2005 et couvre dans un premier temps les USA et le Canada, puis l'Angleterre avant d'être proposé simultanément en France, en Allemagne, en Italie et en Espagne le 25 avril 2006. Aujourd'hui le service est proposé sur toute la surface du globe et est disponible sur le navigateur et application, sur PC et *smartphones*. Les données cartographiques de Google Maps sont utilisées par des millions de sites Web et d'applications mobiles. Pour accéder au service, les applications et les sites Web utilisent l'API proposée par Google. L'API n'est pas gratuite, mais les applications obtiennent un crédit de 200 \$ au début pour essayer le service.

-

<sup>&</sup>lt;sup>7</sup> https://developers.google.com/maps?hl=en

#### **5.2.2 FIREBASE FIRESTORE**

Firebase est un ensemble de services d'hébergement accessible sur internet lancé en 2011 par Andrew Lee et James Templin sous le nom d'Envolve puis racheté par Google en octobre 2014. L'objectif est de faciliter la création d'applications web et mobiles pour tout type de plateforme (iOS, Android, C++, Java, javascript, Php, C#, NodeJs, etc.) en soulageant les entreprises et les développeurs de la complexité de création et de maintenance du backend et en leur fournissant un ensemble de services évolutifs et disponibles à tout moment. Les principaux services sont entre autres les services de base de données NoSQL comme FireStore et Realtime pour les données au format texte, le service d'hébergement de contenu multimédia Firebase Storage, le service Firebase Analytic pour avoir un aperçu de l'utilisation de l'application et de l'engagement des utilisateurs, le service Firebase Cloud Messaging qui est une solution multiplateforme (Android, iOS et Web) pour les messages et les notifications, le service Firebase Auth qui permet d'authentifier les utilisateurs en utilisant uniquement le code client (soit avec Facebook, GitHub, Twitter, téléphone et Google etc.). Pour utiliser ces services, les applications doivent intégrer une bibliothèque permettant les différentes interactions possibles selon le service à utiliser.

Dans le cadre de notre projet, nous avons utilisé le service de base de données Firestore qui est une base de données NoSQL orientée document. Firestore sauvegarde les données sous forme de collections de documents. Les données simples sont stockées dans des documents, qui sont très en fait un JSON, et les données complexes et hiérarchiques sont organisées à l'aide de collections ou de sous-collections de documents. L'avantage de Firestore est qu'il supporte des données complexes et une grande quantité de données, De plus le service est gratuit jusqu'à 1Gib, ce qui correspond parfaitement à nos besoins (https://firebase.google.com/products/firestore).

#### **5.2.3 KAFKA**

Apache Kafka<sup>8</sup> est une plateforme de diffusion d'événements open source écrite en scala, développée par LinkedIn en 2011. Conçu à l'origine comme une file d'attente de messagerie, Kafka est maintenant utilisé dans les entreprises comme un système d'échange d'informations avec l'avènement des architectures de micro-services. Kafka est capable de publier, stocker, traiter et souscrire à des flux d'enregistrement en temps réel. L'un des avantages de Kafka est qu'il peut non seulement fournir des flux de données d'une source à un utilisateur mais également de plusieurs sources à plusieurs autres utilisateurs. Kafka fournit aux applications qui souhaitent intégrer des API de base réparties en cinq catégories : L'API Producer qui permet aux applications d'envoyer des flux de données vers les sujets du cluster Kafka, l'API Consumer qui permet aux applications de lire les flux de données des sujets du cluster Kafka, les l'API Streams qui lui permet de transformer des flux de données en sujets de sortie, l'API Connect qui permet d'implémenter des connecteurs qui récupèrent des données d'un système source ou d'une application vers Kafka ou qui poussent de Kafka vers une application et enfin l'API d'administration ou AdminClient API qui vous permet de gérer et d'inspecter les sujets (ou topics), les courtiers (ou brokers) et d'autres objets Kafka. Dans notre projet, Kafka est utilisé comme pipeline lors de l'enregistrement des points d'accès.

#### **5.2.4 SPRINGBOOT**

Spring est un cadriciel open source pour les applications java, dont il facilite le développement et le test en fournissant aux développeurs un ensemble de modules prêts à l'emploi avec lesquels ils construiront leurs infrastructures applicatives. Cependant, l'utilisation de chaque module nécessite une configuration qui devient un casse-tête au fur et

<sup>8</sup> https://kafka.apache.org/

à mesure qu'on en rajoute. C'est alors que Springboot<sup>9</sup> a été créé pour faciliter la création d'applications Spring autonomes, de qualité de production avec un serveur embarqué pouvant être démarré facilement et incluant un système de configuration automatique. Dans ce projet Springboot est utilisé pour implémenter le backend de l'application.

#### **5.2.5 HEROKU**

Heroku<sup>10</sup> est une plate-forme cloud en tant que service (PaaS) qui permet aux entreprises et développeurs particuliers de déployer, de gérer et de mettre à l'échelle des applications. Heroku propose également des services de base de données, une plateforme de diffusion d'événements comme Kafka et prend en charge plusieurs langages de programmation comme Ruby, Java, Node.js, Scala, Clojure, Python, PHP et Go. Heroku est utilisé dans notre projet pour héberger notre application backend Springboot. La procédure d'hébergement nécessite d'abord la création d'un compte, puis de créer une application sur la plateforme heroku (wifi-location1 pour ce projet) comme le montre la figure 10.

-

<sup>&</sup>lt;sup>9</sup> https://spring.io/projects/spring-boot

<sup>10</sup> https://www.heroku.com/

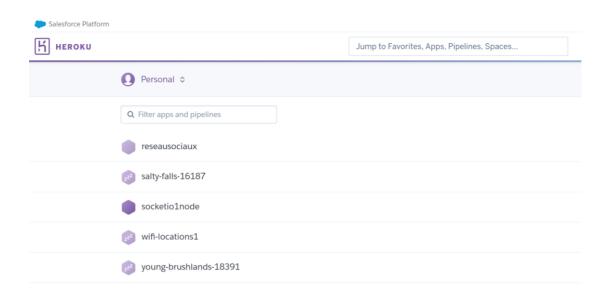


Figure 9 : création de l'application sur heroku.

