
La signalisation des téléphones mobiles Une source adaptée à la mesure des déplacements

Introduction

Ce chapitre présente une première qualification du système de mesures de déplacements à partir des remontées de signalisation des réseaux radio-mobiles. En effet, nous cherchons à reproduire des mesures de déplacements et d'indicateurs qui sont estimables à partir d'autres sources de données et nous les confrontons pour faire ressortir les avantages et limites. Comme exemple d'application, nous avons choisi d'étudier un système de transport en commun ferré : la ligne A du RER dont la section centrale dessert Paris intra-muros. Ce choix est motivé par deux aspects. D'une part, nous voulions réduire la complexité des mécanismes d'exploration de la signalisation. Le « territoire mobile » autour du tronçon central a été construit de manière particulière et assure des déclenchements d'évènements des mobiles pour tout déplacement d'une section souterraine vers une section aérienne et inversement. De plus, la superposition des « territoire mobile » et « territoire géographique » est quasi-directe dans ce cas. D'autre part, La partie 1 nous a montré les difficultés actuelles de l'exploitation des réseaux et les besoins d'une connaissance fine temporellement et spatialement des dynamiques de comportements des usagers qui nous amène à s'intéresser à des mesures de déplacements à partir de sources digitales. Des travaux ont déjà montré que l'emploi de données de type AFC (Automated Fare Collection) permet de mesurer des déplacements et des indicateurs pour des réseaux de transports en commun. Est-ce que la source que nous proposons est également adaptée à cet exercice ?

Nous avons construit ce chapitre de la manière suivante. Nous commencerons par la description dans une première section de la démarche expérimentale que nous avons suivie, puis nous décrirons les données additionnelles dont nous avons disposées pour la

qualification des résultats produits et nous ferons une analyse de la source radio-mobile pour en faire ressortir les mécanismes de fonctionnement dans le cas particulier de ce périmètre d'étude. Cela nous a permis aussi de repérer quelques comportements qui seront plus délicats à interpréter. La deuxième section est consacrée à la construction d'indicateurs de qualité de service de la ligne selon une perception de l'utilisateur qui seront utiles pour les acteurs de l'offre. On construira notamment un indicateur de densité de voyageurs par rames, un indicateur de fiabilité sous la forme de densité des retards sur la ligne, et une mesure des volumes de voyageurs par paire de stations. Ce dernier indicateur est également construit à partir des données billettiques pour alimenter le processus de qualification.

1 Protocole exploratoire

Le protocole exploratoire est présenté ici en trois étapes. Tout d'abord, nous présentons la démarche expérimentale conçue et suivie par l'auteure. Puis, nous proposons une description du périmètre d'étude dans son aspect transport et son aspect télécom grâce à une étude menée sur le terrain. Enfin, nous pourrions décrire les différentes sources de données exploitées dans cette étude.

1.1 Démarche expérimentale

L'objectif de mon travail de thèse est de développer des méthodes de traitement des messages de signalisation collectés sur le réseau de téléphonie mobile d'Orange dans le but de mesurer des déplacements sur les réseaux de transport. Ainsi, ce paragraphe décrit les différentes étapes de la démarche expérimentale que j'ai mise en place, pour proposer les résultats présentés dans ce chapitre. Cette démarche se décompose en trois étapes principales :

- le choix du périmètre d'étude et la compréhension de son fonctionnement d'un point de vue télécoms et d'un point de vue transport,
- la production de premiers indicateurs adaptés aux besoins d'analyse du système de transport et la qualification des méthodes de traitement de la signalisation
- la production d'indicateurs supplémentaires à partir de la mise en place d'une nouvelle campagne de collecte de messages de signalisation. Les méthodes de traitement et le périmètre de collecte ont pu être perfectionnés par la connaissance acquise lors de la production des premiers indicateurs et par la collecte de données de comparaison qui permettent de qualifier nos méthodes.

Le choix du périmètre d'étude repose sur des contraintes relatives à la maîtrise des mécanismes de signalisation, aux besoins des acteurs des transports et à la superposition des observations réalisées sur le « territoire mobile » sur le « territoire géographique ». Tout d'abord, le paragraphe 2.1 du chapitre 4 montre la sophistication de l'architecture des réseaux de télécommunications et des mécanismes de signalisation ainsi que les échelles spatiales et temporelles de déclenchement des événements. J'ai choisi de réduire artificiellement la

complexité du fonctionnement des mécanismes de signalisation en cherchant des espaces du « territoire géographique » couverts par une seule couche technologique et pour lesquels les procédures de signalisation se superposaient à des comportements de mobilité. En transport, ces derniers peuvent être mesurés à partir d'enquête cordons (cf. Chapitre 4 §2.1.1.4). Or, les cellules frontières d'une zone de localisation s'apparentent à un cordon de cette zone. Il fallait chercher un périmètre dont le réseau de télécommunications offrait « un cordon de cellules ». Toutefois, l'architecture du « territoire mobile » n'était pas la seule contrainte fixée. En effet, les besoins des acteurs des transports ont été présentés dans la partie 1. Dans le cadre de la favorisation des transports en commun en ville pour fluidifier la circulation automobile et réduire les impacts environnementaux, l'attractivité de ce mode de transport est un effort à mener en milieu urbain. Je me suis alors intéressée aux indicateurs de qualité de service des systèmes de transport publics du point de vue de l'utilisateur. La dernière étape du choix de périmètre a donc été de trouver un réseau de transport public d'une grande agglomération française dont la couverture par le réseau de téléphonie mobile répondait aux contraintes énoncées. En effet, une des difficultés principales induite par cette méthode de capture, provient des difficultés de projection du « territoire mobile » sur le territoire géographique. Afin de limiter les complications relatives aux projections, une portion du « territoire mobile » dont le découpage cellulaire réduisait les ambiguïtés de superposition des deux territoires a été choisie. Un tel découpage existe sur l'ensemble des sections souterraines du système de transport en commun ferré d'Ile de France. En effet, ce système de transport est couvert par des cellules regroupées en une seule zone de localisation dédiée aux parties souterraines. Les cellules qui couvrent les parties aériennes du système de transport sont regroupées dans des zones de localisation distinctes. Ainsi, un événement sera déclenché à l'entrée et à la sortie de la portion souterraine. Pour la majorité des stations souterraines, un site est affecté à la couverture des plateformes et des tunnels adjacents pour chacune des lignes. Ceci se vérifie particulièrement pour les stations du tronçon central de la ligne A du RER. Alors, les terminaux mobiles générant un événement au niveau des cellules couvrant les stations souterraines, ont une probabilité élevée d'être localisés au niveau de plateformes ou dans les tunnels en étant à bord du train. De plus, au commencement de l'étude, en 2011, les sections souterraines ne proposaient qu'une couverture GSM, bornant ainsi la collecte des événements aux sondes 2G pour l'étude en souterrain. Enfin, les comportements de déplacements au niveau du système de transport ne sont pas équivoques sur ce système de transport. Effectivement, les personnes présentes dans les tunnels du système de transport sont pour une large majorité en train d'utiliser ce réseau.

Le 7 avril 2011, j'ai parcouru ce tronçon avec un équipement de mesure de champ reçu par un mobile – appelé mobile à trace – pour constituer des mesures de l'environnement radioélectriques de déclenchement des événements et constituer une échantillon du comportement d'un mobile à bord d'un train. Le même jour, une première collecte sur le périmètre des cellules couvrant les stations du RERA de Vincennes à La Défense, a été menée. Ce premier jeu de données de signalisation contenait l'activité de signalisation uniquement au niveau de la zone de localisation souterraine. Cette restriction a permis de limiter le volume d'acquisition et donc de simplifier l'exploration du jeu de données. Le

choix du jour d'acquisition s'est porté sur un jour ouvré moyen afin d'observer les migrations pendulaires classiquement étudiées dans les analyses des déplacements. Une enquête terrain s'est également déroulée le même jour et s'est focalisée sur l'activité du territoire géographique. Elle a été réalisée par des observateurs positionnés au niveau des quais des stations et a permis la constitution d'un jeu d'observations centrées sur les passages de train. La confrontation des différentes bases de données a permis de guider et de qualifier les méthodes d'observation des déplacements. Des premiers indicateurs de qualité de service comme le taux de charge des trains ou encore la fiabilité ont pu être produits. Cette première étude a montré le potentiel d'une telle source de données et a guidé la définition d'une nouvelle campagne de mesure sur un « périmètre mobile » contenant également les cellules couvrant les parties aériennes des stations afin d'observer les entrées et les sorties de la portion du système de transport considéré.

Cette deuxième collecte s'est déroulée le jeudi 13 octobre 2011. Les 7 mois qui séparent les deux enquêtes sont imputables aux délais de dépouillement des résultats et de définition des nouveaux paramètres de la collecte, mais surtout au retour d'un jour ouvré moyen, c'est-à-dire éloigné des périodes estivales. J'avais fixé deux objectifs. D'une part, compléter les méthodes de collectes et de traitement et d'autre part qualifier ces méthodologies de mesures de déplacements en comparant les volumes estimés à partir de la signalisation avec des quantités estimées à partir d'une autre source utilisée pour la mesure des déplacements. Une revue de la bibliographie concernant la mesure des déplacements dans les systèmes de transport en commun a montré que les données billettiques sont adaptées à la mesure de flux de déplacements (BERTINI, et al., 2003). Ainsi, nous avons coordonné une nouvelle journée d'acquisition avec le STIF afin de constituer des jeux de données comparables.

1.2 Description de la zone d'étude

Le choix de la zone d'étude a été orienté par le découpage particulier du « territoire mobile » et des comportements non équivoques des déplacements sur le mode de transport. Ce paragraphe vise à présenter de manière plus détaillée, le système de transport en commun ferré d'Ile-de-France sous deux angles, celui des conditions d'opération transport et celui des conditions d'opération télécoms.

1.2.1 Aspect transport de la zone d'étude

Le réseau de transport en commun d'Ile-de-France est composé de sous-systèmes interconnectés : le réseau bus, le réseau de tramway et le réseau ferré regroupant le réseau de chemin de fer métropolitain (métro) et le Réseau Express Régional (RER). Nous nous limiterons à l'étude de ces deux derniers modes dédiés au transport urbain et péri-urbain.

Premièrement, le Métro de Paris dessert principalement Paris et son agglomération. Il est composé de 16 lignes en site propre, pour la plupart souterraines, de 302 stations et s'étend sur plus de 200km de voies. Il est exploité par la RATP. En 2011, 4,13 millions de voyageurs étaient transportés par jour grâce au métro (1.5 milliard par an) (STIF, 2011). Le

réseau métropolitain de Paris a été conçu pour assurer une desserte fine de Paris intra-muros grâce à un quadrillage serré. Les stations dans Paris sont très rapprochées (548 m en moyenne avec une variation selon les lignes – plus d'1km pour la ligne 14). Le service est assuré de 5h30 environ à 1h15 le lendemain matin. Le service est prolongé jusqu'à 2h15 les fins de semaine et certaines veilles de fêtes. La fréquence moyenne est d'une rame toutes les 2 minutes aux heures de pointe et de 4 minutes aux heures creuses. Ces 10 dernières années, le Métro a connu des évolutions relatives à des augmentations du trafic à écouler et aux mutations des comportements des voyageurs. On constate notamment, une augmentation des voyages en fins de semaine, une forte concentration de l'heure de pointe du matin entre 8h30 et 9h00 et un rallongement de la pointe du soir jusqu'à 20h30. Ceci a engendré, entre autre, une dégradation des conditions de transport lors des périodes de pointe. Enfin, le métro ne dessert les proches banlieues que par des prolongements de lignes. De ce fait, un trajet d'une commune de proche banlieue à une autre impose un itinéraire via Paris intra-muros. En raison de la faible vitesse commerciale et de capacité des rames réduites, ce mode ne permet pas la desserte de la grande couronne de Paris.

Le Réseau Express Régional (RER) a été développé pour répondre au besoin de meilleure desserte des banlieues. En 2006, le RER accueillait 2,8 millions de voyageurs par jour dont 35% sur la ligne A et 57% sur les lignes A et B (RAIL PASSION, novembre 2006). Il s'agit d'un moyen de transport ferré hybride qui a permis l'intégration de lignes de banlieue préexistantes à un réseau souterrain à grand gabarit ferroviaire traversant le centre-ville. De multiples connexions avec le Métro sont possibles. Le RER est constitué de cinq lignes et comporte 257 points d'arrêt, dont 33 à Paris, pour 587 km de voies, dont 76,5 km en souterrain. Les sections souterraines sont situées principalement dans Paris intra-muros. Il est co-opéré par la RATP et la SNCF. En heure creuse, la fréquence des rames est de l'ordre de la dizaine de minutes autour de Paris et de l'ordre de la vingtaine de minutes pour les stations éloignées de plus de 40km de Paris. Les fréquences sont généralement doublées aux heures de pointe. La fréquence cible en heure de pointe pour les stations de Paris intra-muros est inférieure à 5 minutes.

Nous nous sommes restreints à l'étude de deux lignes de ce réseau de transport : la ligne A du RER et la ligne 1 du Métro car cette dernière offre un itinéraire alternatif à la ligne A du RER pour la traversée est-ouest de Paris intra-muros. Le choix a été motivé d'une part par le découpage du « territoire mobile » de la ligne A et d'autre part par les conditions difficiles d'opération de la ligne A du RER. En effet, l'importance du trafic est-ouest entraîne des phénomènes de congestion récurrents. Une explication possible d'un tel trafic est le déséquilibre régional entre des emplois majoritairement situés à l'ouest (exemple du quartier de la Défense) et les zones résidentielles en expansion à l'est. Dans la suite du paragraphe, nous présenterons quelques caractéristiques des conditions d'opération de ces deux lignes.

