

Titre: Étude de l'évasion tarifaire dans un réseau de transport collectif à l'aide de données de carte à puce et de comptage à bord
Title: l'aide de données de carte à puce et de comptage à bord

Auteur: Hugo Pourmonet
Author:

Date: 2015

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Pourmonet, H. (2015). Étude de l'évasion tarifaire dans un réseau de transport collectif à l'aide de données de carte à puce et de comptage à bord [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/1695/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/1695/>
PolyPublie URL:

Directeurs de recherche: Samuel Bassetto, & Martin Trépanier
Advisors:

Programme: Génie industriel
Program:

UNIVERSITÉ DE MONTRÉAL

ÉTUDE DE L'ÉVASION TARIFAIRES DANS UN RÉSEAU DE TRANSPORT COLLECTIF À
L'AIDE DE DONNÉES DE CARTE À PUCE ET DE COMPTAGE À BORD

HUGO POURMONET

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INDUSTRIEL)

AVRIL 2015

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉTUDE DE L'ÉVASION TARIFAIRES DANS UN RÉSEAU DE TRANSPORT COLLECTIF À
L'AIDE DE DONNÉES DE CARTE À PUCE ET DE COMPTAGE À BORD

présenté par : POURMONET Hugo

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. AGARD Bruno, Doctorat, président

M. BASSETTO Samuel, Doctorat, membre et directeur de recherche

M. TRÉPANIER Martin, Ph. D., membre et codirecteur de recherche

M. GRONDINES Jocelyn, M. Sc. A., membre

REMERCIEMENTS

Je tiens à remercier chaleureusement mon directeur et mon codirecteur de recherche, Samuel Bassetto et Martin Trépanier, pour m'avoir permis de réaliser ce projet de maîtrise.

Je voudrais leur adresser toute ma gratitude pour leur soutien, leurs conseils avisés ainsi que leur disponibilité. De plus, grâce à eux, j'ai pu acquérir de nouvelles connaissances dans le domaine des transports, du traitement de données, mais plus globalement dans le monde de la recherche. Ce fut une expérience très enrichissante.

Je tiens aussi à remercier aussi la Société des Transports de Montréal et plus particulièrement Jean-François Cantin et Jocelyn Grondines, non seulement pour m'avoir donné accès à leurs données, mais aussi pour leur réactivité et leur disponibilité.

Les remerciements vont aussi aux organismes qui m'ont soutenu financièrement pour ce projet de maîtrise, à savoir le CRSNG (Conseil de recherches en sciences naturelles et en génie du Canada) et l'entreprise Thalès.

Enfin, je réserve une pensée toute particulière à ma famille et mes amis qui m'ont encouragé et soutenu durant cette année et demi à l'école Polytechnique qui restera pour moi une expérience exceptionnelle.

RÉSUMÉ

Les systèmes de transports en commun des grandes villes mondiales ont besoin de s'améliorer en permanence. En effet, qu'il soit question de satisfaction des utilisateurs, d'amélioration de la fluidité des réseaux routiers, ou encore de la réduction des gaz à effet de serre, les enjeux du développement de ces réseaux sont immenses. Dans cette optique, les agences de transport n'ont alors pas d'autres choix que de suivre l'évolution technologique.

On appelle systèmes de transports intelligents (STI), l'ensemble des technologies de communication et d'information que l'on retrouve dans les transports. Depuis quelques années, ces technologies connaissent un véritable essor et ont pour objectif de rendre plus efficaces et plus fiables les infrastructures et systèmes existants. La plupart des outils STI sont basés sur la collecte, le traitement et la diffusion de l'information. Une large quantité de données peut alors être exploitée afin de fournir des informations sur la planification, la gestion du trafic, la sûreté ou encore l'évasion tarifaire.

Les systèmes de carte à puce sont un bon exemple de technologie visant à améliorer les réseaux de transport en commun. En effet, de nombreuses études ont déjà prouvé l'étendue des possibilités que peuvent offrir les cartes à puces. Cependant, il existe d'autres outils STI aussi efficaces. C'est le cas des systèmes de comptage à bord qui peuvent permettre aux agences de transport d'avoir une meilleure connaissance de la fréquentation en vue d'une meilleure gestion de leurs réseaux. C'est dans ce sens que nous développons dans ce mémoire un état de l'art sur ces deux STI afin d'y voir les réels intérêts que les agences peuvent y trouver. La revue de littérature s'oriente ensuite vers l'objectif de cette recherche en faisant un bilan des connaissances sur le phénomène d'évasion tarifaire dans les transports en commun.

Ce mémoire s'inscrit dans un projet de recherche impliquant la compagnie Thalès, grand fournisseur de systèmes de perception par cartes à puces, et la Société de transport de Montréal, le tout cofinancé par le Conseil national de recherche en sciences naturelles et en génie du Canada (CRSNG). Il est basé sur l'étude de données provenant du système de perception par carte à puce et du système automatisé de comptage à bord fournies par la Société des Transports de Montréal (STM). Celle-ci équipe ses autobus de ces technologies depuis 2008 pour la carte à puce et 2013 pour les comptages à bord. L'objectif principal de ce mémoire est **d'exploiter ces données afin d'analyser le phénomène l'évasion tarifaire**. Plusieurs pistes exploratoires ont

alors été étudiées afin de mieux appréhender ce phénomène relativement complexe, qui peut causer des pertes de plusieurs millions de dollars par an pour les agences de transport. Autrement dit, nous nous sommes intéressés aux différents paramètres exploitables grâce aux données qui nous permettraient de mettre en lumière l'évasion tarifaire. Il s'avère que l'on observe des variations du ratio du nombre de validations sur le nombre de montants en fonction des heures de la journée. On ne peut pas affirmer hors de tout doute qu'un ratio plus faible est causé par l'évasion tarifaire, mais cela reste notre hypothèse d'étude. En prenant cela en compte, nous trouvons que les heures creuses semblent alors être plus sensibles à ce phénomène. De plus, les études par arrêts et par ligne nous laissent croire que ce phénomène a réellement une composante récurrente. Globalement, nous trouvons un ratio moyen d'environ 0.97 entre le nombre de validations et le nombre de montants sur l'ensemble du réseau de bus. Cependant, nous avons relevé des écarts entre ce ratio moyen et celui de certaines lignes ou arrêts. En effet, des ratios plus faibles de 4 points de pourcentage par rapport à la moyenne sont mis en évidence à plusieurs reprises sur une même ligne ou un même arrêt. Le deuxième objectif de ce mémoire consiste à étendre l'étude de ce phénomène en proposant un outil dynamique visant à améliorer les contrôles sur le réseau. Cet outil s'appuie à la fois sur le traitement des données, mais aussi sur les observations du personnel de terrain (guichetiers, conducteurs de bus) afin de devenir un véritable outil d'aide à la décision pour le personnel d'inspection.

Les perspectives du projet sont ensuite exposées et résident dans la continuité et l'élargissement de l'outil dynamique à d'autres problématiques comme la maintenance ou la sécurité. Des méthodes mathématiques telles que le Bayésien pourraient aussi être implantées dans l'outil pour lui donner un aspect prévisionnel.

Le transfert de ces résultats de recherche permettra éventuellement à la STM de mettre en place une solution permettant de mieux appréhender ce phénomène. Quant à Thalès, cette recherche pourra leur fournir une première marche vers la mise en place d'une solution clef en main pour la réduction de l'évasion tarifaire. Cette solution pourrait être pour eux un atout non négligeable concernant la vente de leurs technologies aux agences de transports.

ABSTRACT

Public transport systems of major world cities have an ongoing need to improve. Indeed, whether it is a question of user satisfaction, improve the fluidity of road networks, or the reduction of greenhouse gases, the issues of development of these networks are immense. With this in mind, transportation agencies have no other choice but to follow technological developments.

Called intelligent transport systems (ITS), all technologies of communication and information that can be found in transportation. In recent years, these technologies are currently booming and are intended to make existing infrastructure and systems more efficient and reliable. Most tools are based on the collection, processing and dissemination of information. A large amount of data can be exploited to provide information on planning, traffic management, safety or fare evasion.

Smart card systems are a good example of technology to improve public transport services. Indeed, many studies have already demonstrated the extent of the possibilities offered smartcards. However, there are other tools that have the ability to make more efficient transport networks. This is particularly true for automatic passengers counting systems (APC) that can allow transport agencies to have a better understanding of traffic in order to better manage their networks. This is what we develop in this paper a state of the art on both ITS to see the real interests that agencies may have to use them. The literature review then turned toward the goal of this research by making a state of knowledge on the fare evasion phenomenon in transport.

This thesis is part of a research project involving the company Thales, a leading provider of smart card systems for perception and the Montreal transit corporation (STM), all financed by the national research council, the Natural Sciences and Engineering Canada (NSERC). The study of this thesis is performed based on data from the collection system based on smart card and the automatic passenger counting system provided by the Montréal transit corporation (STM), which equips its buses with these technologies since 2008 for the smart card and 2013 for on-board counting. The main goal of this paper is to use these data to analyze the phenomenon of fare evasion. Several exploratory tracks were studied to better understand this fairly complex phenomenon that can cause loss of several million dollars per year for transport agencies. In other words, we looked at different settings through data that allowed us to highlight the fare evasion. It turns out that there are variations concerning the ratio between number of validations and number of passengers depending to time of day. We can not claim that a lower ratio is caused by

the fare evasion, but this is our hypothesis of study. Taking this into account, we find that the peak hours then seem to be more sensitive to this phenomenon. In addition, studies by stops and lines of buses lead us to believe that this phenomenon actually has a recurring component. Indeed, lower ratios of 4 percentage points from the average are repeatedly highlighted several times on the same line or even stop. The second objective of this paper is to extend the study of this phenomenon by offering a dynamic tool to improve controls over the network. This tool is based on data processing, but also on field staff observations (tellers, bus drivers) to become a true decision support tool for the inspection staff.

The project's prospects are then exposed and reside in the continuation and expansion of the dynamic tool to other issues such as maintenance or safety. Mathematical methods such as bayesian could also be implemented in the tool to give it a forward look.

The transfer of research results to the STM will eventually implement a solution to better understand this phenomenon. As for Thales, this research will provide a first step towards the establishment of a turnkey solution for the reduction of fare evasion. This solution could be for them an important asset for the sale of their technologies to transportation agencies.

TABLE DES MATIÈRES

REMERCIEMENTS.....	iii
RÉSUMÉ.....	iv
ABSTRACT.....	vi
TABLE DES MATIÈRES	viii
LISTE DES TABLEAUX.....	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS.....	xiii
LISTE DES ANNEXES.....	xv
CHAPITRE 1 INTRODUCTION.....	1
1.1 Problématique	1
1.2 Objectifs de la recherche	2
1.3 Structure du mémoire	2
CHAPITRE 2 REVUE DE LITTÉRATURE	4
2.1 Les STI au service des transports en commun	4
2.1.1 L'histoire des STI depuis 1970.....	5
2.1.2 La carte à puce	6
2.1.3 Les systèmes de comptage à bord automatique.....	12
2.2 Le phénomène d'évasion tarifaire dans les transports en commun	16
2.2.1 Le phénomène	16
2.2.2 Les méthodes existantes pour réduire l'évasion tarifaire	19
2.3 Synthèse de la revue de littérature.....	20
CHAPITRE 3 MÉTHODOLOGIE.....	22
3.1 Récupération et présentation des données	22
3.2 Processus de traitement.....	25
3.3 Processus et outils d'analyse.....	28
CHAPITRE 4 ANALYSE DES RÉSULTATS ET INTERPRÉTATIONS	31
4.1 Analyse descriptive en fonction des paramètres disponibles	31

4.1.1	Jour de la semaine pour les dates du 7 au 13 avril 2014	31
4.1.2	Heure de la journée.....	32
4.1.3	Différentes lignes de bus et types de bus	34
4.1.4	Titres de transport	38
4.1.5	Quartiers et RTA.....	41
4.1.6	Autour des métros.....	43
4.1.7	Arrêts de bus.....	44
4.2	Analyses comparées entre les deux semaines	47
4.2.1	Jour de la semaine	47
4.2.2	Heure de la journée.....	48
4.2.3	Lignes de bus	49
4.2.4	Arrêts de bus.....	50
4.3	Synthèse	54
CHAPITRE 5 OUTIL DYNAMIQUE D'AIDE A LA DÉCISION		55
5.1	Présentation de l'outil	55
5.1.1	Objectifs et fonctionnement de l'outil.....	55
5.1.2	Affichage des données	56
5.2	Extension pour l'évasion tarifaire observée.....	59
5.3	Utilisation de l'outil au quotidien.....	60
CHAPITRE 6 CONCLUSION		64
BIBLIOGRAPHIE		68
ANNEXES		75

LISTE DES TABLEAUX

Tableau 2-1 : Les cartes à puce dans les transports publics mondiaux.....	9
Tableau 2-2 : Avantages et inconvénients des cartes à puces dans les transports en commun	11
Tableau 2-3 : Types de fraudes.....	17
Tableau 3-1 : Présentation des données de comptage à bord.....	22
Tableau 3-2 : Présentation des données de carte à puce OPUS	24
Tableau 3-3 : Exemple des données jointes au milieu du traitement.....	27
Tableau 3-4 : Fichier de résultat pour analyse	28
Tableau 4-1 : Ratios les plus faibles pour la semaine d'avril.....	35
Tableau 4-2 : Types de lignes de bus sur le réseau de la STM.....	36
Tableau 4-3 : Analyse du ratio validations/montants par arrêt de bus à proximité des métros	44
Tableau 4-4 : Analyse du ratio Validations/Montants par arrêts de bus.....	45
Tableau 4-5 : Ratio Validations/Montants par ligne de bus	49
Tableau 4-6 : Ratio Validations/Montants sur les jours de la semaine d'avril pour 5 arrêts	52
Tableau 4-7 : 15 arrêts ayant les plus faibles ratios pour les 2 semaines.....	53
Tableau 5-1 : Forme des données extraites pour une visualisation sur carte.....	57
Tableau 5-2 : Forme des données récupérées par l'application.....	59

LISTE DES FIGURES

Figure 2-1 : Capacité, fonctionnalité et prix des différentes cartes à puce (tiré de http://www.smartcardbasics.com/smart-card-types.html)	7
Figure 2-2 : Principaux composants d'un système de comptage à bord (tiré de http://www.initag.de/share/news/Themen/APC.pdf).....	13
Figure 2-3 : Les capteurs actifs et passifs (Source STM)	14
Figure 2-4 : Courbe de Gauss du comportement naturel (tiré de http://onlinepubs.trb.org/onlinepubs/circulars/ec112.pdf).....	18
Figure 3-1 : Processus de traitement des données	26
Figure 3-2 : Tableau croisé dynamique	29
Figure 3-3 : Superposition de la couche quartier et des arrêts de bus sur Qgis	30
Figure 4-1 : Ratio Nombre_Validations/Somme_Montants par jour de la semaine (du 7 au 13 avril 2014).....	31
Figure 4-2 : Ratio Nombre_Validations/Somme_Montants en fonction des heures de la journée (lundi 7 au vendredi 11 avril 2014).....	33
Figure 4-3 : Ratio Nombre_Validations/Somme_Montants en fonction des heures de la journée (mardi 8 avril 2014)	34
Figure 4-4 : Résultats par types de lignes de bus pour la première semaine (7 au 13 avril 2014) 37	
Figure 4-5 : Pourcentage validations CPCT, CPO, et CMJ pour les 5 jours de la semaine d'avril (7 au 11 avril 2014)	38
Figure 4-6 : Pourcentage de validations par CPCT, CPO et CMJ pour la fin de semaine d'avril 2014 (12 et 13 avril 2014).....	39
Figure 4-7 : Ratio Validations/Montants en fonction du ratio Validations_CPCT/Validations_Totales pour toutes les lignes de bus de la semaine du 7 au 13 avril 2014	40
Figure 4-8 : Ratio Validations_Montants en fonction du ratio Validations_CPO/Validations_Totales (semaine du 7 au 13 avril 2014)	40

Figure 4-9 : Analyse par RTA des ratios sur l'île de Montréal	42
Figure 4-10 : Analyse par quartier des ratios sur l'île de Montréal.....	43
Figure 4-11 : Arrêts ayant plus de 100 montants gradués par ratio	46
Figure 4-12 : Ratio comparatif Validations/Montants pour les deux semaines.....	47
Figure 4-13 : Ratio Validations/Montants en fonction des heures de la journée pour les deux semaines.....	48
Figure 4-14 : Arrêts de la ligne 11 pour la semaine d'avril 2014.....	50
Figure 5-1 : Explication des étapes du fonctionnement de l'outil dynamique	56
Figure 5-2 : 15 arrêts ayant un faible ratio représentés sur une carte Google Maps.....	58
Figure 5-3 : Carte avec les trois types de données.....	60
Figure 5-4 : Schématisation de l'utilisation de l'outil pour les métros	61
Figure 5-5 : Schématisation de l'utilisation de l'outil pour les bus	62

LISTE DES SIGLES ET ABRÉVIATIONS

AFC Automatic fare collection (perception automatique des titres de transport, des droits de passage)

AMT Agence métropolitaine de transport (Montréal et agglomération)

AOT Autorité organisatrice de transport (en commun)

APC Automatic passenger counting (comptage automatique de passagers)

ATM Automated teller machine (Guichet automatique bancaire)

AVL Automatic vehicle location (localisation automatique du véhicule)

CAP Carte à puce

CSS Cascading Style Sheets (feuille de style en cascade)

CSV Comma separated values

GIS Geographic information system (système d'information géographique)

GPS Global positioning system (système mondial de localisation)

GTFS General transit feed specification (spécification de flux relatifs aux transports en commun)

HTML HyperText Markup Language (langage de balisage hypertexte)

INIT Init Innovation in traffic systems AG, ou ses filiales

IVHS Intelligent vehicle-highway systems

ITS Intelligent Transportation Systems

JS JavaScript (langage programmation de scripts)

MP3 MPEG-1/2 Audio Layer 3 (Fichier de son)

PHP Hypertext Preprocessor (langage de programmation libre)

RTA Région de tri d'acheminement (découpage postal)

RFID Radio-frequency identification

SCAD Système de comptage automatisé et d'acquisition des données

STI Systèmes de transports intelligents

STM Société de transport de Montréal

TCD Tableau croisé dynamique

TCRP Transit Cooperative Research Program

LISTE DES ANNEXES

ANNEXE A - TITRES DE TRANSPORTS	75
ANNEXE B – TABLEAU RÉCAPITULATIF DU RATIO VALIDATIONS/MONTANTS MOYEN, MAX, MIN, MEDIAN SUR LA SEMAINE D'AVRIL 2014 (7 AU 13 AVRIL)	79
ANNEXE C – RATIOS VALIDATIONS/MONTANTS PAR LIGNE SUR LA SEMAINE D'AVRIL 2014 (7 AU 13 AVRIL)	84
ANNEXE D – EXEMPLE FICHIER RÉSULTAT ANALYSE EXCEL	85
ANNEXE E – FONCTIONNEMENT DE L'OUTIL.....	86
ANNEXE F – IMPRESSIONS ECRAN OUTIL	87
ANNEXE G – RÉSUMÉ DES LANGAGES ET Outils UTILISÉS POUR LA RÉALISATION DE L'OUTIL DYNAMIQUE.....	88

CHAPITRE 1 INTRODUCTION

Compte tenu du contexte économique et social que traverse le monde, la tendance est à la valorisation des transports en commun. En effet, nous traversons une période d'urbanisation où les villes s'agrandissent au profit de la désertification des campagnes. Les compagnies de transport se doivent alors de réagir afin de s'adapter à cette demande grandissante. C'est en s'appuyant et en investissant dans les nouvelles technologies que ces compagnies vont pouvoir améliorer leurs services et satisfaire leurs usagers. Les systèmes de cartes à puce et de comptage à bord sont des sources de données potentiellement exploitables pour répondre à des problèmes de planification, gestion de l'achalandage ou encore d'évasion tarifaire. En contrepartie, les agences de transports ont des comptes à rendre aux collectivités locales, aucune perte tarifaire ne doit donc être négligée. En transport collectif, l'évasion tarifaire (communément appelée "fraude") doit être réduite au maximum. Ce phénomène représentait pas moins de 20M\$ de pertes en 2008 pour la Société de transport de Montréal (STM). Depuis 2008, les nouvelles technologies, telles que l'implantation de systèmes de perception par carte à puce OPUS (à Montréal) ont permis de réduire cette perte à 8M\$ en 2011. Mais, même si l'anéantissement complet de ce phénomène paraît impossible, les opportunités pour réduire ces pertes sont potentiellement encore grandes.

1.1 Problématique

À l'aide des données de carte à puce et de comptage à bord, nous nous sommes demandé s'il était possible de mieux caractériser et contrôler l'évasion tarifaire. Comment peut-on utiliser ces nouvelles technologies pour identifier les pertes dues à ce phénomène ?

Théoriquement, la différence entre le nombre de personnes qui entrent dans un bus et le nombre de personnes qui valident leur paiement dans un système de perception nous renseigne sur le taux d'évasion tarifaire (sous réserve que le système enregistre tous les types de paiement, ce qui est le cas à la STM). Toutefois, nous rappelons que nous ne pouvons pas affirmer hors de tout doute que cette différence soit seulement due à l'évasion tarifaire, mais qu'elle peut en être un très bon indicateur. Il s'avère que les données résident dans deux technologies bien différentes, et donc que le lien entre ces sources n'est pas si évident dans la réalité. Les agences de transport caractérisent ce phénomène seulement par des enquêtes de terrain, ou des sondages auprès des utilisateurs. De plus, de nos jours, aucun outil n'est réellement mis en place pour aider l'allocation

des inspecteurs sur le terrain. À Montréal, le personnel d'inspection effectue les contrôles en se basant sur leur expérience du terrain. L'analyse combinée des données de carte à puce et de comptage à bord pourrait alors permettre de mettre à disposition un véritable outil d'aide à la décision pour les inspecteurs visant à réduire ce phénomène d'évasion tarifaire.

1.2 Objectifs de la recherche

L'objectif principal du projet vise à proposer une solution permettant de caractériser ce phénomène d'évasion tarifaire en se basant sur l'exploitation des données fournies par la STM. Il est question ici de tenter de faire apparaître les paramètres principaux influant sur ce phénomène.

Plus précisément les objectifs spécifiques du projet consistent à :

- 1- Développer un algorithme permettant de relier les données de carte à puce et de comptage à bord. Ces données peuvent être liées grâce à plusieurs variables, dont le jour, l'heure, et le numéro du véhicule. La création d'un algorithme pour le traitement de ces données est nécessaire, car les données brutes représentent plusieurs millions d'entrées par semaine.
- 2- Détecter et caractériser l'évasion tarifaire à travers l'analyse des données en calculant un ratio entre le nombre de personnes montant dans le bus et le nombre de personnes qui ont validé. Ce ratio pourra être calculé en fonction des différentes variables disponibles dans les données (jour, heure, numéro du véhicule, ligne, etc ...).
- 3- Développer un outil dynamique permettant d'utiliser les résultats du traitement des données pour faciliter la planification des inspecteurs et avoir une meilleure vision du phénomène sur l'ensemble du réseau au quotidien.
- 4- Étendre l'outil dynamique pour une application dans les bus et les métros en mettant à disposition des guichetiers et conducteurs de bus un outil permettant de remonter une évasion tarifaire observée. Il pourra alors servir de véritable outil d'aide à la décision pour les inspecteurs.

1.3 Structure du mémoire

Pour remplir les objectifs spécifiques visés, une méthodologie bien définie est suivie. Le mémoire vise à rendre compte de cette structure de travail, il se compose de 4 grandes parties.

Le mémoire débute par une revue de littérature permettant de faire un état des connaissances et des recherches effectuées sur les transports en commun, les systèmes intelligents de transport et leurs enjeux, mais aussi sur le phénomène d'évasion tarifaire. Cette partie est le véritable support de ce mémoire. C'est aussi à partir de ces textes que l'on pourra prouver l'originalité de notre recherche et fixer les hypothèses de celle-ci. La partie suivante traite de la méthodologie de recherche mise en place pour répondre aux objectifs visés. Elle détaille le cheminement entre la récupération des données brutes de l'agence de transport, la préparation et la compréhension des données, les premières analyses sur des échantillons de données et le traitement. Cette étape est très importante et c'est à partir d'elle que des résultats peuvent être exploités. La quatrième partie présente les résultats de l'étude, c'est-à-dire la caractérisation indirecte du phénomène d'évasion tarifaire à travers les données de carte à puce et de comptage à bord. Elle permet de discerner et de discuter des paramètres pouvant influencer ce phénomène. Nous verrons que certains paramètres comme l'heure, la ligne de bus, le type de bus, ou encore la proximité avec les métros font partie des paramètres qui permettent de mettre en lumière ce phénomène. C'est clairement l'étape clef du projet, et c'est grâce à celle-ci que nous avons pu étendre la recherche en proposant un outil dynamique visant à mieux contrôler le phénomène. Enfin, la cinquième et dernière partie du projet met en avant l'originalité du projet en proposant un outil dynamique s'appuyant à la fois sur les résultats issus du traitement des données ainsi que sur les observations du personnel de terrain. Cet outil, qui peut être vu comme un véritable aide à la décision, pourrait permettre une meilleure allocation des inspecteurs sur un réseau de transport. De plus, la flexibilité d'un tel outil pourrait être adaptable à toutes sortes d'autres enjeux comme la sécurité ou la maintenance et pourra s'allier à un outil de résolution de problèmes au quotidien.

CHAPITRE 2 REVUE DE LITTÉRATURE

Cet état de l'art permet de mieux situer le projet dans son contexte et de justifier les motivations du projet de recherche. Nous débutons cette revue par les publications et ouvrages pertinents au sujet des transports en commun et plus particulièrement des systèmes de transport intelligents. Nous mettons l'accent sur les recherches associées aux cartes à puce ainsi qu'aux systèmes de comptage à bord. Enfin, nous nous dirigeons vers les recherches effectuées sur l'évasion tarifaire et les méthodes existantes pour contrôler ce phénomène.

2.1 Les STI au service des transports en commun

À la fin des années 1970, alors que l'urbanisation des villes et l'encombrement des routes commencent à poser des problèmes, les deux Français S.Nora et A.Minc introduisent le terme télé-matique, qui est une contraction des mots télécommunication et informatique. Ce terme définit le lien entre la transmission d'information par le biais d'un moyen de communication et le traitement informatique de ces informations (Goel, 2007). C'est durant cette période que B. Marti mettra au point le premier Minitel. Très vite ce terme de télématique va se retrouver dans de nombreux domaines, de l'économie, à la santé en passant par les services postaux.

C'est en 1990 que l'on va commencer à parler plus spécifiquement de télématique dans le domaine des transports (Nowacki, 2012), dont les applications dérouleront vers les Systèmes de Transport Intelligents (STI). Selon la directive 2010/40/EU (7 juillet 2010), les STI sont définis comme un ensemble de systèmes dans lesquels les technologies de communication et d'information sont appliquées au domaine des transports routiers, incluant les infrastructures, les véhicules et les usagers, la gestion du trafic et de la mobilité, ainsi que les interfaces avec d'autres modes de transport. Les STI vont alors suivre les technologies de télécommunication, d'électronique et de micro-électronique, ainsi que d'informatique et bien d'autres domaines pour se mettre au service des transports. Elles finiront par prendre une part non négligeable concernant l'amélioration des transports en termes d'efficacité, de performance environnementale, ou de sécurité.