Ensuite, vous devrez envoyer le projet Springboot déjà hébergé sur GitHub sur la branche Master créée par heroku. Pour cela, vous devrez installer heroku-cli afin de taper les commandes<sup>11</sup> nécessaires, comme indiqué dans la figure 11.

 $<sup>^{11}\ {\</sup>rm https://devcenter.heroku.com/articles/deploying\text{-}spring\text{-}boot\text{-}apps\text{-}to\text{-}heroku}$ 

```
C:\WINDOWS\system32\cmd.exe - ait push heroku master
Microsoft Windows [version 10.0.22579.1]
(c) Microsoft Corporation. Tous droits réservés.
::\Users\nouri>C:\Users\nouri\IdeaProjects\dbscan
C:\Users\nouri\IdeaProjects\dbscan' n'est pas reconnu en tant que commande interne
ou externe, un programme exécutable ou un fichier de commandes.
:\Users\nouri>cd C:\Users\nouri\IdeaProjects\dbscan
C:\Users\nouri\IdeaProjects\dbscan>heroku login
neroku: Press any key to open up the browser to login or q to exit:
Opening browser to https://cli-auth.heroku.com/auth/cli/browser/e3ce595f-ac7f-401f-a6fc-72fb96d506
tE4HRqJs-U
ogging in... done
ogged in as nouridine27041998@gmail.com
:\Users\nouri\IdeaProjects\dbscan>git checkout master
Switched to branch 'master'
:\Users\nouri\IdeaProjects\dbscan>git commit -am "deploy on heroku"
On branch master
nothing to commit, working tree clean
C:\Users\nouri\IdeaProjects\dbscan>git push heroku master
Enumerating objects: 31, done.
Counting objects: 100% (31/31), done.
Delta compression using up to 8 threads
Compressing objects: 100% (14/14), done.
Writing objects: 100% (17/17), 1.33 KiB | 1.33 MiB/s, done.
Total 17 (delta 8), reused 0 (delta 0), pack-reused 0
remote: Compressing source files... done.
remote: Building source:
emote:
remote: ----> Building on the Heroku-20 stack
remote: ----> Using buildpack: heroku/java
emote: ----> Java app detected
emote: ----> Installing JDK 1.8... done
remote: ----> Executing Maven
remote: $ ./mvnw -DskipTests clean dependency:list install
emote:
               [INFO] Scanning for projects...
               [INFO]
emote:
emote:
                [INFO]
                              -----dbscan >-----
               [INFO] Building dbscan 0.0.1-SNAPSHOT
emote:
               [INFO] -----[ jar ]-----
emote:
```

```
maven-jar-plugin:3.2.0:jar (default-jar) @ dbscan
                [INFO] Building jar: /tmp/build_ce1b0078/target/dbscan-0.0.1-SNAPSHOT.jar
remote:
remote:
                INFO
                       --- spring-boot-maven-plugin:2.6.2:repackage (repackage) @ dbscan ---
remote:
                [INFO]
                      Replacing main artifact with repackaged archive
remote:
                [INFO]
remote:
                [INFO]
                       --- maven-install-plugin:2.5.2:install (default-install) @ dbscan
remote:
                INFO]
                [INFO] Installing /tmp/build ce1b0078/target/dbscan-0.0.1-SNAPSHOT.jar to /tmp/codon/
remote:
                [INFO] Installing /tmp/build_ce1b0078/pom.xml to /tmp/codon/tmp/cache/.m2/repository
remote:
remote:
                [INFO]
                [INFO] BUILD SUCCESS
remote:
remote:
                [INFO]
                [INFO] Total time: 6.315 s
remote:
remote:
                [INFO] Finished at: 2022-03-21T02:20:25Z
remote:
                [INFO]
               Discovering process types
remote:
               Procfile declares types
                                            -> (none)
remote:
               Default types for buildpack -> web
remote:
remote:
               Compressing...
remote:
               Done: 105.3M
remote:
remote:
               Launching...
               Released v5
remote:
               https://wifi-locations1.herokuapp.com/ deployed to Heroku
remote:
remote:
remote: Verifying deploy... done.
To https://git.heroku.com/wifi-locations1.git
   1920c36..189d75f master -> master
C:\Users\nouri\IdeaProjects\dbscan>
```

Figure 10 : Déploiement de l'application Springboot sur heroku.

#### **5.2.6 CONFLUENT**

Confluent¹² est une plate-forme basée sur le cloud qui fournit une distribution complète de Kafka, développée par les développeurs originaux d'Apache Kafka. Confluent Platform améliore Kafka avec des fonctionnalités communautaires et commerciales supplémentaires conçues pour améliorer l'expérience de streaming pour les opérateurs et les développeurs en production, à grande échelle. Nous avons utilisé confluent pour avoir une distribution Kafka sans se soucier de la gestion d'un serveur. De plus confluent offre 300\$ utilisables en trois mois pour tester le service. Nous avons donc d'abord créé un compte, puis créé un cluster Kafka nommé "cluster0" ayant un seul topic (considérez-le comme un journal vers lequel les messages seront envoyés et lus) nommé "accesspointTopic" comme indiqué dans les figures 12 et 13.

<sup>12</sup> https://confluent.cloud/

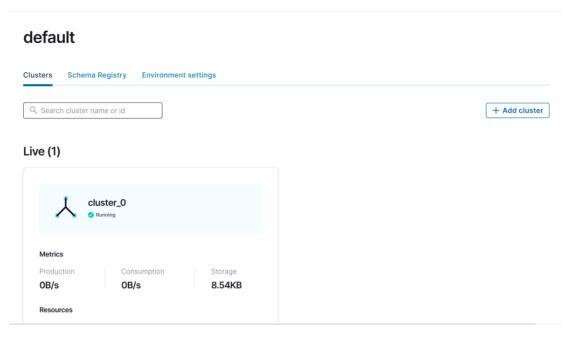


Figure 11 : Création du Cluster Kafka.

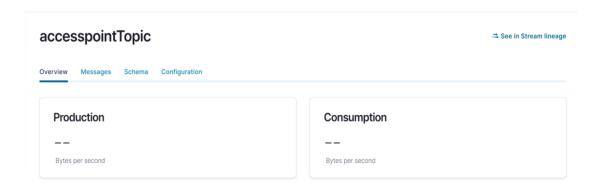


Figure 12 : Création du Topic Kafka.