La ligne A du RER

L'agglomération parisienne est traversée d'Est en Ouest par la ligne A du RER (RERA) et du Nord au Sud par la ligne B (RERB). La ligne A est de loin la plus chargée du

réseau (voir Figure 30). En effet, on comptabilise 1.14 millions de voyageurs par jour ouvrable, soit près de 310 millions de voyageurs au total en 2011. Ce nombre conséquent de voyageurs amène régulièrement la ligne dans une situation proche de la saturation (STIF, 9 février 2011). Ainsi, la ligne A fait partie des lignes les plus denses d'Europe et du monde. En deuxième position, la ligne B transportait 900 000 voyageurs par jour ouvrable en 2009 (STIF, 3 novembre 2009). On peut comparer le trafic de la ligne B à celui des lignes de métro les plus chargées (STIF, 2005). Depuis les années 2000, la saturation récurrente de ces deux lignes principales entraîne une forte dégradation de leur fonctionnement. En effet, l'indice de régularité pour ces deux lignes est passé de 90.1% en 2007 à 88.4% en 2008¹³. Cet indice de régularité se définit pour le RER comme le pourcentage de voyageurs n'ayant pas subi de retard sur l'ensemble de la journée ou ayant subi un retard inférieur à 5 minutes sur le nombre total de voyageurs. En 2008, parmi les 10 des stations RER les plus fréquentées, 7 sont des stations du RERA et 3 du RERB¹⁴. La station de Châtelet–les Halles est comptée à la fois pour la ligne A et pour la ligne B car elle permet leur interconnection.

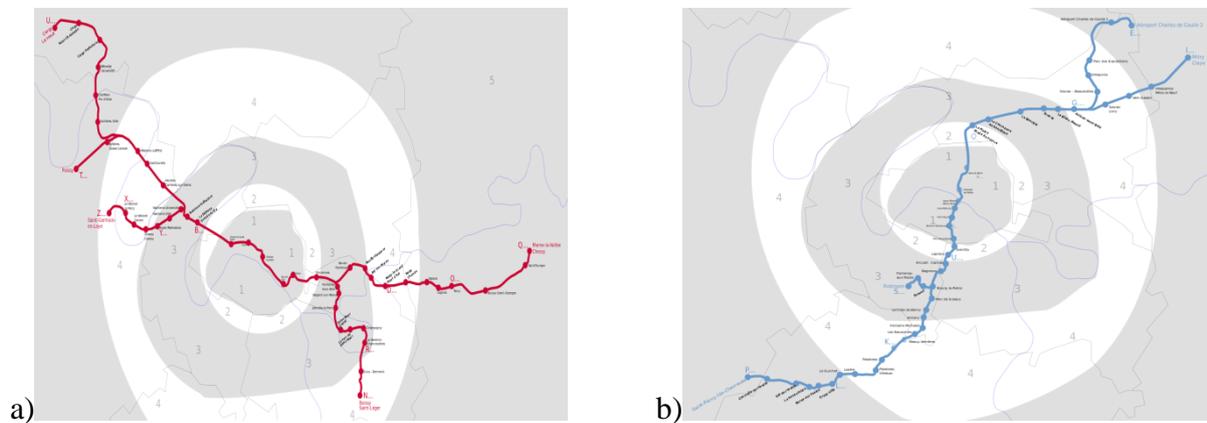


Figure 30: Tracé des lignes A (a) et B (b) du RER (auteur Benoit Mortagt)

Aux extrémités de la section centrale qui traverse Paris, une division en branches permet la desserte de Saint-Germain-en-Laye (A1), Poissy (A5) et Cergy (A3) à l'Est et celle de Boissy-Saint-Léger (A2) et Marne-la-Vallée-Chessy (A4) à l'Ouest. Elle dessert ainsi 41 communes de 5h00 à 1h20 le jour suivant. La ligne A dispose de 46 points d'arrêts -35 opérés par la RATP et 11 par la SNCF-, sur 109 km total de voies dont 26 km en souterrain. Ces sections sont situées majoritairement au niveau du tronçon central. La section du tronçon centrale entre Châtelet-les Halles et Auber est la plus chargée avec près de 50 000 voyageurs par heure et par sens pendant les heures de pointe. La fréquence des trains visée par l'opérateur est d'un train toutes les 2 minutes dans un sens, c'est-à-dire 30 trains par heure (STIF, 31 janvier 2008). Par la suite, nous appellerons tronçon central du RERA les stations

¹³ www.paris.fr, les transports en commun, 2008

¹⁴ www.paris.fr, les transports en commun, 2008

de la ligne comprises entre Vincennes et La Défense, ces dernières étant incluses. Toutes ces stations et les sections inter-stations sont souterraines à l'exception des sections qui entourent Vincennes. La liste de ces stations est énoncée pour un sens de parcours de la ligne Est-Ouest :

- Vincennes
- Nation, connectée au réseau métropolitain par les lignes 1, 2, 6, et 9
- Gare de Lyon, connectée au réseau métropolitain par les lignes 1 et 14
- Châtelet-les-Halles, connectée au réseau métropolitain par les lignes 1, 4, 7, 11 et 14
- Auber, connectée au réseau métropolitain par les lignes, 3, 7, 8 et 9
- Charles-de-Gaulle-Etoile, connectée au réseau métropolitain par les lignes 1, 2 et 6
- La Défense, connectée au réseau métropolitain par la ligne 1

La ligne A est pratiquée pour différents motifs de déplacements. En effet, elle est empruntée dans le cadre des migrations pendulaires en reliant les banlieues résidentielles à Paris, et notamment au quartier de la Défense. Mais elle dessert également le parc Disneyland Paris et divers complexes commerciaux comme « Val d'Europe » ou les « 4 temps ». Ainsi, depuis 2002, les trains du week-end sont systématiquement prolongés à *Marne-la-Vallée-Chessy* pour faire face à l'important trafic que draine le nouveau quartier du *Val d'Europe* (METROPOLE, 29 août 2002). La Figure 31 illustre l'évolution du nombre moyen de voyages effectués en RER un jour ouvrable (bleu) et du nombre moyen de voyages effectués en RER le samedi et le dimanche cumulé (rouge)¹⁵. On remarque une augmentation significative du nombre de voyages de fin de semaine. Cette tendance a fortement réduit l'écart entre les volumes des jours ouvrés et du cumul des volumes du week-end.

¹⁵ www.paris.fr, les transports en commun, 2008

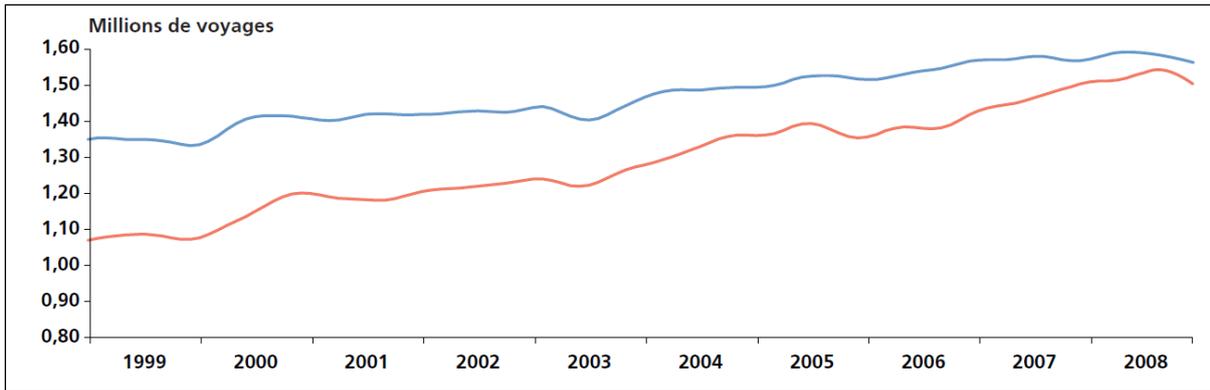


Figure 31: évolutions du nombre de voyages effectués un jour ouvrable en bleu et en fin de semaine en rouge (les volumes de trafic du samedi et du dimanche sont cumulés) de 1999 à 2008 d'après un recensement RATP.¹⁶

Ligne 1 du Métro : une alternative pour la desserte Est-Ouest de Paris

Notre choix s'est porté sur la ligne 1 car elle traverse Paris d'Est en Ouest en longeant la ligne A du RER. D'ailleurs ces deux lignes partagent 5 stations (Nation, Gare de Lyon, Châtelet-les Halles, Charles-de-Gaulle-Etoile et la Défense). Ainsi la question de la compétition d'itinéraire est posée (STIF, 8 février 2012). Cette idée est renforcée par des propositions de prolongement de la ligne 1 vers Val de Fontenay à l'Est et vers Nanterre à l'Ouest (METROPOLE) (STIF). La ligne 1 du réseau métropolitain fait partie des lignes les plus chargées avec 4 de ses stations appartenant aux 10 stations de Métro les plus fréquentées en 2008¹⁷. En 2012, le trafic quotidien atteint une moyenne de 715 000 voyageurs par jour ouvré.

Elle comprend 25 stations sur une longueur de voies de 16km. Il faut environ 36 minutes pour parcourir la ligne en entier. La fréquence des rames varie de 4 à 8 minutes selon le moment de la journée. En heure de pointe, une rame d'une capacité de 720 voyageurs, dont 242 assis, passe toutes les 1 min 45. Néanmoins, la ligne reste particulièrement chargée aux heures de pointe.

1.2.2 Aspect télécom de la zone d'étude

1.2.2.1 Un découpage particulier du « territoire mobile »

Les parties souterraines du réseau de transport en commun ferré sont couvertes par un « territoire mobile » particulier. Ces particularités sont à l'origine du choix de ce premier périmètre d'étude. En effet, chaque station est équipée de sites dont le groupe de cellules couvrent les quais et les tunnels adjacents des lignes à l'étude. La proportion des portions de

¹⁶ www.paris.fr

¹⁷ www.paris.fr, les transports en commun, 2008

tunnels couvertes, dépend de la distance inter-station, du rayon de courbure du tunnel, de la présence ou non de répéteurs dans le tunnel, de l'emplacement des sites et de la puissance d'émission de l'antenne. Les stations offrant de multiples connections avec le réseau RER et/ou le réseau métropolitain, sont construites sur plusieurs étages afin de séparer chaque ligne. Dès lors, pour des raisons de forte atténuation du signal radio par les étages, chaque étage est équipé de sites. Ainsi, selon les stations, un groupe de cellules est dédié à une ligne spécifique de la station. L'association groupe de cellules/lignes est unique pour chaque station du tronçon central du RERA, à l'exception de La Défense. De plus, toutes ces cellules dédiées aux stations souterraines font partie de la même zone de localisation que nous nommerons **LA_souterraine** par la suite. Ces conditions permettent l'observation des événements de mobilité en entrée et en sortie du réseau souterrain. La dernière particularité de ce « territoire mobile » est qu'il est, au moment de l'étude, composé d'une seule « couche technologique », la couche 2G.

Ce découpage particulier offre alors deux avantages principaux pour la mesure des déplacements. D'une part, il se produira un déclenchement systématique d'un événement de mobilité de type **LAUN** à chaque transition extérieur/souterrain. Ceci permet notamment de mesurer les entrées et sorties au niveau du tronçon central mais aussi d'avoir des points de mesures intermédiaires lorsque le train change de type de section (souterraine/aérienne). D'autre part, la précision spatiale à la cellule de notre système de capture ne sera pas une limitation. En effet, un événement localisé au niveau d'une cellule de la **LA_souterraine** ne laisse que deux possibilités pour la position du mobile : il est soit à quai, soit dans un des tunnels adjacents. Par la suite, nous ne ferons pas la distinction et indiquerons seulement que le mobile était présent au niveau de la station S de la ligne L.

1.2.2.2 Nature des événements de signalisation

Les particularités de « territoire mobile 2G » dédié au système de transport souterrain d'Ile de France nous incitent à nous concentrer uniquement sur les événements de mobilité de type **LAUN**. Afin de confirmer cette hypothèse, nous avons tout d'abord mesuré la distribution des types d'évènement sur l'ensemble des cellules souterraines qui couvrent la ligne RERA. En effet, sur l'ensemble des 2 millions d'évènements qui ont été capturés le 7 avril 2011, 77% étaient des LAUN, 8% des appels ou des SMS reçus (COM0), 6% des HO, 5% des SMS envoyés et 4% des évènements autres (ATT0, ATT1, LAUP). Afin de mieux comprendre les conditions de déclenchement de ces évènements, nous avons procédé à des mesures de champ reçu par un mobile à bord d'un train circulant sur la ligne A sur RER.

1.2.2.3 Conditions radioélectriques d'un train parcourant la ligne A

Protocole

Nous cherchons à mesurer le champ reçu par un mobile installé dans le train à différents emplacements (tête, milieu et queue de rame). Nous avons alors parcouru trois fois le tronçon central Vincennes- la Défense du RERA en heure creuses le jeudi 7 avril 2011.

Nous nous sommes munis du matériel suivant pour les mesures des signaux radioélectriques:

- Un mobile à traces synchronisé temporellement sur le réseau mobile. Il permet d'enregistrer le champ reçu à chaque instant par le mobile
- Une caméra dont l'horloge est synchronisée avec le mobile qui permet une capture de l'environnement et surtout d'identifier les portions souterraines et les quais.

Parallèlement, afin d'acquérir une meilleure compréhension des conditions de déclenchement des événements de signalisation dans lesquelles se trouve un mobile voyageant à bord du RERA, nous avons comptabilisé les LAUN qui se sont produits au niveau de cellules souterraines par plage d'une seconde. Ces valeurs sont obtenues à partir du jeu de données collectées le 7 avril 2011 que nous avons mentionné dans le paragraphe 1.2.2.2. Nous allons observer la confrontation des mesures de champs reçus par un mobile à bord du train avec l'ensemble des événements LAUN capturés par notre système de collecte qui ont été générés par des mobiles allumés et autorisés à se connecter au réseau opéré par Orange, et dont le mobile à traces fait partie.

Résultat

La position dans le train n'entraîne pas de modifications significatives sur le champ reçu par le mobile ni sur le mécanisme de déclenchement des événements.