2.1.1 L'histoire des STI depuis 1970

G. Nowacki (2012) a discerné trois phases dans l'histoire des STI. Une première phase qui s'étale entre 1970 et le début des années 1980. Celle-ci se base sur des systèmes complexes de guidage routier mis au point autour des technologies de radiocommunications. L'Allemagne, les États-Unis et la France ont été les premiers pays à s'intéresser à de telles solutions durant cette décennie, mais il faut noter qu'aucune solution importante ne sera mise en pratique, elles resteront toutes au niveau de la recherche (Nowacki, 2012).

La deuxième phase s'étalant jusqu'en 1994 correspond à l'étape la plus significative dans l'histoire des STI. Les avancées technologiques en matière de stockage d'informations vont permettre de lancer de grands programmes d'innovation en matière de STI. Le programme "pour un système routier européen ayant une meilleure efficacité et une sécurité sans précédent", dont l'abréviation anglophone est PROMETHEUS, a été lancé en 1986. Il concerna 18 compagnies automobiles et 40 centres de recherches dans le but de stimuler les recherches en terme de télécommunication, de transfert d'information et de technologies de transport (Nowacki, 2012). Ce projet européen permettra notamment d'améliorer l'information disponible pour les conducteurs, de développer les premiers systèmes de coopération entre les conducteurs afin de partager les informations importantes ou encore de mettre au point des systèmes améliorant la gestion du trafic sur les autoroutes (Woodrow Barfield & Thomas A. Dingus, 1997). D'autres projets de grande envergure ont aussi vu le jour aux États-Unis avec IVHS (Intelligent Vehicle Highway Systems) ou RACS (Road/Automobile Communication System) au Japon. L'issu des études de RACS sera d'ailleurs la base de nos actuels systèmes de navigations routières (Nowacki, 2012).

En 1994, le début de la troisième phase correspond au moment où l'on commence à entendre parler de STI à travers les termes d'intermodalité dans les transports. Les STI ne se focalisent plus que sur les automobiles, mais à l'ensemble des moyens de transport. Les projets vont alors prendre des dimensions encore plus grandes et des collaborations vont naître entre les pays de manière à accélérer les phases de recherche et développement.

De nombreuses solutions vont voir le jour de manière à améliorer l'efficacité du système de transport intermodal mondial et nous allons nous attarder sur les cartes à puce et les systèmes de comptage à bord.

2.1.2 La carte à puce

La carte à puce est de nos jours un des moyens de perception le plus répandu dans les transports collectifs (Pelletier et al, 2011) et nous allons faire un retour sur cette technologie.

Le japonais Arimura et le français R. Moreno seront les premiers à déposer des brevets sur cette technologie respectivement en 1970 et 1974 (Shelfer & Procaccino, 2002). Il faudra attendre 3 ans pour qu'en 1977, Motorola et Bull produisent les premières cartes à puce qui seront utilisées dans les systèmes bancaires français pour accentuer la sécurité des transactions (Pelletier et al, 2011). Les applications vont alors se multiplier durant les années 1980 en France dans les services téléphoniques et postaux, et pour les cartes d'assurance maladie en Allemagne (Attoh-Okine et Shen, 1995). À la fin des années 1980, l'utilisation de cette carte s'est réellement banalisée dans de nombreux secteurs comme la santé, les banques, les télécommunications et les transports¹.

De la taille d'une carte de crédit, elles sont facilement utilisables et durables (Lu, 2007). Elles sont toujours constituées d'une mémoire pour stocker les informations et/ou d'un processeur pour les traiter. La carte à puce peut être lue avec, sans contact ou les deux (Blythe, 2004).

Pour la lecture avec contact, la carte va s'insérer dans un appareil afin que les contacts électriques de la puce se connectent avec le lecteur. C'est le type de carte le plus connu, et le plus répandu notamment grâce à son application pour les services de paiement avec une carte de crédit. Les cartes sans contact, de leur côté, utilisent la technologie *Radio-Frequency Identification* (RFID) pour communiquer avec le lecteur à une fréquence comprise entre 13,5 MHz et 960Mhz suivant les modèles et les applications². Dans les transports en commun, du fait de la rapidité nécessaire pour la validation, elles communiquent à 13,56 MHz en conformité avec la norme ISO 14443 standard (Mac Donald, 2000). Elles permettent une validation à environ 10 cm ce qui facilite la

¹Smart Card Alliance. « Smart Cards Applications ». <http://www.smartcardalliance.org/pages/smart-cards-applications>, page consulté le 15 janvier 2015.)

² Smart Card Basics. «Smart Cards Types » <http://www.smartcardbasics.com/smart-card-types.html>, page consultée le 15 janvier 2015

communication avec le lecteur.³ Certaines cartes possèdent seulement une carte mémoire qui leur permet juste de stocker des informations (Shelfer & Procaccino, 2002). Ces cartes ne peuvent pas traiter l'information à la différence des cartes possédant un microcontrôleur ou un microprocesseur. Ces dernières peuvent alors gérer et traiter les informations stockées dans la partie mémoire. On les nomme alors des cartes multi-fonctions⁴. Elles accordent aussi une sécurité plus importante que les simples cartes mémoires de base.

Tout bien considéré, le choix du type de cartes dépend vraiment de son application et de plusieurs paramètres comme la dimension dynamique (vitesse de communication par exemple), les infrastructures du marché ou encore le prix⁵. Les cartes multi-fonctions sont par exemple plus coûteuses, mais plus performantes. Voici ci-dessous un schéma confrontant le prix des différents types de cartes et leurs capacités.

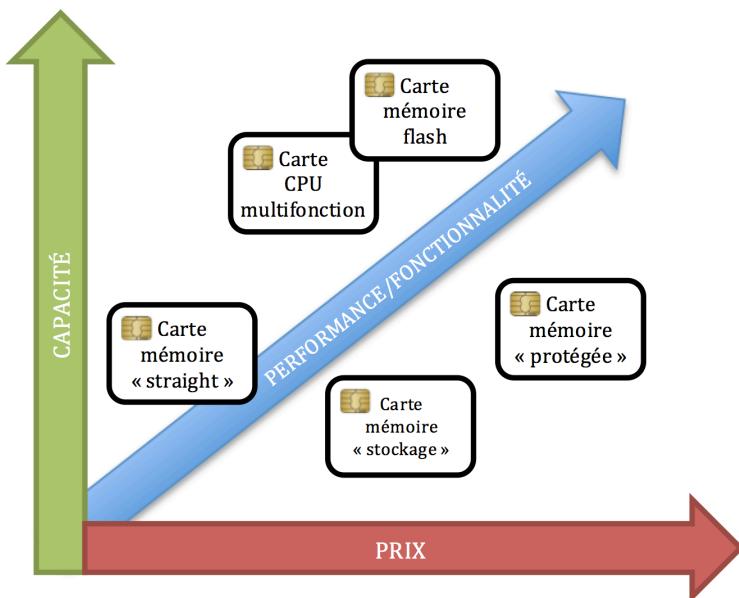


Figure 2-1 : Capacité, fonctionnalité et prix des différentes cartes à puce (tiré de <http://www.smartcardbasics.com/smart-card-types.html>)

³ Smart Card Basics. «Smart Cards Standards » <http://www.smartcardbasics.com/smart-card-standards.html>, page consultée le 15 janvier 2015

⁴ Smart Card Basics. «Smart Cards Types » <http://www.smartcardbasics.com/smart-card-types.html>, page consultée le 15 janvier 2015

⁵ Smart Card Alliance. « Smart Cards Into Primer ». <http://www.smartcardalliance.org/smart-cards-into-primer>, page consulté le 15 janvier 2015.)

Les cartes mémoires « Straight » sont les cartes à puce de base servant à stocker une toute petite quantité de données, elles concernent souvent la catégorie des cartes jetables. Les cartes mémoires dites de stockage sont munies d'une mémoire plus importante et sont souvent rechargeables tout comme les cartes protégées qui en plus sont cryptées. Les cartes multifonctions sont munies d'un microcontrôleur et sont la plupart du temps sans contact, tout comme les cartes flash que l'on retrouve de nos jours sur les cartes de débit pour payer à distance.

Shelfer et Procaccino (2002) ont identifié trois catégories d'applications distinctes pour lesquelles la carte à puce se révèle utile :

- L'authentification : Les cartes à puce permettent de stocker des données pouvant servir à identifier une personne. On peut alors prendre l'exemple d'un passeport, ou d'un permis de conduire qui pourrait avoir la forme d'une simple carte à puce.
- L'autorisation : Les cartes à puce facilitent aussi le transfert d'information (images et textes) entre deux personnes ou systèmes. Dans le domaine de la santé, cette application des cartes à puce prend une place de plus en plus importante. Dans les systèmes de santé français et allemand, une standardisation des informations sur les patients a pu être possible grâce à ce dispositif. Il facilite par exemple le lien entre les médecins et les pharmacies concernant les médicaments.
- La transaction : La carte à puce a su trouver sa place dans les transactions de biens et services en pouvant prendre le rôle de ce qu'on appelle les cartes de crédit. Elle permet d'enregistrer et d'effectuer un paiement, ou de retirer de l'argent comptant à des guichets automatiques bancaires (GAB, ou *ATM* en anglais).

Concernant les transports en commun, la carte à puce trouvera son utilité dans l'enregistrement d'une transaction. Il faudra cependant attendre quelques années avant que ces cartes se généralisent dans les réseaux de transport et deviennent une solution viable de paiement (Blythe 2004).

Le tableau ci-après présente un petit tour d'horizon des grandes villes mondiales utilisant un système de carte à puce dans leurs transports publics :

Tableau 2-1 : Les cartes à puce dans les transports publics mondiaux

Villes et Systèmes	Année	Caractéristiques
Hong-Kong Octopus	Sept. 1997	<ul style="list-style-type: none"> - Carte sans contact, rechargeable. - Environ 4,8 millions d'utilisateurs quotidiens (dec 2014)⁶ - Carte utilisable dans les parkings, épicerie, restaurants rapides, cafés - Système double-tap (validation à l'entrée et à la sortie)
Gatineau (Canada, QC) Multi (depuis 2013) et Passe-Partout Plus avant.	1998	<ul style="list-style-type: none"> - Carte sans contact, rechargeable - Première ville d'Amérique du Nord à utiliser les cartes à puces
Washington (US) SmarTrip	1999	<ul style="list-style-type: none"> - Carte sans contact RFID, rechargeable, prépayée - Implanté en 2002 dans les métros - 1,8 million de cartes en circulation (2010)⁷
Shanghai (Chine) SPTC	1999	<ul style="list-style-type: none"> - Carte sans contact, rechargeable, et peut servir de carte de paiement. - Carte utilisable dans de nombreux endroits dans Shanghai (métros, bus, parking, station-service, centres touristiques, etc.) - Solde de la carte consultable en ligne⁸

⁶ <http://www.mtr.com.hk/en/corporate/investor/patronage.php>

⁷ Washington Metropolitain « History », http://www.wmata.com/about_metro/docs/history.pdf page consultée le 16 janvier 2015

⁸ Shanghai Public Transport Card « Key Figures» <http://www.sptcc.com> page consultée le 16 janvier 2015

Tableau 2-1: Les cartes à puce dans les transports publics mondiaux (suite et fin)

Paris (France) Navigo	2001	- Carte sans contact, rechargeable - Carte utilisable pour métros, trains, tramways, bus, vélo - 4 536 000 passes Navigo en circulation (2009) ⁹ - Validation seulement à l'entrée
Londres (Royaume-Uni) Oyster Card	2003	- Carte sans contact, rechargeable - Validation double tap dans les métros, mais seulement à l'entrée dans les bus et trains
Tokyo (Japon) Pasmo et Suica	2007	- Carte sans contact, rechargeable, prépayée - Plus de 7 millions de cartes Pasmo vendues durant l'année 2007 ¹⁰
Montréal (Canada) Opus	2008	- Carte à puce sans contact, rechargeable - Utilisable sur l'île de Montréal et en périphérie, et même à Québec. - 3,5 millions de cartes vendues entre 2008 et 2012 ¹¹

La plupart des auteurs traitant des cartes à puce sont d'accord pour dire que cette technologie a révolutionné la collecte des titres de transport. Cependant même si de nombreux avantages sont facilement distinguables, les systèmes de carte à puce présentent malgré tout quelques inconvénients. En nous appuyant sur la revue de littérature au sujet des cartes à puce dans les transports en commun de M.-P. Pelletier et al. (2011), nous pouvons dresser le tableau suivant des principaux avantages et inconvénients :

⁹ Wiki Ratp « Nouvelles technologies»

<http://archive.wikiwix.com/cache/?url=http://www.ratp.fr/corpo/service/6technologies.shtml#navi&title=Site%20officiel%20RATP> page consultée le 17 janvier 2015

¹⁰ East Japan Railway Company <http://www.jreast.co.jp/e/press/20080101/>, page consultée le 17 janvier 2015

¹¹ Radio Canada, « Remplacement des cartes Opus » <http://ici.radio-canada.ca/regions/Montreal/2012/02/29/003-carte-opus-replacement.shtml> page consultée le 18 janvier 2015

Tableau 2-2 : Avantages et inconvénients des cartes à puces dans les transports en commun

Avantages	Inconvénients
- Les coûts sont réduits, le service et la gestion des revenus améliorés, plus de flexibilité dans la tarification (McDonald, 2000)	- Aucune d'information sur le but et le lieu du voyage de l'utilisateur ainsi que sur l'évaluation du service ne peut être fourni en dépit de méthode pour en déduire (Bagchi et White, 2005)
- Caractérisation du comportement des utilisateurs plus facile en étudiant les données disponibles (Bagchi et White, 2005)	- Les coûts de recherche et développement et d'implémentation sont élevés (Deakin et Kim, 2001)
- La commodité et l'utilisation du système de perception sont améliorées pour les utilisateurs (Bagchi et White, 2005 & Ibrahim, 2003)	- Il y a un risque important dans l'investissement dans les systèmes de carte à puce, et le temps d'acceptation par les utilisateurs peut être long (Deakin et Kim, 2001)
- Les mécanismes de paiement et de flux d'informations sont améliorés (Blythe, 2004)	-Il n'y a aucune garantie d'amélioration de la rentabilité et le succès de la mise en œuvre dépend souvent des utilisateurs (McDonald, 2000)
- Le temps d'embarquement de l'utilisateur est réduit et la charge de travail du conducteur est aussi réduite (Chira - Chavala et Coifman , 1996)	-Plus la carte est complexe, moins sa fiabilité est garantie (Blythe, 2004)
- La durée de vie des cartes à puces est plus longue que celle des cartes traditionnelles. Utsunomiya et al., 2006)	-La confidentialité et la protection de la vie privée sont limitées (Dempsey, 2008; Trépanier, et al., 2004)

Les données de carte à puce peuvent être exploitées pour trois principaux usages selon M.-P. Pelletier et al. :

- Niveau stratégique : Les données peuvent être utilisées pour mieux cerner le comportement des utilisateurs (Agard et al., 2006 and Bagchi and White, 2005) à des fins marketing par exemple.
- Niveau opérationnel : Morency et al. (2007) et Trépanier et al. (2009) ont travaillé sur la génération d'indices de performance basés sur le traitement des données de carte à puce dans

les bus de manière à améliorer la conformité des horaires. D'autres chercheurs comme Chapleau et Chu (2007) ont eux travaillé sur la remontée, la quantification et la correction d'erreurs dans les données de transaction.

- Niveau tactique : C'est le niveau pour lequel le traitement des données de carte à puce est le plus répandu. Munizaga et al. (2010) ont utilisé les données de validations (Date, Heure, Lieu) pour générer des matrices origine-destination. Bagchi et White (2005) ont aussi travaillé sur ce sujet et cela permet de mieux adapter le service, en améliorant par exemple la coordination entre les bus et les métros.

Pour conclure cette partie sur les cartes à puce dans les transports, la littérature nous confirme que cette technologie est le STI qui s'est le plus imposé durant les deux dernières décennies. Elle devrait continuer à se banaliser durant les prochaines années en ce qui concerne sa fonction de perception des titres de transport. En termes de recherches, les possibilités restent encore très larges en ce qui concerne les trois niveaux d'exploitation des données en raison de l'avancée des technologies sur le stockage et le traitement des « big data ».

2.1.3 Les systèmes de comptage à bord automatique

Le comptage à bord communément appelé APC (*Automatic Passenger Counting*) est une technologie datant des années 1970. Ce système permettant de compter le nombre de personnes montant et descendant d'un véhicule a su adapter sa technologie au fil des années. En effet, les premiers systèmes étaient mécaniques et c (Schiavone, 1999). Puis l'électronique avec les capteurs infrarouges a réellement pris les devants (Boyle, 1998). Dans cette partie de littérature, nous nous focaliserons sur le fonctionnement de cette technologie, son rôle dans les transports en commun, sa fiabilité et sa précision de comptage, ainsi que les données issues de cette technologie et leur potentiel.

De nos jours, les systèmes de comptage à bord sont un ensemble de plusieurs technologies qui se décompose en trois principaux composants, les capteurs, une unité d'analyse, ainsi qu'un micro-ordinateur comme l'illustre la **Figure 2-2**.

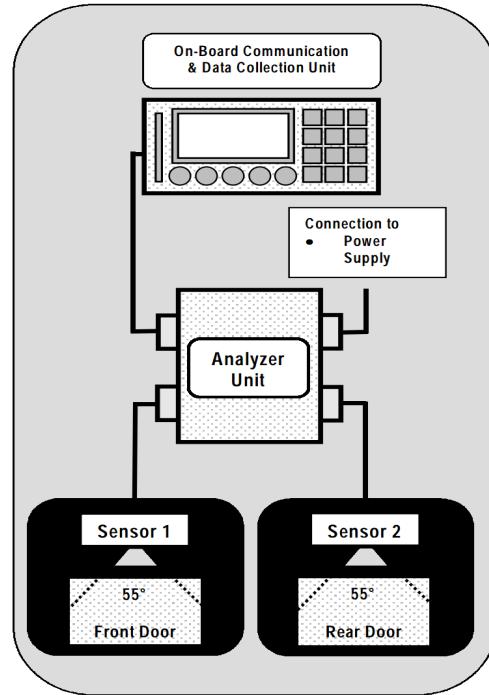


Figure 2-2 : Principaux composants d'un système de comptage à bord (tiré de <http://www.initag.de/share/news/Themen/APC.pdf>)

Les capteurs sont les éléments qui vont permettre de détecter la présence d'un individu. De nos jours, plusieurs technologies de capteurs existent. Les premières technologies étaient des capteurs au sol qui détectaient la pression d'un individu lors de la montée, ils sont désormais devenus obsolètes¹². Ensuite, les capteurs infrarouges ont fait leur apparition, d'abord positionnés latéralement à l'entrée du véhicule, puis en hauteur au-dessus des portes pour une meilleure efficacité. Ce sont les technologies les plus répandues. Bien souvent, les capteurs infrarouges sont composés de deux modules, un actif permettant de détecter la coupure d'un rayon, et un passif permettant de détecter une source de chaleur comme l'illustre la **Figure 2-3**.

¹² Automatic Passenger Counting « APC Report » http://www.itscanada.ca/files/MOBILE-APC_Cuta_final1.pdf, page consultée le 20 janvier 2015

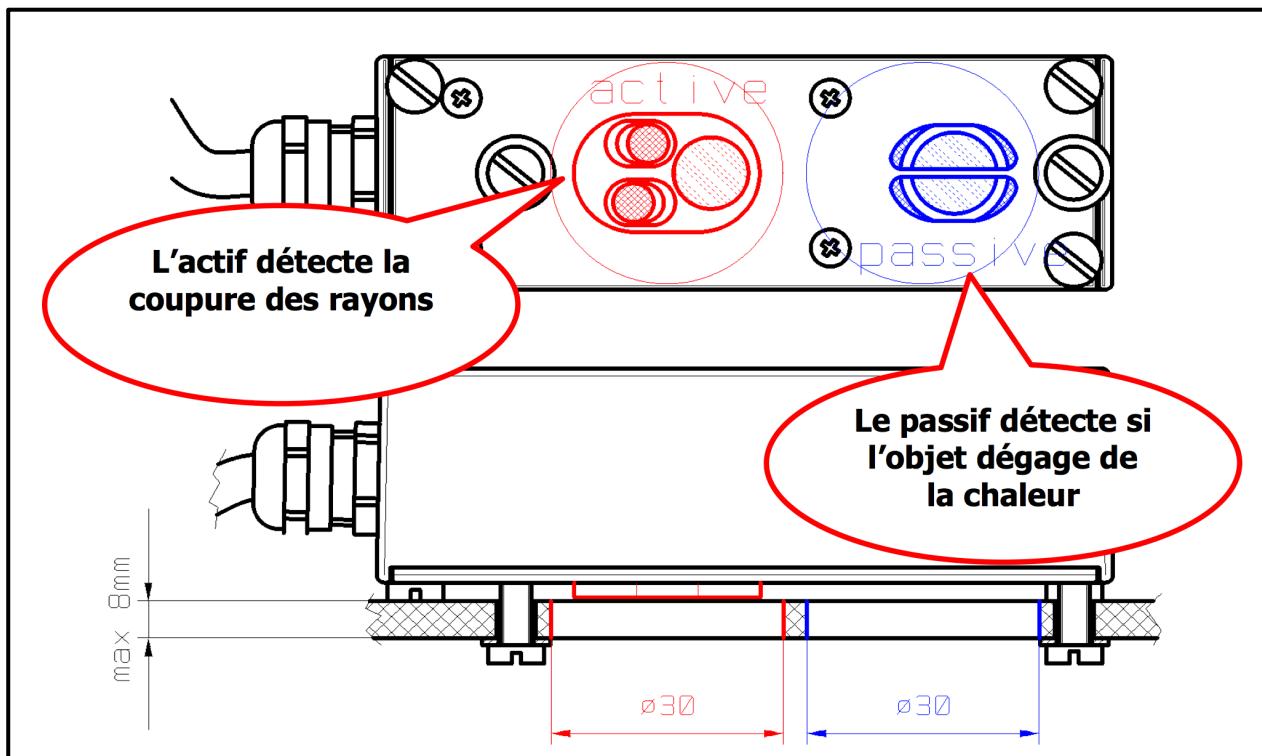


Figure 2-3 : Les capteurs actifs et passifs (Source STM)

Cette technologie permet alors de compter plusieurs personnes rentrant simultanément, ou d'éviter les comptages d'objets tels qu'une valise ou une poussette. De nouvelles technologies optiques commencent à faire leur apparition, comme des caméras stéréoscopiques¹³ qui viennent s'ajouter à la technologie infrarouge pour de meilleurs résultats.

L'unité d'analyse va permettre de récupérer les données des différents capteurs installés dans le véhicule pour les transmettre à l'ordinateur de bord.

L'ordinateur de bord va pouvoir récupérer l'ensemble des données des capteurs, pour les transmettre à ce qu'on appelle l'ordinateur de garage¹⁴. Il se compose d'une mémoire permettant le stockage des informations. Cet ordinateur de bord va pouvoir communiquer avec l'ordinateur

¹³ HELLA Introduces "People Counter" Systems at Two European Shows. <http://www.prnewswire.com/news-releases/hella-introduces-people-counter-systems-at-two-european-shows-171021501.html> page consultée le 20 janvier 2015

¹⁴ Infodev « System Installation & Data Collection » <http://www.infodev.ca/vehicles/vehicle-installations-data-collection.html>

du garage de deux manières, manuellement grâce à un petit boîtier portatif appelé collecteur de données que l'on branche chaque jour à l'ordinateur de garage ou directement en temps quasi réel par réseau internet Wi-Fi ou infrarouge. L'ordinateur de bord est souvent équipé d'un module AVL (*Automatic Vehicle Location*) permettant de suivre le parcours du véhicule, d'un odomètre pouvant compter le nombre de kilomètres parcourus, et d'un émetteur MP3 servant à annoncer le nom du prochain arrêt. Il est intéressant de noter que dans le cas des APC l'ordinateur est relié au système d'ouverture des portes de manière à synchroniser les données avec l'unité d'analyse à la fermeture des portes. Grâce à ces modules rajoutés sur l'ordinateur de bord, les données transmises au garage sont enrichies et permettent un traitement plus intéressant.

De nombreuses études ont examiné les impacts et avantages que pouvait avoir cette technologie sur les transports en commun. Boyle et Casey (1998, 1999) font partie des premiers à avoir étudié cette technologie. Boyle confirme les propos de Casey en disant que cette technologie permet de récolter un grand nombre de données rapidement et à moindre coût. Il insiste aussi sur le fait qu'il est important de consacrer du temps et des efforts pour assurer le succès de l'implantation d'une telle technologie. D'autres auteurs, tel que James G. Strathman a étudié l'impact des APC sur le Tri-Met (Tri-County Metropolitan Transportation District of Oregon), et a démontré une amélioration de 9% de la performance globale et une réduction de 18% des variations de temps de transports. En 2008, D. Boyle revient sur sa précédente étude en prenant en compte l'avancée technologique. Ses conclusions restent néanmoins très similaires, et il insiste toujours sur l'importance d'une gestion efficace et une bonne implication lors de l'implantation.

Quelques études ont aussi été menées sur la fiabilité et la précision de cette technologie. En 2002, la précision des dispositifs de comptage à bord était d'environ 95% (Kimpel et Strathman, 2002). Désormais grâce aux avancées de la technologie, la précision se situe plutôt autour de 97% voire 98% selon les vendeurs¹⁵.

Pour conclure cette revue de littérature au sujet des APC, on peut dire que la littérature insiste beaucoup sur la quantité de données disponible grâce à cette technologie qui en fait son atout le plus important. Ces données peuvent alors être étudiées à des fins de planification, de gestion

¹⁵ IRIS Infrared Intelligent Sensors« Product Irma 3d » <http://www.irisgmbh.de/products/irma-3d>

opérationnelle ou encore de gestion financière (Boyle, 2008). En termes de gestion financière, une des applications peut être l'estimation des pertes engendrées par l'évasion tarifaire.

2.2 Le phénomène d'évasion tarifaire dans les transports en commun

2.2.1 Le phénomène

L'évasion tarifaire (ou fraude tarifaire), lorsque l'on parle de transport en commun, concerne le fait d'utiliser ces services sans posséder de ticket ou avec des tickets contrefaits. Pour les entreprises gérant ces services, qu'elles soient publiques ou privées, ce phénomène qui émergea vraiment dans les années 1960 en Europe puis aux US (Diebel 1981), est un problème préoccupant. En effet, l'évasion tarifaire implique à la fois des pertes financières pour ces entreprises, mais elle a aussi des effets négatifs sur la sécurité et l'image publique de ces transports.