Une fois le cluster et le topic créés, il faudra l'intégrer à l'application Springboot en suivant la procédure d'intégration du client Springboot proposée par la plateforme confluent comme indiquée sur la figure 14.

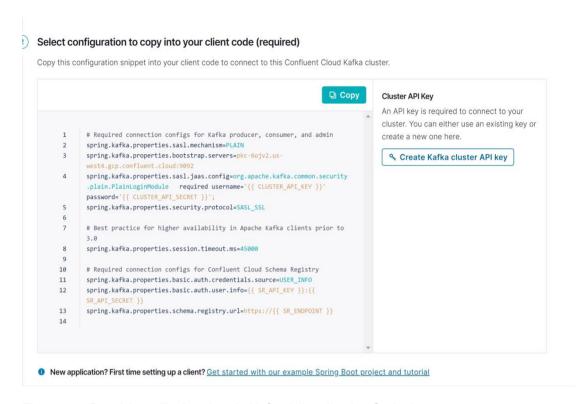


Figure 13 : Procédure d'intégration de Kafka à l'application Springboot.

#### **5.2.7 DBSCAN**

DBSCAN signifie Density-Based Spatial Clustering of Applications with Noise et est un algorithme de clustering non supervisé proposé par Martin Ester, Hans-Peter Kriegel, Jörg Sander et Xiawei Xu en 1996. L'algorithme DBSCAN se base sur deux paramètres pour former les groupes, la distance D correspondant au rayon de la zone de couverture de chaque groupe et le minimum de points Mpts que chaque cluster doit comporter. La formation des clusters se fait en parcourant la liste des points et pour chaque point on recherche ses voisins soit ceux qui sont à une distance inférieure ou égale à D. si le nombre de voisins atteint Mpts alors le point est considéré comme partie d'un cluster. On parcourt ensuite pas à pas chacun de ses voisins afin de trouver l'ensemble des points formant le groupe. La figure 15 illustre la formation des groupes suite à l'application de l'algorithme DBSCAN.

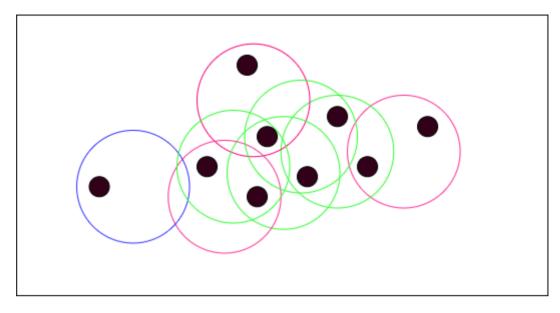


Figure 14: Illustration des clusters DBSCAN.

© Oumarou Nouridine, 2022.

L'algorithme comprend trois types de points : les points centraux qui ont un voisinage dense, les border points qui sont des points appartenant au voisinage d'un point central et enfin les points aberrants qui sont les autres points du jeu de données qui n'appartiennent à aucun cluster. Comme mentionné, le voisinage entre les points est déterminé mathématiquement par des méthodes de calcul de distance.

Cet algorithme a été utilisé pour détecter les points d'accès mobiles car il existe aujourd'hui plusieurs routeurs installés dans les transports en commun ou des *smartphones* avec des capacités de point d'accès. Ces types d'appareils ne peuvent pas être utilisés comme références et doivent être détectés et supprimés de nos données. On regroupe donc d'abord les points d'accès selon un critère de proximité, plus précisément ceux situés dans un rayon de 100 mètres qui correspond à la couverture d'un point d'accès Wifi, puis on enregistre ces groupes (en anglais clusters) dans une base de données. Donc quand un point d'accès est scanné on va récupérer le cluster auquel il appartient, s'il appartient à un

cluster on va calculer la distance entre ce point d'accès et ces voisins, si le nouveau cluster contient 75% de ces anciens voisins alors on en déduit qu'il est fixe sinon nous considérons qu'il s'agit d'un point d'accès mobile alors il est supprimé de notre base de données. Si ce point d'accès n'appartient à aucun groupe (ou est un point aberrant) nous calculerons la distance avec ses anciennes coordonnées si cette distance est inférieure à 200 mètres nous considérons qu'il est fixe, sinon nous le supprimons de notre base de données.

Pour les paramètres de l'algorithme, comme mentionné précédemment nous considérerons les points d'accès proches s'ils sont dans un rayon d=100 qui correspond à la zone de couverture d'un point d'accès Wifi et nos Mpts égal à trois. En ce qui concerne la distance, puisque nous ne sommes pas dans un système de positionnement à l'échelle de la société ou sur le globe terrestre, nous utiliserons la formule de *haversine* qui détermine la distance orthodromique entre deux points sur une sphère compte tenu de leurs longitudes et latitudes. Pour l'implémentation de l'algorithme DBSCAN nous avons adapté l'implémentation trouvée sur GitHub<sup>13</sup>.

### 5.3 ARCHITECTURE DE L'APPLICATION

L'application est réalisé selon l'architecture N-tiers plus précisément à trois niveaux. Dans l'architecture à trois niveaux, ou trois tiers), les trois couches applicatives sont séparées et s'exécutent sur des machines différentes, à savoir : Un client, c'est-à-dire le smartphone demandant des ressources, équipé d'une interface utilisateur chargée de la présentation. Le serveur d'application (également appelé middleware), chargé de fournir la ressource au client et le serveur de données, fournissant au serveur d'application les données dont il a besoin. L'architecture à trois niveaux permet une plus grande flexibilité et agilité, une sécurité accrue puisque la sécurité peut être définie indépendamment pour

13 https://github.com/chrfrantz/DBSCAN/

\_

chaque niveau, et enfin permet de meilleures performances compte tenu du partage des tâches entre les différentes parties. La figure 16 décrit l'architecture de l'application.