La Figure 32 montre le niveau de champ enregistré par le mobile lors d'un trajet de Vincennes vers La Défense sur la ligne A. Nous observons la puissance du signal radio émise par les différentes antennes au long du trajet et mesurée par le mobile. Lorsque la rame circule, le mobile va changer de cellule selon un mécanisme basé sur la meilleure serveuse, c'est-à-dire l'antenne qui offre un niveau de champ reçu par le mobile le plus haut. Ce mécanisme de rattachement à la « meilleure cellule » se déclenche lorsque la cellule serveuse ne permet plus de garantir le service demandé. Les traits noirs séparent les parties souterraines des parties aériennes. Ainsi, en lisant la Figure 32 de gauche à droite, on a :

- l'évolution du champ émis par l'antenne qui couvre la station Vincennes et ses tunnels en rose
- l'évolution du champ émis par les 2 antennes qui couvrent la station Nation et ses tunnels en violet
- l'évolution du champ émis par les 2 antennes qui couvrent la station Gare de Lyon et ses tunnels en vert

- l'évolution du champ émis par les 4 antennes qui couvrent la station Châtelet-les-Halles et ses tunnels en bleu.
- l'évolution du champ émis par les 2 antennes qui couvrent la station Auber et ses tunnels en orange

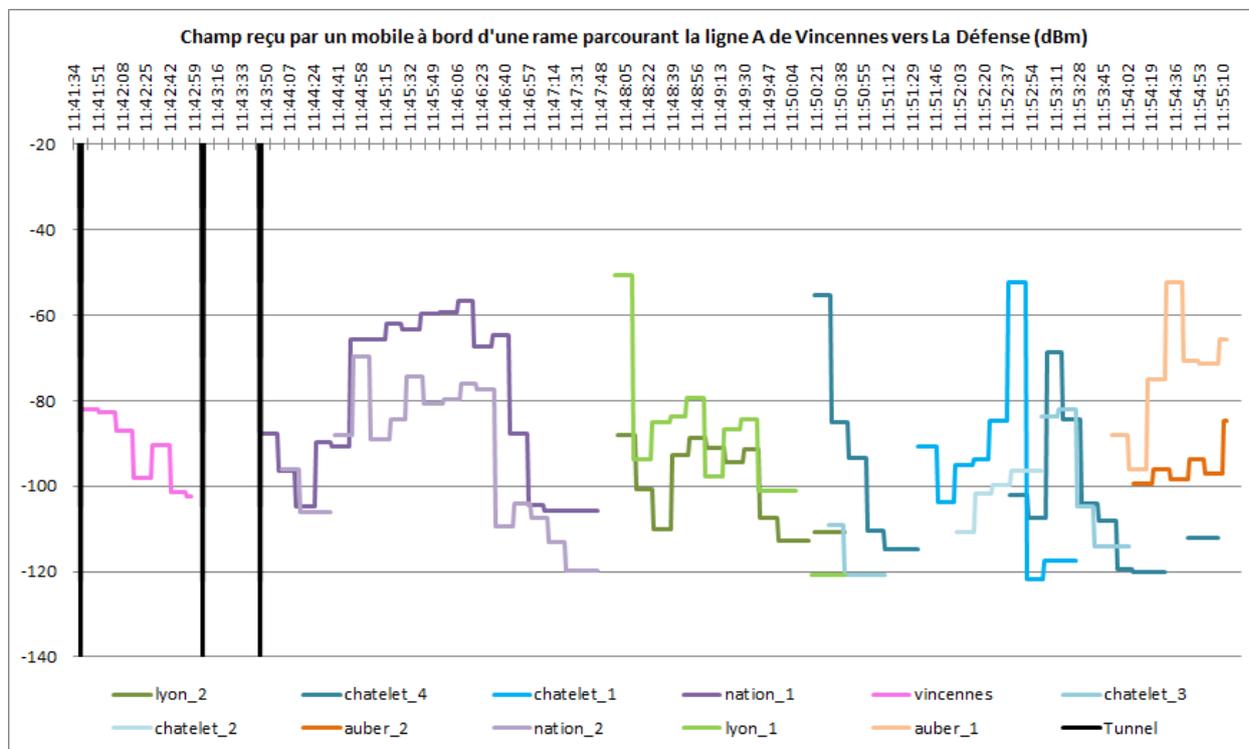


Figure 32 : champ reçu par un mobile présent à bord d'une rame de RER parcourant la ligne A de Vincennes vers Auber en heure creuse

Il semble, dans le cas de la station de Châtelet-les-Halles que les antennes fonctionnent par paires. En effet, une paire couvre le quai et une autre paire les tunnels. On considère qu'un mobile reçoit un niveau de champ suffisant d'une antenne pour se connecter à sa cellule si le niveau de champ est supérieur à -110dBm. On constate alors une discontinuité dans le niveau de champ minimum reçu par le mobile. Quelles en sont les conséquences au niveau des évènements ? Nous précisons que la discontinuité entre la deuxième et la troisième ligne verticale noire n'est qu'une discontinuité de couverture par des cellules souterraines. En effet, cette section est aérienne. Pour faciliter la lecture, nous avons choisi de ne pas représenter les champs reçus par les antennes aériennes.

Nous confrontons ici les distributions de LAUN autour des passages des trains au niveau de deux stations de configuration différente en terme de nature des sections voisines : Vincennes et Châtelet. La Figure 33 montre la distribution des LAUN autour du passage d'un train à Vincennes qui est une station souterraine entourée de portions de voies aériennes. La Figure 34 montre la distribution des LAUN autour du passage d'un train à

Châtelet qui est une station souterraine entourée de portions de voies souterraines et dont les stations voisines sont également entourées de voies souterraines.

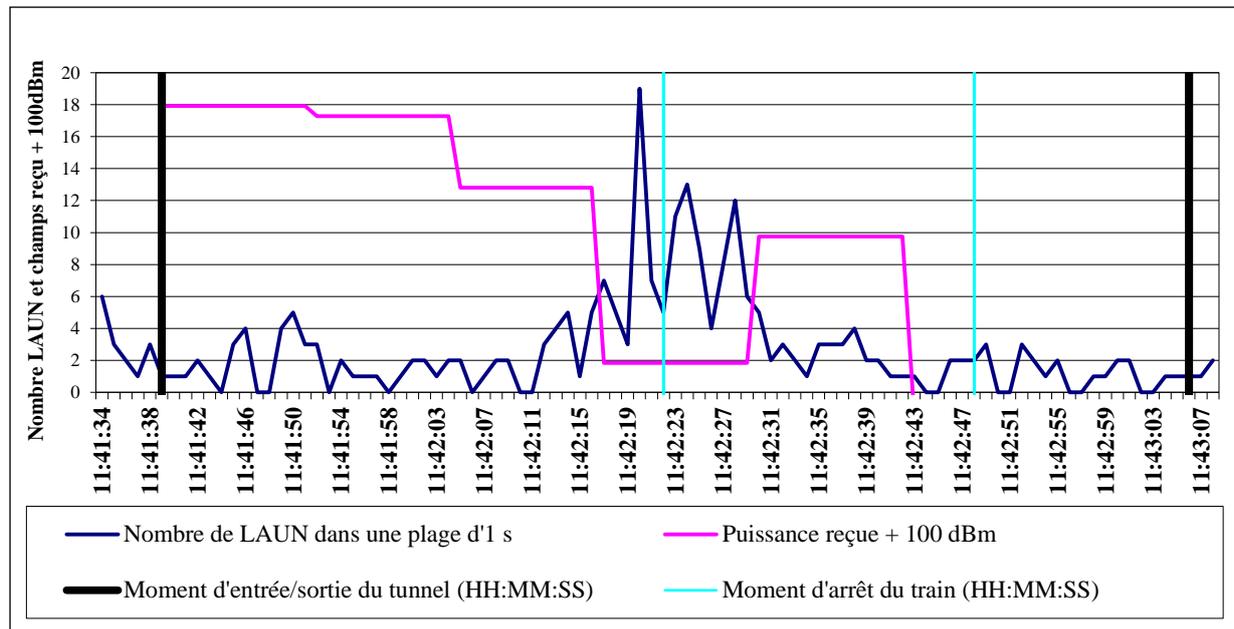


Figure 33 : Corrélation entre le champ reçu par un mobile et le déclenchement d'évènements LAUN par l'ensemble des mobiles connectés au niveau de la cellule qui couvre la station Vincennes (dBm +100)

Dans le cas de Vincennes, la plage de temps d'observation s'étend de 11h41 à 11h44. Les portes sont restées ouvertes pendant 26s. Sur une plage de temps de 60s autour de l'ouverture des portes, il s'est produit un total de 190 LAUN.

Dans le cas de Châtelet, la plage de temps d'observation s'étend de 11h50 à 11h55. Les portes sont restées ouvertes pendant 29s. Sur une plage de temps de 60s autour de l'ouverture des portes, il s'est produit un total de 160 LAUN.

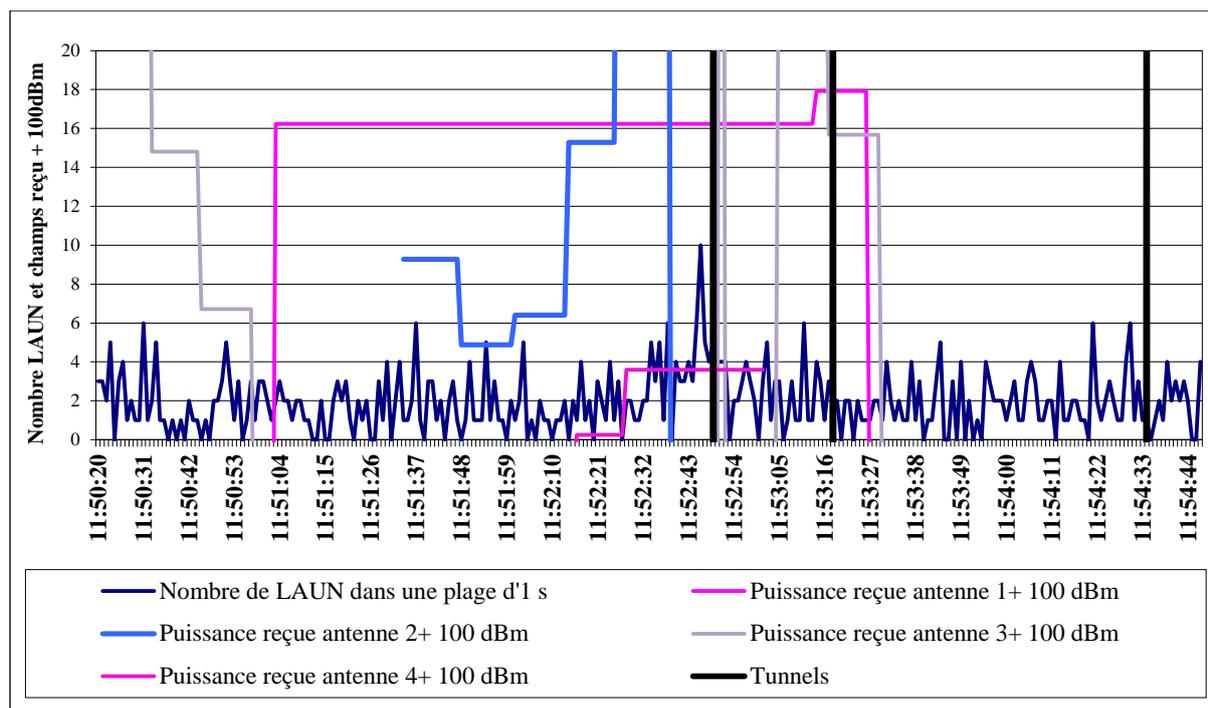


Figure 34 : Corrélation entre le champ reçu par un mobile et le déclenchement d'évènements LAUN par l'ensemble des mobiles connectés au niveau des cellules qui couvrent la station Châtelet (dBm +100)

Pour ces deux stations, on observe une distribution des LAUN qui se décompose en un « bruit de LAUN » et en un pic de LAUN autour de l'ouverture des portes du train. Le "bruit de LAUN" continu que l'on observe en dehors de l'arrivée du train en station est commun aux deux stations. Il correspond aux personnes entrant dans la station en mode piéton, que ce soit pour débiter un trajet ou pour effectuer une correspondance. Il est caractérisé par un étalement homogène et semble indépendant du passage des trains. Ceci est cohérent avec les résultats de l'étude menée par le STIF qui a montré qu'à partir d'une offre d'un train toutes les 8 minutes, plus de 50 % des voyageurs arrivent à la gare sans référence à l'horaire des trains (ROQUIGNY, 2013). En théorie, le pic de LAUN que l'on observe devrait correspondre à un important volume, par rapport au "bruit ambiant" de mobiles qui déclenchent un évènement dans une plage de temps resserré. Il est donc censé se produire de manière synchrone avec le passage des trains. C'est ce que l'on observe au niveau des deux stations autour de l'ouverture des portes. Ce résultat était attendu pour Vincennes en raison de sa configuration de station souterraine entourée de sections aériennes. En effet, tous les mobiles à bord du train vont signaler ce changement de zone de localisation par un LAUN dans une certaine plage de temps. Cette explication ne s'applique pas à Châtelet car il n'y a pas de changement de zone de localisation pour un mobile qui entre ou sort de cette station à bord d'un train. Cela montre qu'il existe un phénomène de déclenchement d'évènements supplémentaires par rapport aux attentes théoriques. Ce dernier peut s'expliquer par des besoins de renvoi de localisation des mobiles au réseau en raison des conditions de

couverture en souterrain. Enfin, l'ordre de grandeur semblable entre les pics de LAUN au niveau des deux stations tient de la fréquentation plus importante de la station Châtelet sur une plage horaire 11h-12h et non d'un déclenchement de l'ensemble des mobiles d'un évènement de localisation.

Conclusion

Ces premières mesures nous ont montré la faisabilité de capturer différents aspects de la mobilité des usagers des transports communs ferrés au niveau des zones RATP 1 et 2, c'est-à-dire Paris et proche banlieue. En effet, des évènements de mobilité sont déclenchés dans certaines situations particulières :

- L'entrée/sortie dans le système de transport au niveau des gares souterraines
- La circulation d'un train avec variation de type de sections (extérieur/souterrain)
- La circulation d'un train sans variation de type de sections.