Dans la littérature, on remarque facilement que les transports en commun de type bus, métro et tramways sont les plus touchés par le phénomène, contrairement aux autres services de transport tels que les autoroutes à péage, les transports aérien ou ferroviaire, qui sont beaucoup moins touchés.

Même si depuis la fin des années 1980, les technologies naissantes (tels que les systèmes de vente automatique de tickets, les barrières automatiques ou les systèmes de carte à puce) ont permis une baisse de la fraude, les chiffres restent tout de même très importants et préoccupants.

Basée sur des données provenant de 31 compagnies de transport en commun de 19 pays différents, représentant 300 000 employés et 20 milliards de passagers par an, une étude affirme que la perte pour ces compagnies due à la fraude s'élève à 1 milliard d'euros par an (Bonfanti et Wagenknecht 2010). À Montréal, la fraude représente moins de 1% (STM 2014), mais elle atteint jusqu'à 27,7% dans des villes telles que Santiago de Chile (Torres Montoya, 2013). À titre d'exemple, un audit sur l'évasion tarifaire de l'agence de transport de Victoria en Australie a

permis d'estimer que l'évasion représentait environ 11,9% sur l'ensemble des passagers rien que pour le métro, causant des pertes de plus de 40 millions de dollars en 2011¹⁶.

Des chercheurs ont alors étudié qui étaient ces personnes fraudeuses et comment procèdent-elles.

Comme le montre le **Tableau 2-3** inspiré de l'article « Transport ticketing security and fraud controls » de K. Mayes (2009) il existe deux types de fraudes qui se déclinent suivant le mode de validation.

Tableau 2-3 : Types de fraudes

Type Fraude / Type Validation	Non Technique	Technique
Ticket	<ul style="list-style-type: none"> -Sans ticket (porte ouverte, saut de barrière ...) -Avec ticket non validé 	<ul style="list-style-type: none"> -Falsification de ticket, réutilisable à l'infini
Carte à puce	<ul style="list-style-type: none"> -Sans Carte à puce ou Carte à puce non rechargeée (porte ouverte, saut de barrière, ou ouverture par l'opérateur) -Sans carte à puce, mais validation avec celle d'un autre usager -Avec carte à puce valide d'un autre usager 	<ul style="list-style-type: none"> -Falsification de carte à puce

¹⁶ Victorian Auditor « Fare Evasion on Public Transport » Audit <http://www.audit.vic.gov.au/publications/20120829-Fare-Evasion/20120829-Fare-Evasion.pdf>

Lorsqu'elle est technique, la fraude est volontaire cependant la difficulté pour falsifier un ticket est bien plus faible qu'une carte à puce (K. Mayes, 2009). Lorsqu'elle est non technique, elle peut être souhaitée ou non. Le fraudeur peut être opportuniste (porte ouverte par exemple), ou en avoir la réelle intention (saut de barrière).

Suivant le mode de validation, mais aussi l'équipement permettant le passage (tourniquet, porte, grilles), la fraude est bien évidemment plus ou moins difficile. Cependant, ces technologies ont un prix, et tous les réseaux ne peuvent pas se permettre d'investir dans le renouvellement de leur mode de validation et de leurs équipements.

Des sondages internationaux ont montré que l'attitude de la population vis-à-vis de la fraude dans les transports en commun suivait une courbe de Gauss (**Figure 2-4**). Une grande majorité des fraudeurs sont jugés opportunistes, ils estiment préalablement le risque et passe à l'acte (Dauby et Kovacs, 2006) et très peu sont des fraudeurs endurcis.

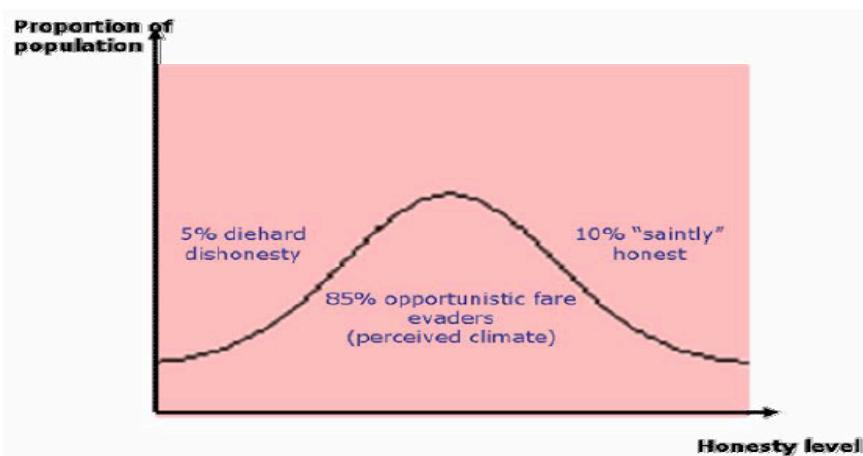


Figure 2-4 : Courbe de Gauss du comportement naturel (tiré de <http://onlinepubs.trb.org/onlinepubs/circulars/ec112.pdf>)

Plusieurs études ont été réalisées de manière à savoir quel était le profil de ces fraudeurs. A. Reddy a étudié et montré qu'à New York dans le métro, le passage des enfants en dessous des tourniquets représentait 43% des fraudeurs, et que le reste se divisait entre la violation des tourniquets et le passage par les portes. Dans les bus, le passage par la deuxième porte est la façon la plus courante. Dans ce même article, Ready a permis de mettre en évidence qu'une saisonnalité existait pour ce phénomène avec une augmentation du taux d'évasion tarifaire en été.

Il explique aussi que l'heure de la journée a un impact sur le taux d'évasion tarifaire. Durant les heures de pointe du matin et de la fin d'après-midi, les taux sont plus faibles. Ceci peut s'expliquer par le fait que le fraudeur doit être moins à l'aise de frauder devant le regard des autres usagers. En milieu d'après-midi les taux sont plus élevés, il y a moins de monde donc l'opportunité est plus forte. Ready explique aussi qu'à New York entre 15h et 16h, les étudiants quittent l'université ou l'école, et il constate un pic d'évasion tarifaire à cette heure.

2.2.2 Les méthodes existantes pour réduire l'évasion tarifaire

Les compagnies mettent en place de plus en plus d'inspections et de systèmes visant à lutter contre ce phénomène, cependant cela a un coût. Bien que la carte à puce n'a pas comme unique but de réduire l'évasion tarifaire, la STM a déboursé 174 M\$ en 2008 pour se doter de cette technologie, un somme qui n'est pas à la portée de n'importe quelle agence de transport. Cependant les résultats sont bien là, depuis l'implantation de cette technologie, la STM estime que la carte OPUS aurait permis de diminuer incontestablement la fraude et que 13 M\$ seraient sauvés chaque année¹⁷. Les entreprises doivent alors trouver un équilibre pour que les systèmes mis en place pour réduire l'évasion tarifaire soient les mieux adaptés.

Jusqu'à aujourd'hui, plusieurs articles traitent du phénomène de fraude dans les transports publics en étudiant ses impacts (R. Clarke 1992, K. Mayes, 2009) ou comment la mesurer (Reddy et al. , 2011). Les impacts sécuritaires de ce phénomène sont plus souvent étudiés que l'aspect économique.

Mais quelques articles présentent des méthodes pour lutter contre ce phénomène. Barabino et al (2013), ont étudié un modèle d'optimisation de profit basé sur le niveau d'inspection optimum. Ils ont testé ce modèle en Italie, et trouvent un taux d'inspection optimum de 4,5%, c'est-à-dire un contrôle de 45 personnes sur 1000. Ce taux d'inspection est évident directement lié au nombre d'inspecteurs. Comme l'affirme G. Bonfanti (2010) le nombre de fraudeurs est inversement proportionnel au nombre d'inspecteurs. La présence d'un contrôleur dans un bus ou près des

¹⁷ Journal Métro « La fraude en chute dans le réseau de la STM »

<http://journalmetro.com/actualites/montreal/429948/la-fraude-en-chute-dans-le-reseau-de-la-stm/> page consultée le 21 janvier 2015

bornes de validation empêche ou diminue en tout cas très fortement la fraude. Mais ils ne peuvent pas être partout et à tout moment. Les sociétés de transports ne peuvent pas se permettre d'employer trop de personnel, elles doivent trouver le juste milieu pour que ce soit le plus rentable possible. Thorlacius et Clausen (2010) se sont eux plutôt orientés vers l'allocation temporelle des inspecteurs en fonction des heures de pointe ou des heures creuses dans une même optique d'optimisation de profit.

D'autres auteurs ont proposé de mettre en place des partenariats entre plusieurs institutions (police, justice, éducation) (Bonfanti et Wagenknecht 2010, Dauby et Kovacs 2006) de manière à réduire ce phénomène. À Rouen en France, la mise en place de sensibilisation à lutter contre ce phénomène dans les écoles, d'équipes de polices déployées dans les réseaux de transports, ou de jugements rapides permettant de condamner des fraudes répétées par des travaux d'intérêts généraux par exemple, sont des très bons appuis aux sociétés de transport (Dauby et Kovacs, 2006)

Un des enjeux des sociétés de transport en commun est donc de trouver des moyens pour réduire ces opportunités qui peuvent se présenter aux utilisateurs. Lorsque les réseaux possèdent des installations récentes avec des caméras, un système de validation difficile à enfreindre physiquement, ou encore des opérateurs vigilants, ces opportunités diminuent, cependant il sera impossible de les réduire à néant.

Comme nous le disions auparavant, ce phénomène étant très ciblé dans l'espace et dans le temps, l'allocation des inspecteurs est très difficile à prévoir. Mais avec les avancées de la technologie, et la mise en place des systèmes des cartes à puces, il est désormais possible de caractériser des flux de personnes, de contrôler le nombre de validations en temps réel. À Montréal, depuis l'implantation de cette technologie en 2008, l'évasion tarifaire a fortement chuté en incitant la plupart utilisateurs à prendre des abonnements qui ne leur donnent plus l'opportunité de frauder pour économiser de l'argent.

2.3 Synthèse de la revue de littérature

À travers cette revue de littérature, nous avons constaté que les données de carte à puce et de comptage à bord étaient issues de technologies de pointe, et que l'implantation de ces systèmes était relativement récente dans les réseaux de transport en commun. De plus, la quantité de

données disponibles grâce à ces technologies représente une véritable mine d'or pour les agences de transports en terme de planification globale, de coordination des bus, d'analyse du comportement des usagers et évidemment d'évasion tarifaire. Il est certain que la liste des possibilités offertes par ces données ne s'arrête pas là. L'avancée de la technologie dans ces domaines nous pousse à croire que les données vont continuer à s'enrichir et devenir de plus en plus complètes, et qu'elles seront vraiment indispensables aux agences de transport.

Concernant le phénomène d'évasion tarifaire, nous avons relevé que de nombreuses études avaient été réalisées à ce sujet depuis plusieurs décennies. Cependant, dans la plupart des études les auteurs se basent sur des données issues de sondages ou d'enquêtes de terrain pour caractériser ce phénomène. Ils étudient souvent les conséquences financières ainsi que les modes opératoires des fraudeurs. Aucune étude n'avait réellement été réalisée en se basant sur les données des systèmes de carte à puce et de comptage à bord pour caractériser ce phénomène, d'où l'enjeu de notre étude.

CHAPITRE 3 MÉTHODOLOGIE

L'étude présentée dans ce mémoire se base sur les données de carte à puce et de comptage à bord récupérées auprès de la Société de transport de Montréal. Dans cette partie, nous présentons le contexte de récupération de ces données, puis nous détaillons leur architecture. Ensuite, nous expliquons comment le traitement et l'analyse de ces données ont été réalisés, c'est-à-dire la logique mise en place ainsi que les outils utilisés.

3.1 Récupération et présentation des données

Dans le cadre de cette recherche, nous avons pu accéder aux données provenant du système de perception par carte à puce et de comptage à bord de la Société de transport de Montréal pour la semaine du 7 au 13 avril 2014. Il s'avère qu'à la suite des premières analyses, nous avons voulu comparer cette semaine avec une autre. La STM nous a donc également fourni les données du 6 au 12 octobre 2014. Mentionnons ici que les données provenant du système de perception par carte à puce sont complètement anonymes, et qu'il n'existe aucune façon de pouvoir relier les données au détenteur d'une carte, puisque les numéros de carte sont encryptés à l'aide d'un algorithme irréversible.

Le tableau présente ci-dessous les données brutes de comptage à bord appelé SCAD (système de comptage automatisé et d'acquisition des données).

Tableau 3-1 : Présentation des données de comptage à bord

SCAD	Exemple	Explications
Date_Sortie	10/04/2014	Date du jour de la sortie
Numero_Vehicule	27018	Numéro du véhicule
Numero_Ligne	32	Numéro de la ligne de bus
Numero_Arret	112502	Numéro de l'arrêt de bus
Desc_Arret_Client	Saint-Joseph / De Bullion	Nom de l'arrêt de bus
Lat_GPS	45,524102	Latitude
Lng_GPS	-73,590067	Longitude
Heure_Depart_Arret	0,37418	Heure numérique du départ de l'arrêt

Tableau 3-1 : Présentation des données de comptage à bord (suite et fin)

SCAD	Exemple	Explications
Montants_Porte_1	4	Nombre de personnes montantes porte 1 entre l'ouverture et la fermeture des portes
Descendants_Porte_1	0	Nombre de personnes descendantes porte 1 entre l'ouverture et la fermeture des portes
Montants_Porte_2	0	Nombre de personnes montantes porte 2 entre l'ouverture et la fermeture des portes
Descendants_Porte_2	1	Nombre de personnes descendantes porte 2 entre l'ouverture et la fermeture des portes
Montants_Porte_3	0	Nombre de personnes montantes porte 3 entre l'ouverture et la fermeture des portes (pour les bus à 3 portes)
Descendants_Porte_3	0	Nombre de personnes descendantes porte 3 entre l'ouverture et la fermeture des portes (pour les bus à 3 portes)

Il est important de préciser que l'heure est au format décimal et que l'unité est la journée, c'est-à-dire que 24h = 1, et par exemple 0,371418 = 8:58:49.

Le système de comptage à bord équipe 312 bus de la STM sur les 1900 qui parcourrent le réseau représentant environ 950 000 montants par semaine dans ces bus répartis sur 171 lignes différentes. Concernant les données, pour une semaine, le fichier récupéré comporte 782 765 enregistrements.

Ci-dessous voici le tableau explicatif concernant les données de carte à puce.

Tableau 3-2 : Présentation des données de carte à puce OPUS

OPUS	Exemple	Explications
Id_Support	764	Id du support
Date_Oper	2014-04-11	Date de la validation
Heure_Oper	09:53:12	Heure de la validation
Id_Deplacement	1999674614	Id du déplacement (prend en compte les correspondances)
Id_Validation	1999485881	Id de la validation (unique)
Desc_Type_Support	CPCT	Description du type de support (Carte à puce commune, Carte à puce magnétique jetable, Carte Puce occasionnelle)
desc_Type_Validation	Correspondance	Type de trajet
Desc_Produit	STM-CAM MENSUELLE-O	Description du type de produit (abonnement mensuel, journalier, hebdomadaire, trajet unique, etc...)
Code_Ceb	32BPB--022371----01	Numéro du véhicule
Code_Ligne	018	Numéro de la ligne de bus

Concernant les données de carte à puce, le format de l'heure est en format date classique. Le Code_Ligne correspond au numéro de la ligne, cependant un 0 se trouve devant chaque numéro. Le Code_Ceb permet de distinguer de quel véhicule il s'agit, il contient le numéro du véhicule à l'intérieur, ici c'est le bus numéro 22371. Les données de carte à puce récupérées comportent aussi l'ensemble des validations dans les métros, on distingue notamment s'il s'agit d'un percepteur OPUS de bus ou de celui d'un métro grâce au Code_Ceb. Une ligne de donnée correspond à une validation, ce qui représente 10 502 628 entrées pour la semaine d'avril 2014 pour l'ensemble bus et métros.

Le système de carte à puce est lui présent sur tous les bus et métros de l'île de Montréal, cela représente 465 collecteurs OPUS dans les métros de Montréal, en plus des 1900 bus. Le système enregistre tous les types de paiement, soit ceux effectués avec le carte OPUS, les billets et le paiement comptant, pour lequel un billet unitaire est émis par le valideur.

3.2 Processus de traitement

Le but du traitement des données a été de relier ces deux bases de données relativement volumineuses de manière à comparer le nombre d'usagers montants dans les bus et ceux qui valident. Un logiciel classique de manipulation de données comme Excel ne permet pas d'ouvrir de tels fichiers et d'en faire un traitement efficace. Nous avons plutôt utilisé R, un logiciel *Open-Source* de traitement de données et d'analyses statistiques. Le but a été de faire la somme des validations entre chaque ouverture et fermeture de porte, c'est-à-dire pouvoir comparer la somme des montants et des validations à chaque arrêt de bus.

Tout d'abord, il a fallu passer par une phase de prétraitement des données. En effet, les données sont issues de deux technologies différentes, ce qui implique des différences dans la mise en forme des données. Ce prétraitement permet de joindre les deux tables de données plus facilement. L'étape de jointure est l'étape clef du processus. Il a fallu faire correspondre les validations et les comptages entre chaque ouverture et fermeture de porte. Enfin, une étape de post traitement a été nécessaire pour corriger les erreurs liées aux données, enrichir les données et préparer l'exportation. La **Figure 3-1** présente le processus de traitement que nous détaillerons ensuite :

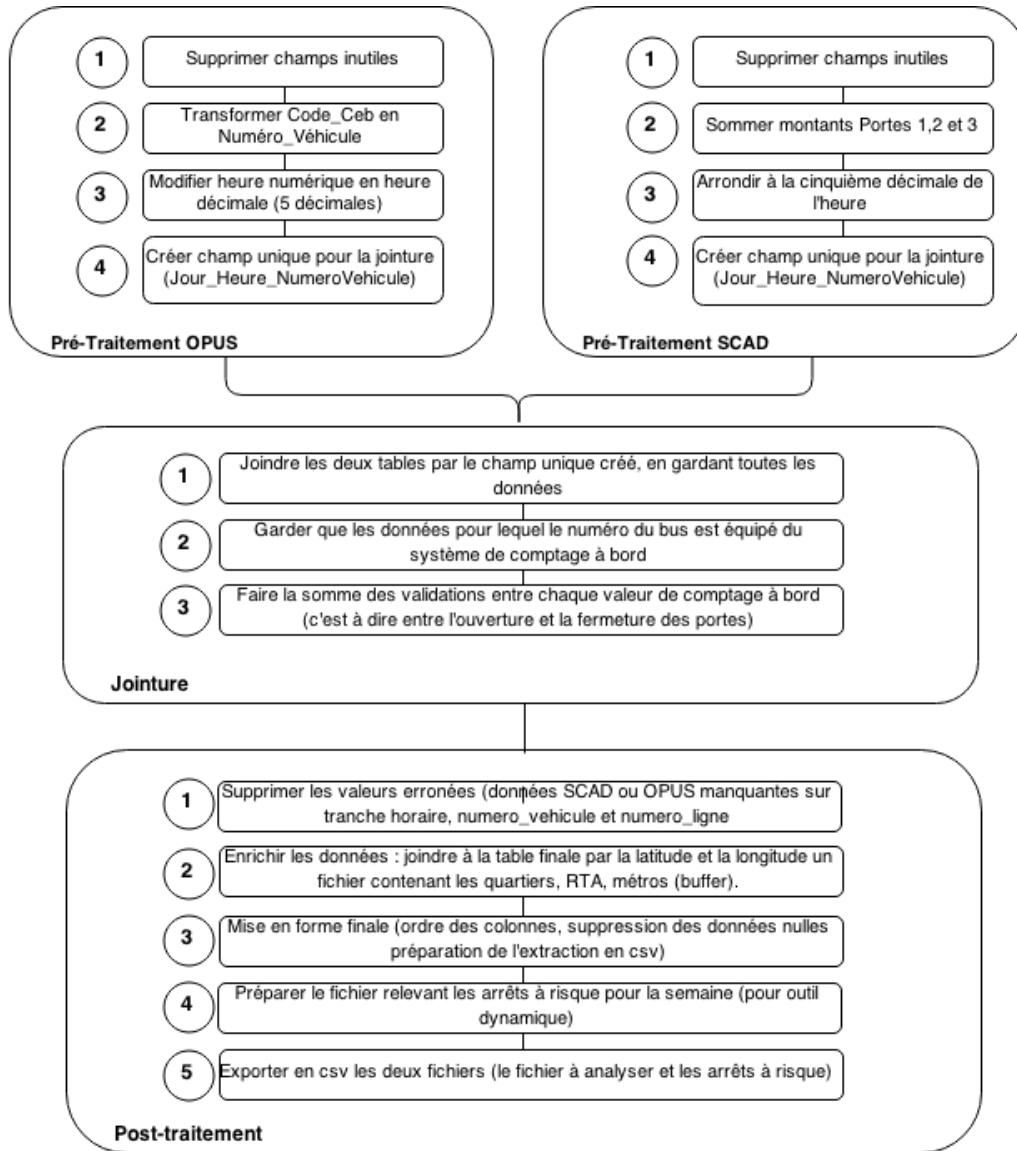


Figure 3-1 : Processus de traitement des données

Concernant les pré-traitements, il s'agit la plutôt d'une étape de mise en forme des données. La création d'un champ unique pour chaque entrée avec la date du jour, l'heure et le numéro du véhicule va nous permettre de joindre les données SCAD et OPUS. En effet, pour la jointure le but a été d'apparier les deux tables par ce champ unique. Évidemment, toutes les données sont gardées lors de la jointure et le tableau suivant nous présente l'état des données après l'étape 2 de la jointure.

Tableau 3-3 : Exemple des données jointes au milieu du traitement

DATE_SORTIE	NUMERO_VEHICULE	HEURE_DEPART_ARRET	DESC_ARRET_CLIENT	SOMME_MONTANTS	NOMBRE_VALIDATIONS
2014-04-07	22040	0.54499	Arrêt A		1X
2014-04-07	22040	0.54545	X	X	1
2014-04-07	22040	0.54549	X	X	1
2014-04-07	22040	0.54552	X	X	1
2014-04-07	22040	0.54556	X	X	1
2014-04-07	22040	0.5456	X	X	1
2014-04-07	22040	0.54561	Arrêt B		5X

Il faut noter que nous avons simplifié le tableau issu du traitement à des fins d'explication, toutes les colonnes ne sont pas présentes. Ce tableau nous dit qu'à l'arrêt B le système SCAD a détecté 5 personnes montantes à l'heure numérique 0,54561 (13:05:40), le lundi 7 avril 2014 pour le bus numéro 22040. De plus, on note que 5 validations ont été effectuées entre 0,54545 (13:05:26) et 0,54560 (13:05:39), pour ce même bus ce jour-là. L'étape 3 du traitement a été d'effectuer des sommes de ces validations entre deux entrées SCAD pour faire mettre en vis à vis Validations et Montants.

Une fois les données Validations et Montants en vis à vis, nous avons pu commencer l'étape de post-traitement. Les étapes étant relativement bien explicites sur la **Figure 3-1**, nous nous attarderons seulement sur l'étape 1. Il s'est avéré qu'après la jointure, nous nous sommes aperçus que de nombreuses données aberrantes résidaient dans ces données. Nous avons remarqué à plusieurs reprises un grand nombre de montants pour un arrêt, mais aucune validation à ce moment-là. Ces données aberrantes qui proviennent selon nous d'une défaillance du collecteur ou de sa désactivation par le conducteur rendaient les ratios totalement incohérents sur certains bus ou ligne. Nous avons donc décidé d'effectuer des tests sur chaque entrée de données pour supprimer ces données qui empêchaient toutes analyses.

Concernant l'étape d'enrichissement nous la détaillerons un peu plus tard avec l'utilisation de l'outil QGIS, les autres étapes sont relativement explicites et ne nécessitent pas une explication plus détaillée.

Le **Tableau 3-4** présente le fichier de résultats créé à la suite des étapes précédentes et permettant l'analyse.

Tableau 3-4 : Fichier de résultat pour analyse

OPUS/SCAD	Exemple
DATE_SORTIE	2014-04-07
HEURE_DEPART_ARRET	0.28162
LAT_GPS	45.485393
LNG_GPS	-73.627853
QUARTIER	Cote des Neiges-Notre Dame de Grace
RTA	H3W
METRO	Station Snowdon
NUMERO_ARRET	131692
DESC_ARRET_CLIENT	Station Snowdon (Queen-Mary/Westbury)
NUMERO_LIGNE	51
NUMERO_VEHICULE	26077
NOMBRE_VALIDATIONS	8
SOMME_MONTANTS	8
SOMME_DESCENDANTS	2

Il faut noter que les lignes Quartier, RTA (région de tri d'acheminement, les trois premières lettres du code postal) et Métro proviennent de l'enrichissement des données grâce à un logiciel GIS, nous expliquerons la méthode un peu après. La ligne Métro correspond au fait que l'arrêt de bus soit à proximité d'une station de métro ou non.

C'est donc sur deux fichiers finaux de respectivement 288 102 entrées et 273 733 entrées de ce type pour les semaines d'avril et octobre que je me suis appuyé pour analyser les données.

3.3 Processus et outils d'analyse

Le fichier de résultat issu du traitement (présenté au **Tableau 3-4**) étant un fichier CSV ne dépassant pas les capacités d'Excel, nous avons décidé d'utiliser ce tableur pour effectuer l'analyse à travers des tableaux et graphiques croisés dynamiques. C'est un bon outil d'analyse relativement simple d'utilisation et permettant une maniabilité des paramètres très efficace. Voici

une petite présentation de l'outil TCD (tableau croisé dynamique) d'Excel appliquée aux fichiers de résultats.

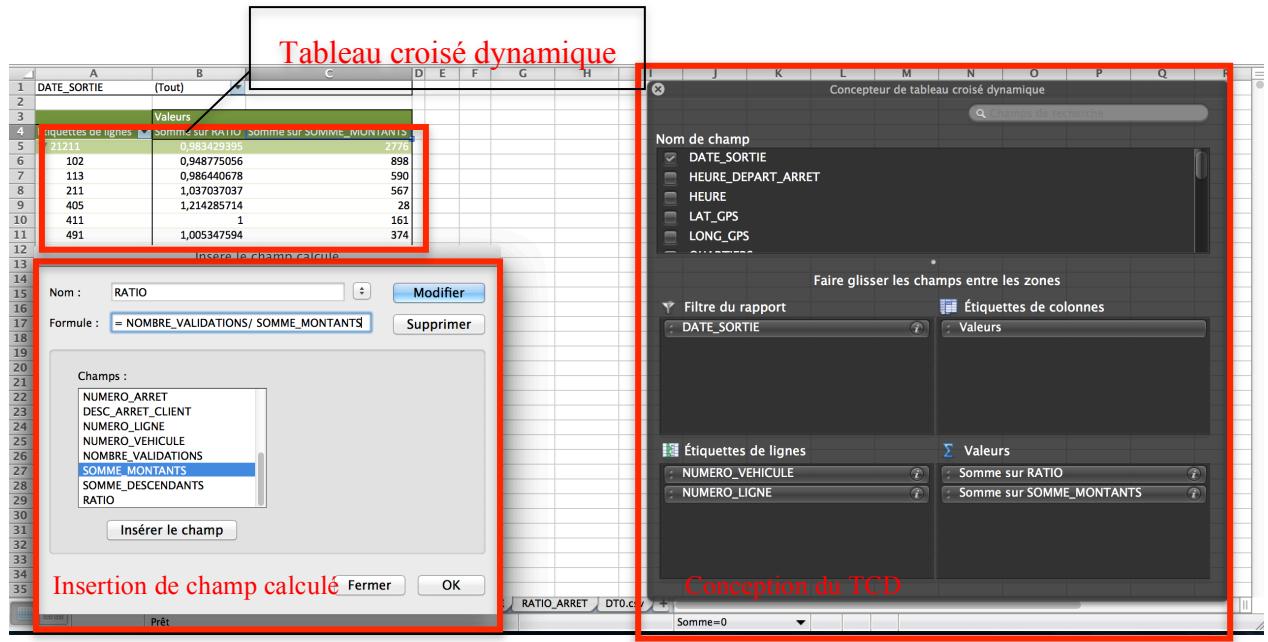


Figure 3-2 : Tableau croisé dynamique

Cette dernière figure présente la démarche de création d'un TCD, avec son concepteur et son module de création de champs pour faire calculer le ratio entre le nombre de validations et le nombre de montants.