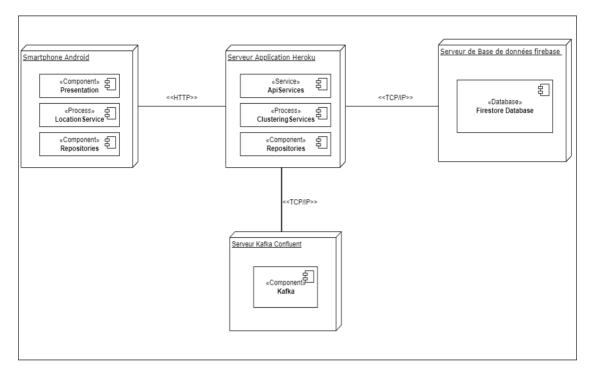


Figure 15 : Diagramme de déploiement de l'application.

© Oumarou Nouridine, 2022.

# **5.3.1 LA COUCHE PRÉSENTATION**

La couche de présentation est réalisée via l'application mobile Android qui fournit une interface à l'utilisateur pour effectuer des actions spécifiques. Il est composé de trois modules : le module de présentation, le module de localisation et le module de données ou *repositories*.

Le module de présentation se charge d'afficher les écrans et contient donc les différentes activités et fragments de l'application. C'est aussi cette partie qui s'occupe de récupérer les différents points d'accès environnants scannés par l'application d'analyse Wifi du *smartphone*. En effet, Android propose une API à partir de laquelle les applications

peuvent accéder à presque toutes les informations d'une connexion Wifi comme la vitesse de liaison du réseau connecté, l'adresse IP, l'état de la négociation, la force du signal reçu et d'autres informations réseau. Les applications peuvent également analyser, ajouter, enregistrer, mettre fin et initier des connexions Wifi via cette API.

Ce module s'abonne également au module de localisation pour avoir accès aux informations de localisation GPS. Chaque fois qu'un balayage Wifi est effectué, il vérifie le temps qui s'est écoulé avec la dernière position GPS. Si le temps est inférieur au temps souhaité, soit cinq secondes dans notre cas, il appelle le service API pour envoyer ces points d'accès au serveur d'application afin qu'il les ajoute à la base de données.

Le module de localisation est un service qui s'exécute en arrière-plan et donne au composant qui y est abonné l'accès à la dernière position GPS du *smartphone*. En effet Android propose une API qui permet à l'application disposant des autorisations nécessaires d'accéder aux données de localisation GPS.

Le module de données s'occupe de la communication avec les différentes API proposées par l'application Springboot. Pour effectuer les requêtes, nous avons utilisé la bibliothèque *Volley* qui est une bibliothèque http pour les applications Android développée par Google en 2013. Volley vous permet d'effectuer des requêtes réseau rapidement et facilement.

## **5.3.2 LA COUCHE MÉTIER**

La couche métier est représenté par l'application Springboot qui fournit une api a l'application cliente pour enregistrer par exemple les points d'accès scannés. Cette couche est composée de quatre modules distincts. Le premier est le module web qui offre des services d'api a l'application Android. Le deuxième module est le module de clustering et de

traitement continu qui implémente les services consommateur et producteur Kafka et enfin le dernier module est le module de données ou repositories qui envoie et récupère des données sur Firestore.

Lorsque le module Web reçoit une demande d'ajout de point d'accès, il appelle le service Producteur pour ajouter le point d'accès dans le pipeline. Puis lorsque le consommateur récupère ce point d'accès, il vérifie d'abord s'il ne s'agit pas d'un point d'accès mobile en vérifiant que 75% de ses anciens voisins sont toujours ses voisins. Si tel est le cas, il calcule la moyenne de ses nouvelles coordonnées avec ses anciennes coordonnées (afin d'améliorer progressivement la localisation du point d'accès) puis l'ajoute à la base de données. Si moins on en déduit que le point d'accès est mobile et est supprimé de la base de données.

L'utilisation de Kafka permet d'éviter les cas où plusieurs utilisateurs voudront ajouter le même point d'accès simultanément et de pouvoir détecter rapidement les points d'accès (PAs) mobiles. Mais surtout ne pas charger le serveur et permettre des vérifications pour chaque point d'accès avant de l'ajouter à la base de données.

Le module de clustering se charge de créer des clusters en récupérant tous les points d'accès via le module de données, puis d'appliquer le modèle de clustering à ces données. Cette tâche est effectuée par un Job Scheduler qui est planifié chaque jour. Cependant la fréquence peut être modifiée lorsque le volume de la base de données augmente beaucoup pour éviter de charger le serveur d'application de travail très lourd.

# **5.3.2 LA COUCHE DE DONNÉES**

La couche de données est le serveur de base de données, dans notre cas, elle est représentée par le service cloud Firestore de Firebase. La base de données contient trois collections. La première, la collection des points d'accès contient les points d'accès avec leurs coordonnées géographiques comme indiqué sur la figure 17.

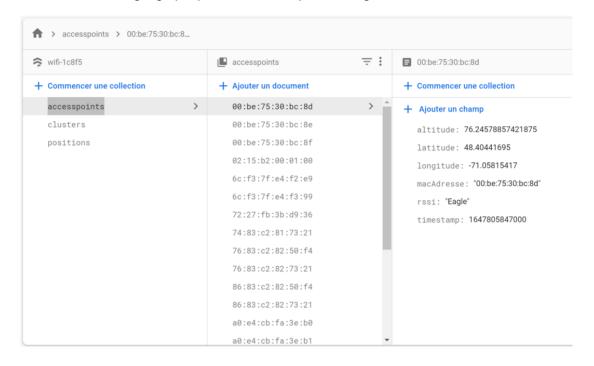


Figure 16 : Collection des points d'accès scannés.

La seconde est la collection de positions qui contient les informations de déplacements des utilisateurs. Chaque document de cette collection contient l'identifiant de l'utilisateur, l'heure de localisation, les coordonnées de localisation et les adresses mac des points d'accès balayés à cet emplacement. C'est de cette collection que l'on déduit l'historique des déplacements de l'utilisateur. La figure 18 montre cette collection.

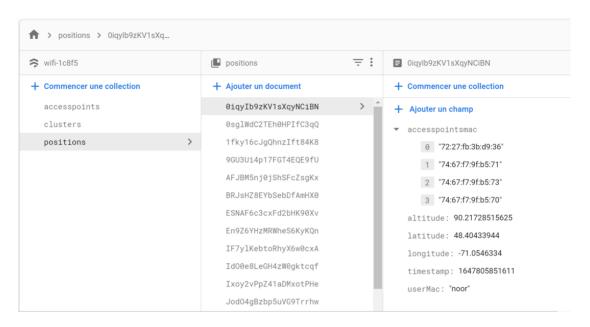


Figure 17 : Collections de l'historique de déplacement des utilisateurs.