Le premier comportement est illustré sur les figures précédentes au travers du "bruit" récurrent de LAUN. Les deux autres comportements s'observent au travers des pics de LAUN au moment des arrêts en station et d'ouverture des portes. Comment distinguer ces comportements du flux d'évènement capturé ? Pour cela, il faut prendre en considération quelques limitations du système de capture. En effet, la Figure 32, montre des zones sans couverture radio dans les tunnels. Ainsi, certains évènements peuvent être retardés, après rejet ou non, voire être déclenchés sans changement de zone de localisation mais après une perte trop longue de connexion au réseau impliquant une mise à jour de zone de localisation. En effet, le réseau informe le mobile de sa localisation dans le « territoire mobile » (numéro de zone de localisation) et ce dernier communique régulièrement sa dernière position enregistrée. Ainsi, si le réseau n'a pu « atteindre » le mobile pour lui communiquer la zone de localisation à enregistrer, celui-ci par défaut sauvegarde un numéro de zone de localisation invalide. Lorsque la connexion sera disponible à nouveau, le mobile cherchera à remettre sa zone de localisation à jour en demandant au réseau une mise à jour de zone de localisation. Ces mécanismes de retard d'évènements ou de création d'évènements supplémentaires ne concernent pas tous les mobiles. En effet, cela dépend des caractéristiques techniques du terminal mobile, notamment de la sensibilité au champ reçu, de l'environnement radio et du temps d'exposition à cet environnement. Typiquement, un train arrêté ou ralenti entre deux stations, dans une zone où la couverture des cellules de la station amont et de la station aval est faible aura une probabilité plus forte de générer des évènements de mise à jour de zone de localisation. Ce cas de figure apparaît également lors d'une connexion entre deux lignes du réseau de transport en commun au niveau d'une station. En effet, tous les tunnels piétons ne sont pas couverts avec la même puissance de signal, et l'on recrée ainsi la conjoncture des sections de tunnel. Le nombre important d'évènements collectés en chaque station offre un avantage statistique dans la mesure des déplacements si l'on est capable d'isoler les conditions qui les ont déclenchés.

1.3 Description des sources de données exploitées

Ce paragraphe permet de présenter les données que nous avons utilisées pour qualifier nos méthodes de mesures de déplacements à partir des messages de signalisation. Ces descriptions servent également à introduire les notations pour la suite du chapitre. Nous débuterons par décrire les sources complémentaires car elles sont nécessaires pour développer une première analyse de la source radio-mobile.

1.3.1 Jeux de données complémentaires

Nous présentons dans ce paragraphe, les jeux de données annexes constitués pour valider les résultats produits à l'aide de notre système de capture.

1.3.1.1 Observation à quai

Le 7 avril 2011, pendant l'heure de pointe du matin, de 7h15 à 9h03, 6 observateurs se sont tenus au centre des plateformes des stations Vincennes à Charles de Gaulle-Etoile de la ligne A. Ainsi, pour tous les trains circulant dans le sens Est vers Ouest, ils ont relevé le type de train, une estimation qualitative du taux d'occupation de la voiture en face d'eux à l'instant de l'arrêt du train, l'instant d'arrêt total du train, les instants d'ouverture/fermeture des portes et l'instant de départ du train. Le type de train est une donnée importante pour estimer le taux d'occupation. En effet, trois types de train sont opérés sur la ligne A faisant varier la capacité totale d'une rame entre 1760, 1900 et 2580 voyageurs pour une densité de voyageurs nominale de 4 personnes/m². Les capacités de places assises respectives sont de 432, 600 et 1056. L'échelle qualitative du taux d'occupation de la voiture est la suivante :

- **Faible (*low*)**: places assises disponibles, aucun voyageur n'est debout
- **Moyenne (*medium*)** : toutes les places assises sont occupées et quelques voyageurs sont debout
- **Elevée (*high*)** : beaucoup de voyageurs sont debout
- **Très élevée (*very high*)** : les voyageurs debout sont pressés contre les portes.

1.3.1.2 Données billettiques

1.3.1.2.1 Présentation du jeu de données

Le jeu de données billettiques, fourni par le STIF, est composé de toutes les validations des cartes à puces d'abonnement au niveau de barrières de péage permettant l'accès à la ligne 1 du métro et au niveau des valideurs permettant l'accès et la sortie de la ligne A du RER, dans sa totalité, le jeudi 13 octobre 2011. Le STIF considère qu'en heure de pointe un jour ouvré moyen, 90% des voyageurs possèdent une carte d'abonnement (« pass Navigo »). Les conditions de déclenchement d'un événement billettique alimentant le jeu de données sont le passage de la carte devant une barrière de péage. Les règles de validation sont différentes pour l'utilisation du Métro et du RER. Dans le cas du RER, une validation est

requis à l'entrée, à la sortie et lors d'une connexion Métro vers RER. Dans le cas du Métro, une validation est requise seulement à l'entrée. Aucune validation n'est nécessaire dans le cas d'une connexion RER vers RER. À chaque validation, une ligne est ajoutée à la base de données billettiques. Elle contient un identifiant unique de carte d'abonnement *id*, le code de la station *cs* et le moment de validation *h* (en seconde écoulées depuis minuit). Le code station est composé d'un identifiant de ligne, permettant de distinguer RER et Métro et d'un identifiant de station.

1.3.1.2.2 Sources d'erreurs du système de capture

Le système de mesure des déplacements est sujet également à la production d'erreurs d'interprétation des déplacements. On considère trois principales sources d'erreur : une source statistique, une source relative aux équipements d'accès et de sortie et une source relative au fonctionnement du système de transport. Premièrement, la représentativité des abonnés Navigo parmi les usagers du réseau de transport en commun ferré varie dans la journée et selon la paire Origine/Destination considérée. Néanmoins, l'approximation d'un taux de 90% d'usagers munis de « pass Navigo » en heure de pointe sur des paires Origine-Destination domicile – travail n'est pas irréaliste. Deuxièmement, le déclenchement d'un événement billettique se produit lorsqu'une carte d'abonnement est présentée devant une des bornes de validation. L'horloge des bornes n'est pas synchronisée sur l'ensemble du réseau, voire même au niveau d'une station. Ces différences d'horloge entraînent une erreur à compenser lors de la production de temps de parcours. De plus, chaque borne contient l'information de station et de ligne. Les stations dont l'offre de transport est composée de plusieurs lignes sont équipées de plusieurs barrières contenant l'information de ligne. En pratique, les bornes qui forment les barrières produiront une information de ligne relative à celle qui est la plus proche en distance. Or, il est possible pour un usager d'entrer par une borne proche d'une ligne qui n'est pas celle qu'il va utiliser. Cette erreur quant à la station d'origine est quantifiée par l'observation de paires Origine/Destination. Cela permet la distinction entre Métro et RER et entre les lignes de RER. En effet, selon les informations de station d'origine et de station de destination, l'utilisateur est affecté à la ligne la plus probable. En pratique, plus les stations sont éloignées, plus l'ensemble des lignes empruntées se réduit. Néanmoins, on ne peut établir de distinction entre les lignes de Métro par défaut d'évènement de sortie du réseau métropolitain. Troisièmement, il arrive que le fonctionnement du système de capture soit mis en défaut lorsque, pour des raisons d'opération, les barrières de sortie du RER sont « ouvertes » pendant une période donnée. Cela peut être le cas au niveau de stations « chargées » pour faciliter le flux de sortie de la station. D'après l'analyse des données AFC, aucune baisse significative de flux de sorties des stations n'a été observée le 13 octobre 2011.

1.3.2 Exploration des événements de signalisation

Ce paragraphe est consacré à une exploration préliminaire du jeu de données radio-mobile du 7 avril. Nous avons cherché à caractériser les possibilités de ce système pour la mesure de déplacements. Pour qualifier le système de capture, nous avons observé des trajectoires spatio-temporelles comprenant plus de 3 points. Néanmoins, ces trajectoires ne permettent pas de constituer des données personnelles car les points pris en considération sont restreints au tronçon central du RERA pendant l'heure de pointe du matin. Ainsi, étant donné le nombre de voyageurs (50 000 voyageurs/heure/sens) circulant sur ce tronçon pendant cette période, il est impossible de remonter à la notion d'individu.

Le but de cette analyse préliminaire est de confronter trois aspects du système de mesure: l'activité des mobiles, l'activité au niveau des stations et l'activité par paires de stations constituant les trajectoires des trains. Pour faciliter la lecture de ce paragraphe, nous faisons de temps à autre, l'amalgame entre station et groupe de cellules de la zone de localisation souterraine couvrant la ligne A du RER au niveau de cette station.

1.3.2.1 Activité des mobiles

Nous noterons **E_CRERA** l'ensemble de cellules couvrant les stations du tronçon central - entre Vincennes et La Défense - de la ligne du RER A. A partir des flux de signalisation « écoutés » par la sonde au niveau du MSC contenant les cellules de **E_CRERA**, nous avons compté le nombre d'évènements de type **LAUN** déclenchés au niveau de ces cellules lors de la journée du 7 avril 2011. Parmi, l'ensemble des mobiles ayant déclenché au moins un évènement sur **E_CRERA**, 11% ont généré 1 seul évènement, 6% 2 évènements, 3% 3 évènements et 1,6% plus de 10 évènements, pour une moyenne de 3,4 évènements générés par mobile au niveau sur **E_CRERA**. Ceci constitue un écart significatif avec les résultats théoriques. En effet, un évènement **LAUN** sur **E_CRERA** correspond à un déplacement sur le « territoire mobile » d'une des zones de localisation extérieures vers la zone de localisation souterraine. Dès lors, un évènement correspond soit à une entrée dans le système de transport au niveau du tronçon central, soit à un déplacement à bord d'un train arrivant depuis une section aérienne de la ligne. Sur la Figure 35, on trouve la configuration des trajets réalisables sur le RERA autour de notre périmètre d'étude entre La Défense et Vincennes. La nature des sections – souterraine et aérienne – est représentée pour notre périmètre.

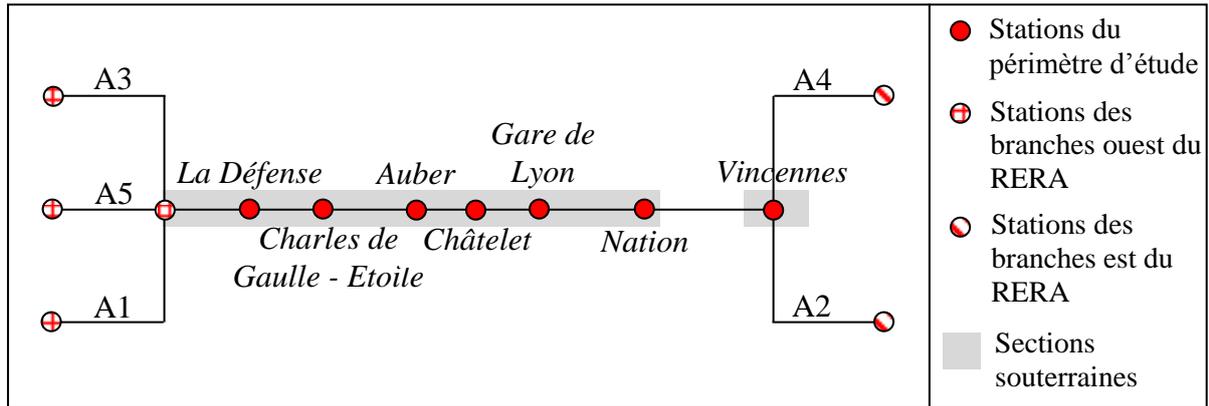


Figure 35 : schéma simplifié de la configuration des trajets réalisables par les usagers de la ligne A

A partir de cette Figure 35, nous proposons un tableau du nombre d'évènements LAUN qui seront générés en théorie pour des trajets O-D. Nous distinguons 5 groupes de stations pour constituer ces O-D. Le premier groupe, noté $\{A1, A3, A5\}$ contient l'ensemble des stations des branches ouest du RERA ainsi que la station Nanterre-Préfecture. Le deuxième groupe, noté $TC \setminus \{Vinc\}$, contient les stations du tronçon central, représentées par un cercle rouge plein sur la figure, à l'exception de Vincennes qui constitue le troisième groupe. Le quatrième groupe de station est composé des stations des branches Est et est noté $\{A2, A4\}$. Précisons que nous considérons ces origines et destinations comme les points d'entrée et de sortie du système de transport en commun ferré. Nous considérerons alors un dernier groupe de correspondances au niveau des stations de $TC \setminus \{Vincennes\}$, noté Corresp. Nous reportons dans le Tableau 3, le nombre théorique d'évènements LAUN attendus au niveau des cellules de E_CRERA pour des trajets entre les différents groupes de stations proposés.

O\D	$\{A1, A3, A5\}$	$TC \setminus \{Vinc\}$	Vincennes	$\{A2, A4\}$	Corresp
$\{A1, A3, A5\}$	X	0	1	1	0
$TC \setminus \{Vinc\}$	1	X	2	2	1
Vincennes	2	2	X	1	2
$\{A2, A4\}$	2	2	1	X	2
Corresp	0	0	1	1	X

Tableau 3 : nombre d'évènements LAUN attendus au niveau des cellules de E_CRERA pour différents trajets sur la ligne A

Par groupe de deux cases du tableau, nous faisons apparaître le nombre de LAUN théoriques attendus pour des migrations pendulaires sur le réseau de transport. Le nombre d'évènements attendus au niveau des cellules de E_CRERA se situe entre 0 et 4 évènements.

Alors comment expliquer la distribution du nombre d'évènements observée par mobiles qui montre que la majorité des mobiles a généré entre 4 et 9 évènements ?

Ce décalage entre la théorie et la pratique peut s'expliquer par des hypothèses erronées. Les hypothèses que nous avons faites portent sur deux aspects. Premièrement, nous avons fait des hypothèses relatives à une connaissance à priori des comportements de déplacements au niveau du tronçon central de la ligne A du RER un jour ouvré moyen. Nous avons supposé d'une part que la majorité des déplacements étaient des migrations pendulaires domicile-travail (**H1**) et d'autre part que le classement des poids des stations ne comptait pas Vincennes comme l'une des stations générant le plus de flux (**H2**). Deuxièmement, nous avons supposé un fonctionnement théorique du système de capture basé sur le découpage particulier du « territoire mobile » décrit ci-avant (**H3**).

La vérification des hypothèses passe par une confrontation avec la source de données décrite en 2.3.1.2. Il s'agit de données de validations billettiques collectées par le STIF. La date de collecte n'est pas la même pour les deux sources. Cependant, elles ont toutes deux été produites un jour ouvré moyen. On supposera que les comportements globaux de déplacements pour motif domicile-travail sont comparables entre le 7 avril et le 13 octobre 2011.