C'est à partir de cet outil que nous avons pu analyser le fichier brut des résultats et comparer les paramètres pertinents permettant de caractériser l'évasion tarifaire. Cependant, pour une meilleure visibilité et compréhension, des graphiques croisés dynamiques ont aussi été réalisés. De plus nous avons utilisé le logiciel *Open-Source QGIS* qui a permis dans un premier temps d'enrichir les données et ensuite de représenter les résultats en fonction des quartiers, zones, métros et arrêts.

Pour l'enrichissement des données, nous avons extrait les coordonnées GPS de chaque arrêt de bus du réseau de la ville de Montréal puis superposé des couches géographiques de type *shapefile* des quartiers de Montréal, des régions de tri d'acheminement (RTA, les trois premières lettres du code postal) et des stations de métro. Nous avons ensuite pu associer à chaque arrêt son quartier,

sa RTA et sa présence ou non dans un rayon de 75 m autour d'un métro. Ci-dessous en **Figure 3-3**, la couche des quartiers avec l'ensemble des arrêts de bus de l'île de Montréal superposés.



Figure 3-3 : Superposition de la couche quartier et des arrêts de bus sur Qgis

Pour ce qui est de la représentation des ratios par zone, quartier ou arrêt, nous avons extrait les données à partir des analyses des TCD, puis créé des couches sur QGIS. Nous verrons ces résultats dans la partie IV.

Des régressions linéaires ont aussi été effectuées avec R pour analyser l'influence de certains paramètres sur l'évasion tarifaire. Nous détaillerons aussi ces régressions dans la partie IV.

CHAPITRE 4 ANALYSE DES RÉSULTATS ET INTERPRÉTATIONS

Cette partie a pour objectif de présenter les résultats et de les analyser. Pour cela nous nous sommes appuyés sur les deux semaines de données disponibles et nous avons étudié les différents paramètres dans le but de caractériser l'évasion tarifaire de façon indirecte à l'aide du ratio entre les transactions de carte à puce et les montants. La première partie concerne l'étude de ces différents paramètres tandis que la deuxième sera plutôt une analyse comparative des deux semaines de données.

4.1 Analyse descriptive en fonction des paramètres disponibles

4.1.1 Jour de la semaine pour les dates du 7 au 13 avril 2014

La première analyse que nous allons aborder correspond au jour de la semaine. L'évasion tarifaire est-elle caractérisable en fonction des différents jours de la semaine ? Pour répondre à cela, nous avons analysé sur l'ensemble d'une semaine, et sur l'ensemble des bus du réseau de la STM, le ratio entre le nombre de personnes ayant validé leur titre et le nombre de personnes qui sont montés à bord. Voici ci-dessous un graphique représentant le ratio en fonction des différents jours de la semaine. Un ratio inférieur à 1 signifie qu'il y a eu plus de montées détectées que de transactions dans le système de perception par cartes à puce.

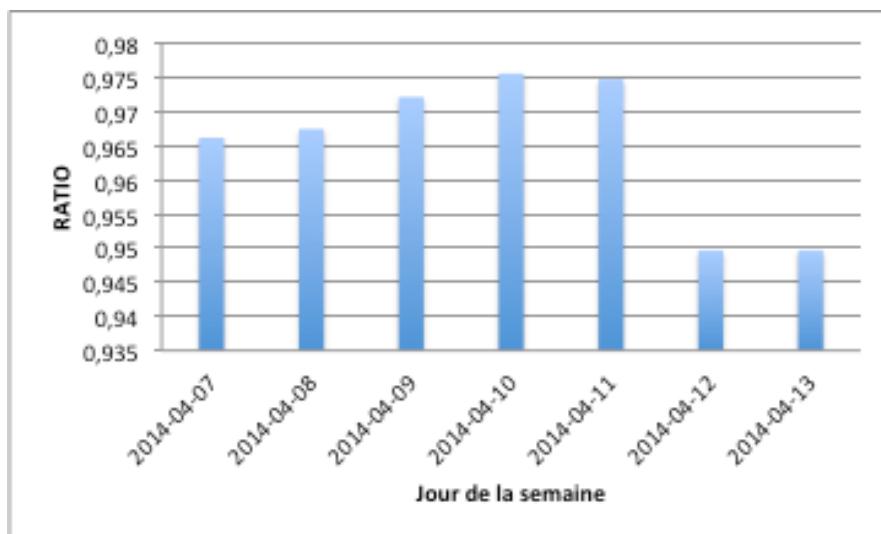


Figure 4-1 : Ratio Nombre_Validations/Somme_Montants par jour de la semaine (du 7 au 13 avril 2014)

À des fins d'interprétation, il est important de noter qu'il s'agit d'une semaine complète du lundi au dimanche. On remarque alors que lors de la fin de semaine le ratio est plus faible que la semaine d'environ 1,5% (en points de pourcentage). Théoriquement, cela signifie donc qu'il y a moins de personnes validant en fin de semaine que la semaine par rapport au nombre de personnes montantes dans les bus. Cependant, il s'avère qu'à Montréal, les enfants de moins de 12 ans ne paient pas les transports en commun la fin de semaine. Cela peut alors justifier cet écart par rapport au reste de la semaine, et on ne peut donc pas vraiment en tirer de conclusions sur le phénomène d'évasion tarifaire en fin de semaine. Pour la suite des analyses, nous éviterons donc de prendre en compte les ratios des fins de semaine. Cependant les autres jours de la semaine, on remarque que le ratio est plus faible au début de la semaine qu'au milieu. Il sera alors intéressant de comparer cela à une autre semaine, c'est ce que nous ferons dans l'analyse comparée en partie 4.2.

4.1.2 Heure de la journée

Les transports en commun sont soumis à des variations importantes de fréquentation, appelées heures de pointe et heures creuses. En général, les heures de pointe correspondent aux moments où les usagers se rendent au travail et en sortent, c'est à dire le matin entre 6h et 9h, et en fin d'après-midi 15h-18h. Dans la littérature nous avons lu que l'évasion tarifaire était souvent plus faible en heure de pointe due au regard des autres personnes et au manque d'opportunité. Nous avons jugé utile de voir si nous pouvions confirmer ou réfuter cela.

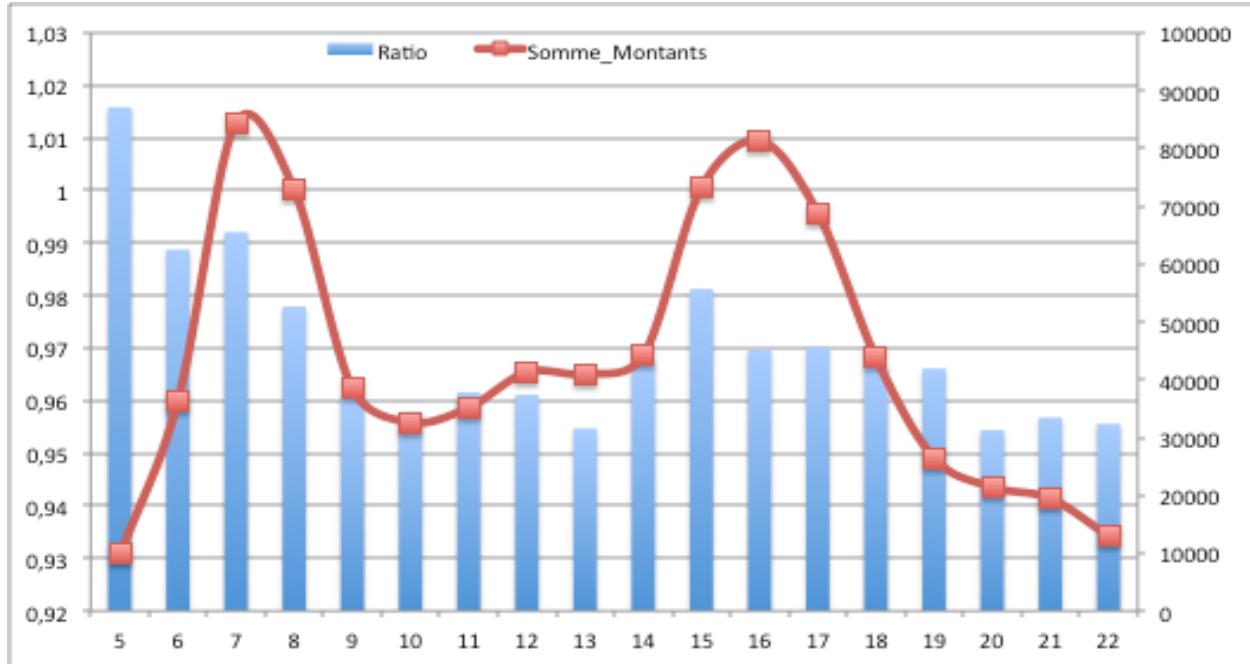


Figure 4-2 : Ratio Nombre_Validations/Somme_Montants en fonction des heures de la journée (lundi 7 au vendredi 11 avril 2014)

La courbe rouge nous indique la somme de l'achalandage dans les bus pour les 5 premiers jours de la semaine. Nous avons volontairement enlevé la fin de semaine pour bien identifier les heures pointes et les heures creuses en semaine. On distingue alors assez facilement les deux périodes d'heures de pointe qui se situe entre 7h et 9h le matin, et entre 15h et 18h. Si nous nous intéressons au ratio pour ces deux périodes, on peut noter que celui-ci est plus élevé en pointe que pendant la période creuse de 9h à 14h. Cela nous permet de croire que l'évasion tarifaire serait bien sujette au phénomène d'heures creuses et d'heures de pointe, tel que le rapporte Ready. Pour appuyer cela, nous avons isolé une journée, comme présenté à la **Figure 4-3**.

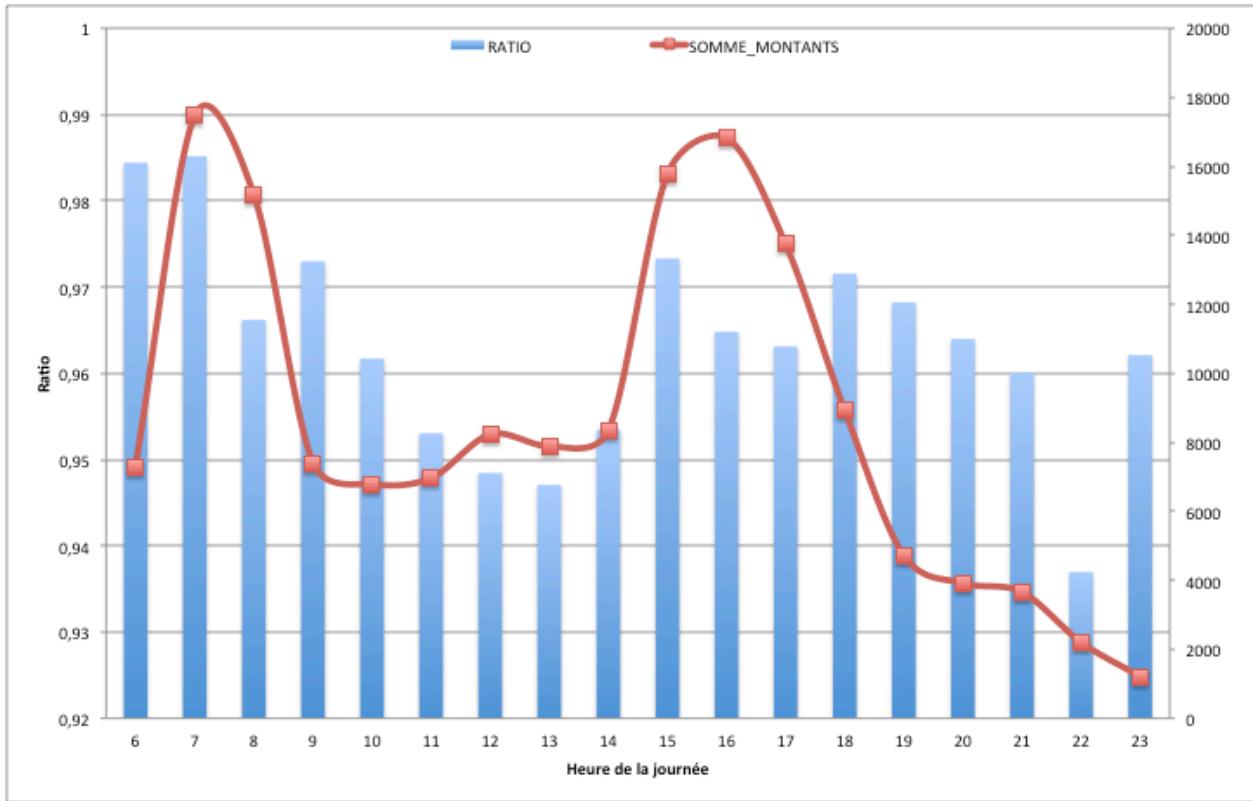


Figure 4-3 : Ratio Nombre_Validations/Somme_Montants en fonction des heures de la journée (mardi 8 avril 2014)

On observe le même phénomène pour le mardi que pour les 5 jours de la semaine rassemblés. Il semblerait donc bien que les heures creuses soient potentiellement plus soumises à l'évasion tarifaire que les heures de pointe.

4.1.3 Différentes lignes de bus et types de bus

La fréquentation des lignes de bus est souvent bien différente en fonction de son trajet. L'évasion tarifaire comme nous l'avons vu dans la revue de littérature peut-être volontaire et quotidienne. On peut alors se demander si l'on peut caractériser ce phénomène en fonction des différentes lignes qui composent le réseau. Le but étant de constater ou non le faible ratio Validations/Montants sur une certaine ligne de bus et de voir si ce phénomène est récurrent. Ce raisonnement permet alors de distinguer et de catégoriser les lignes de bus en fonction de leur risque potentiel d'évasion tarifaire. Le **Tableau 4-1** nous montre les douze lignes ayant le plus faible ratio sur la semaine complète du 7 au 13 avril 2014.

Tableau 4-1 : Ratios les plus faibles pour la semaine d'avril

NUMERO_LIGNE	RATIO	SOMME_MONTANTS	INDICE_NBRE_MONTANTS
40	0,88387097	155	1
877	0,8843302	2023	5
39	0,88915663	415	2
11	0,90134228	1490	4
747	0,9054887	3407	7
216	0,91111111	360	2
185	0,91499756	2047	5
102	0,91932527	2727	6
355	0,92101341	671	3
365	0,92105263	114	1
195	0,92492492	4995	7
30	0,93017456	2005	5

Pour pouvoir analyser ce tableau, il faut auparavant connaître la valeur moyenne du ratio sur l'ensemble du réseau pour la semaine. Il s'avère que ce ratio moyen se trouve à 0,967 et l'on peut imaginer que cela peut correspondre au niveau de fiabilité des capteurs infrarouges des APC. Cependant, les dix premiers ratios sont bien inférieurs à cette moyenne, et nous faisons donc l'hypothèse qu'au moins une partie de cet écart est due à l'évasion tarifaire. Toutefois, nous ne pouvons pas décrire à l'aide de ces résultats que la ligne 40 est la ligne la plus à risque de Montréal. En effet, comme nous l'avons présenté en partie 3, seulement 16% des bus sont équipés du système APC sur l'île de Montréal. De plus, chaque bus n'est pas associé à une ligne en particulier. Il s'avère que la ligne 40 a été parcourue que par 2 bus équipés du système de comptage à bord durant 4 trajets pendant la semaine. Cela représente un échantillon de seulement 155 montants. Il faut donc être vigilant et considérer un échantillon plus élevé de passagers pour conclure que c'est une ligne ayant un risque d'évasion tarifaire potentiellement plus élevé. C'est pourquoi nous avons jugé utile de mettre un indice entre 1 et 10 représentant le nombre de montants pour ce bus. Pour déterminer cet indice, nous avons divisé l'ensemble des 172 lignes en 10 quartiles en fonction de leurs nombres de montants sur la semaine d'avril ici. Cela permet de donner plus ou moins d'importance aux données remontées par ligne de bus. L'étude comparative avec la deuxième semaine de données pourra aussi nous confirmer ces lignes dites à risque.

Nous avons aussi étudié les différents types de lignes. À Montréal, il y a 6 types de lignes résumées dans le **Tableau 4-2** ci-dessous.

Tableau 4-2 : Types de lignes de bus sur le réseau de la STM

Types de lignes	Numéros des lignes	Description
Réseau local et 10 minutes max	1 à 249	Les bus du réseau local sont partout. Les bus 10 minutes max passent toutes les 10 minutes ou moins.
Réseau de nuit	350 à 382	Les bus de nuits fonctionnent après minuit 7 nuits/7.
Réseau express	401 à 496	Ces lignes font moins d'arrêts et circulent sur des voies réservées.
Navettes	715, 747 et 777	Ces navettes desservent l'aéroport et certaines destinations touristiques
Navettes temporaires	800 et plus	Ces navettes sont mises en place en cas de panne de métro ou de travaux sur les lignes.
Navettes OR	252 à 300	Les navettes OR sont mises en place pour répondre au besoin des aînés.

Nous avons donc pu étudier le ratio Validations/Montants pour les différents types de bus, comme le présente la **Figure 4-4** qui correspond aux résultats issus du logiciel R.

Call : lm(formula = RATIO ~ TYPES_DE_LIGNES)

Valeurs résiduelles :

Min	1Q	Median	3Q	Max
-0,2417	-0,0382	-0,0104	0,0095	5,8073

Coefficients :

	Valeur pente	Erreur type	t value
Réseau Local (référence)	0,972054	0,040215	24,172
Réseau de nuit	0,001979	0,299734	0,009
Réseau express	0,220663	0,093162	2,369
Navette	-0,027455	0,327941	-0,084
Navette temporaire	-0,041377	0,268768	-0,154

R carré	0,03364
---------	---------

Figure 4-4 : Résultats par types de lignes de bus pour la première semaine (7 au 13 avril 2014)

À présent, nous allons commenter ces résultats, mais nous ne nous intéresserons pas à tous les termes de ce résumé. Les valeurs résiduelles nous informent sur la qualité de la régression linéaire. Cependant nous ne cherchons pas à prouver ici que nos résultats suivent un modèle linéaire, donc nous ne nous pencherons pas sur ces valeurs. Les valeurs des coefficients sont cependant plus importantes et significatives pour nous. La valeur de référence correspond au réseau de bus locaux en termes de ratio. On peut ainsi comparer les valeurs des ratios des types de lignes par rapport à cette référence. On observe donc que les bus de nuit ont un ratio quasiment identique, alors que les bus du réseau express ont un ratio nettement supérieur, donc potentiellement moins à risque. À l'inverse, les navettes et navettes temporaires ont un ratio de respectivement 0,9445 et 0,9306. Les autres coefficients ne nous intéressent pas pour l'analyse. La valeur de multiple R-carré étant très faible, nous pouvons conclure qu'il n'y a pas de corrélation entre les différents types de bus, ce qui semble logique.

Concernant l'interprétation de ces résultats, le fait que les navettes et navettes temporaires aient des ratios plus faibles ne semble pas incohérent. En effet, la navette pour l'aéroport (747) a un ratio relativement faible, et selon les dires du personnel de la STM que nous avons rencontré, ils s'avèrent que des personnes viennent à entrer dans le bus pour aider à porter les valises, ou la personne rentre et sort plusieurs fois si elle transporte plusieurs bagages. Le comptage à bord devient alors moins significatif pour une telle ligne. Les navettes mises en place en cas de panne

ou de travaux sur les lignes de métro possèdent elles aussi un ratio bien plus faible que la moyenne des bus du réseau local. Cependant l'achalandage de ces navettes est parfois tellement important que le chauffeur empêche la validation pour sauver du temps.

Pour conclure cette petite étude sur les différentes lignes de bus, nous pouvons dire que l'évasion tarifaire est un phénomène relativement isolé avec une variabilité importante entre les différentes lignes et/ou types de lignes. Il est donc plus difficile d'appréhender l'évasion tarifaire à ce niveau-ci.

4.1.4 Titres de transport

Les données de carte à puce nous renseignent sur le type de titre utilisé par les utilisateurs, c'est-à-dire le support (carte à puce OPUS ou papier), ainsi que les différents abonnements ou titres (abonnement mensuel, pass 24h, pass week-end, etc ...). Nous avons donc cherché à savoir si l'on pouvait caractériser l'évasion tarifaire à travers les différents types de titre de transport. Autrement dit, on a cherché à regarder si le fait d'avoir plus de validations par OPUS ou par ticket papier dans un bus pouvait influencer le ratio Validations/Montants.

Avant de présenter les résultats de cette étude, il faut tout de même noter que sur le réseau de la STM environ 90% des validations dans les bus se font par la carte OPUS comme nous le présente la **Figure 4-5**.

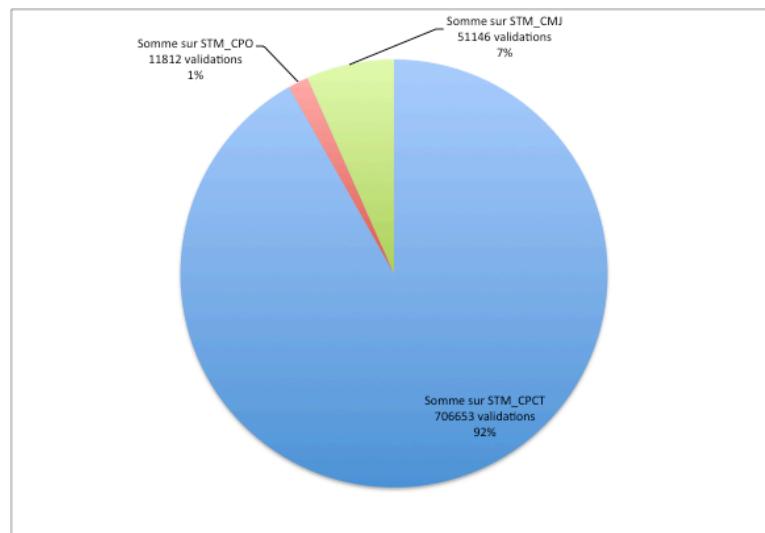


Figure 4-5 : Pourcentage validations CPCT, CPO, et CMJ pour les 5 jours de la semaine d'avril (7 au 11 avril 2014)

Concernant la fin de semaine, des passes week-end sont disponibles sur le réseau de la STM via un carte occasionnelle (CPO), le pourcentage de validation par CPO augmente donc en fin de semaine, mais il y a toujours 90% de validations par CPCT comme nous le montre la **Figure 4-6**.

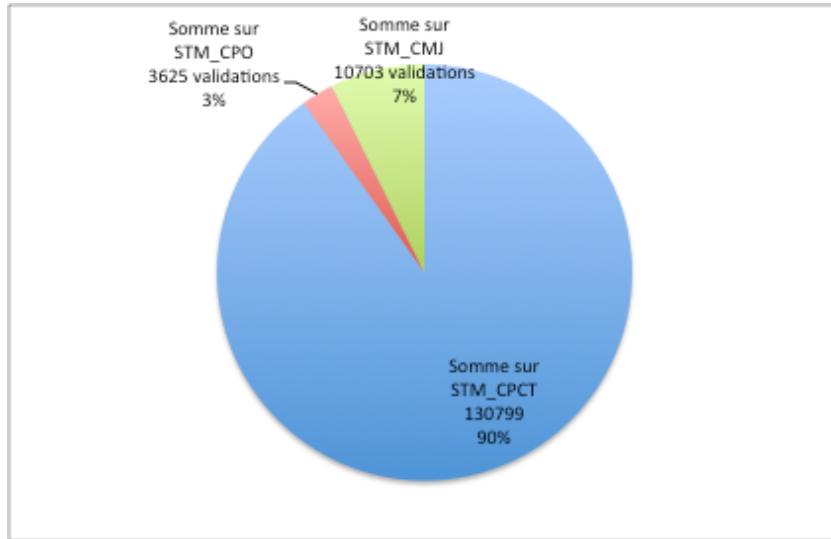


Figure 4-6 : Pourcentage de validations par CPCT, CPO et CMJ pour la fin de semaine d'avril 2014 (12 et 13 avril 2014)

En annexe A nous pouvons trouver un tableau récapitulatif des différents supports et des différents abonnements. Plus de 120 titres différents existent sur le réseau de Montréal. Nous les avons décrits suivant 2 critères : l'avantage au niveau du prix (régulier ou réduit) et le type de support (carte à puce commune CPCT, carte à puce occasionnelle CPO et carte magnétique CMJ). Concernant le type de population (enfant, étudiant, adulte, aîné) il est impossible de pouvoir le déterminer à partir des données de carte à puce. En effet, les étudiants et les aînés peuvent avoir accès à un tarif réduit sur la carte OPUS, mais l'appellation est la même pour ces deux types de population (STM-ABONNEMENT RED). Il n'est donc pas possible de connaître les différents types de population à travers les données OPUS.

Cependant, nous avons pu tout de même regarder si le pourcentage de validation par carte OPUS dans un bus influait sur le ratio Validations/Montants. Pour cela nous avons tracé sur la **Figure 4-7** ci après, le ratio Validations/Montants en fonction du ratio Somme_Validations_CPCT/Somme_Validations_Total pour l'ensemble des lignes de bus sur la semaine d'avril. On rappelle que CPCT correspond en fait à la carte à puce rechargeable c'est-à-dire la carte OPUS à Montréal.