La dernière est la collection clusters qui contient les différents clusters de notre jeu de données comme indiqué sur la figure 19.

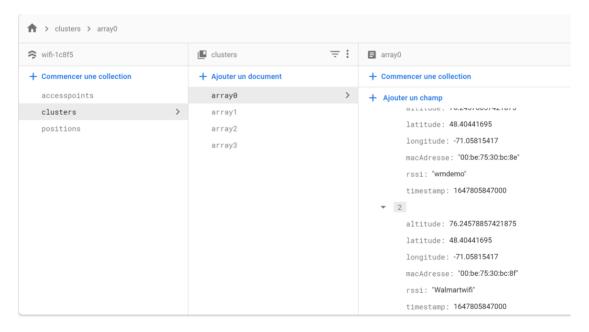


Figure 18: Collections des clusters.

## **5.4 PROCÉDURE DU TEST**

Pour tester l'application nous avons utilisé une tablette Samsung Galaxy Tab A et avons procédé en deux étapes, la première consistait à acquérir les données de point d'accès et de localisation, alors nous nous sommes déplacés vers les endroits les plus fréquentés et avons scanné les points d'accès qui s'y trouvent, en commençant par les bâtiments de l'UQAC puis des centres commerciaux situés à la Place du Royaume et au centre-ville. La seconde, qui est la phase de localisation, consiste simplement à se rendre sur les lieux déjà scannés et à voir si l'on peut être correctement localisé. Cependant, les deux parties étant implémentées dans la même application, il est possible de réaliser les deux étapes en même temps, c'est-à-dire dès que les points d'accès ont été localisés, vous pouvez directement vous situer dans le lieu à l'aide de notre système positionnement.

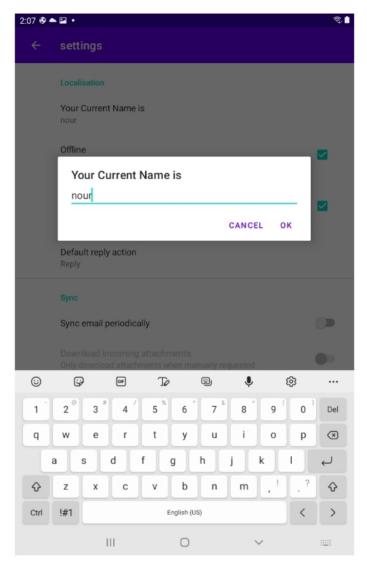
## **5.5 RÉSULTAT**

La mise en place de la solution nous ramènera à présenter quelques écrans de l'application, puis nous ferons quelques observations sur les résultats obtenus.

#### 5.5.1 CAPTURE DE L'APPLICATION

#### Ecrans d'embarquements

Lors de la première installation de l'application, des écrans d'accueil sont présentés à l'utilisateur pour lui montrer comment fonctionne l'application, c'est-à-dire comment il doit l'utiliser. La figure 20 présente les trois écrans d'embarquements.

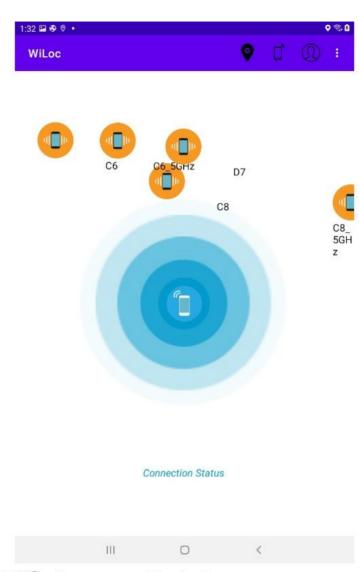


# **Set User Name**

Renseigner votre non d'utilisateur afin que votre position soit enregistrer.

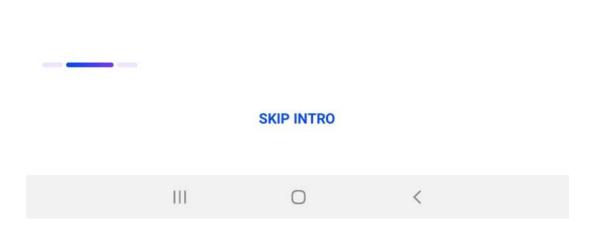


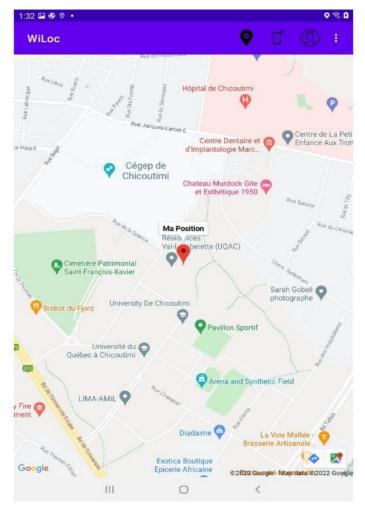




# **Scan Wifi Access Points**

Activer votre wifi et scanner les points d'acces environnant.





## **Get Your Position**

Visualiser votre position actuel.

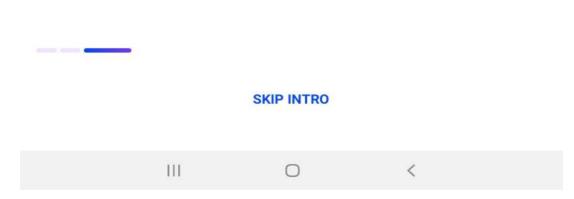
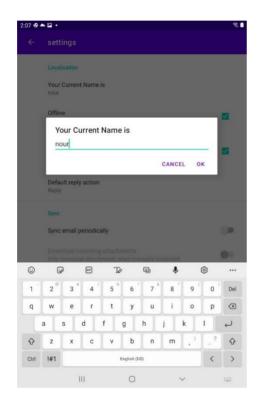


Figure 19: Ecrans d'embarquement de l'application Android.