Validation de H1

Testons tout d'abord l'hypothèse de comportement de déplacement. Est-il raisonnable d'assumer que la majorité des déplacements observés un jour ouvré moyen sur le tronçon central sont des migrations pendulaires ? Nous nous sommes donc intéressés aux plages horaires dans lesquelles la majorité des déplacements se sont produits. Nous définissons pour cela deux plages horaires. La première de 7h à 9h correspond à l'heure de pointe du matin (**HM**) et la deuxième de 17h à 20h correspond à celle du soir (**HS**).

Premièrement, nous quantifions la part des déplacements effectués dans ces plages horaires à partir du jeu de données billettique (voir paragraphe 1.3.1.2). La Figure 36 illustre la distribution par tranche de 30 minutes des validations sur l'ensemble des stations du tronçon central. Sur le tronçon central, 61% des validations ont lieu dans les plages horaires **HM** et **HS**. Il existe une symétrie dans le nombre de validations entre les deux périodes pour toutes les stations. Cependant, chacune des stations du tronçon central possède un fonctionnement propre. Par exemple, les stations Châtelet et Gare de Lyon génèrent moins de 60% de leur validations pendant les plages horaires **HM** et **HS**. A l'inverse, Auber, Charles de Gaulle-Etoile et La Défense comptabilisent plus de 65% de leurs validations pendant ces périodes. On trouve une explication dans le fait que Châtelet et Gare de Lyon sont des stations de connexion avec le réseau de transport en commun ferré ou d'autres modes ferrés (Gare de Lyon). De plus, Châtelet est une station centrale dans le réseau de transport car elle dessert une zone d'effervescence « d'activité de loisirs et d'achats » parisienne.

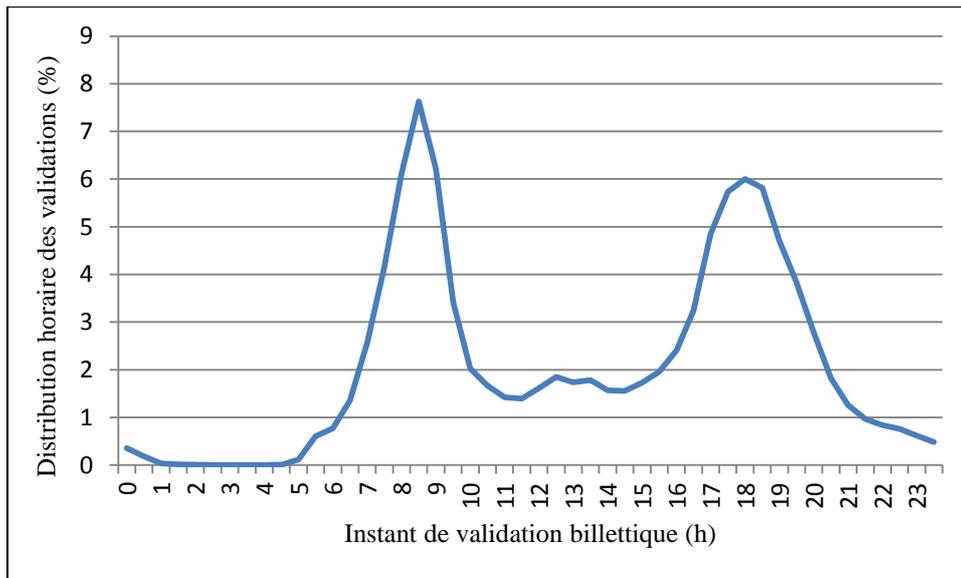


Figure 36 : distribution par tranche de 30 minutes des validations sur l'ensemble des stations du tronçon central obtenues à partir des données billettique

Deuxièmement, nous vérifions que la distribution horaire obtenue avec notre système de capture est similaire. Nous observons que 67% des **LAUN** au niveau des stations du tronçon central ont lieu dans les plages horaires **HM** et **HS**. La Figure 37 illustre la distribution des **LAUN** par tranche de 30 minutes sur le tronçon central à partir des données du 7 avril 2011. On observe également un fonctionnement propre à chacune des stations. En effet, les stations Châtelet, Gare le Lyon, Vincennes et La Défense déclenchent moins de 70% de leur **LAUN** pendant les plages horaires **HM** et **HS**. A l'inverse, Nation, Auber et Charles de Gaulle-Etoile produisent plus de 75% de leur **LAUN** pendant les périodes **HM** et **HS**. De plus, la symétrie du nombre de validations entre les deux périodes n'est plus vérifiée que pour Charles-de-Gaulle-Etoile.

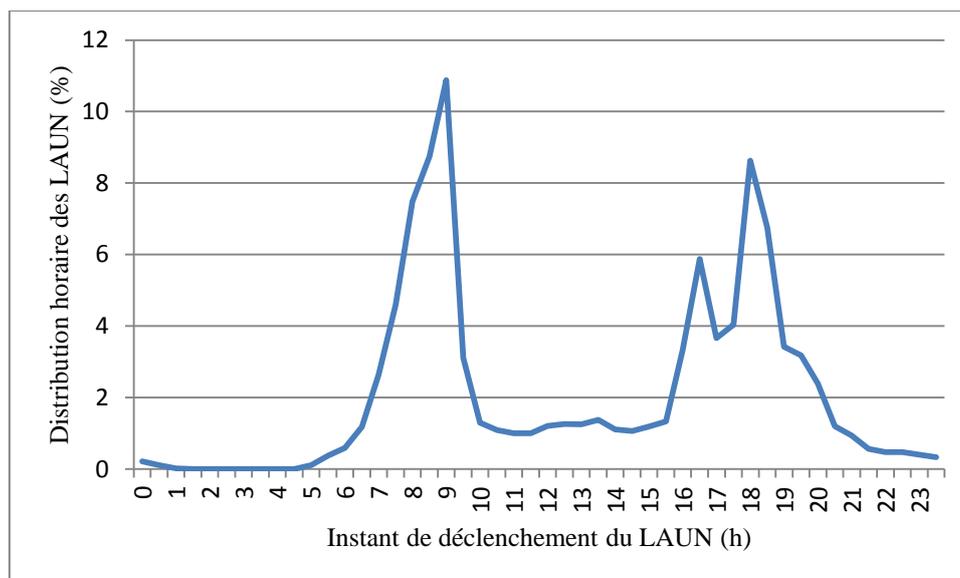


Figure 37 : Distribution des **LAUN** par tranche de 30 minutes sur le tronçon central à partir des données du 7 avril 2011

Ces deux distributions varient de manière similaire avec une concentration autour des heures de pointe. Les deux concordent avec l'hypothèse des migrations pendulaires. Les écarts entre les deux systèmes de mesure s'expliquent d'une part par le fait que la proportion d'abonnés Orange et celle d'abonnés STIF parmi les voyageurs ne varient pas de la même manière entre heures creuses et heures de pointe, et d'autre part par les méthodes de collecte et leur technologies qui diffèrent.

Validation de H2

Nous étudions l'hypothèse relative au comportement des stations quant à la production de déplacement. Est-il raisonnable d'assumer que le poids de la station de Vincennes dans la génération des déplacements est inférieur ou égal aux autres stations? Pour cela, nous nous sommes intéressés au nombre de validations par station dans le jeu de données billettique. Ainsi, nous avons produit le « poids naïf » de chacune des stations du tronçon central. Le qualificatif « naïf » est utilisé car nous considérons toutes les validations billettiques, ces dernières pouvant être soit des entrées, soit des sorties du système. Le Tableau 4, montre que la station Vincennes est celle qui a généré le moins de validation du tronçon central. Ces résultats « naïfs » seraient à redresser pour connaître la proportion d'entrées dans le système au niveau de chaque station. Cependant, ce tableau reflète une qualification naïve du poids des stations et nous permet de valider l'hypothèse **H2**. La station Vincennes n'occupe pas la place la plus structurante dans la production des flux de déplacement du tronçon central.

Stations - S	Pourcentage du nombre de validations au niveau de S parmi toutes les validations du tronçon central
Vincennes	2,99 %
Nation	10,5 %
Charles De Gaulle – Etoile	11,1%
Auber	14,2%
La Défense	18,7%
Gare de Lyon	20,7%
Châtelet - Les Halles	21,8%

Tableau 4 : répartition des validations billettiques sur les stations du tronçon central

Validation de H3

Nous testons ici notre hypothèse de cohérence du système de capture radio-mobile par rapport au fonctionnement théorique de la signalisation au niveau d'un « territoire mobile » connu à priori. Plusieurs facteurs peuvent mettre en défaut cette hypothèse. Par exemple, la différence d'architecture des stations (taille, nombre de connexions ferroviaires et piétonnes) entraîne à la fois une hétérogénéité des comportements de mobilité et des conditions dans lesquelles les ondes radioélectriques se propagent. Afin d'observer les phénomènes à l'œuvre, nous nous sommes intéressés à deux points qui feront l'objet des deux prochains paragraphes. Nous caractériserons l'activité au niveau des stations du tronçon central de la ligne A du RER puis l'activité au niveau de paires de stations.

1.3.2.2 Activité aux stations

Premièrement, nous avons regardé la distribution du nombre de **LAUN** par stations à partir du jeu du 7 avril 2011 (Figure 38). Cette figure montre un déséquilibre du « taux de capture » entre les stations.

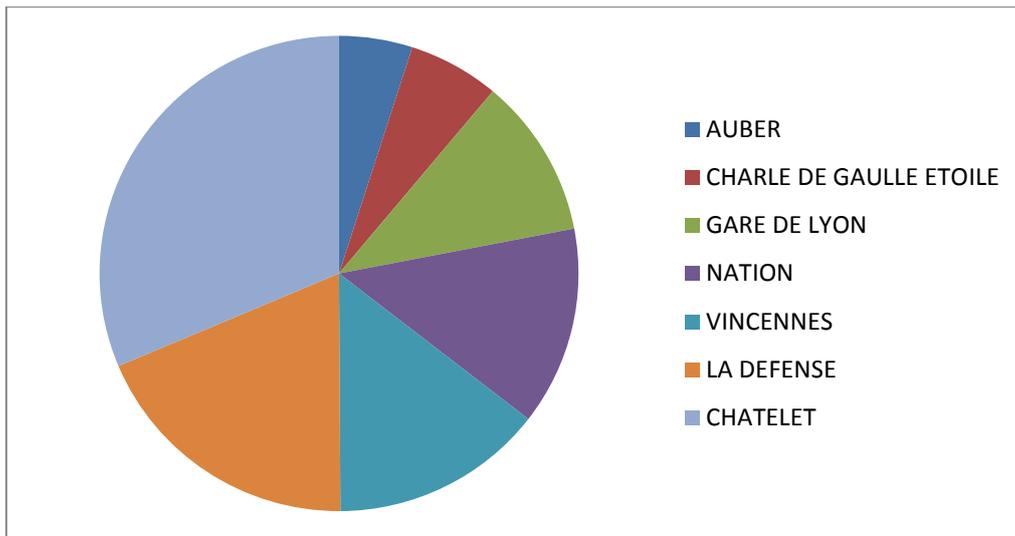


Figure 38 : distribution des LAUN collectés selon les stations du tronçon central

Ces écarts étaient prévisibles car la fréquentation des stations n'est pas homogène. Cependant, le Tableau 4 produit à partir des données billettiques, propose un tout autre classement de fréquentation. Deux phénomènes expliquent cette distribution. D'une part, les comportements de déplacement, et donc la densité de voyageurs varie selon les stations et d'autre part, le système de capture est sensible à la station considérée. En effet, la station Vincennes étant la seule station souterraine entourée de sections aériennes, tout mobilet passant par Vincennes (à bord d'un train ou en entrant dans le système de transport) va déclencher un événement **LAUN**. La station Nation est également particulière car elle est entourée d'une section de rail aérienne à l'est et d'une section souterraine à l'ouest. Ainsi, à Nation, les événements **LAUN** devraient se produire pour les mobiles à bord d'un train en provenance des branches Est du RERA et lors de l'entrée dans le système de transport. Le fonctionnement théorique des autres stations est le suivant : un LAUN est généré si et seulement si un mobile entre dans le système de transport. La Figure 38 montre que 6 fois moins de personnes « entrent » dans le système de transport par Auber que par Châtelet. Ceci ne reflète pas le comportement transport théorique de ces deux stations. On peut supposer qu'un autre mécanisme relatif à notre compréhension du système de capture radio-mobile est à l'œuvre.

Deuxièmement, nous avons tracé l'évolution de cette distribution du nombre de **LAUN** par tranches temporelles de 30 minutes afin d'analyser les phénomènes temporels sous-jacents (Figure 39).