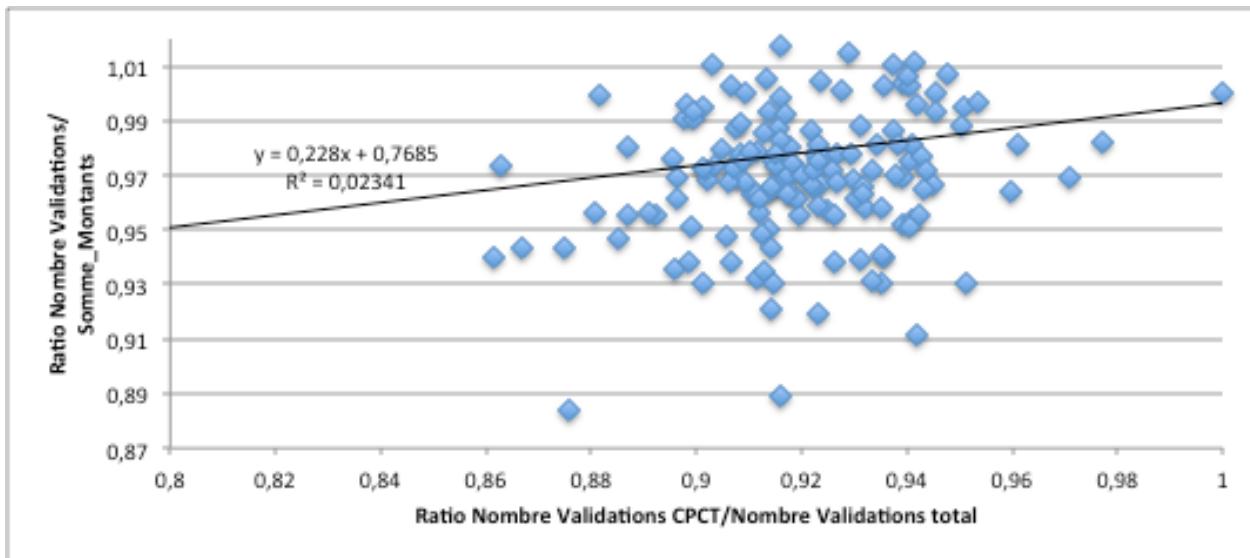


Figure 4-7 : Ratio Validations/Montants en fonction du ratio Validations CPCT/Validations_Totales pour toutes les lignes de bus de la semaine du 7 au 13 avril 2014

On observe assez bien qu'il n'y a pas une réelle corrélation sous la forme d'une droite entre le ratio de validation OPUS/Papier et le ratio Validations/Montants. De plus le coefficient de détermination R^2 proche de zéro nous le confirme. Ce paramètre n'est donc pas très pertinent pour l'étude de l'évasion tarifaire.

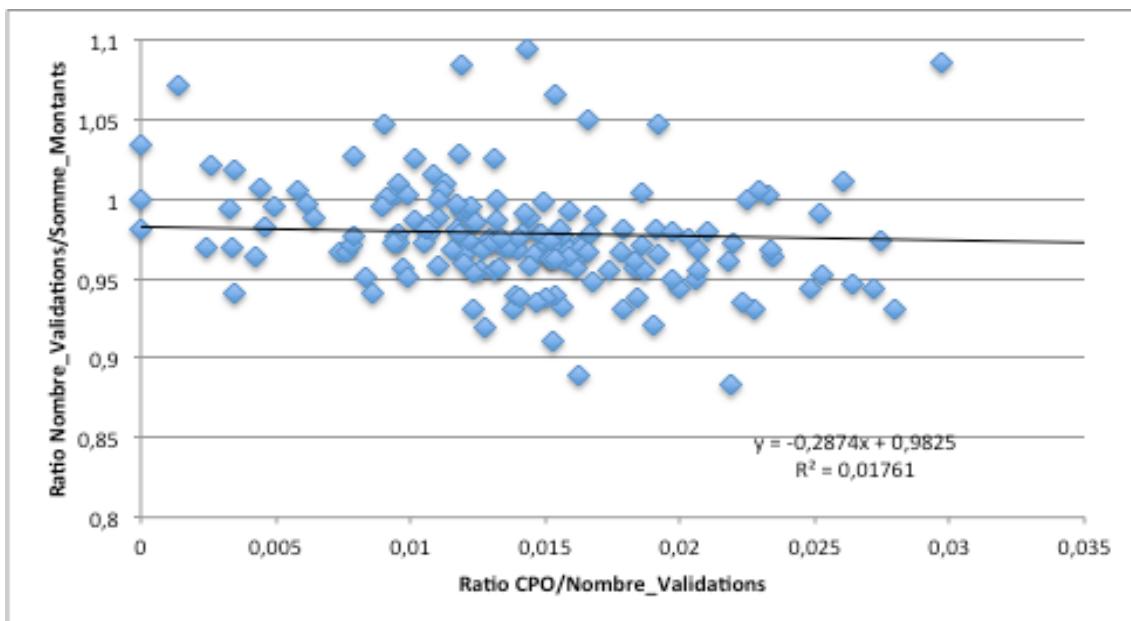


Figure 4-8 : Ratio Validations_Montants en fonction du ratio Validations_CPO/Validations_Totales (semaine du 7 au 13 avril 2014)

A des fins d'exploration, nous avons aussi jugé utile de confronter le pourcentage de validation par CPO (Carte Puce Occasionnelle) au ratio Validations_Montants.

Une fois de plus, on n'observe pas une réelle corrélation entre ces deux paramètres, cependant on peut remarquer une légère tendance à la baisse du Ratio Validations_Montants lorsque le pourcentage de validations par CPO augmente. Cela reste juste une interprétation de cette figure, et nous ne pouvons pas en tirer de réelles conclusions. De plus, une valeur a été supprimée à des fins de lisibilité, il s'agit du bus 747 qui a un ratio de validation par CPO bien supérieur à la moyenne (0,24).

Pour conclure, on peut dire qu'il est difficile de caractériser une éventuelle tendance à l'évasion tarifaire à travers les titres.

4.1.5 Quartiers et RTA

Pour avoir une meilleure vision de l'évasion tarifaire sur le réseau de l'île de Montréal, durant les phases d'analyse, nous avons jugé utile de représenter ce phénomène sur une carte (**Figure 4-9**). Les données ont été enrichies avec les quartiers et RTA (régions de tri d'acheminement) de l'île. Nous avons alors pu représenter les ratios Validations/Montants grâce au logiciel QGIS.

Les RTA sont colorées du rouge au bleu suivant leur ratio, plus le quartier tend vers le rouge, plus son ratio est faible. Le nombre à l'intérieur de ces zones correspond à la somme des montants pour la semaine. Nous remarquons alors que les zones du centre-ville de Montréal ont un ratio relativement plus faible que le sud-ouest de Montréal par exemple. Cela nous donne une vision plus globale de l'évasion tarifaire sur l'île. Il y a cependant des zones pour lesquelles le nombre de montants n'est pas assez significatif pour déclarer que c'est une zone à risque. C'est le cas de l'île Bizard par exemple, à l'ouest de Montréal. La deuxième zone la plus à l'ouest correspond au RTA comprenant l'aéroport Montréal Trudeau. On peut alors faire le lien avec la partie précédente sur les types de bus et les navettes de l'aéroport qui ont pour la majorité un ratio plus faible que la moyenne des autres bus.

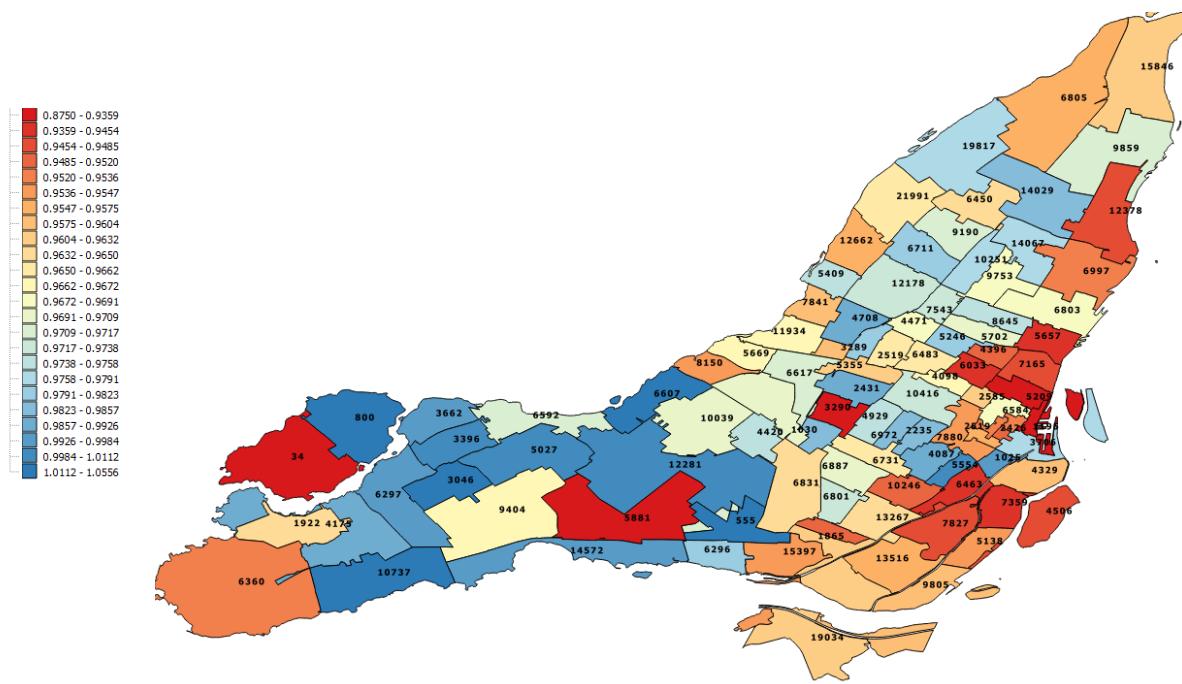


Figure 4-9 : Analyse par RTA des ratios sur l'île de Montréal

Ci-dessous (**Figure 4-10**), l'étude avec les quartiers propose une vision à plus petite échelle étant donné que les quartiers sont plus grands que les RTA.

Ces études cartographiques permettent d'avoir une meilleure vision du phénomène d'évasion tarifaire. Cependant elles ne permettent pas une réelle aide en termes d'allocation d'inspection, car elle n'est pas assez précise. C'est pourquoi nous allons nous concentrer désormais sur des zones bien plus ciblées, en l'occurrence, les zones autour des métros et les arrêts de bus.

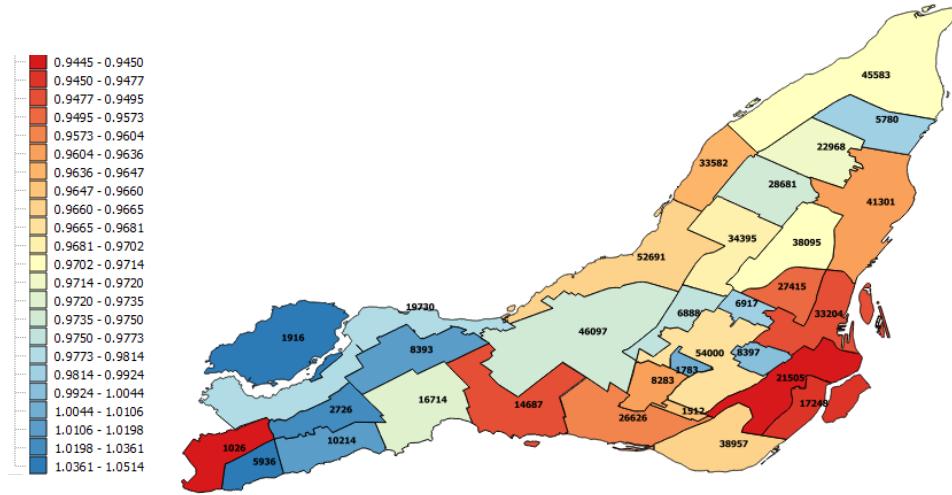


Figure 4-10 : Analyse par quartier des ratios sur l'île de Montréal

4.1.6 Autour des métros

Dans la plupart des villes se composant d'un réseau de bus et d'un réseau de métro, de nombreux arrêts de bus se retrouvent alors à proximité des stations de métro pour faciliter les correspondances. L'affluence dans les bus autour de ces stations de métro est donc souvent importante. Nous avons alors voulu vérifier si cette affluence autour des stations de métro pouvait avoir un impact sur l'évasion tarifaire. Pour réaliser cela, nous avons utilisé QGIS avec lequel des entités de 75 m de rayon ont été tracées autour des 68 stations de métro de Montréal. Ensuite grâce à ces zones tampons autour des métros, on a donc pu associer le nom de la station aux arrêts de bus à proximité. Ci-dessous, un tableau présentant les ratios pour les arrêts autour des stations de métro et leurs nombres de montants par semaine. Ce tableau ne présente que quelques stations qui ont un ratio relativement faible.

Tableau 4-3 : Analyse du ratio validations/montants par arrêt de bus à proximité des métros

METRO	RATIO	SOMME_MONTANTS
Station Place-d'Armes	0,7973	148
Station Acadie	0,8523	325
Station Place-Saint-Henri	0,8756	426
Station McGill	0,8976	1660
Station Plamondon	0,9112	4652
Station Laurier	0,9126	5663
Station Verdun	0,9152	401
Station Beaubien	0,9163	4765
Station De l'Eglise	0,9195	683
Station Du College	0,9195	870
Station Sauve	0,9209	4502
Station LaSalle	0,9263	1018
Station Joliette	0,9283	2761
Station Outremont	0,9299	1312
Station Sherbrooke	0,9345	2611
Station Angrignon	0,9366	9037
Station Cadillac	0,9410	4067
Station Rosemont	0,9425	3441
Station Charlevoix	0,9431	1266
Station Place-des-Arts	0,9455	2916

Si on analyse ces résultats qui correspondent aux 5 jours de la semaine, on remarque rapidement que les ratios sont plus faibles que pour les lignes complètes. Sur l'ensemble des arrêts à proximité des stations le ratio Validations/Montants est de 0,9452 comparativement au 0,9714 sur l'ensemble du réseau pour les 5 jours de la semaine. Nous pouvons donc en déduire que les stations proches des métros sont globalement plus à risque que les autres stations. Nous allons pouvoir vérifier cela en étudiant les ratios de l'ensemble des arrêts de bus de l'île de Montréal.

4.1.7 Arrêts de bus

La finesse du traitement des données nous a permis de lier les tables de données SCAD et OPUS jusqu'aux arrêts. À des fins d'aide à la décision pour les inspecteurs du réseau, l'étude de l'évasion tarifaire par arrêt semble être la plus cohérente. Elle pourrait alors permettre de cibler potentiellement les arrêts les plus à risques de manière à y déployer des inspecteurs. Comme nous l'avons vu auparavant, les ratios pouvant être moins significatifs la fin de semaine, nous

étudierons une nouvelle fois le ratio sur les arrêts durant les 5 jours de la semaine. Voici ci-dessous, un tableau de quelques arrêts ayant un risque potentiel d'évasion tarifaire plus élevé.

Tableau 4-4 : Analyse du ratio Validations/Montants par arrêts de bus

Arrêts	Ratio	Somme_Montants
Papineau / Rachel	0,8283	530
Station Jean-Talon Sud / De Chateaubriand	0,8319	571
Saint-Laurent / Saint-Zotique	0,8355	608
Beaubien / De Normanville	0,8586	587
Honore-Beaugrand / Hochelaga	0,8659	798
Station Beaubien / de Saint-Vallier	0,8701	585
Station Sauve Sud (Sauve / Saint-Denis)	0,8750	816
Des Ormeaux / Georges	0,8752	689
Henri-Bourassa / du Bois-de-Boulogne	0,8836	2234
Jean-Talon / Valdombre	0,8838	809
Station Square-Victoria OACI / Viger	0,8884	905
Station Atwater (Tupper / Atwater)	0,8891	911
Gouin / Alexander	0,8905	776
Beaconsfield / Saint-Charles	0,8920	713
Saint-Michel / Rosemont	0,8925	893
De Serres / Grenet	0,8947	731
Kildare / Cavendish	0,8965	831
Graham / Dunkirk	0,9007	836
Jean-Talon / 24e Avenue	0,9017	1577
Pierrefonds / Saint-Jean	0,9022	501
Leger / Rolland	0,9022	542
Station McGill (Union / De Maisonneuve)	0,9031	1630
Henri-Bourassa / Seguin	0,9033	703
Garon / Henri-Bourassa	0,9045	1236
Station Plamondon Sud (Van Horne / Victoria)	0,9048	4235

Il faut noter que pour extraire ce tableau, nous avons choisi volontairement de mettre un seuil de montants minimum de manière à rendre les résultats cohérents. Un échantillon de 100 personnes par jour au minimum a été fixé. Si l'on ne fixe pas un minimum, on peut alors analyser des résultats provenant de seulement 1 ou 2 trajets de bus sur cette ligne sur la semaine, ce qui n'est pas pertinent.

En analysant de plus près ces résultats, nous observons que quelques arrêts proches des stations de métro (Ex : Station Jean-Talon Sud/De Chateaubriand ou Station Beaubien / de Saint-Vallier) se retrouvent parmi les autres dans les stations aux plus faibles ratios. Cela confirme donc notre analyse faite précédemment en 4.1.5. Cependant il y a bien d'autres arrêts qui ne sont pas forcément proches des stations de métro et qui ont un ratio faible. Il serait alors intéressant de comparer ces résultats avec la deuxième semaine disponible pour confirmer que certains arrêts peuvent être réellement catégorisées comme arrêt à risque d'évasion tarifaire élevé.

Pour illustrer le fait qu'un nombre pertinent de montants par semaine doit être fixé avant d'étudier chaque arrêt, j'ai représenté sur une carte (**Figure 4-11**) tous les arrêts ayant plus de 100 montants sur la semaine.

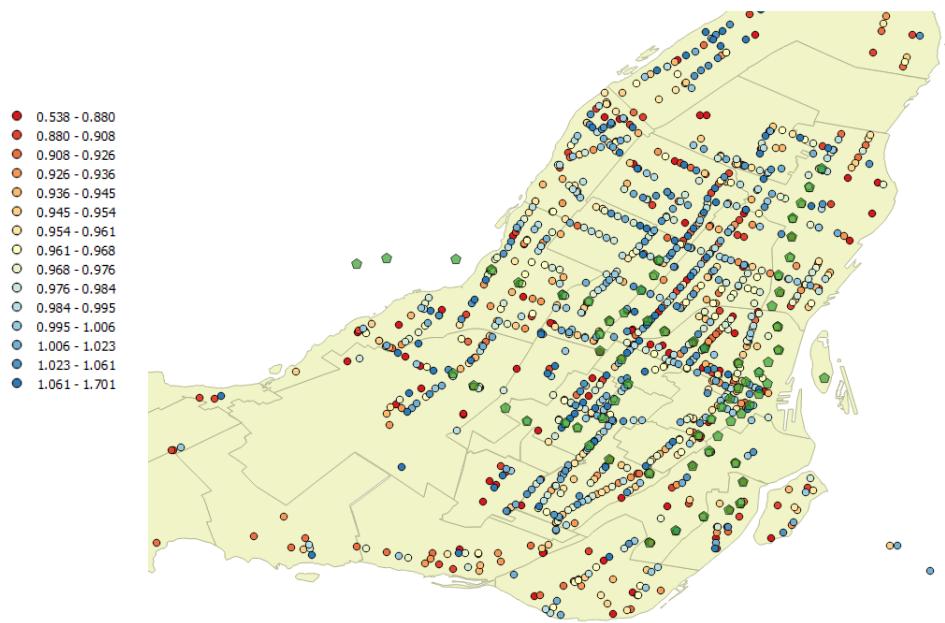


Figure 4-11 : Arrêts ayant plus de 100 montants gradués par ratio

On observe que la variabilité du ratio est plus grande lorsque l'on diminue le seuil (0,538-1,701) de montants. De plus on observe aussi grâce aux marqueurs verts qui représentent les stations de métros, que les arrêts de bus ayant un faible ratio ne se trouvent pas forcément proche d'une station.

Pour conclure cette analyse quantitative du ratio montants/transactions sur le réseau, nous pouvons dire que l'on peut caractériser ce phénomène à travers de nombreux paramètres comme l'heure, le quartier ou encore l'arrêt de bus. Toutefois, même si les données d'une semaine sont

relativement conséquentes, et explicites pour certains paramètres, une analyse comparative avec une deuxième semaine ne peut que donner de la véracité à cette première analyse. Ce sera l'objet de la prochaine partie.

4.2 Analyses comparées entre les deux semaines

Dans cette deuxième partie de l'analyse des résultats, nous allons comparer les résultats issus de la deuxième semaine d'avril (du 7 au 13 avril 2014) avec la deuxième semaine d'octobre (du 6 au 12 octobre 2014). Cette analyse comparative va alors permettre de confirmer ou éventuellement de réfuter certaines avancées émises lors de la première partie.

4.2.1 Jour de la semaine

Dans la première partie des analyses, nous avions vu que le ratio Validations/Montants était bien plus faible la fin de semaine, mais que cela pouvait s'expliquer par le fait que les enfants n'est pas besoin de valider sur ces deux jours. Toutefois, nous avions aussi remarqué que le ratio semblait augmenter en fonction des jours de la semaine. Cela signifierait que le début de semaine pourrait être plus sujet à l'évasion tarifaire que le reste de la semaine. Nous allons pouvoir vérifier si cette hypothèse est plausible avec la deuxième semaine de données (**Figure 4-12**).

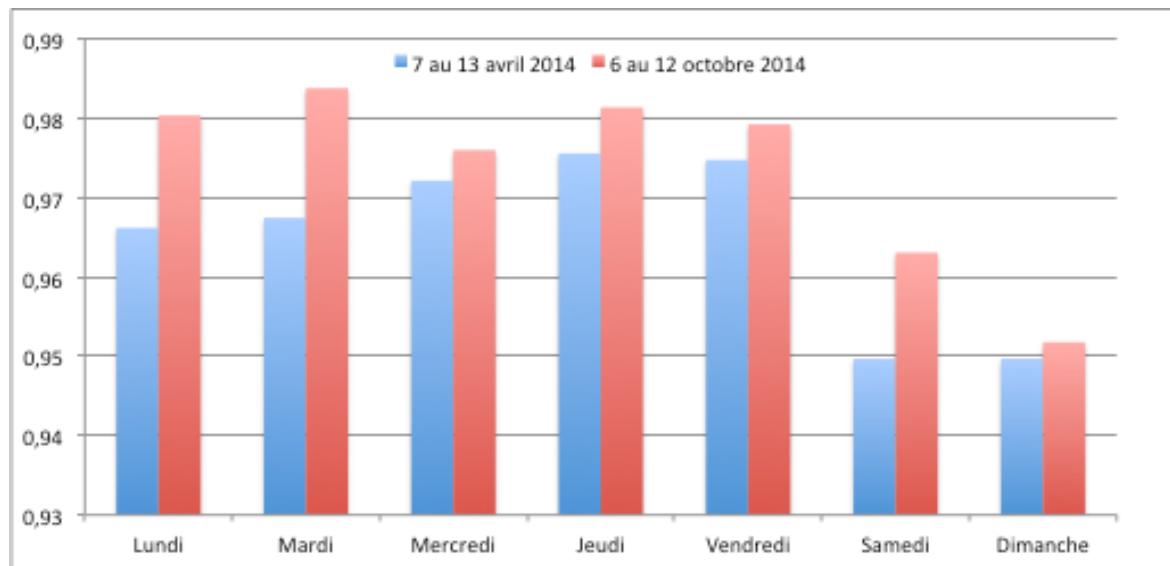


Figure 4-12 : Ratio comparatif Validations/Montants pour les deux semaines

En analysant cette figure, on constate que pour la semaine d'octobre, le ratio est aussi plus faible en fin de semaine. Cependant à la différence de la première semaine, le ratio n'est pas plus faible

en début de semaine. On peut dire qu'il n'y a pas vraiment de jour dans la semaine où le ratio est beaucoup plus faible mis à part la fin de semaine. De plus, il faut noter que sur l'ensemble de la semaine d'octobre le ratio moyen est de 0,976 comparativement au 0,967 de la semaine d'avril.

4.2.2 Heure de la journée

Comme nous l'avons vu en 4.1.1, l'évasion tarifaire semble être soumise au phénomène d'heures creuses et d'heures de pointe. Il s'avérait que le ratio Validations/Montants était plus faible lors des périodes creuses de la journée. Autrement dit, l'évasion tarifaire pour les 5 jours semaine du 7 au 11 avril 2014 était plus faible lors des heures de pointe que lors des heures creuses. Nous pouvons alors peut-être appuyer cela avec le graphique suivant (**Figure 4-13**).

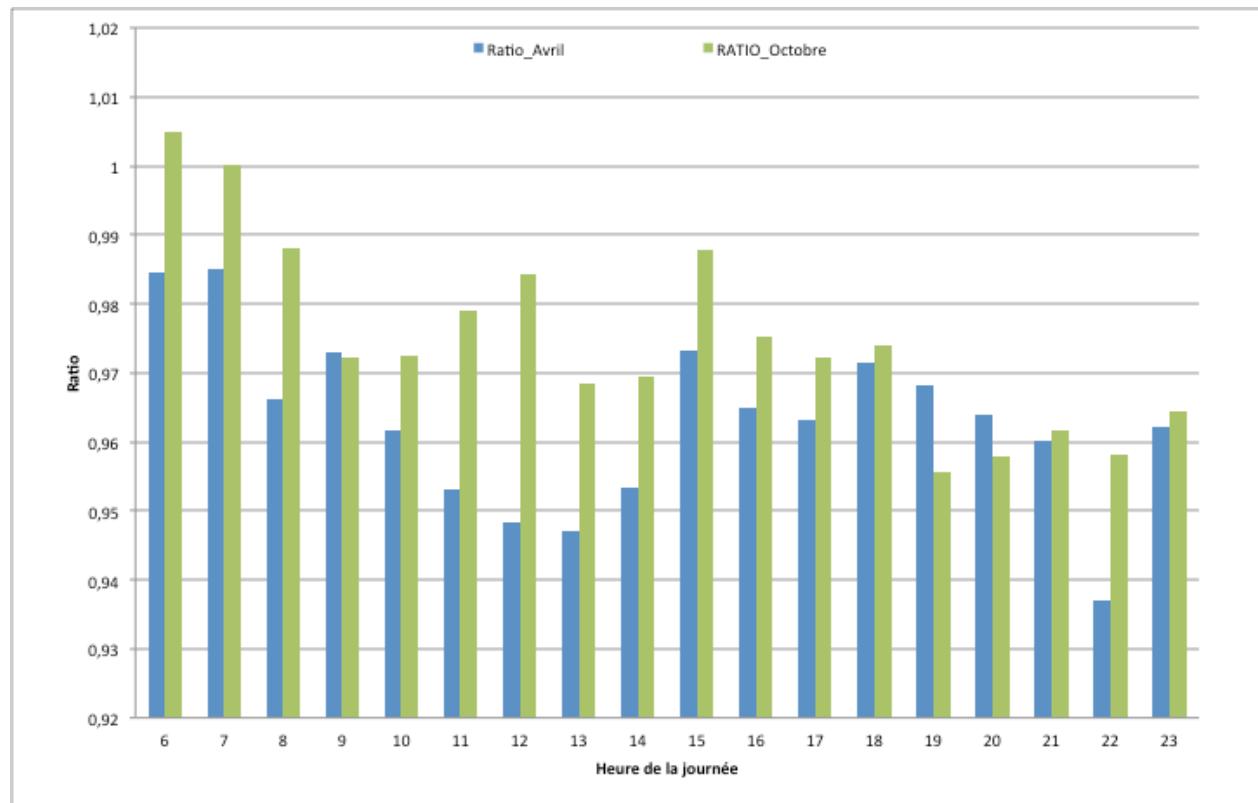


Figure 4-13 : Ratio Validations/Montants en fonction des heures de la journée pour les deux semaines

Ce graphique qui représente le ratio Validations/Montants pour les 5 premiers jours des semaines d'avril et d'octobre (7 au 11 avril 2014 et 6 au 10 octobre 2014) appuie notre première conclusion

vis-à-vis de l'évasion tarifaire en heures creuses. On retrouve des ratios plus faibles aux heures creuses pour la semaine d'octobre comme par exemple le soir après 19h ou entre 13h et 14h. Nous pouvons donc confirmer l'hypothèse émise en 4.1.1.