#### ♣ Page de renseignement du nom d'utilisateur

Une fois les pages d'embarquement passées, l'utilisateur doit suivre la procédure indiquée, donc renseigner son identifiant à partir duquel son historique de voyage pourra être récupéré. Pour cela, il doit se rendre dans les paramètres comme on peut le voir sur les images de la figure 21.



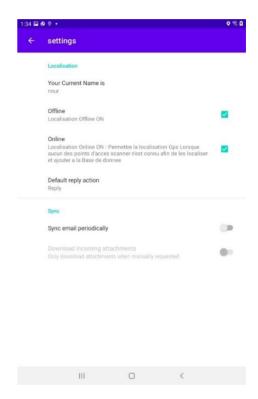


Figure 20: Renseignement du nom d'utilisateur.

#### Page de scannage des points d'accès

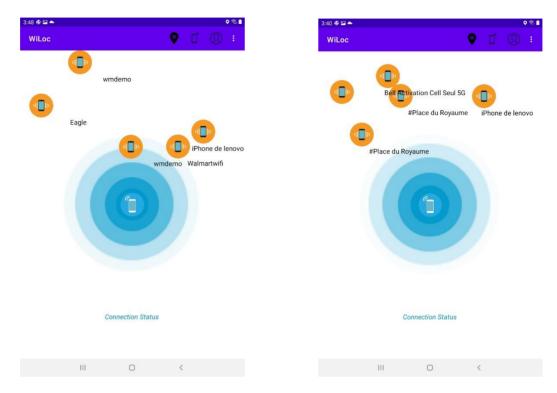


Figure 21 : Ecran de scannage des points d'accès.

#### Page de localisation

Lorsque le lieu a une fois été scanné, l'utilisateur peut alors être localisé. Cependant, si l'emplacement n'a pas été scanné auparavant et que le mode en ligne est activé, l'emplacement sera scanné et enregistré dans la base de données. En revanche si le mode en ligne a été désactivé un message de localisation inconnue lui serait affiché. La figure 23 indique deux images où l'utilisateur a été localisé.



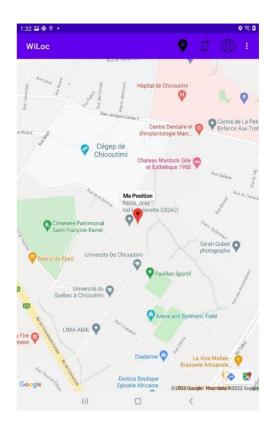


Figure 22 : Ecran de localisation d'utilisateur.

♣ Page de paramétrage de l'application

Pour pouvoir scanner l'utilisateur dois activer le mode offline comme indiqué sur la figure

24.

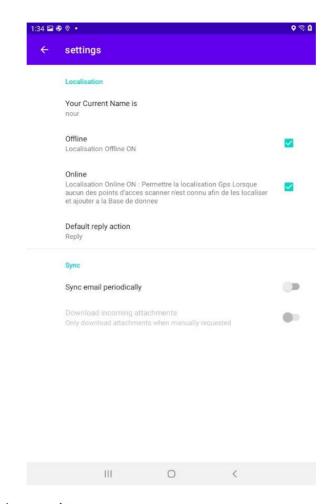


Figure 23 : Ecran de paramètre.

Si le mode hors ligne n'est pas activé, un message s'affichera pour que l'utilisateur le fasse, comme indiqué dans la figure 25.



Figure 24: Message d'activation du mode Offline.

Le mode en ligne lorsqu'il est activé permet de localiser la position de l'utilisateur par GPS et balayer les points d'accès environnants, lorsque l'utilisateur tente de se localiser dans des endroits inconnus. Cependant, lorsqu'il est désactivé, un message d'emplacement inconnu s'affiche pour l'utilisateur.

### Page d'historique

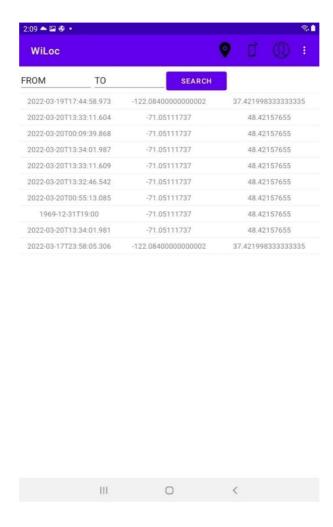
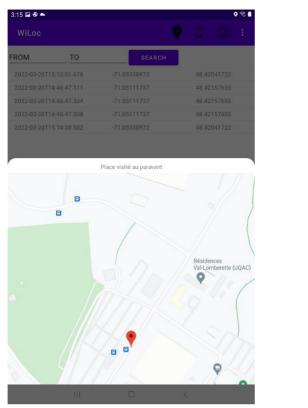


Figure 25 : Historique de déplacement de l'utilisateur.

Il est possible de visualiser la localisation de chaque lieu visité en cliquant sur l'historique comme indiqué dans la figure 27.



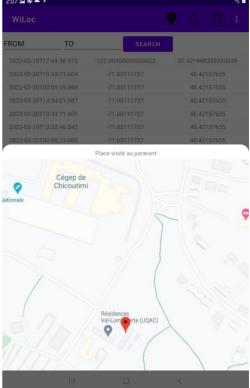


Figure 26: Visualisation d'un historique.

#### **5.6 APPRÉCIATION**

Á partir des écrans présentés, on voit l'efficacité du système à pouvoir localiser les utilisateurs en fonction des points d'accès des lieux où un scan a été effectué et à retracer leur historique de déplacement. Cependant, lors des tests, nous remarquons souvent un déphasage entre le temps du scan Wifi et celui de la localisation GPS, qui dépasse l'intervalle acceptable (c'est-à-dire entre 0 et 5 secondes dans notre cas), ce qui signifie qu'il n'est pas toujours garanti qu'un lieu a été correctement scanné dès le premier scan. Pour garantir l'enregistrement des points d'accès d'un lieu il faudrait effectuer au moins trois scans Wifi, en espérant qu'une localisation GPS se produise en moins de cinq secondes à la suite d'un scan Wifi, car les périodes de Wifi et de scan de localisation ne sont pas identiques et sont variables. Cependant, on remarque qu'il y a plus de chance d'être dans cet intervalle

quand on lance l'application, car à ce moment-là la localisation et le scan se font presque simultanément. Nous pensons que l'ensemble des données recueillies nous permettra de mieux définir ce paramétrage qui peut aussi être influencé de nombreux éléments comme : la puissance de calcul, le niveau de charge de la batterie, les applications tournant en arrière-plan.