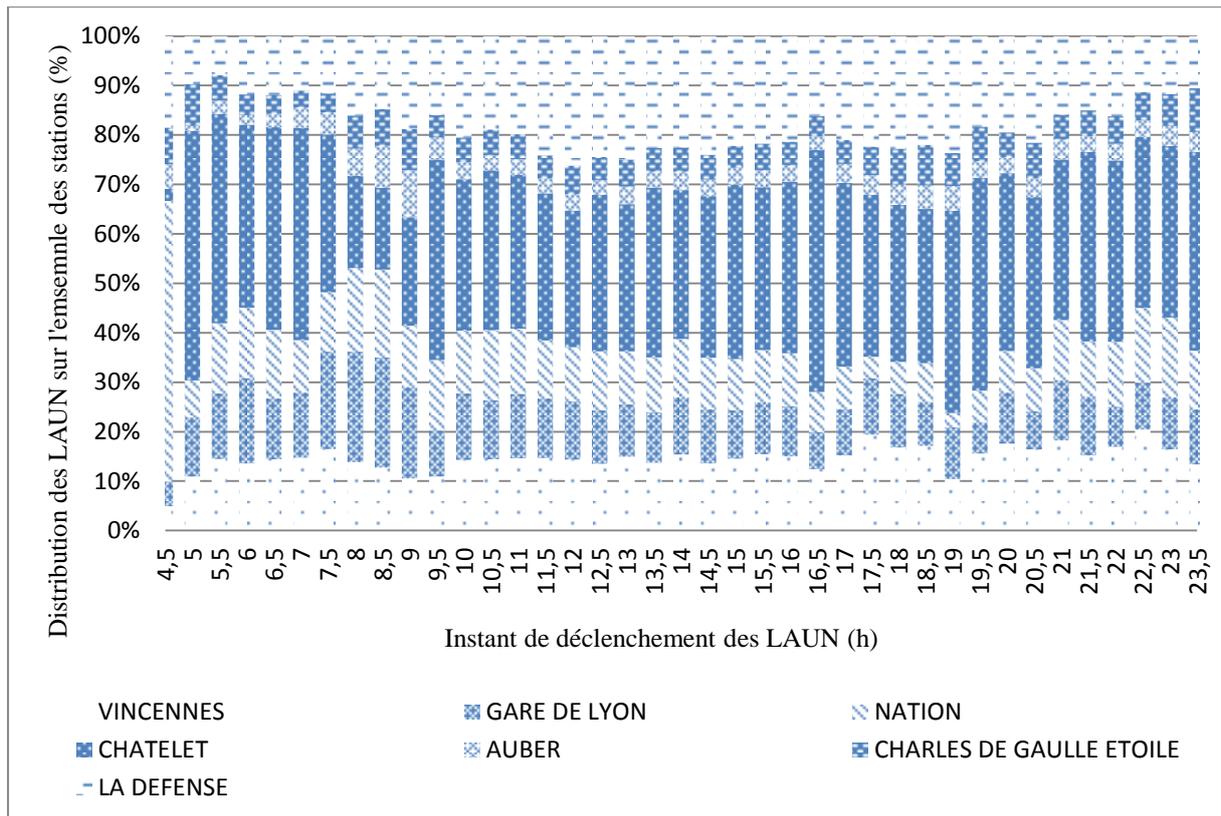


Figure 39 : évolution de la distribution des LAUN parmi les stations du tronçon central au cours de la journée du 7 avril 2011

La distribution varie avec le moment de la journée. Ces variations s'expliquent de manière différente selon les stations. En effet, on remarque que le « taux de capture » de La Défense diffère entre les périodes **HM** et **HS** tout en augmentant en heure creuse. On trouve deux explications à ce phénomène. D'une part, les antennes couvrant la ligne A du RER couvrent également la ligne 1 du métro au niveau de la station La Défense. D'autre part, cet écart entre **HM** et **HS** vient du fait que le système capture les entrées dans le système de transport souterrain et non les sorties. Ainsi, lorsque la destination du trajet domicile-travail est la Défense, nous mesurerons une activité au niveau de la Défense plus élevée lors du trajet retour du soir que lors du trajet aller du matin. Cette explication de variation de « taux de capture » entre **HM** et **HS** s'applique de manière identique au cas de la station Châtelet. On observe le phénomène inverse pour les stations Nation et Gare de Lyon dont le « taux de capture » est bien plus élevé le matin que le soir. Dans le cas de Nation, on remarque qu'après 20h, le « taux de capture » augmente significativement symétriquement à Vincennes. Il s'agit des trajets de retour depuis la banlieue Est. Quant aux stations Auber et Vincennes, elles conservent un « taux de capture » quasi-constant dans la journée.

Troisièmement, on observe la distribution selon le nombre d'événements réalisés au niveau d'une station. Afin d'observer la symétrie des migrations pendulaires, pour chacune des stations, on a tracé sur la Figure 40, la proportion d'événements ayant été répétés n fois

(pour n de 1 à >10) au niveau d'une même station parmi tous les événements s'étant produits au niveau de la station. Par exemple, sur tous les événements qui ont eu lieu à Châtelet, 18% sont des événements répétés deux fois par un même mobile.

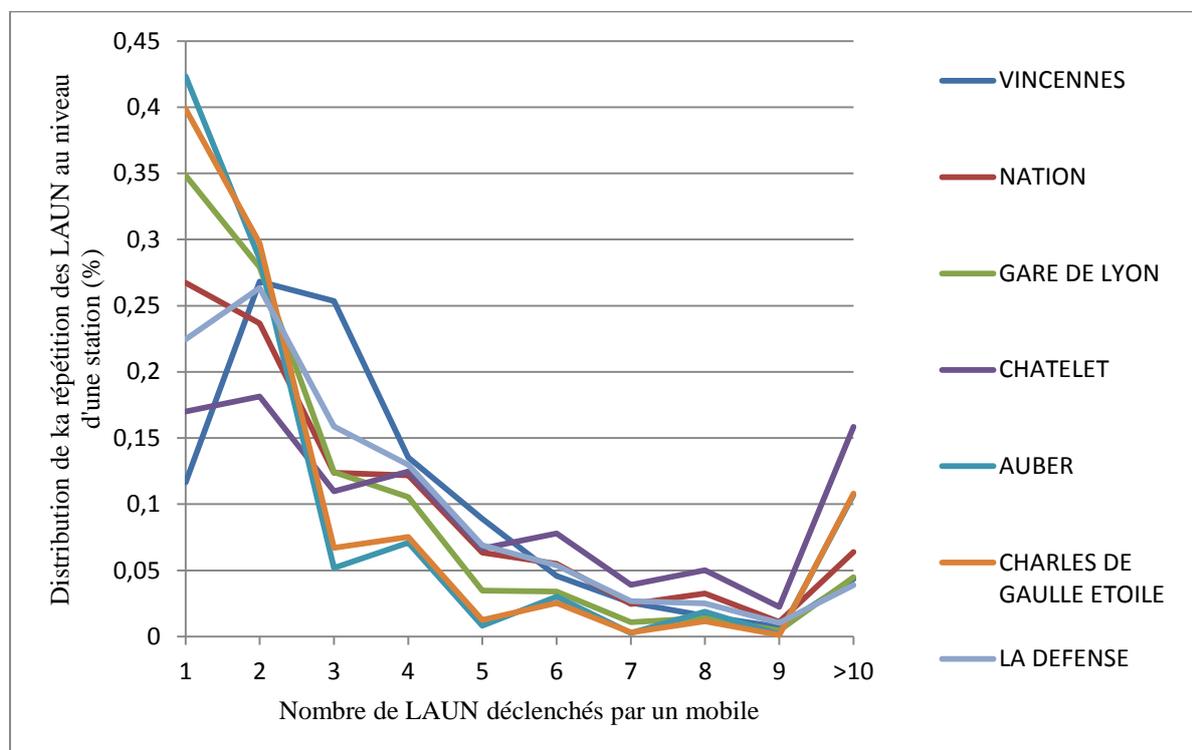


Figure 40 : distribution de la répétition des événements au niveau d'une station

On s'attend à une allure de courbe décroissante avec un maximum pour 1 seul événement réalisé. En effet, nous mesurons les entrées dans le système via une station d'origine. Dans le cas de trajets domicile-travail, deux entrées sont à mesurer : l'entrée proche du domicile le matin et l'entrée proche du travail le soir. Ceci se vérifie pour l'ensemble des stations à l'exception de Vincennes, Châtelet et La Défense qui présentent un maximum pour 2 événements réalisés. L'explication pour le cas de Vincennes est évidente. Il s'agit d'une station souterraine entourée de portions aériennes, ainsi, tout accès vers Vincennes déclenche un LAUN. Cette explication ne tient pas pour Châtelet et La Défense entourées de portions souterraines. Sa seule particularité est la non-dissociation sur le « territoire mobile » entre la ligne de RER et la ligne de métro. Observons alors la distribution de l'écart temporel inter-événements pour les mobiles ayant déclenché exactement 2 LAUN au niveau de la même station. Considérons deux classes de distance temporelle entre les événements, la classe des événements se produisant à moins d'une heure d'intervalle et celle de ceux qui se produisent à plus d'une heure d'intervalle. Pour l'ensemble des stations, parmi les événements qui ont une distance temporelle inférieure à une heure, plus de 95% se produisent avec un écart inférieur ou égal à 15 minutes. Ce sont des événements supplémentaires par rapport à ceux attendus en théorie. Une nouvelle

procédure de **LAUN** peut être demandée en réaction à l'échec d'une première tentative de mise à jour de localisation ou en cas de perte d'information de zone de localisation suite à une dégradation des conditions de couverture. Les tunnels des lignes de train ainsi que les tunnels de connexion piétonnes offrent des conditions de couverture radio particulièrement dégradées. Cela explique le pourcentage élevé de mobiles ayant réalisés plus de 10 **LAUN** à Châtelet. La proportion des mobiles qui ont déclenché exactement deux évènements au niveau d'une même station dans un intervalle de 15 minutes est de 32% à Vincennes alors qu'elle est comprise entre 80% et 90% pour les autres stations hormis La Défense. En effet, cette station se distingue des autres avec un taux de 70%. Intéressons-nous à la deuxième classe de mobile dont les 2 évènements générés sont espacés temporellement de plus d'une heure. Sur la Figure 41, on représente la distribution des distances temporelles inter-évènement par tranche d'une heure pour différentes stations lorsqu'un mobile a généré deux évènements **LAUN** au niveau de la même station dans un intervalle de temps supérieur à 15 minutes.

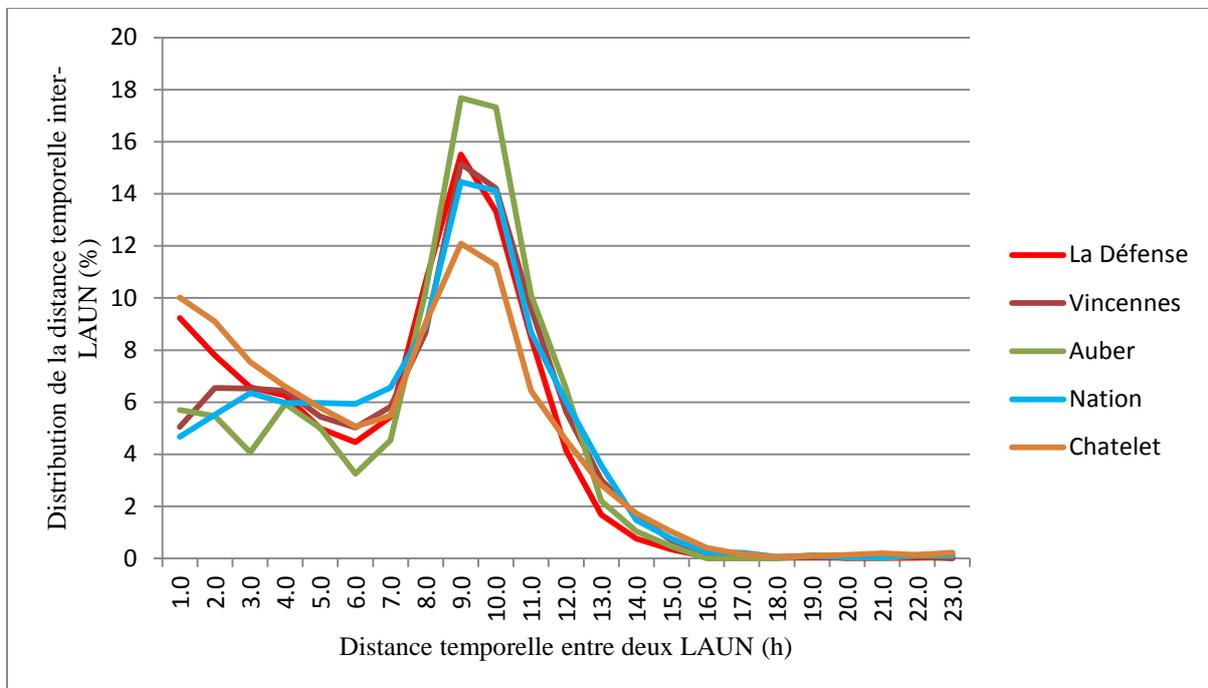


Figure 41: distribution des distances temporelles inter-évènements pour différentes stations

La majorité des mobiles qui ont réalisé deux évènements au niveau d'une même station, les ont espacés d'une journée de travail (entre 7h et 11h). On remarquera que les stations Châtelet et La Défense ont un pic pour une distance temporelle entre 1h et 2h. Ceci correspond plutôt à un comportement achats/loisirs. Les activités localisées à proximité d'une station influent sur les comportements de déplacements à destination de cette station.

Pour conclure cette analyse de l'activité au niveau des stations, nous constatons une dépendance de notre système de capture à la station considérée et au moment d'observation. Notre système de capture repose sur le déclenchement d'évènements lors de déplacements du téléphone mobile sur le « territoire mobile ». Ainsi, on peut distinguer deux causes creusant l'écart entre la théorie et la pratique. Tout d'abord, le « territoire mobile » est très sensible aux conditions de propagation des ondes radio-électriques. Ces conditions se dégradent facilement dans des environnements tels que des tunnels. Ensuite, certains comportements de mobilité vont placer des abonnés dans ces parties aux conditions de propagation dégradée. Typiquement, les couloirs piétons assurant les connexion entre les lignes du réseau de transport en commun et les sections ferrées entre deux stations sont des zones où la probabilité d'avoir une qualité de réception du signal radio-électrique détériorée est plus forte. Le fait que les comportements de mobilité varient selon le moment la journée entraînent une variation du « taux de capture » avec le temps. La conséquence directe est la déclenchement de **LAUN** supplémentaires par rapport à ceux attendus en pratique. Ces déclenchements « inattendus » dépendent eux même de la station. Plus précisément, il s'agit d'une dépendance à la localisation exacte, ou la succession de localisations, dans la station sur toute la durée de la procédure de **LAUN**. Ainsi, des **LAUN** peuvent se produire sans qu'ils ne correspondent à une entrée dans la zone de localisation souterraine. Dans certains cas, on observe une répétition de l'évènement au niveau d'un même groupe de cellules couvrant une ligne d'une station. Le déclenchement de ces évènements supplémentaires est un avantage car il permet d'augmenter la taille de l'échantillon. Cependant, la répétition de **LAUN** en situation de mobilité pourrait entraîner un décalage spatial dans le cas où le délai de succès de la procédure excède le temps de présence dans le groupe de cellules associées à une station. En effet, quelles sont les conséquences de la dépendance du système de capture aux stations sur l'étude des déplacements d'une station vers une autre ? La section suivante traite de l'impact de la paire de stations considérée dans l'étude des trajectoires par paires de stations.