Comme nous avions présenté dans la revue de littérature, des auteurs se sont déjà penchés sur une meilleure allocation des inspections en fonction des heures de la journée. Des résultats comme ceux-là, peuvent permettre de faciliter et d'améliorer la planification des inspections pour cibler les heures les plus sensibles à l'évasion tarifaire.

4.2.3 Lignes de bus

Lors de la première analyse concernant les lignes de bus, nous avons vu que certaines lignes étaient potentiellement plus soumises à l'évasion tarifaire que d'autres. Il serait alors intéressant de savoir s'il existe une forme de récurrence à long terme dans ce phénomène. Pour cela nous allons comparer quelques lignes à risque de la semaine d'avril avec celle d'octobre.

Tableau 4-5 : Ratio Validations/Montants par ligne de bus

Du 7 au 13 avril 2014

NUMERO_LIGNE	RATIO	MONTANTS	INDICE
40	0,8839	155	1
877	0,8843	2023	5
39	0,8892	415	2
11	0,9013	1490	4
747	0,9055	3407	7
216	0,9111	360	2
185	0,9150	2047	5
102	0,9193	2727	6
355	0,9210	671	3
365	0,9211	114	1
195	0,9249	4995	7
30	0,9302	2005	5

Du 6 au 12 octobre 2014

NUMERO_LIGNE	RATIO	MONTANTS	INDICE
747	0,8491	3969	7
19	0,8716	257	1
31	0,9032	124	1
420	0,9033	631	3
411	0,9142	664	3
53	0,9145	433	2
188	0,9225	2155	5
191	0,9226	5659	8
102	0,9266	1607	4
777	0,9272	2197	5
487	0,9327	609	2
11	0,9338	4216	7

Il faut noter qu'à des fins d'explication du raisonnement d'analyse nous allons juste comparer les 12 lignes aux plus faibles ratios. Nous observons à travers ces tableaux que les lignes ayant un faible ratio en avril n'ont pas forcément un ratio aussi faible en octobre. C'est le cas par exemple de la ligne 40, qui pour la période d'avril présente un ratio de 0,8839 alors qu'elle a un ratio de 0,9594 en octobre. Cependant, comme nous l'avons déjà vu auparavant pour cette ligne le nombre de montants n'est que de 155 personnes sur la semaine d'avril et 370 sur celle d'octobre. L'échantillon paraît donc un peu faible pour tirer des conclusions sur cette ligne de bus. Par contre, pour la ligne 11, celle-ci présente un ratio faible en avril et aussi en octobre. De plus, son

nombre de passagers n'étant pas négligeable (1490 et 4216 personnes), il serait plus « juste » de catégoriser la ligne 11 comme une ligne à risque plutôt que la 40. C'est dans ce mode de raisonnement la que nous pourrions analyser l'ensemble des 172 lignes du réseau. Les indices du nombre de montants peuvent permettre de donner ou non une meilleure pertinence aux ratios. Cependant, on peut alors se demander si ces lignes ayant des faibles ratios sont seulement influencées par un seul arrêt ou par plusieurs. Pour représenter cela, la **Figure 4-14** illustre les arrêts de la ligne 11 pour l'ensemble de la semaine d'avril 2014.

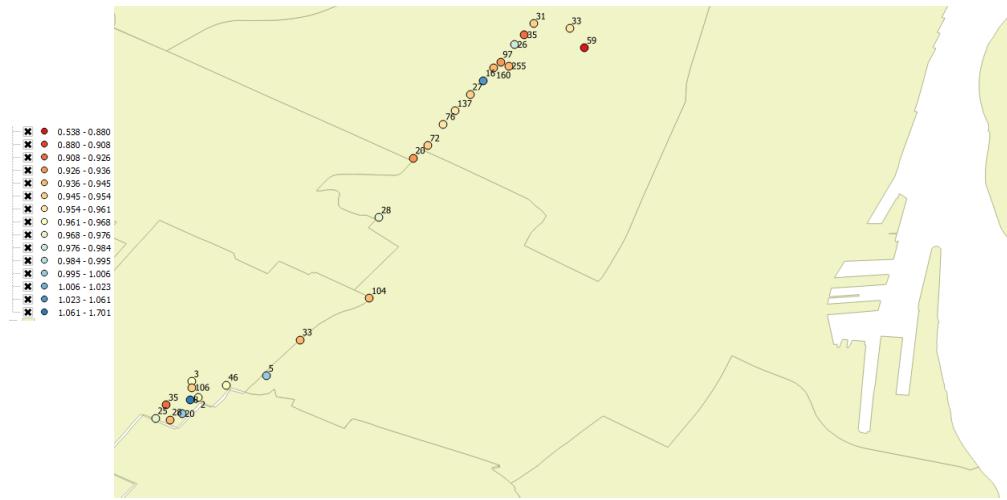


Figure 4-14 : Arrêts de la ligne 11 pour la semaine d'avril 2014

Sur la figure ci dessus, on observe que certains arrêts ont un ratio très faible (0,538) mais que globalement la plupart des arrêts ont un ratio inférieur à 0,95. Pour cette ligne, il semblerait donc que le ratio faible ne soit pas seulement dû qu'à un arrêt.

4.2.4 Arrêts de bus

Dans la partie 4.1.6, nous avons analysé les arrêts à risque pour la semaine entière d'avril 2014. Cependant, il nous a paru intéressant de savoir si à ce niveau-là il existe aussi une récurrence de l'évasion tarifaire sur le court terme et sur le long terme. Concernant le court terme, nous avons comparé les jours de la semaine entre eux pour une même semaine. Quant au long terme, il s'agira d'observer ou non des arrêts ayant un faible ratio aussi bien en avril qu'en octobre.

Pour ce qui est de la comparaison à court terme des arrêts de bus, nous nous sommes confrontés à un problème. Il s'avère que comme toutes les lignes ne sont pas équipées du système de comptage à bord, les lignes de faible affluence ne sont pas forcément parcourues par un bus

équipé par jour. C'est-à-dire que pour certains arrêts la comparaison avec les autres jours de la semaine est impossible. Elle l'est seulement pour les lignes à plus forte affluence qui présentent plus de bus par jour, donc un pourcentage de chance d'avoir un bus équipé plus élevé. Comme nous avons voulu tout de même faire cette analyse comparative entre les jours de la semaine, nous avons dû choisir des arrêts ayant des affluences supérieures à la moyenne. En l'occurrence, les arrêts à forte affluence sont souvent près des métros, notre analyse à court terme sur ces arrêts ne prend donc pas en compte tous les arrêts, il est important de le noter avant de faire des conclusions.

En préanalysant les données, nous avons établir un seuil d'environ 1500 passagers sur les 5 jours de la semaine pour un arrêt, celui-ci était parcouru au minimum une dizaine de fois par jour et au moins 4 jours par semaine. Si nous sélectionnons des arrêts avec des valeurs en dessous de ce seuil, certains auront alors par exemple 200 montants le lundi, mais seulement 30 le mardi, car moins de bus équipés sont passés le mardi. Nous avons donc essayé de trouver des arrêts ayant un nombre de montants relativement constant et non négligeable sur la semaine.

Le but étant d'observer une récurrence ou non pour ce phénomène d'évasion tarifaire, ce n'est pas de présenter ici les arrêts les plus à risque de Montréal. Pour mettre en lumière cela, nous avons choisi de regarder les 5 arrêts ayant le plus faible ratio sur les 5 jours de la semaine avec plus de 1500 montants et d'observer sur ce ratio était aussi faible sur l'ensemble des jours ou si c'était qu'un seul jour. Voici le tableau récapitulatif ci-dessous :

Tableau 4-6 : Ratio Validations/Montants sur les jours de la semaine d'avril pour 5 arrêts

RATIO (Somme_Montants)	Henri-Bourassa / du Bois-de-Boulogne	Jean-Talon / 24e Avenue	Station McGill (Union / De Maisonneuve)	Station Plamondon Sud (Van Horne / Victoria)	Station Laurier Sud (Berri / Gilford)
Lundi	0,9491 (118)	0,8016 (620)	0,8577 (436)	0,8762 (614)	0,8897 (499)
Mardi	0,9123 (479)	0,9888 (180)	X	0,8945 (929)	0,8927 (261)
Mercredi	0,8726 (573)	0,9668 (241)	0,8965 (392)	0,8498 (766)	0,9526 (570)
Jeudi	0,8707 (851)	0,9520 (125)	0,9872 (508)	0,9240 (1383)	0,9359 (703)
Vendredi	0,8638 (213)	0,9610 (411)	0,8707 (294)	0,9834 (543)	0,8652 (920)
Lundi au Vendredi	0,8836 (2234)	0,9017 (1577)	0,9030 (1630)	0,9048 (4235)	0,9055 (2953)

À travers ce tableau, on peut déjà observer que ces arrêts ont un ratio d'environ 0,90 sur les 5 jours, ce qui est beaucoup plus faible que l'ensemble des arrêts. Cependant si l'on regarde l'arrêt Jean-Talon/24^e Avenue, il présente un très faible ratio le lundi (0,8016) et un ratio proche de la moyenne les autres jours. On peut alors penser que ce lundi fut une exception et que ce n'est pas forcément un arrêt à risque, mais qu'un évènement inconnu a fait chuter le ratio. À l'inverse, l'arrêt Henri-Bourassa/Bois de Boulogne présente lui un ratio très faible 4 jours/5, tout comme Station Plamondon (Van Horne/Victoria). On peut donc faire l'hypothèse que l'évasion tarifaire a bien une composante récurrente, mise en évidence ici.

L'autre étude à ce sujet a été de comparer les arrêts ayant un faible ratio sur la semaine entre avril et octobre pour voir si on y retrouve des arrêts sensibles. Ci-dessous, voici les 15 arrêts ayant le plus faible ratio et plus de 1500 passagers pour chacune des semaines.

Tableau 4-7 : 15 arrêts ayant les plus faibles ratios pour les 2 semaines

ARRETS (du 7 au 11 avril)	RATIO	MONTANTS	ARRETS (du 6 au 10 octobre)	RATIO	MONTANTS
Henri-Bourassa / du Bois-de-Boulogne	0,8836	2234	Lakeshore / Sainte-Anne	0,8955	1665
Jean-Talon / 24e Avenue	0,9017	1577	/ Terminus Macdonald	0,8988	3400
Station McGill (Union / De Maisonneuve)	0,9031	1630	Gilford / Station Laurier Sud (Berri / Gilford)	0,9115	1605
Station Plamondon Sud (Van Horne / Victoria)	0,9048	4235	Maurice-Duplessis / Armand-Bombardier	0,9162	1718
Station Laurier Sud (Berri / Gilford)	0,9055	2953	Crémazie / Station Crémazie Nord (Crémazie / Berri)	0,9163	1673
Station Lionel-Groulx / Atwater	0,9102	1770	Atwater / Station Lionel-Groulx / Atwater	0,9235	2221
Terminus Fairview	0,9102	5468	Rosemont / Station Rosemont / Terminus	0,9241	4322
Station Villa-Maria / Terminus	0,9178	4321	privé Honoré-Beaugrand sud / Station Honoré-Beaugrand	0,9295	1546
Station Sauve Nord (Sauve / Berri)	0,9196	2153	Queen-Mary / Station Snowdon (Queen-Mary / Circle)	0,9296	2799
Station Sauve Sud (Sauve / Berri)	0,9239	2366	Villa-Maria / Station Villa-Maria / Terminus	0,9305	2978
Terminus de la gare Dorval	0,9244	4618	Jarry / Station Jarry (Jarry / Berri)	0,9316	2221
Station Saint-Michel Est	0,9254	2239	Henri-Bourassa / du Bois-de-Boulogne	0,9336	1778
Station Cremazie Nord (Cremazie / Berri)	0,9269	1546	Cartier / Station Papineau (Cartier / Sainte-Catherine)	0,9347	1516
Station Sherbrooke Est (Berri / Cherrier)	0,9269	2190	Jean-Talon / 24e Avenue	0,9377	3308
Station Snowdon (Queen-Mary / Circle)	0,9273	1788	Sauvé / Station Sauvé Nord (Sauvé / Berri)	0,9384	1639
			Rosemont / Station Rosemont (Rosemont / De Saint-Vallier)		

Il faut noter qu'encore une fois nous avons dû fixer un seuil minimum de montants et que nous ne comparons donc pas tous les arrêts, mais seulement ceux qui ont 1500 passagers sur les 5 jours de la semaine. Cela représente environ une centaine d'arrêts de bus sur l'île. Cependant, si l'on regarde ces 15 arrêts, on remarque très vite, grâce au code de couleur que sur les 15 arrêts, 8 se retrouvent à la fois en avril et en octobre. De plus tous ces arrêts ont des ratios relativement faibles comparés à l'ensemble du réseau. Cette petite analyse nous montre encore une fois que l'évasion tarifaire peut être potentiellement mise en évidence ici de cette façon et qu'une forme de récurrence semble réellement exister dans ce phénomène.

4.3 Synthèse

Pour conclure cette partie sur l'analyse des résultats, nous pouvons dire que le ratio Validations/Montants nous laisse suggérer que l'évasion tarifaire peut réellement être caractérisable à travers les données de cartes à puce et de comptage à bord. Nous avons étudié plusieurs paramètres, il s'avère que certains sont plus pertinents que d'autres pour mettre en évidence ce phénomène. Cependant, comme nous l'avons mentionné à plusieurs reprises, les bus ne sont pas tous équipés du système de comptage à bord ce qui rend l'analyse et l'interprétation un peu plus complexe. L'analyse comparée nous a permis de prouver que des arrêts ou lignes à risque élevé d'évasion tarifaire pouvaient être caractérisés en comparant les données de plusieurs semaines ou de plusieurs jours. En outre, on peut aussi dire que les lignes et les arrêts de bus semblent être la source la plus facilement utilisable pour proposer un outil d'aide à la décision pour le personnel d'inspection, c'est dans cette optique-là que nous avons continué le projet.

CHAPITRE 5 OUTIL DYNAMIQUE D'AIDE A LA DÉCISION

Dans ce chapitre, nous allons présenter l'outil dynamique que nous avons développé dans le but de procurer une aide à la décision pour le personnel d'inspection. Nous allons évidemment nous baser sur les résultats issus de l'analyse et plus particulièrement sur les arrêts de bus. Cet outil dynamique a pour objectif d'automatiser le traitement des données et de représenter sur une carte les arrêts dits à risque de manière à aider le personnel d'inspection. Comme nous l'avons vu dans la revue de littérature, il faut rappeler que les faibles ratios Validations/Montants ne traduisent pas que de l'évasion tarifaire volontaire. Un faible ratio signifie en tout cas qu'un certain nombre de personnes sont dans l'incapacité de valider leurs titres de transport que ce soit de manière volontaire ou non. L'outil aura donc pour objectif de mettre en évidence ces arrêts pour que le personnel d'inspection puisse éventuellement examiner la cause de ce mauvais fonctionnement.

5.1 Présentation de l'outil

5.1.1 Objectifs et fonctionnement de l'outil

Nous parlons volontairement d'outil dynamique, car le but étant de s'appuyer sur le traitement des données, mais en l'automatisant pour faciliter son utilisation. Une fois les données traitées, une extraction des arrêts à risque dans un fichier CSV (toujours à partir de R) permettra déjà une consultation relativement simple. Le processus de traitement sera donc exactement le même que celui détaillé dans la partie méthodologie, cependant il comporte une partie en plus permettant d'extraire un fichier comportant les arrêts à risque. Nous discuterons des critères de ces arrêts à risque par la suite. Toutefois, le but étant une bonne visualisation de ces arrêts, l'envoi de ce fichier sur un serveur et le chargement dans une base de données en ligne nous permettra de représenter ces arrêts sur une carte de type Google Maps.

La **Figure 5-1** explique plus en détail les étapes de cet outil tandis que l'annexe G résume les langages utilisés pour réaliser cet outil en fonction des différentes étapes.

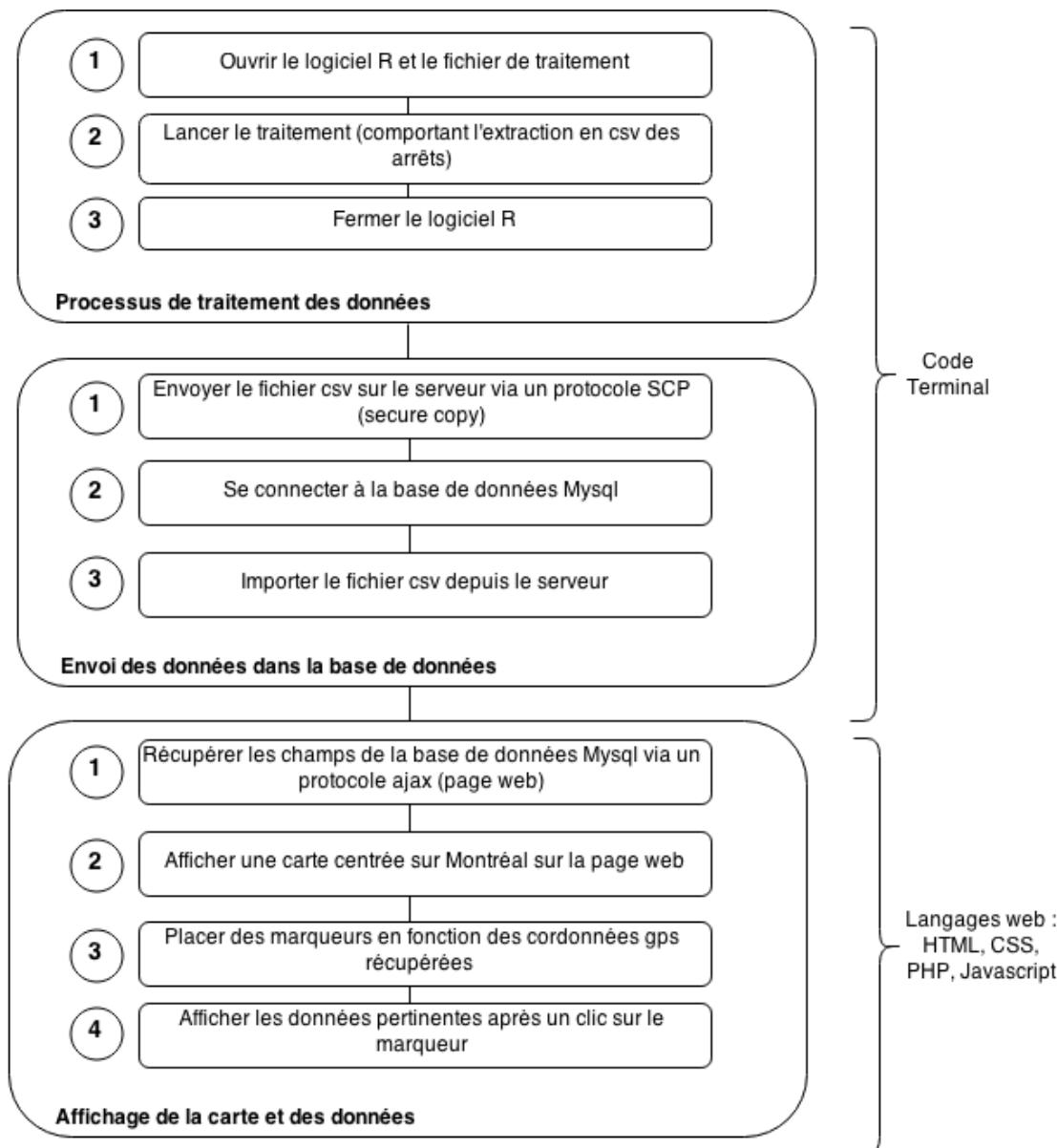


Figure 5-1 : Explication des étapes du fonctionnement de l'outil dynamique

5.1.2 Affichage des données

Les données transmises comportent alors plusieurs champs permettant de récupérer toutes les informations pour l'affichage sur la carte. Le **Tableau 5-1** présente ces différents champs :

Tableau 5-1 : Forme des données extraites pour une visualisation sur carte

Données_Sortie_Traitement	Exemple	Explications
Date_Sortie	13/04/2014	Date des données qui correspond toujours à la veille
Heure_Depart_Arret	13:34:45	Heure de fin du traitement (utile seulement à des fins de programmation)
Lat_GPS	45,55452	Latitude
Long_GPS	-73,65743	Longitude
Desc_Arret_Client	Saint-Joseph / De Bullion	Nom de l'arrêt
Type	Données	Type de source (données ou observations terrains)
Ratio	0,897635	Ratio Validations/Montants pour l'arrêt

La date de sortie correspond en fait à la date de la veille lorsqu'on utilisera l'outil. En effet, nous avons fait l'hypothèse que les données étaient exploitables dès le lendemain et qu'au lancement de l'outil, le traitement s'effectuerait sur les données de la veille ou de la dernière semaine. Concernant l'heure nous l'avons retransformée en heure numérique pour une meilleure compréhension. Le type va lui permettre de différencier sur la carte les « marqueurs » issus des données ou des autres sources comme nous le verrons en 5.2.

Le choix des critères permettant de relever les arrêts est variable. Nous avons choisi de remonter les 15 arrêts les plus sensibles ayant plus de 100 passagers par jour, mais ces critères sont modifiables très facilement à partir du code source et pourront être adaptables en fonction des besoins du personnel d'inspection. Nous avons fixé un nombre de passagers minimum pour les mêmes raisons que celles mentionnées auparavant.

Sur la **Figure 5-2** ci-dessous, on retrouve un imprimé-écran de l'outil web avec les marqueurs correspondant aux 15 arrêts de bus ayant le plus faible ratio sur l'ensemble de la semaine d'avril.

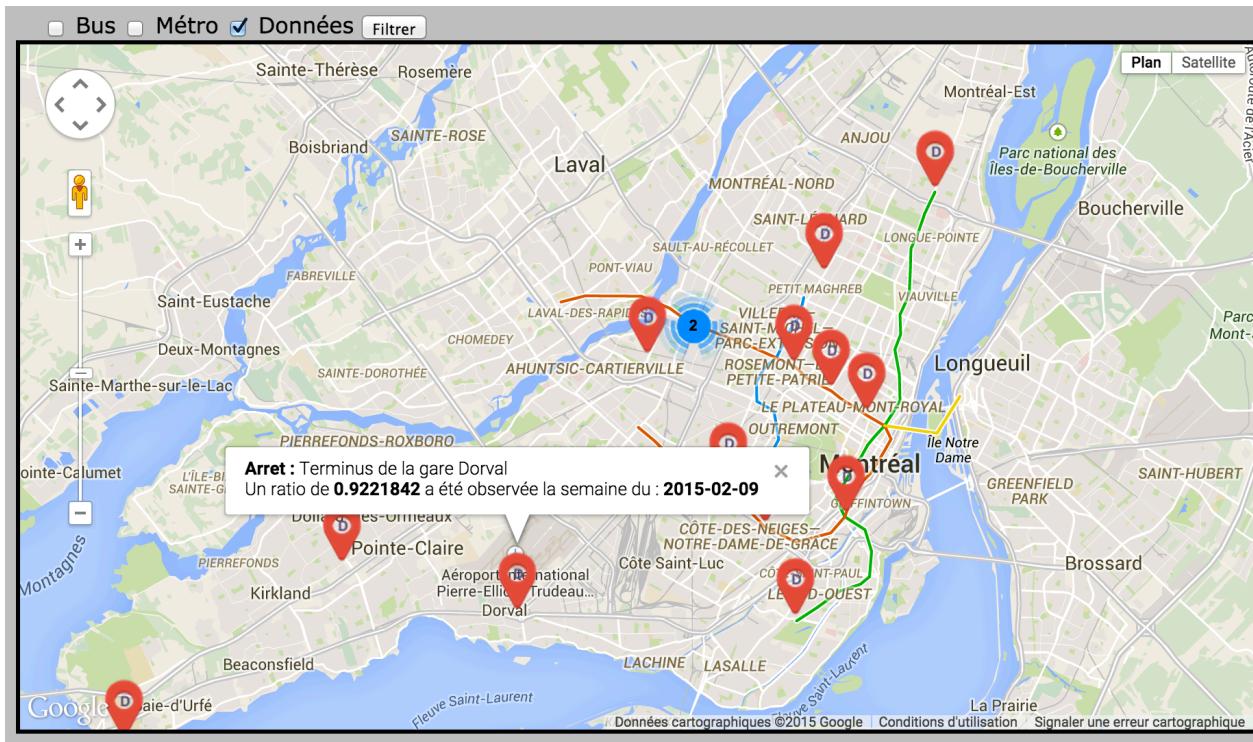


Figure 5-2 : 15 arrêts ayant un faible ratio représentés sur une carte Google Maps

On peut observer sur la figure 17 un « cluster » bleu, celui-ci correspond en fait à un regroupement de deux marqueurs, dans ce cas cela signifie que deux arrêts sont relativement proches et cela améliore la lisibilité.

Il faut noter que ces arrêts sont basés sur les données d'avril 2014, cependant à des fins d'adaptabilité de l'outil, la date a dû être modifiée pour faciliter une éventuelle implémentation. Le but est de récupérer sur les 7 derniers jours et d'en extraire les 15 arrêts à risque. Cependant, étant donné que l'ensemble des données sera gardé dans la base de données, des filtres par date pourraient éventuellement être mis en place pour faire afficher par exemple, les arrêts à risque du mois dernier, ou seulement ceux de la veille.

Des contraintes sont tout de même à relever au niveau de la mise à jour automatique d'un tel outil. En effet, il s'avère que pour le moment, les données brutes doivent être anonymisées et extraites d'un logiciel de gestion avant d'être utilisables. De plus, la remontée des données depuis le bus jusqu'au logiciel de gestion peut prendre plusieurs heures, voir plusieurs jours. Cependant, cette étude reste à des fins exploratoires et à pour but de montrer l'étendue des possibilités de

l'exploitation des données de carte à puce et de comptage à bord pour caractériser l'évasion tarifaire dans les bus.

Cette partie de l'outil se basant uniquement sur les données, nous avons jugé utile de le compléter avec les données que l'on peut éventuellement remonter du terrain. Ce sera l'enjeu de la partie 5.2.