Comme indiqué, l'un des critères importants de tout système de positionnement est sa capacité à couvrir la zone concernée. Il est donc important de noter que ce système est loin d'avoir une couverture complète. Pour avoir une couverture totale, il faudrait parcourir tout le territoire concerné (la ville de Chicoutimi dans notre cas) et scanner les points d'accès situés à chaque endroit, ce qui n'est pas facile à faire et demande un temps considérable. Cependant, utiliser une voiture permettrait de gagner du temps. En revanche, nous ne serons pas assurés d'avoir un scan Wifi et une localisation GPS dans l'intervalle prévu à chaque endroit. De plus, cet intervalle devrait être réduit à une seconde compte tenu de la vitesse de déplacement de la voiture, car un intervalle de cinq secondes pourrait entraîner une mauvaise précision du système. Soulignons qu'il existe des sites web qui rendent les coordonnées des points d'accès accessibles pour le public comme Mozilla<sup>14</sup>. Toute fois ils demandent tout un processus de vérification avant d'avoir l'autorisation d'utilisation de l'API. De plus le mécanisme de fonctionnement n'est pas connu.

Ainsi compte tenu des observations, pour avoir une meilleure fiabilité et une couverture optimale du système, la phase Offline, c'est-à-dire l'enregistrement des points d'accès, doit être réalisée à pied afin d'augmenter la probabilité qu'une localisation GPS soit effectuée à la suite d'un balayage Wifi, et l'application doit être installée par plusieurs utilisateurs d'intérêts différents. L'utilisation par plusieurs utilisateurs permettrait d'avoir une couverture optimale rapidement. De plus les gens vont généralement au même endroit et même lorsqu'une personne change d'environnement pour aller dans un nouveau lieu, ce nouveau lieu est

-

<sup>14</sup> https://location.services.mozilla.com/

familier à d'autres (Sapiezynski et al., 2015). En revanche, un tel scénario implique une contrainte de temps qui va au-delà de ce qui est possible dans le cadre d'un travail de mémoire. Ainsi, cette phase d'expérimentation est prévue après.

#### CONCLUSION

Le *smartphone* est devenu un compagnon inséparable de l'Homme du 21<sup>ième</sup> siècle grâce aux innombrables possibilités qu'il offre à travers les différentes applications. Or, l'utilisation de ces applications est à l'origine de la génération d'une quantité massive de données dont l'utilité intéresse à la fois les entreprises et les chercheurs. Ainsi la recherche de l'utilisation de ses données a donné lieu à plusieurs études d'exploration, dont celle relative à la localisation des utilisateurs. Dans cette catégorie d'exploration, c'est-à-dire celle sensible à la localisation, l'une des données utilisées pour déterminer la position des utilisateurs est celle des signaux des réseaux Wifi, en raison de leur fort déploiement. Ainsi, les points d'accès Wifi déployés partout sont utilisés comme balises de localisation. Cependant les mécanismes exacts ne sont connus que des grandes entreprises et sont très peu documentés.

Ainsi dans ce travail nous avons proposé un système de positionnement extérieur à grande échelle en utilisant les réseaux Wifi. Nous avons d'abord décrit les données présentes sur les *smartphones* ainsi que leurs sources, puis nous avons détaillé les systèmes de positionnement dans le mobile afin de voir leurs avantages et inconvénients. Par la suite, nous avons fait un survol sur les systèmes de positionnement basés sur le Wifi pour enfin proposer un modèle permettant un positionnement de l'utilisateur à l'échelle de la société.

Les résultats ont montré l'efficacité du système à pouvoir positionner les utilisateurs dans des endroits connus et à retracer leur historique de déplacement. Cependant, nous sommes loin d'avoir une couverture optimale. Le système nous permet également d'avoir un historique des mouvements de l'utilisateur. A partir de ces données de mouvement, nous pourrions à l'avenir être en mesure de prédire les mouvements de l'utilisateur grâce à des algorithmes de machine Learning.

En somme, les données générées par nos *smartphones* peuvent servir de base à la création de nouveaux services, avec notamment la possibilité de fournir un système de positionnement grâce aux signaux de requête Wifi émis par les *smartphones* lors de leurs tentatives de connexion à différents points d'accès (PAs).

#### LISTE DE RÉFÉRENCES

- Abidin, D. Z., Nurmaini, S., Erwin, Rasywir, E., Pratama, Y., & Li, Y. (2021). Indoor positioning system in learning approach experiments. *Journal of Electrical and Computer Engineering*, 2021, 1-16. <a href="https://doi.org/10.1155/2021/6592562">https://doi.org/10.1155/2021/6592562</a>
- Androulidakis, I. I. (2012). *Mobile phone security and forensics : a practical approach*. Springer. <a href="https://doi.org/10.1007/978-1-4614-1650-0">https://doi.org/10.1007/978-1-4614-1650-0</a>
- Bahl, P., & Padmanabhan, V. N. (2000, 26-30 March 2000). *Radar: An in-building rf-based user location and tracking system*. Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064). <a href="https://doi.org/10.1109/INFCOM.2000.832252">https://doi.org/10.1109/INFCOM.2000.832252</a>
- Bassoni, M., & Weygand, F. (2011). Les enjeux économiques de la géolocalisation pour les réseaux sociaux numériques. *Hermès*,  $n^{\circ}$  59(1). https://doi.org/10.3917/herm.059.0137
- Bhavnani, S. P., Narula, J., & Sengupta, P. P. (2016). Mobile technology and the digitization of healthcare. *European Heart Journal*, *37*(18), 1428-1438. <a href="https://doi.org/10.1093/eurheartj/ehv770">https://doi.org/10.1093/eurheartj/ehv770</a>
- Carreras, I., Matic, A., Saar, P., & Osmani, V. (2012, 19-23 March 2012). *Comm2sense: Detecting proximity through smartphones.* 2012 IEEE International Conference on Pervasive Computing and Communications Workshops. <a href="https://doi.org/10.1109/PerComW.2012.6197489">https://doi.org/10.1109/PerComW.2012.6197489</a>
- Chan, S., & Sohn, G. (2015). Indoor localization using wi-fi based fingerprinting and trilateration techiques for lbs applications. *The international archives of the photogrammetry, remote sensing and spatial information sciences*, 4(C26), 1-5. https://doi.org/10.5194/isprsarchives-XXXVIII-4-C26-1-2012
- Desai, N., & D'Mello, L. (2014). An overview on mobile data mining. *International journal of engineering research* & technology, 3(9). <a href="https://doi.org/IJERTV3IS090857">https://doi.org/IJERTV3IS090857</a>
- Du, H., Zhang, C., Ye, Q., Xu, W., Kibenge, P. L., & Yao, K. (2018). A hybrid outdoor localization scheme with high-position accuracy and low-power consumption. *EURASIP Journal on Wireless Communications & Networking*, 2018(1).
- Francillette, Y., Menelas, B. A. J., Bouchard, B., Bouchard, K., Gaboury, S., Kingsbury, C., St-Amour, S., Romain, A. J., & Bernard, P. (2021, 2021//).