1.3.2.3 Etude des trajectoires

Tout d'abord, définissons ce que nous appellerons la trajectoire d'un mobile. La trajectoire d'un mobile m est constituée de n points datés et localisés (h,s) avec $n \in [2, 7]$ car le périmètre de notre étude est constitué de 7 stations. Dans notre cas, la localisation est à la station s . Le temps h est celui du premier évènement **LAUN** au niveau de la station. Les points de la trajectoire sont ordonnés dans le temps. Ainsi, une trajectoire de taille n est constituée de la manière suivante :

$$\forall i \in [0, 7], \quad \forall h_n - h_1 \in [20, 60] \text{ minutes},$$

$$(m_1 = \dots = m_i = \dots = m_n) \ \&\& \ (h_1 < \dots < h_i < \dots < h_n) \ \&\& \ (s_1 < \dots < s_i < \dots < s_n)$$

(15)

L'ordre des stations dépend du sens de parcours de la ligne. L'intervalle de temps pour le parcours de la ligne depuis Vincennes vers La Défense est de 23 minutes selon le site de l'opérateur RATP. Nous gardons une marge de temps afin de prévenir des décalages entre notre système de capture d'entrée en souterrain, des comportements d'arrivée en station avancée par rapport au passage du train et des phénomènes perturbant la ligne qui augmentent les temps de parcours.

Observation de la trajectoire des trains

Nous débutons cette étude par l'observation de trajectoires de taille 2 entre les stations Vincennes (S_V) et Nation (S_N), c'est-à-dire l'ensemble des trajectoires des mobiles telles que

$$\forall |h_N - h_V| < 24 \text{ heures},$$

$$(m_N = m_V) \& \{ (h_N < h_V) \& \& (s_N < s_V) \ || \ (h_N > h_V) \& \& (s_N > s_V) \} \quad (16)$$

Nous observons alors les trajectoires entre ces deux stations pour les deux directions de Vincennes vers Nation et de Nation vers Vincennes à l'échelle de la journée. Nous représentons ces trajectoires sur la Figure 42 qui présente en abscisse l'instant de l'évènement à Vincennes et en ordonnée l'instant de l'évènement à Nation.

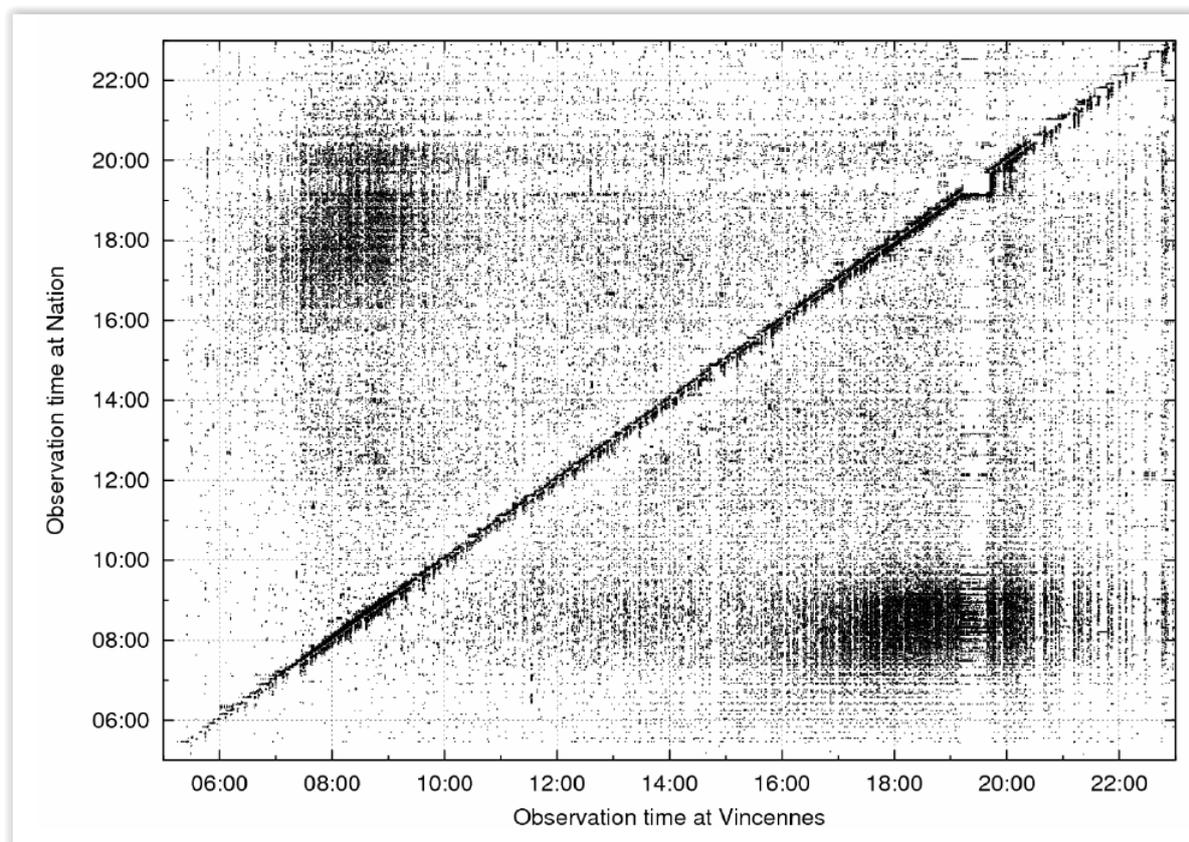


Figure 42 : trajectoires de taille 2 entre Vincennes et Nation à l'échelle de la journée

Cette construction de trajectoires permet notamment d'observer les comportements de déplacements des usagers de la ligne A à l'échelle des migrations pendulaires. Cette figure confirme l'hypothèse 1 énoncée au début du paragraphe 1.3.1.1 sur les motifs de déplacement. En effet, les zones elliptiques correspondent à l'observation des migrations pendulaires, c'est-à-dire une entrée au niveau d'une des deux stations le matin et une entrée au niveau de l'autre station le soir.

La diagonale correspond au passage des trains. On peut observer la signature particulière du mode de transport ferré sous forme de « paquets ». La Figure 43 représente la distribution des **LAUN** entre Vincennes et Nation entre 8h00 et 9h00 le 7 avril 2011 et propose un zoom sur cette signature particulière des trains sur la signalisation. Les carrés gris au-dessus de la diagonale sont les trains circulant de Vincennes vers Nation. La partie verticale des carrés est à assimiler au temps à quai et la partie horizontale au temps d'arrivée en station avant le passage du train. Des croix rouges sont placées selon les indications d'heure d'arrivée des trains à Nation observées par l'équipe d'enquête terrain (voir 2.3.2.1 pour une description de la source). On capture bien le passage de trains, et donc la fréquence de desserte de la ligne. Néanmoins, on constate également sur cette figure, l'impact de la différence de qualité de capture entre les stations Nation et Vincennes. La densité des carrés sous la diagonale (sens Nation vers Vincennes) est en effet plus faible qu'au-dessus (sens

Vincennes vers Nation). Ceci s'explique par la dysmétrie de la station Nation, en terme de type de section permettant l'accès à la station à bord de la rame.

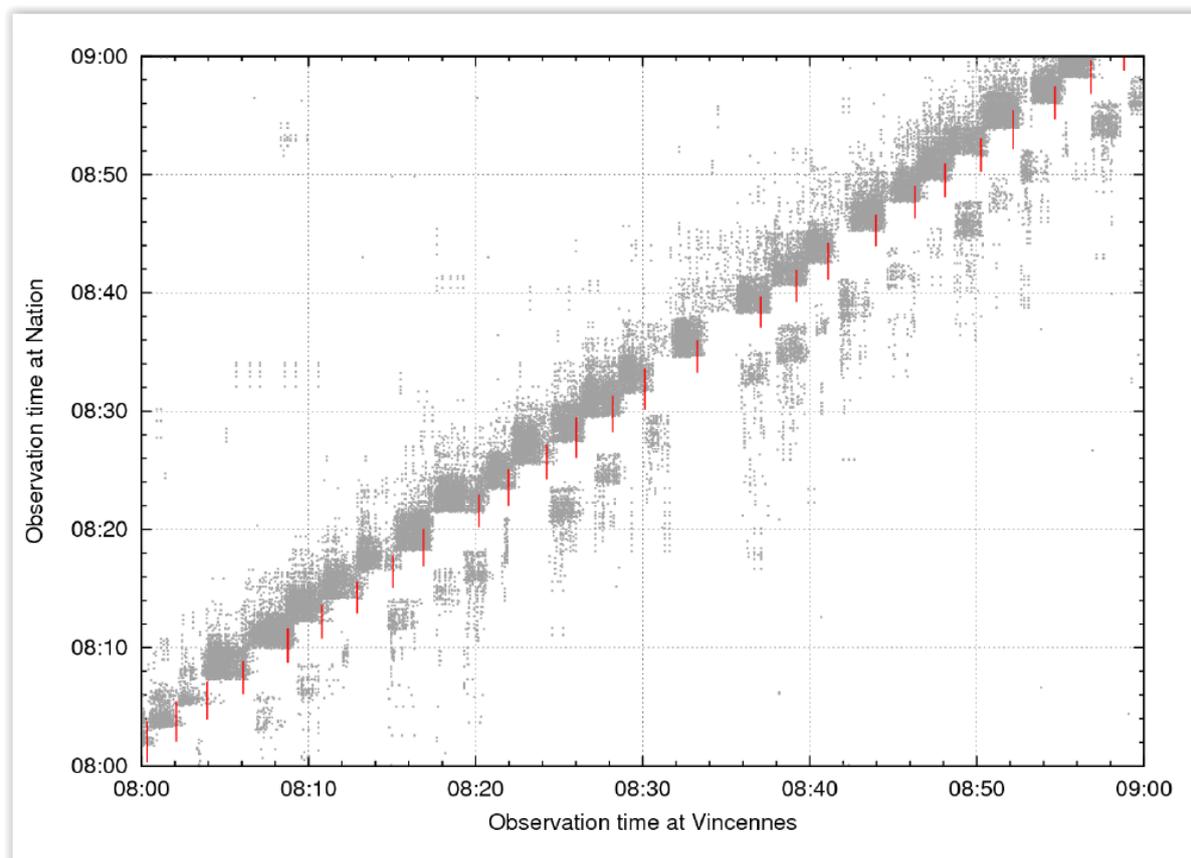


Figure 43 : trajectoires de taille 2 entre Vincennes et Nation entre 8h et 9h le 7 avril 2011

L'ensemble des trajectoires des mobiles permet de reconstruire la trajectoire des trains circulant sur la ligne A du RER. On observe sur la Figure 44 des trajectoires de taille 6 et des trajectoires de taille 3 lorsque la ligne est parcourue dans le sens Vincennes vers La Défense. Les traits noirs représentent les trajectoires des trains produites à partir des observations terrains et les grises celles produites à partir de notre système de capture. On remarque que plus n est grand, plus le nombre de trajectoires est faible. Ainsi, il ne sera pas utile pour nous de reconstruire des trajectoires de plus de 3 points. Les trajectoires mesurées suivent assez fidèlement les trajectoires observées depuis les quais avec un décalage temporel imputable à la taille des cellules qui couvrent également une portion des tunnels et une éventuelle désynchronisation entre les deux méthodes.

On observe une quantité de trajectoires grises qui peuvent être considérées comme des « sauts de trains » sur la figure du dessous. Cela ne correspond pas à une déficience du système de capture mais à un comportement de changements de missions. En effet, le RERA possède 3 branches à l'ouest qui sont desservies selon une permutation. Toutes les stations ne proposent pas le choix complet de mission. Ainsi, un usager sera amené à prendre un train

dont la branche de destination n'est pas celle désirée, puis descendre au niveau d'une autre station proposant sa mission d'intérêt et changer de train. Notre système de capture permet d'observer les stations privilégiées pour un changement de mission.

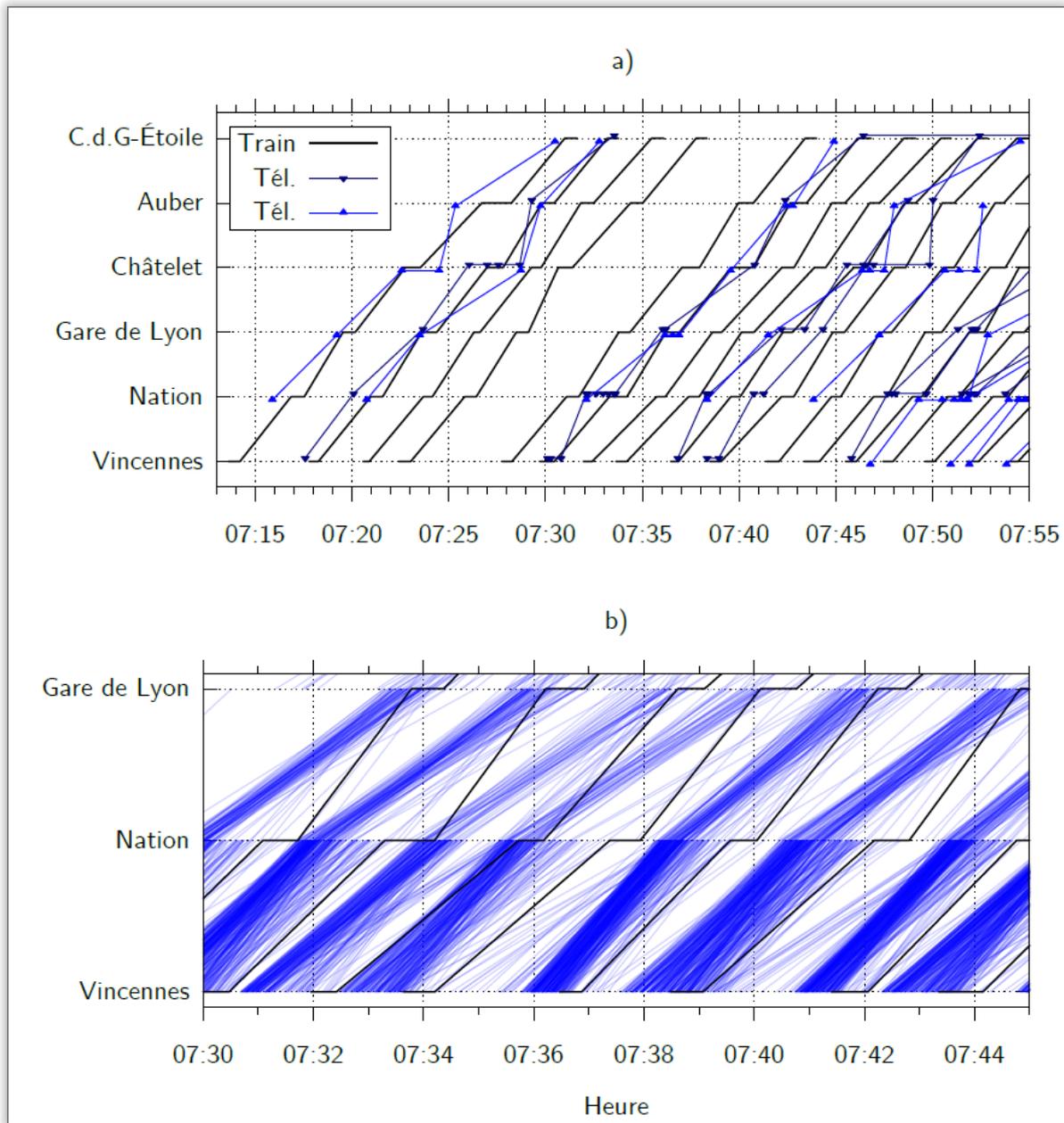


Figure 44 : trajectoires de taille 6 (a) et trajectoires de taille 3 (b)