5.2 Extension pour l'évasion tarifaire observée

Pour compléter cet outil basé uniquement sur les données, nous avons décidé de le compléter par les observations provenant des guichetiers et des chauffeurs de bus. Pour cela, nous avons développé un prototype d'application mobile permettant de remonter ces informations et de les envoyer vers la même base de données. Une fois dans la base de données, elles pourront être affichées sur la carte.

L'utilisation de l'application est relativement simple. L'utilisateur choisit d'abord s'il se trouve dans un bus ou dans un métro. Puis, il notifie soit la ligne de bus ou l'arrêt de métro, et envoie son observation lorsqu'il est témoin d'une fraude (Annexe F). Les données envoyées sont alors complétées par les informations disponibles sur l'appareil (Heure, Coordonnées GPS). Le **Tableau 5-2** nous montre les données envoyées à la base de données via cette application.

Tableau 5-2 : Forme des données récupérées par l'application

Données_Sortie_Traitements	Exemple	Explications
Date_Sortie	09/02/2015	Date du jour de l'observation
Heure_Depart_Arret	14:05:34	Heure de l'observation
Lat_GPS	45.56652	Latitude
Long_GPS	-73.59131	Longitude
Type	Bus	Type de sources (données, observations dans un bus ou un métro)
Ligne	141	Numéro de la ligne (bus)
Sens	Ouest	Sens de la ligne (bus)
Station		Station (métro)

C'est ici que l'on comprend l'utilité de la variable « Type » dans les données. On aura alors trois types de marqueurs sur la carte :

- les marqueurs sur les arrêts à risque issus du traitement des données;
- les marqueurs sur les stations de métro où une évasion tarifaire a été observée;
- les marqueurs sur les arrêts de bus où une évasion tarifaire a été observée.

À noter que le paramètre « station » est rempli que lorsqu'il s'agit d'un type Métro, et les paramètres « lignes » et « sens » que lorsqu'il s'agit d'un type Bus. La **Figure 5-3** nous montre à quoi pourrait ressembler la carte avec cette extension pour les évasions tarifaires observées.

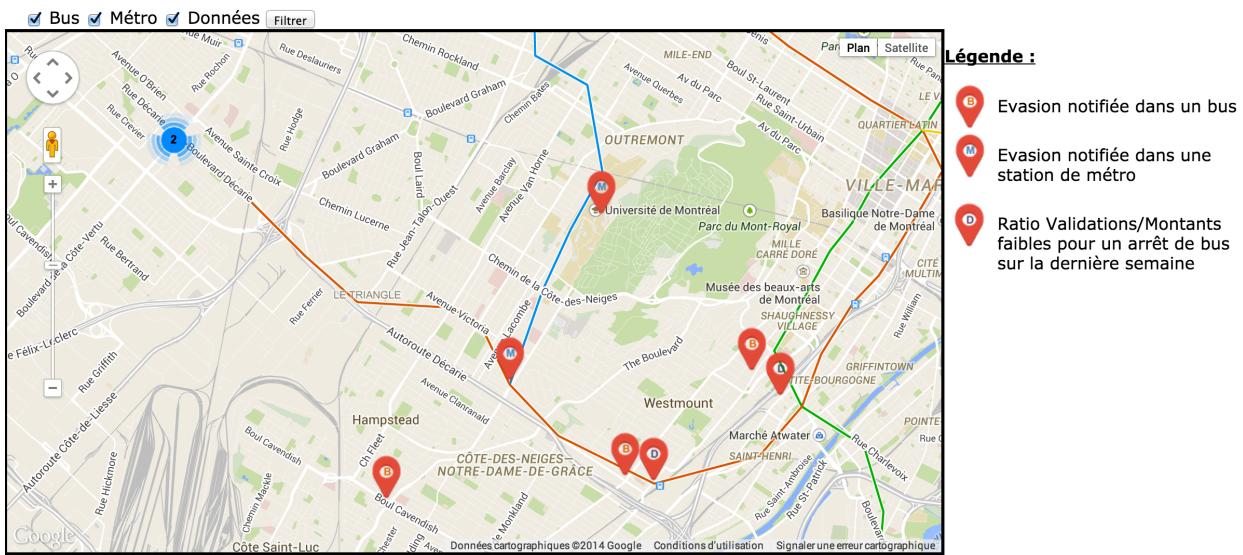


Figure 5-3 : Carte avec les trois types de données

On observe donc sur cette figure les trois types de marqueurs, et la possibilité de filtrer par les différents types. Cet outil qui peut être affichable sur n'importe quel support ayant un accès à internet. L'annexe E récapitule le fonctionnement de cet outil d'une manière plus visuelle tandis que la partie 5.3 va expliquer l'utilisation potentielle de cet outil tel que nous le voyons.

5.3 Utilisation de l'outil au quotidien

L'objectif, comme nous l'avons déjà mentionné auparavant, a été de proposer un outil d'aide à la décision pour le personnel d'inspection, qui à ce jour ne peut se baser que sur leurs propres connaissances du terrain. Pour ce qui est de son utilisation au quotidien, nous voyons cet outil comme un support pour les inspecteurs, ils ne pourront pas se baser uniquement sur celui-ci, mais

il peut aider la planification des inspections de la journée. De plus, sa consultation étant relativement facile à l'aide d'un ordinateur ou d'un téléphone intelligent, un déplacement des inspecteurs proches des évasions observées par les guichetiers ou les conducteurs peut-être envisageable. Grâce à un tel outil, la planification des inspections pourrait donc être facilitée, et leur réactivité sur le terrain améliorée.

Cependant pour éviter un phénomène de profilage qui pourrait se traduire par l'utilisation d'un tel outil nous avons pensé séparer les ressources d'inspection en deux. Nous préconisons par exemple de garder une partie des inspecteurs qui n'utiliseraient pas l'outil, c'est à dire qui ne ciblerait pas les arrêts à risques et qui continuerait à effectuer des patrouilles plus ou moins aléatoirement. L'autre partie pourrait alors utiliser un tel outil de manière à être réactif aux observations relevées par les guichetiers et conducteurs et prendre leurs décisions en fonction de cela. Cependant le temps de déplacement des inspecteurs doit être à prendre en compte. Pour les bus, nous pouvons imaginer que si une évasion est notifiée sur un arrêt quelconque sur la ligne, les inspecteurs devraient alors se déplacer vers l'arrêt où il y a le plus de descente (souvent le terminus, proche d'un métro par exemple). Concernant les métros, la réactivité paraît plus complexe pour arriver à intercepter le fraudeur. Toutefois, comme pour les bus nous pouvons imaginer que les inspecteurs pourraient cibler les stations ayant un achalandage plus élevé encadrant la station où la fraude a été observée. Les 2 figures ci-après illustrent ces différents cas :

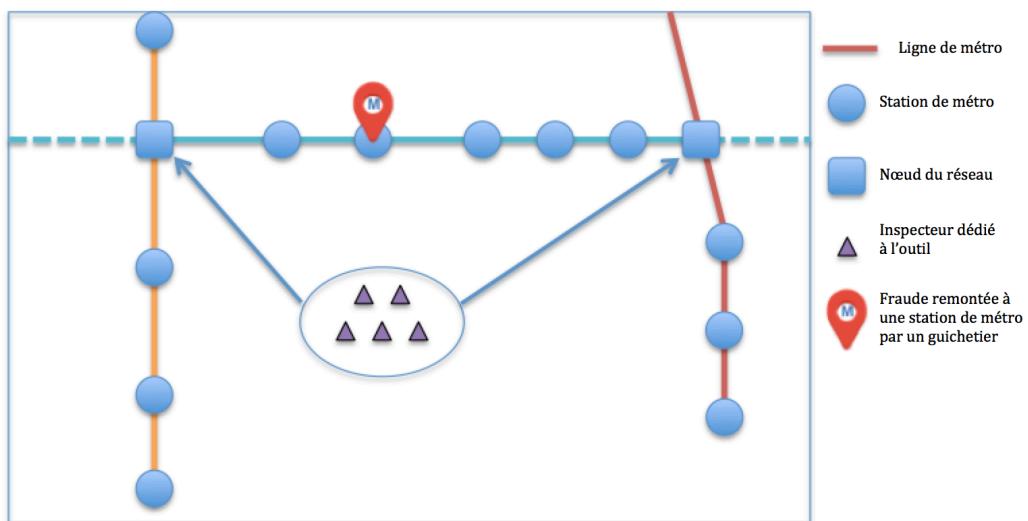


Figure 5-4 : Schématisation de l'utilisation de l'outil pour les métros

La **Figure 5-4** nous montre une utilisation potentielle de l'outil pour les métros. Si une fraude est observée dans une station de métro entre deux nœuds du réseau et que des inspecteurs sont relativement proches de ces nœuds, une intervention à la sortie de ces stations peut être mise en place. Toutefois, cette utilisation de l'outil dans les métros demandera à ce que l'outil devienne vraiment communicatif avec le personnel de contrôle, et l'on peut aussi imaginer de rajouter des calculs du temps de déplacement entre la position actuelle et les nœuds.

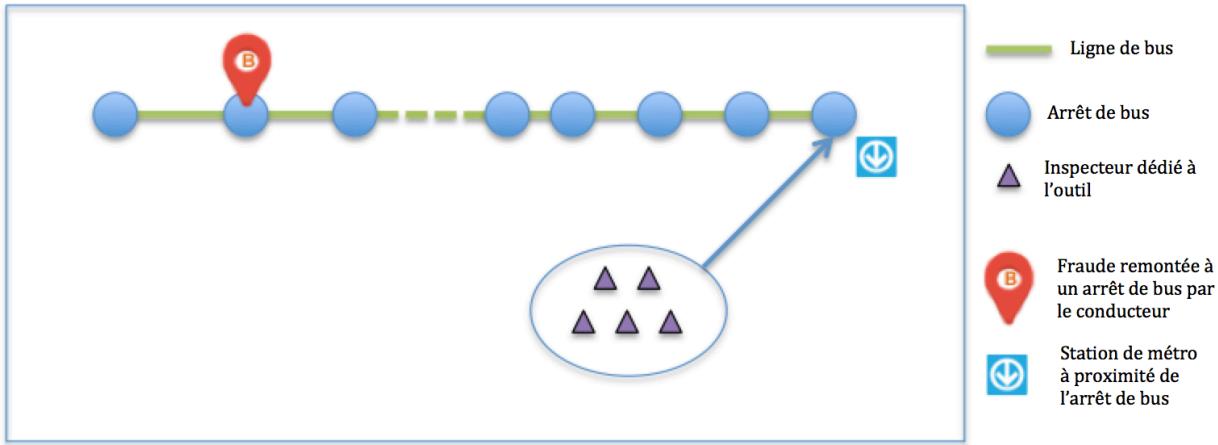


Figure 5-5 : Schématisation de l'utilisation de l'outil pour les bus

La **Figure 5-5** schématisse l'utilisation pour les bus. Le principe serait le même que pour les métros et serait de cibler les arrêts où le fraudeur est le plus à même de descendre.

Il faut noter que ces utilisations sont des cas isolés et l'outil ne permettra pas une intervention du personnel de contrôle à chaque observation. Cependant, étant donné que toutes les observations sont sauvegardées dans la base de données, cela pourra leur permettre de distinguer les arrêts à risques selon les observations et de conforter ou non la pertinence des informations remontées par le traitement. De plus, pour une mise en application d'un tel outil, l'ajout de fonctions comme la communication entre agents ou des alertes deviendraient alors nécessaire. Cela pourrait être l'objet d'une prochaine étude.

Pour conclure cette partie sur l'outil dynamique, nous pouvons dire que notre prototype fonctionnel d'application pourrait devenir une véritable aide pour le personnel d'inspection. Cependant, il faut de même bien noter que les inspecteurs possèdent déjà un appareil pour

vérifier la validité des titres de transport, et que le déploiement de téléphones intelligents aurait un coût très important. Nous pouvons alors peut-être suggérer le développement d'un outil similaire sur l'appareil permettant de vérifier la validité à des fins de praticité et d'efficacité.

CHAPITRE 6 : CONCLUSION

Synthèses des travaux de recherche

Ce projet a permis de mettre en évidence le potentiel des données de carte à puce et de comptage à bord comme point de départ pour la caractérisation de l'évasion tarifaire dans les bus.

D'abord, à travers la revue de littérature, nous avons constaté que les données de carte à puce et de comptage à bord étaient issues de technologies de pointe, et que l'implantation de ces systèmes était relativement récente dans les réseaux de transport en commun. De plus, la quantité de données disponibles grâce à ces technologies représente une véritable mine d'or pour les agences de transports en terme de planification globale, de coordination des bus, d'analyse du comportement des usagers et évidemment d'évasion tarifaire. Il est certain que la liste des possibilités offertes par ces données ne s'arrête pas là. L'avancée de la technologie dans ces domaines nous pousse à croire que les données vont continuer à s'enrichir et devenir de plus en plus complètes, et qu'elles seront vraiment indispensables aux agences de transport.

Concernant le phénomène d'évasion tarifaire, nous avons relevé que de nombreuses études avaient été réalisées à ce sujet depuis plusieurs décennies. Cependant, dans la plupart des études les auteurs se basent sur des données issues de sondages ou d'enquêtes de terrain pour caractériser ce phénomène. Ils étudient souvent les conséquences financières ainsi que les modes opératoires des fraudeurs. Aucune étude n'avait réellement été réalisée en se basant sur les données des systèmes de carte à puce et de comptage à bord pour caractériser ce phénomène. De plus, il s'avère que très peu d'outils sont disponibles sur le terrain pour faciliter le déploiement du personnel d'inspection et leur réactivité sur le terrain.

Après pour ce qui est du traitement, ce fut la partie la plus délicate du projet. En effet, la jointure entre les deux sources de données ne fut pas évidente, mais avec de la minutie pour manipuler les données nous avons pu réaliser un petit programme permettant de récupérer les fichiers sources et de les joindre. Cependant, de nombreuses disparités existaient entre ces deux sources de données, une étape de prétraitement a été nécessaire avant cette jointure. De plus, une étape de post-traitement a dû être réalisée pour supprimer les données inutiles ou erronées ainsi que pour enrichir les données.

Une fois le traitement réalisé, c'est grâce à l'étape d'analyse que nous avons pu répondre à la problématique de ce mémoire au sujet de la caractérisation de l'évasion tarifaire par les données de carte à puce (CAP) et de comptage à bord (APC). Pour cela, de nombreux paramètres ont pu être étudiés comme les jours de la semaine, les types de bus, les titres de transport ou encore les arrêts. Nous nous sommes alors basés sur le ratio du nombre de validations par rapport au nombre de montants en fonction de ces différents paramètres sur le réseau de l'île. Cependant comme tout appareil, les capteurs de comptages à bord ont une certaine fiabilité qu'il a fallu prendre en compte, c'est pourquoi nous avons choisi de toujours nous reporter au ratio moyen qui se situe autour de 0,97. Pour caractériser l'évasion tarifaire, nous avons donc regardé si en fonction des paramètres comme par exemple une ligne de bus, des ratios pouvaient être bien inférieurs à ce ratio moyen. Dans ce cas là, on a alors fait l'hypothèse que l'écart provient au moins en partie de la fraude. Après une analyse comparée avec une deuxième semaine de données, nous avons pu mettre en évidence une certaine récurrence dans ce phénomène avec des lignes ou des arrêts plus sensibles que d'autres qui se retrouvaient aussi bien en avril qu'en octobre. De plus, une étude sur l'heure de la journée nous a permis de constater que les heures creuses étaient plus sujettes au phénomène d'évasion tarifaire que les heures pleines.

Après avoir réalisé cette analyse, nous avons jugé intéressant de proposer un outil dynamique d'aide à la décision pour le personnel d'inspection basé sur ces résultats. Pour cela, nous avons automatisé le processus et proposé une visualisation cartographique du phénomène. Grâce à cet outil, le personnel d'inspection peut aussi bien s'appuyer sur les données issues des systèmes APC et CAP, mais aussi sur les fraudes observées sur le terrain. Étant donné que l'ensemble de ces données sont stockées dans une base de données, un historique est alors consultable. Cet outil pourrait donc être une bonne ressource en terme de planification des inspecteurs sur le terrain, afin de mieux cibler les arrêts et stations à risques.

Limitations des solutions proposées

À travers les différentes étapes du projet, nous avons rencontré plusieurs limitations ou contraintes. La première que nous avons déjà introduite à plusieurs reprises concerne l'implantation de la technologie APC sur le réseau, qui ne représente que 16% des bus à ce jour. Une analyse sur l'ensemble des bus aurait pu donner peut-être une meilleure vision et un meilleur suivi de ce phénomène par jour de la semaine. De plus, il s'est avéré que certaines données

n'étaient pas toujours fiables, ce fut le cas par exemple pour la ligne de bus issus des données de carte à puce qui ne correspondait pas toujours avec celle des données APC. L'étape de post traitement pour corriger un certain nombre d'erreurs de ce type fut nécessaire. Cependant, nous ne pouvons pas assurer non plus que toutes les erreurs sur les données aient été supprimées à travers ce post-traitement. Une autre limite réside dans le fait que l'on ne peut pas savoir à 100% que la différence entre un ratio qui se situe dans la moyenne et un ratio bien plus faible est due au phénomène d'évasion tarifaire. Cela reste une supposition, et des enquêtes sur le terrain ou un test de l'outil dynamique sur le terrain pourraient alors confirmer cela.

Une autre limitation concerne l'implantation d'un outil comme celui que nous avons pu réaliser. Nous vivons actuellement dans une période où l'information dans les entreprises est centralisée grâce à de lourds logiciels appelés PGI (progiciel de gestion intégrée) plus connu sous son abréviation anglaise ERP (Enterprise Resource Planning). Un outil comme le nôtre ne respecterait alors pas vraiment les chartes de centralisation de l'information dans les entreprises.

Perspectives et améliorations futures

La fin du projet d'implantation des APC est prévue fin 2015 par la STM, et fait partie du grand projet « ibus » représentant plus de 200M\$ et s'inscrivant dans le plan stratégique 2020. À ce titre, il pourra être intéressant de reconsidérer une telle étude sur un réseau entièrement équipé d'APC. On peut tout de même noter que ce projet « ibus » a notamment pour but de faciliter l'information pour les utilisateurs avec par exemple la mise en place d'afficheurs et de systèmes d'annonces à bord des bus. Ces évolutions technologiques nous laissent entendre que la quantité de données disponibles sur les réseaux de transport ne va faire qu'augmenter. À plus long terme, nous pouvons donc espérer que les données de collecte des titres et de comptages à bord soient directement liées. Des études plus poussées concernant la prédiction d'un tel phénomène avec des méthodes mathématiques comme le Bayésien pourraient aussi être sujet à des futures études.

Mais, à court terme, nous pouvons penser qu'un perfectionnement de notre outil à des fins d'un test réel pourrait être plus qu'envisageable. De plus, l'implantation d'un outil comme le nôtre dans un logiciel de résolution de problèmes pourrait aussi être une opportunité pour un projet futur. En effet, l'évasion tarifaire n'est au final qu'une non-conformité du réseau, et d'autres mesures peuvent être mises en place pour la résoudre. Concernant la représentation cartographique de l'outil d'aide à la décision, nous pouvons aussi prétendre que d'autres données

pourraient être liées à cet outil comme les données de maintenance ou de sécurité. Les personnels de terrains de ces domaines pourraient aussi s'appuyer sur une représentation cartographique dynamique pour gérer leur planification. Cela permettrait ainsi une centralisation de l'information.

BIBLIOGRAPHIE

- Attoh-Okine, N. O., & Shen, L. D. (1995). Security issues of emerging smart cards fare collection application in mass transit. In *Vehicle Navigation and Information Systems Conference, 1995. Proceedings. In conjunction with the Pacific Rim TransTech Conference. 6th International VNIS. "A Ride into the Future"*, 523–526. <http://doi.org/10.1109/VNIS.1995.518887>
- Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, 12(5), 464–474. <http://doi.org/10.1016/j.tranpol.2005.06.008>
- Barabino, B., Salis, S., & Useli, B. (2013). A modified model to curb fare evasion and enforce compliance: Empirical evidence and implications. *Transportation Research Part A: Policy and Practice*, 58, 29–39. <http://doi.org/10.1016/j.tra.2013.10.007>
- Barfield, W., & Dingus, T. A. (Eds.). (1997). *Human Factors in Intelligent Transportation Systems* (1 edition). Mahwah, N.J. : Psychology Press.
- Bernini, N., Bombini, L., Buzzoni, M., Cerri, P., & Grisleri, P. (2014). An embedded system for counting passengers in public transportation vehicles. Communication présentée au *10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, du 10-12 septembre 2014 (p. 6 pp.). <http://doi.org/10.1109/MESA.2014.6935562>
- Blythe, P. T. (2004). Improving public transport ticketing through smart cards. *Proceedings of the ICE - Municipal Engineer*, 157(1), pp 47–54. <http://doi.org/10.1680/muen.2004.157.1.47>
- Bonfanti, G., & Wagenknecht, T. (2010). Human Factors Reduce Aggression and Fare Evasion. *Public Transport International*, 59(1), 28–32.

- Borndörfer, R., Omont, B., Sagnol, G., & Swarat, E. (2012). A Stackelberg Game to Optimize the Distribution of Controls in Transportation Networks. In V. Krishnamurthy, Q. Zhao, M. Huang, & Y. Wen (Eds.), *Game Theory for Networks*, 224–235. Springer Berlin Heidelberg.
- Boyd, C., Martini, C., Rickard, J., & Russell, A. (1989). Fare Evasion and Non-Compliance : a simple model. *Journal of Transport Economics and Policy*, 23(2) 189–197.
- Boyle, D.K., (1998) *TCRP Synthesis 29: Passenger Counting Technologies and Procedures*, Transportation Research Board, National Research Council, Washington, D.C.
- Boyle, D. K. (2009). *TCRP Synthesis 77 Passenger Counting Systems*. S.l.: s.n. Transportation Research Board, National Research Council, Washington, D.C.
- Bucciol, A., Landini, F., & Piovesan, M. (2012). *Unethical Minds: Individual Characteristics that Predict Unethical Behavior* (Working Paper No. 24/2012). Université de Vérone, Italie, Département d'économie.
- Casey, R. 1999. Advanced Public Transportation Systems: Deployment in the United States, Cambridge, MA: John A. Volpe National Transportation Systems Center. Report No. DOT-VNTSC-FA-99
- Chira-Chavala, T., & Coifman, B. (1996). Impacts Of Smart Cards On Transit Operators: Evaluation Of I-110 Corridor Smart Card Demonstration Project. *California Partners for Advanced Transit and Highways (PATH)*, 14-36.
- Clarke, R. V., Contre, S., & Petrossian, G. (2009). Deterrence and fare evasion: Results of a natural experiment. *Security Journal*, 23(1), 5–17. <http://doi.org/10.1057/sj.2009.15>
- Clarke, R. V., (1993). Fare evasion and automatic ticket collection on the london underground. *Crime Prevention Studies*. Monsey, N.Y. (Vol 1), 135–146.
- Dauby, L., & Kovacs, Z. (2006). Fare Evasion in Light Rail Systems : A World of Applications and Opportunities. *Joint International Light Rail Conference*, 230–246.

- Deakin, E., & Kim, S. (2001). Transportation Technologies: Implications for Planning. Université de Califonie Transportation Center, CA. 10–14.
- Dempsey, S.P (2008). Privacy issues with the use of smart cards. Legal Research Digest 25.
- Department of Transport, P. T. V. (n.d.). Fare Evasion on Public Transport. Tiré de http://www.audit.vic.gov.au/reports_and_publications/latest_reports/2012-13/20120829-fare-evasion.aspx
- De Potter, P., Belet, P., Poppe, C., Verstockt, S., Lambert, P., & Van De Walle, R. (2012). Passenger counting in public rail transport: Using head-shoulder contour tracking. In *International Conference on Computer Vision Theory and Applications, VISAPP 2012*, 24 au 26 février 2012 (Vol. 1, pp. 705–708). Inst. for Syst. and Technol. of Inf. Control and Commun.
- Descoimps, E. (2011). *Analyse des données issues d'un système de perception par carte à puce d'une société de transport en commun : Normalité des déplacements et influence des conditions météorologiques* (Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, QC).
- Diebel, L. E. (1981). The introduction of European fare collection techniques in United States. *Journal of Advanced Transportation*, 55–60.
- Directive, 2010 Directive, 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport (2010).
- Durango, P. L., & Madanat, S. M. (2002). Optimal maintenance and repair policies in infrastructure management under uncertain facility deterioration rates: an adaptive control

- approach. *Transportation Research Part A: Policy and Practice*, 36(9), 763–778.
[http://doi.org/10.1016/S0965-8564\(01\)00038-6](http://doi.org/10.1016/S0965-8564(01)00038-6)
- Goel, A. (2007). *Fleet Telematics: Real-time management and planning of commercial vehicle operations*. USA : Springer Science & Business Media.
- Heij, C., Bijwaard, G. E., & Knapp, S. (2011). Ship inspection strategies: Effects on maritime safety and environmental protection. *Transportation Research Part D: Transport and Environment*, 16(1), 42–48. <http://doi.org/10.1016/j.trd.2010.07.006>
- Ibrahim, M. F. (2003). Improvements and integration of a public transport system: the case of Singapore. *Cities*, 20(3), 205–216. [http://doi.org/10.1016/S0264-2751\(03\)00014-3](http://doi.org/10.1016/S0264-2751(03)00014-3)
- Joslin, A., & Morris, W. P. (2013). Ridership Impacts of South Florida's EASY Smart Card. *National Center for Transit Research*. Rapport No. BDK85 977-42. Tampa, FL : Florida Department of Transportation, University of South Florida.
- Kimpel, T., Strathman, J., Griffin, D., Callas, S., & Gerhart, R. (2003). Automatic Passenger Counter Evaluation : Implications for national transit database reporting. *World Transit Research*. Transportation Research Record, 1835, 93-100.
- Kooreman, P. (1993). *Fare evasion as a result of expected utility maximization: Some empirical support*. Tilburg University. 1-5.
- Lomone, A. (2014). *Exploration et traitement multidonnées appliqués à des corridors d'autobus*. (Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, QC).
- Lu, H. K. (2007). Network Smart Card Review and Analysis. *Comput. Netw.*, 51(9), 2234–2248.
<http://doi.org/10.1016/j.comnet.2007.01.009>
- Mayes, K. E., Markantonakis, K., & Hancke, G. (2009). Transport ticketing security and fraud controls. *Information Security Technical Report*, 14(2), 87–95.
<http://doi.org/10.1016/j.istr.2009.06.003>

- McDonald, N. (2000). Multipurpose Smart Cards in transportation : Benefits and barriers to use. University of California Transportation Center Research, 630, 27p.
- Morency, C., Trépanier, M., & Agard, B. (2007). *Measuring transit use variability with smart-card data*. *Transport policy*, 14(3), 193–203.
- Munizaga, M. A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies*, 24, 9–18. <http://doi.org/10.1016/j.trc.2012.01.007>
- Nowacki, G. (2012). Development and Standardization of Intelligent Transport Systems. *TransNav, International Journal on Marine Navigation and Safety Od Sea Transportation*, 6(3), 403–411.
- Pelletier, M.-P., Trepanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557–568. <http://doi.org/10.1016/j.trc.2010.12.003>
- Perron, B. (2013). *Rapport d'inspection concernant le système de carte OPUS du réseau de transport de la capitale* (Rapport Inspection No. 11 01 68). Commission d'accès à l'information du Québec. Tiré de http://www.cai.gouv.qc.ca/documents/CAI_RI_110168.pdf
- Reddy, A. V., Kuhls, J., & Lu, A. (n.d.). Measuring and Controlling Subway Fare Evasion: Improving Safety and Security at New York City Transit Authority. *Transportation Research Record*, 2216, 85–99.
- Société de transport de Montréal. (2013a). R-105 : Règlement concernant les conditions au regard de la possession et de l'utilisation de tout titre de transport émis par la Société de transport de Montréal. *Règlements*. Tiré de <http://www.stm.info/fr/a-propos/informations-entreprise-et-financieres/les-reglements-et-politiques/reglement-r-105>
- Société de transport de Montréal. (2013b). Rapport d'activités 2012. Tiré de

<http://www.stm.info/sites/default/files/pdf/fr/ra2012.pdf>

Sasaki, Y. (2014). Optimal choices of fare collection systems for public transportations: Barrier versus barrier-free. *Transportation Research Part B: Methodological*, 60, 107–114.