- Development of a Mobile Exergame to Implement Brief Interventions to Increase Physical Activity for Adults with Schizophrenia. Serious Games, Cham. Springer International Publishing.
- Fulara, K. (2018). *Big data development driven by your smartphone*. Repéré le 27/06/1018 à <a href="https://www.sparkbit.pl/big-data-development-driven-smartphone/">https://www.sparkbit.pl/big-data-development-driven-smartphone/</a>
- IDC. (2021). Smartphone market share. <a href="https://www.idc.com/promo/smartphone-market-share">https://www.idc.com/promo/smartphone-market-share</a>
- Johnson, N. E., Mandiola, P., Blankinship, C., Bonczak, B., & Kontokosta, C. E. (2019, 9-12 Dec. 2019). *Validating the use of wi-fi signals to estimate hyperlocal urban populations*. 2019 IEEE International Conference on Big Data (Big Data). https://doi.org/10.1109/BigData47090.2019.9006517
- Köbben, B., Bunningen, A. v., & Muthukrishnan, K. (2006). Wireless campus lbs: Building campus-wide location based services based on wifi technology. Dans *Geographic Hypermedia* (pp. 399-408). <a href="https://doi.org/10.1007/978-3-540-34238-0\_22">https://doi.org/10.1007/978-3-540-34238-0\_22</a>
- Li, B., Quader, I., & Dempster, A. G. (2008). On outdoor positioning with wi-fi. *Journal of Global Positioning Systems*, 7(1), 18-26. <a href="https://doi.org/10.5081/jgps.7.1.18">https://doi.org/10.5081/jgps.7.1.18</a>
- Li, B., Wang, Y., Lee, H. K., Dempster, A., & Rizos, C. (2005). Method for yielding a database of location fingerprints in WLAN. *IEE Proceedings Communications*, 152(5), 580 586. <a href="https://doi.org/10.1049/ipcom:20050078">https://doi.org/10.1049/ipcom:20050078</a>
- Lin, Y., & Huang, W. (2017). Behavior analysis and individual labeling using data from wi-fi ips.
- Marketing-Professionnel. (2012). *Les entreprises face au défi de la fidélisation client*. <a href="https://www.marketing-professionnel.fr/tribune-libre/fidelisation-satisfaction-client-entreprises-defi-201203.html">https://www.marketing-professionnel.fr/tribune-libre/fidelisation-satisfaction-client-entreprises-defi-201203.html</a>
- Menelas, B.-A. J., & Otis, M. J. D. (2012, 2012//). A Serious Game for Training Balance Control over Different Types of Soil. Serious Games Development and Applications, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Messaoudi, M. D., Menelas, B.-A. J., & McHeick, H. (2020). Autonomous smart white cane navigation system for indoor usage. *Technologies*, 8(3). https://doi.org/10.3390/technologies8030037

- Planétoscope. (2012). *Ventes mondiales de smartphones*. <a href="https://planetoscope.com/electronique/728-ventes-mondiales-de-smartphones.html">https://planetoscope.com/electronique/728-ventes-mondiales-de-smartphones.html</a>
- Roos, T., Myllymäki, P., Tirri, H., Misikangas, P., & Sievänen, J. (2002). A probabilistic approach to wlan user location estimation. *International Journal of Wireless Information Networks*, *9*(3), 155-164. <a href="https://doi.org/10.1023/A:1016003126882">https://doi.org/10.1023/A:1016003126882</a>
- Safei, S. (2012). Indoor position detection using wifi and trilateration technique. <a href="https://www.researchgate.net/publication/230771403">https://www.researchgate.net/publication/230771403</a> INDOOR POSITION DETECTION USING WIFI AND TRILATERATION TECHNIQUE
- Saint-Louis, H., & Ménélas, B.-A. J. (2021, 2021//). *User Perceptions of Security and Privacy Risks with Contact Tracing Apps.* HCI International 2021 Late Breaking Papers: HCI Applications in Health, Transport, and Industry, Cham. Springer International Publishing.
- Sapiezynski, P., Stopczynski, A., Gatej, R., & Lehmann, S. (2015). Tracking human mobility using wifi signals. *PloS one*, *10*(7), e0130824. https://doi.org/10.1371/journal.pone.0130824
- Soundararaj, B., Cheshire, J., & Longley, P. (2020). Estimating real-time high-street footfall from Wi-Fi probe requests. *International Journal of Geographical Information*Science, 34(2). https://doi.org/https://doi.org/10.1080/13658816.2019.1587616
- Taboubi, M. (2014). *Indoor localization using wi-fi signals*. Université De Montréal. <a href="https://publications.polymtl.ca/1483/1/2014\_MohamedTaboubi.pdf">https://publications.polymtl.ca/1483/1/2014\_MohamedTaboubi.pdf</a>
- Yao, Y., Su, X., & Tong, H. (2018). *Mobile data mining* (Vol. 1). Springer. https://doi.org/https://doi.org/10.1007/978-3-030-02101-6