Ces dernières figures montrent des groupes de trajectoires confinées dans un espace temporel qui semblent permettre une identification des trains. Il faut cependant prendre garde au recouvrement des « paquets » et au changement de mission. La Figure 45 présente une manière de représenter les trajectoires de taille 2 entre Vincennes et Nation par des points dans un espace à deux dimensions temporelles (moment du LAUN à Nation en ordonnée et

moment de LAUN à Vincennes en abscisse). Dans le cas de LAUN répétés au niveau de l'une ou l'autre des stations, le point représentant la trajectoire d'un mobile est choisi de telle sorte que l'écart temporel soit le minimum. La figure est constituée de carrés de 10 secondes de largeur. Un carré apparaît s'il contient au moins une trajectoire. Un code de couleur est appliqué en fonction du débit de mobiles. Le bleu correspond à moins d'un mobile par seconde, jaune correspond à un nombre de mobile entre 1 et 2 par seconde et le rouge signifie que plus de deux mobiles ont leur trajectoire contenue dans une seconde. Les segments verticaux proviennent des observations à quai. Leurs extrémités sont l'instant où le train part de Vincennes et l'instant où le train quitte Nation.

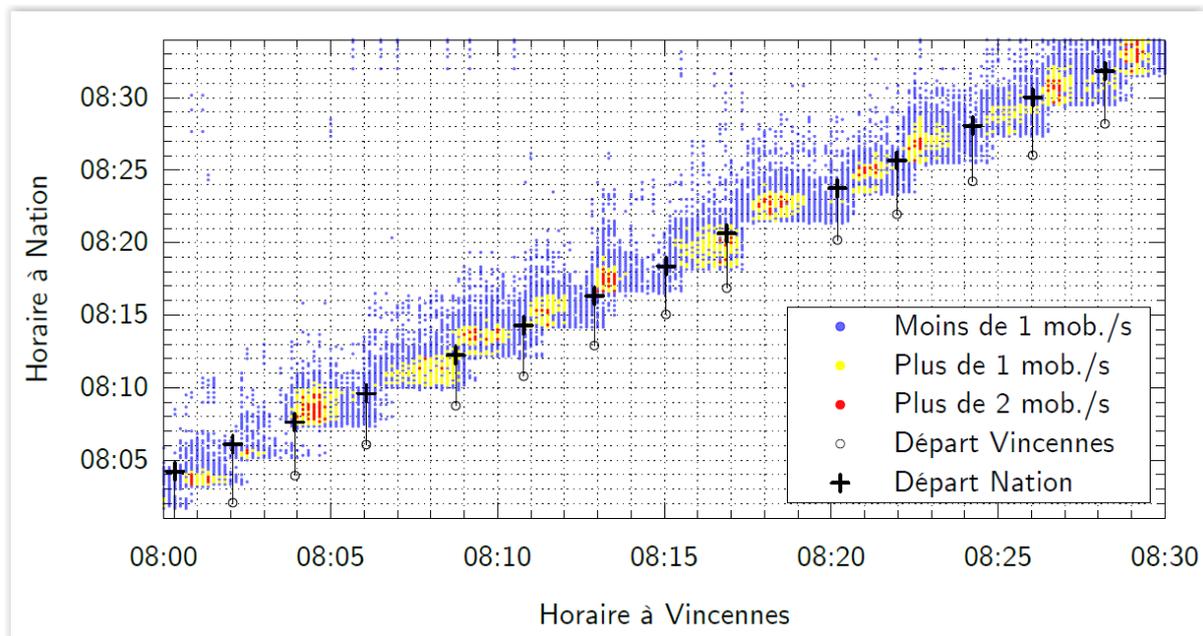


Figure 45: densité d'évènements pour des trajectoires de taille 2 entre Vincennes et Nation entre 8h et 9h le 7 avril 2011

Des techniques standard de segmentation, comme des méthodes de croissance des régions (region growing) ou des algorithmes de détection des contours à gradient (gradient based directed edge detection), permettraient le partitionnement de la carte des densités des mobiles ayant une trajectoire de taille 2. Chaque contour ainsi obtenu pourrait alors être associé à un train. Enfin, à chaque train T est rattaché un nombre de mobiles ayant produit une trajectoire entre deux stations consécutives ($n_{T,s}$). Néanmoins, nous avons vu en 2.3.1.1.2 que le taux de capture dépend fortement de la station considérée ainsi que du moment de la journée. Le paragraphe suivant propose une estimation du nombre de mobiles à bord d'un train T , notée $N_{T,s}$ ainsi qu'un estimateur de la probabilité de capture de trajectoires.

Estimateur de probabilité de capture

Posons tout d’abord les notations utilisées dans ce paragraphe ainsi que les hypothèses que nous ferons. Soit une station s , un point de mesure en s signifie qu’un **LAUN** a été déclenché au niveau d’une des cellules de **E_CRERA**. Soit s^+ et s^- les stations aval et amont de s pour un sens de parcours de la ligne donnée. Notons $P(s)$ la probabilité de déclencher un **LAUN** en s . Nous supposons que $P(s)$, $P(s^+)$ et $P(s^-)$ sont indépendantes. Nous supposons également que ces probabilités dépendent du temps, bien que pour alléger les notations, cette dépendance n’apparaisse pas dans les transcriptions.

Soit les notations suivantes, $n_{T,s}$ est le nombre de trajectoires de taille 2 vers la station s qui ont été associées au train T et $N_{T,s}$ est le nombre de mobiles à bord du train T .

$$\text{Alors : } N_{T,s} = \frac{n_{T,s}}{P(s)*P(s^-)} \quad (17).$$

Pour produire un estimateur de $P(s)$ nous utilisons l’hypothèse d’indépendance qui nous permet d’écrire de formule de Bayes pour les probabilités conditionnelles.

$$P(s) = P(s|s^+ \cap s^-) = \frac{P(s^+ \cap s \cap s^-)}{P(s^+ \cap s^-)} \quad (18)$$

En utilisant le jeu de trajectoires de taille 3 et dont les stations sont consécutives, on compte le nombre de mobiles ayant produit au moins un **LAUN** au niveau des 3 stations ($n_{s^+ \cap s \cap s^-}$) ainsi que le nombre de mobiles ayant produit au moins un **LAUN** au niveau des stations amont et aval ($n_{s^+ \cap s^-}$). Alors, $P^* = \frac{n_{s^+ \cap s \cap s^-}}{n_{s^+ \cap s^-}}$ est un estimateur de $P(s)$, la probabilité pour un mobile de déclencher un **LAUN** au niveau de s . l’intervalle de confiance de cet estimateur est noté $P_{95}(s)$:

$$P_{95}(s) = P^* \pm 1.96 \sqrt{\frac{P^*(1-P^*)}{n_{s^+ \cap s^-}}} \quad (19).$$

Cet estimateur est fortement dépendant de s , du sens de parcours de la ligne et du moment de la journée. La Figure 46 montre les variations de $P_{95}(s)$ par tranche de 30 minutes sur la plage matinale [7h, 12h] pour un sens de parcours Est-Ouest de la ligne (Figure 46.a) et par tranche de 30 minutes pour la plage vespérale [16h, 21h] pour le sens de parcours opposé (Figure 46.b). Nous avons considéré les stations Nation, Gare de Lyon, Châtelet et Auber. On observe un comportement différent entre le matin dans le sens est-ouest et le soir dans le sens ouest-est. En effet, le matin, l’intervalle de confiance de $P(s)$ reste autour de 1 pour la station Nation. Ceci est cohérent avec le fonctionnement théorique car Nation est séparée de sa station amont par une portion aérienne. Pour Auber et Gare de Lyon, $P_{95}(s)$ est plus élevé pendant l’heure de pointe du matin avec un plat remarquable pour Auber autour de 0,6. L’intervalle de confiance de l’estimateur a le comportement inverse pour Châtelet avec une amplitude des variations entre 0,4 et 0,6. Ainsi, des éventuels ralentissements du temps de parcours pendant l’heure de pointe impacte plus les stations Auber et Gare de Lyon que Châtelet. Le soir, pour un sens de parcours inverse, $P_{95}(s)$ reste également quasi-constant pour Nation, mais sa valeur a chuté à 0,4. En effet, dans ce cas, la station amont est connectée par une section souterraine. Cette constance s’applique également à Auber avec une valeur

autour de 0,2 plus faible que le matin. $P_{95}(s)$ chute en heure de pointe pour Châtelet et Gare de Lyon.

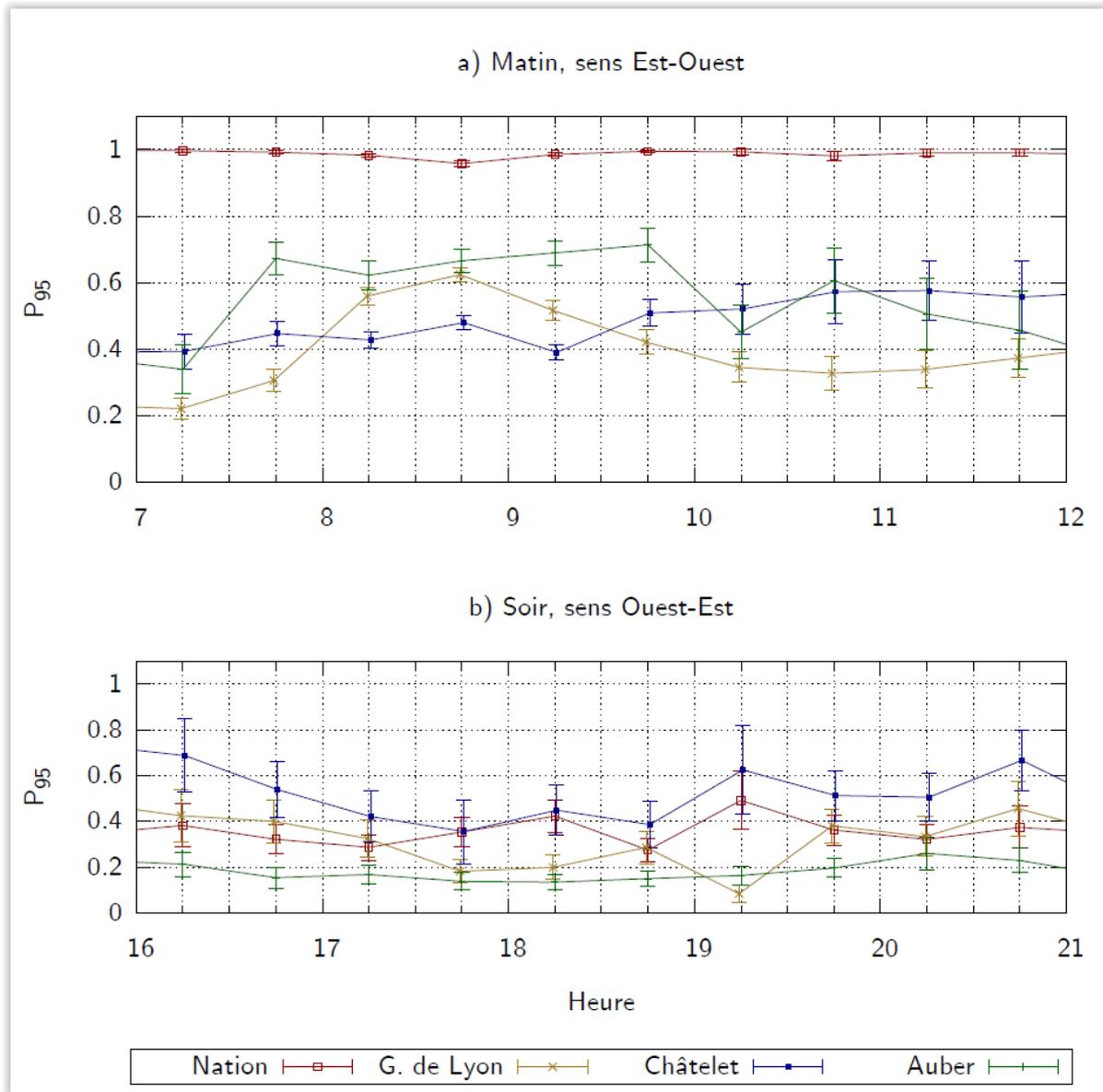


Figure 46 : variation de $P_{95}(s)$ pour différentes stations le matin pour un parcours de la ligne d'est en ouest (a) et le soir pour un parcours d'ouest en est

1.3.2.4 Conclusion

Pour conclure, le système de capture n'est pas parfait. Cependant, cette analyse préliminaire sur le jeu de données du 7 avril montre que ce système est perfectible. En effet, il est possible de corriger la dépendance de capture à la station, au sens de parcours de la ligne et au moment de la journée via un estimateur de $P(s)$ en considérant des trajectoires de mobiles.