<http://doi.org/10.1016/j.trb.2013.12.005>

Schiavone, J. J. (1999). Understanding and Applying Advanced On-board Bus Electronics. Transportation Research Board (TCRP) Rapport 43.

Shelfer, K. M., & Procaccino, J. D. (2002). Smart Card Evolution. *Commun. ACM*, 45(7), 83–88.

<http://doi.org/10.1145/514236.514239>

Standaert, F.-X. (2008). *Smart Card Research and Advanced Applications: 8th IFIP WG 8.8/11.2 International Conference, CARDIS 2008, London, UK, September 8-11, 2008, Proceedings*. Springer Science & Business Media.

Strathman, J. G. (2002). *Tri-Met's experience with automatic passenger counter and automatic vehicle location systems*. Portland, OR : Portland State University.

Thorlacius, P., Clausen, J., Thorlacius, P., & Clausen, J. (2010). Scheduling of inspectors for ticket spot checking in urban rail transportation. Communication présentée au *Trafikdage På Aalborg Universitet*, Aalborg, Danemark, 25-26 août 2008.

Torres-Montoya, M. (2014). Tackling Fare Evasion in Transantiago: An Integrated Approach. Communication présentée au Transportation Research Board 93rd Annual Meeting. USA, Washington, D.C.

Trepanier, M., Barj, S., Dufour, C., & Poilpre, R. (2004). Examination of potentials in the data analysis of a smart card payment system for urban transport. Communication présentée au *Annual Conference of the Transportation Association of Canada: Transportation Innovation* du 22-24 Septembre 2004, Québec, QC.

- Utsunomiya, M., Attanucci, J., & Wilson, N. (2006). Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning. *Transportation Research Record: Journal of the Transportation Research Board*, 1971, pp 119-126.
- Wang, J. (2012). *A traffic accident risk mapping framework* (Mémoire de thèse, University of Calgary, Calgary, AB).
- Yoh, A. C. (2008). *Institutional Issues in the Adoption of Smart Card Systems Among U.S. Transit Agencies for Fare Collection*. (Mémoire de thèse, University of California, USA, Los Angeles, CA).

ANNEXES

ANNEXE A - TITRES DE TRANSPORTS

Description_Produit	Usager	Support	Type
EMPLOYEE/RETRAITE-AMT	Autre	CPCT	Abonnement
EMPLOYEE/RETRAITE-OPT	Autre	CPCT	Abonnement
OPUS-TRAM1-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM2-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM3-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM4-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM5-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM6-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM7-MENSUEL-E	Réduit	CPCT	Abonnement
OPUS-TRAM8-MENSUEL-E	Réduit	CPCT	Abonnement
TRAM1-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM2-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM3-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM4-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM5-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM6-ABONNEMENT-E	Réduit	CPCT	Abonnement
TRAM7-ABONNEMENT-E	Réduit	CPCT	Abonnement
2 PASSAGES-R	Réduit	CMJ	Unitaire
OPUS-TRAM1-BILLET-R	Réduit	CPCT	Unitaire
OPUS-TRAM1-CARNET-R	Réduit	CPCT	Unitaire
OPUS-TRAM1-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM2-BILLET-R	Réduit	CPCT	Unitaire
OPUS-TRAM2-CARNET-R	Réduit	CPCT	Unitaire
OPUS-TRAM2-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM3-BILLET-R	Réduit	CPCT	Unitaire
OPUS-TRAM3-CARNET-R	Réduit	CPCT	Unitaire
OPUS-TRAM3-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM4-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM5-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM6-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM7-MENSUEL-R	Réduit	CPCT	Abonnement
OPUS-TRAM8-MENSUEL-R	Réduit	CPCT	Abonnement
STM 10 PASSAGES RED.	Réduit	CPCT	Unitaire
STM- 6 PASSAGES RED.	Réduit	CPCT	Unitaire

STM-1 PASSAGE-R	Réduit	CPCT ou CMJ	Unitaire
STM-10 PASSAGES-R	Réduit	CPCT	Unitaire
STM-2 PASSAGES-R	Réduit	CPCT ou CPO	Unitaire
STM-ABONNEMENT RED.	Réduit	CPCT	Abonnement
STM-CAM 4 MOIS-R	Réduit	CPCT	Abonnement
STM-CAM HEBDO-R	Réduit	CPCT	Abonnement
STM-CAM MENSUELLE-R	Réduit	CPCT	Abonnement
TRAM1-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM2-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM3-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM4-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM5-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM6-ABONNEMENT-R	Réduit	CPCT	Abonnement
TRAM7-ABONNEMENT-R	Réduit	CPCT	Abonnement
2 PASSAGES-O	Régulier	CPCT ou CPO	Unitaire
GRAT8-ABONNEMENT-O	Régulier	CPCT	Abonnement
OPUS-AQLPA-TR	Régulier	CPCT	Abonnement
OPUS-COURTOISIE-B	Régulier	CPCT ou CMJ	Abonnement
OPUS-TRAM1-BILLET-O	Régulier	CPCT	Unitaire
OPUS-TRAM1-CARNET-O	Régulier	CPCT	Unitaire
OPUS-TRAM1-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM2-BILLET-O	Régulier	CPCT	Unitaire
OPUS-TRAM2-CARNET-O	Régulier	CPCT	Unitaire
OPUS-TRAM2-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM3-BILLET-O	Régulier	CPCT	Unitaire
OPUS-TRAM3-CARNET-O	Régulier	CPCT	Unitaire
OPUS-TRAM3-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM4-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM5-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM6-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM7-MENSUEL-O	Régulier	CPCT	Abonnement
OPUS-TRAM8-MENSUEL-O	Régulier	CPCT	Abonnement
SOLO- 1 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 2 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 3 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 4 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 5 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 6 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 8 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO- 9 COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire
SOLO-10COURTOISIE-B	Régulier	CPCT ou CMJ	Unitaire

SOLO-TRAM1-BILLET-O	Régulier	CMJ	Unitaire
SOLO-TRAM1-CARNET-O	Régulier	CMJ	Unitaire
SOLO-TRAM2-BILLET-O	Régulier	CMJ	Unitaire
SOLO-TRAM2-CARNET-O	Régulier	CMJ	Unitaire
SOLO-TRAM3-BILLET-O	Régulier	CMJ	Unitaire
SOLO-TRAM3-CARNET-O	Régulier	CMJ	Unitaire
STM - 1 PASSAGE ORD.	Régulier	CMJ	Unitaire
STM - CIRCU MB 3 PAS	Régulier	CPCT ou CPO	Unitaire
STM 1 JOUR 0\$	Régulier	CPCT ou CPO	Abonnement
STM 7 JOURS 0\$	Régulier	CPCT	Abonnement
STM ABONNEMENT ORD.	Régulier	CPCT	Abonnement
STM COURTOISIE BUS	Régulier	CMJ	Unitaire
STM COURTOISIE METRO	Régulier	CMJ	Unitaire
STM- 6 PASSAGES ORD.	Régulier	CPCT ou CPO	Unitaire
STM-1 JOUR	Régulier	CPCT ou CPO	Abonnement
STM-1 PASSAGE PROMO	Régulier	CPCT ou CMJ	Unitaire
STM-1 PASSAGE-O	Régulier	CPCT ou CMJ	Unitaire
STM-10 PASSAGES-O	Régulier	CPCT ou CPO	Unitaire
STM-2 PASSAGES	Régulier	CPCT ou CPO	Unitaire
STM-2 PASSAGES-O	Régulier	CPCT ou CPO	Unitaire
STM-2X1 PASSAGE	Régulier	CMJ	Unitaire
STM-3 JOURS	Régulier	CPCT ou CPO	Abonnement
STM-7 JOURS CONGRES	Régulier	CPCT	Abonnement
STM-ABON.+MAESTRO	Régulier	CPCT	Abonnement
STM-ABON+MAESTRO+COM	Régulier	CPCT	Abonnement
STM-ABONNEMENT	Régulier	CPCT	Abonnement
STM-ABONNEMENT 48	Régulier	CPCT	Abonnement
STM-ABONNEMENT-O	Régulier	CPCT	Abonnement
STM-ABONNEMENT-O12	Régulier	CPCT	Abonnement
STM-ABRI VELO	Régulier	CPCT	Abonnement
STM-AEROPORT 747	Régulier	CPCT ou CMJ ou CPO	Unitaire
STM-CAM HEBDO-O	Régulier	CPCT	Abonnement
STM-CAM MENSUELLE-O	Régulier	CPCT	Abonnement
STM-CAM-OECH	Régulier	CPCT	Abonnement
STM-COURT. STAT.MTL	Régulier	CPCT	Abonnement
STM-GAGNANT A VIE	Régulier	CPCT	Abonnement
STM-GROUPE 6-13ANS	Régulier	CPO	Abonnement
STM-SOIREE ILLIMITEE	Régulier	CPCT ou CPO	Abonnement
STM-WEEK-END ILLIM.	Régulier	CPCT ou CPO	Abonnement
TRAM1-ABONNEMENT-O	Régulier	CPCT	Abonnement

TRAM2-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM3-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM4-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM5-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM6-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM7-ABONNEMENT-O	Régulier	CPCT	Abonnement
TRAM8-ABONNEMENT-O	Régulier	CPCT	Abonnement
EVSMV-BILLET	Régulier	CMJ	Unitaire

**ANNEXE B – TABLEAU RÉCAPITULATIF DU RATIO
VALIDATIONS/MONTANTS MOYEN, MAX, MIN, MEDIAN SUR LA
SEMAINE D'AVRIL 2014 (7 AU 13 AVRIL)**

LIGNES	MOYENNE	MAX	MIN	MEDIANE	SOMME_MONTANTS
11	0,901	0,930	0,863	0,926	1490
12	0,932	0,933	0,931	0,932	1713
13	1,034	1,034	1,034	1,034	87
14	0,935	0,956	0,924	0,936	2978
15	0,938	0,952	0,922	0,937	1101
16	0,990	1,244	0,965	0,973	923
18	0,981	1,004	0,960	0,993	36916
19	0,967	0,967	0,967	0,967	270
22	0,941	1,000	0,667	0,931	393
24	0,956	0,996	0,913	0,957	20772
25	0,981	0,981	0,981	0,981	105
26	0,971	1,058	0,960	0,984	1577
27	0,963	0,963	0,963	0,963	1995
28	0,938	0,987	0,858	0,933	2954
29	1,003	1,098	0,951	1,024	1164
30	0,930	0,933	0,928	0,930	2005
32	0,970	0,983	0,958	0,970	12042
33	0,972	1,044	0,957	0,969	32367
34	0,944	0,963	0,911	0,944	4463
36	0,968	1,022	0,896	0,967	3410
37	0,961	0,994	0,914	0,970	8626
39	0,889	1,000	0,739	0,869	415
40	0,884	0,905	0,793	0,849	155
41	1,001	1,023	0,958	0,996	4822
43	0,944	1,016	0,894	0,950	6300
44	0,959	0,975	0,932	0,969	13761
45	0,970	0,984	0,954	0,969	12686
47	0,963	0,979	0,944	0,963	700
48	0,973	1,015	0,891	0,982	12178
49	0,973	1,009	0,944	0,971	16818
51	0,963	0,985	0,924	0,965	48997
53	0,955	0,996	0,925	0,961	1248
54	1,003	1,003	1,003	1,003	591
55	0,969	0,999	0,952	0,963	16035
56	0,982	0,988	0,978	0,983	221
57	0,947	0,953	0,937	0,944	2762

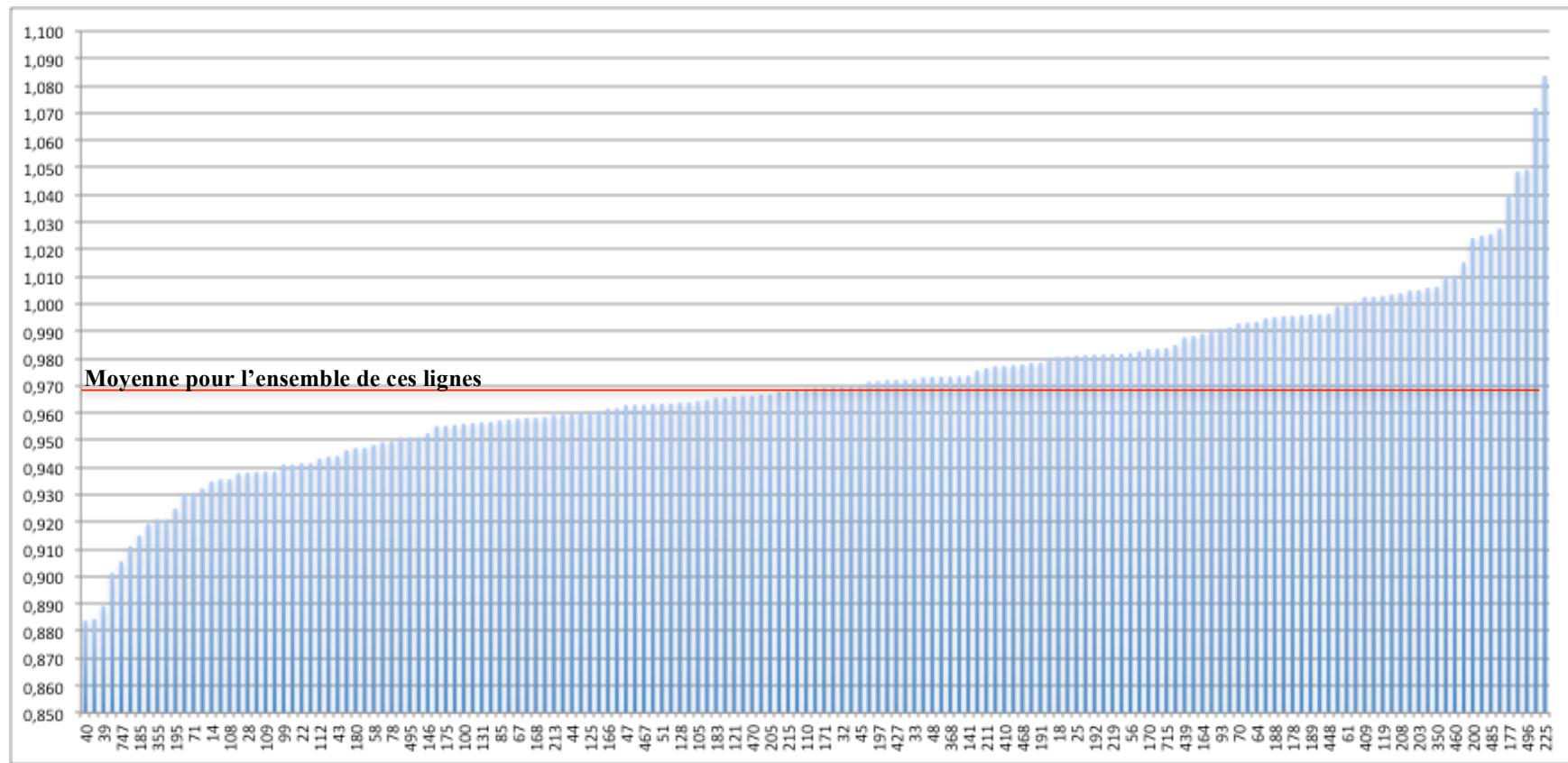
58	0,948	1,059	0,921	0,958	3053
61	1,000	1,032	0,959	0,986	1243
63	0,958	0,996	0,928	0,957	2148
64	0,993	1,073	0,975	1,000	3032
67	0,958	0,977	0,926	0,961	13300
68	0,993	1,034	0,939	0,982	8406
69	0,962	0,980	0,933	0,957	23966
70	0,993	1,058	0,984	1,021	1394
71	0,931	0,931	0,931	0,931	577
72	1,015	1,106	0,963	1,002	2626
74	1,094	1,094	1,094	1,094	64
78	0,950	0,956	0,943	0,948	4299
80	0,972	0,988	0,955	0,967	16398
85	0,957	0,977	0,930	0,964	4513
86	0,941	1,250	0,888	0,967	1829
90	0,968	0,979	0,865	0,968	10439
93	0,991	1,026	0,935	1,003	3288
94	0,941	0,955	0,924	0,940	6594
95	0,985	1,010	0,935	0,970	6075
97	0,946	0,972	0,927	0,941	14261
99	0,941	0,941	0,941	0,941	339
100	0,956	1,034	0,917	0,958	3116
101	1,086	1,086	1,086	1,086	93
102	0,919	0,940	0,908	0,938	2727
103	0,982	1,101	0,965	0,973	5685
104	0,963	0,985	0,956	0,960	4222
105	0,964	0,985	0,932	0,969	17985
106	0,964	1,003	0,953	0,957	13714
107	0,951	1,000	0,925	0,964	6176
108	0,936	0,967	0,914	0,946	4123
109	0,938	0,952	0,870	0,911	2417
110	0,969	1,001	0,932	0,962	13430
112	0,943	1,017	0,921	0,955	2884
113	0,996	1,138	0,925	0,989	1878
115	0,991	2,500	0,955	1,023	1850
117	0,988	1,005	0,975	0,985	2118
119	1,003	1,078	0,950	1,001	2406
121	0,966	0,988	0,944	0,977	26741
123	0,938	0,960	0,911	0,934	3913
124	0,959	1,000	0,903	0,938	2950
125	0,961	0,967	0,941	0,950	5901

128	0,964	1,072	0,917	0,948	3727
129	0,977	1,008	0,917	0,972	5960
131	0,957	1,059	0,944	1,001	1357
136	0,982	1,018	0,944	0,979	7647
138	0,965	0,998	0,928	0,951	3251
139	0,963	1,009	0,942	0,954	32814
140	0,966	1,004	0,903	0,957	6028
141	0,974	0,981	0,958	0,974	40967
144	0,958	0,990	0,928	0,943	3754
146	0,953	0,964	0,943	0,953	6616
150	0,957	0,962	0,955	0,959	762
160	0,936	0,959	0,000	0,946	1943
161	0,980	1,025	0,947	0,986	16773
162	0,949	0,960	0,893	0,954	4708
164	0,989	1,032	0,947	0,984	13645
165	0,966	0,977	0,938	0,967	26231
166	0,962	1,029	0,922	0,955	2209
168	0,958	0,984	0,938	0,958	7445
170	0,983	1,106	0,926	0,975	1025
171	0,969	1,022	0,842	0,988	11561
175	0,955	1,151	0,855	1,017	960
177	1,040	1,258	0,991	1,008	452
178	0,996	1,115	0,932	1,023	224
179	0,983	1,022	0,961	0,983	6339
180	0,947	1,039	0,931	0,954	4971
183	0,966	0,966	0,966	0,966	29
185	0,915	1,026	0,873	0,926	2047
186	0,960	1,034	0,901	0,985	3338
187	0,978	1,012	0,955	0,973	5948
188	0,995	1,006	0,955	0,980	204
189	0,996	1,063	0,906	0,995	4275
191	0,978	1,020	0,925	0,978	2770
192	0,981	1,052	0,963	0,977	6056
193	0,958	0,992	0,933	0,965	26531
195	0,925	1,018	0,874	0,937	4995
196	1,025	1,040	0,998	1,030	2525
197	0,972	1,004	0,941	0,977	5887
200	1,024	1,106	0,491	1,032	2370
201	1,010	1,021	0,995	1,012	3380
202	0,995	1,032	0,950	0,997	5813
203	1,005	1,209	0,927	1,018	1974

204	0,995	1,080	0,925	0,963	1553
205	0,967	0,992	0,903	0,969	3306
206	0,996	0,996	0,996	0,996	1537
207	1,018	1,018	1,018	1,018	281
208	1,004	1,010	0,993	0,994	779
209	0,981	1,016	0,914	0,968	2799
211	0,976	0,990	0,965	0,975	17365
213	0,959	0,994	0,417	0,946	785
215	0,968	1,002	0,930	0,989	4199
216	0,911	1,000	0,842	0,945	360
217	1,028	1,136	0,966	1,016	617
219	0,982	0,982	0,982	0,982	433
225	1,084	1,213	0,965	1,150	466
350	1,006	1,155	0,920	1,026	647
355	0,921	0,921	0,921	0,921	671
356	1,049	1,200	1,047	1,123	453
365	0,921	1,071	0,842	0,935	114
368	0,973	1,032	0,932	0,982	75
401	1,000	1,000	1,000	1,000	7
405	0,973	1,001	0,930	0,962	3115
406	0,981	1,283	0,877	0,986	1598
407	1,034	1,076	0,977	1,014	386
409	1,002	1,092	0,985	1,039	402
410	0,977	1,012	0,955	0,982	2195
411	0,978	1,000	0,968	0,984	535
419	1,072	1,153	1,002	1,082	1336
420	0,970	1,051	0,898	0,993	1225
425	0,973	1,036	0,950	0,972	1185
427	0,972	1,000	0,956	0,968	3515
430	0,951	0,955	0,946	0,953	2776
432	0,972	0,994	0,943	0,977	2292
435	0,981	1,149	0,957	0,982	7224
439	0,988	1,024	0,897	0,995	2525
440	0,976	1,024	0,931	0,953	613
444	1,011	1,099	0,943	0,995	456
448	0,996	1,037	0,979	1,000	817
449	1,005	1,063	0,976	1,000	2037
460	1,010	1,022	0,997	1,012	8811
467	0,963	0,987	0,962	0,975	5615
468	0,978	1,011	0,966	0,987	6839
469	0,999	1,012	0,986	0,999	998

470	0,966	1,019	0,945	0,970	13357
475	1,066	1,066	1,066	1,066	61
485	1,026	1,500	0,919	0,970	1055
487	1,006	1,032	0,945	0,981	681
491	0,956	0,981	0,929	0,955	2587
495	0,951	0,971	0,937	0,952	8375
496	1,049	1,396	0,986	1,191	345
715	0,984	1,202	0,945	0,969	798
747	0,905	0,950	0,854	0,920	3407
804	0,938	0,992	0,911	0,952	2320
877	0,884	0,899	0,864	0,881	2023
935	0,969	1,000	0,950	0,973	424

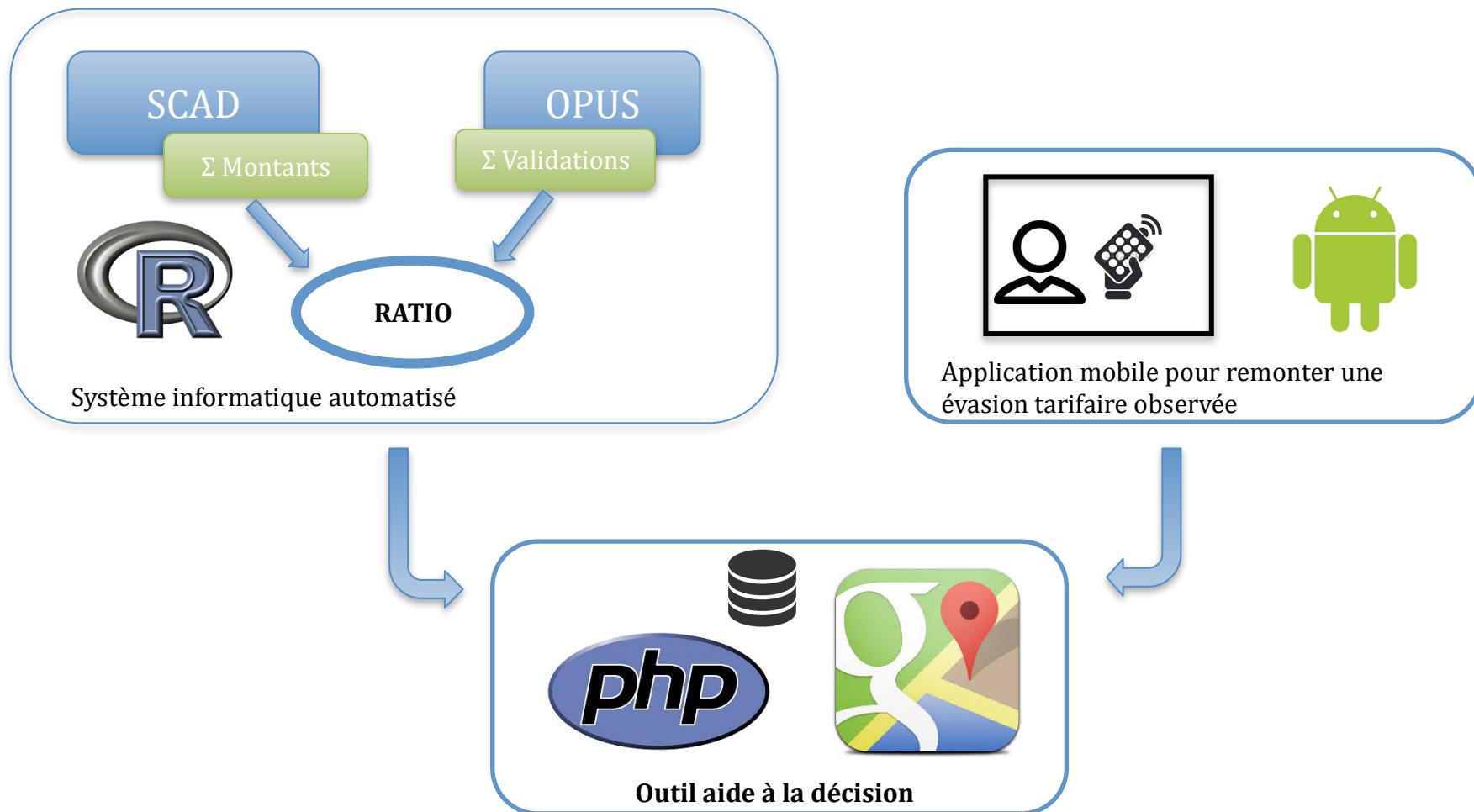
**ANNEXE C – RATIOS VALIDATIONS/MONTANTS PAR LIGNE SUR LA SEMAINE D'AVRIL 2014 (7
AU 13 AVRIL)**



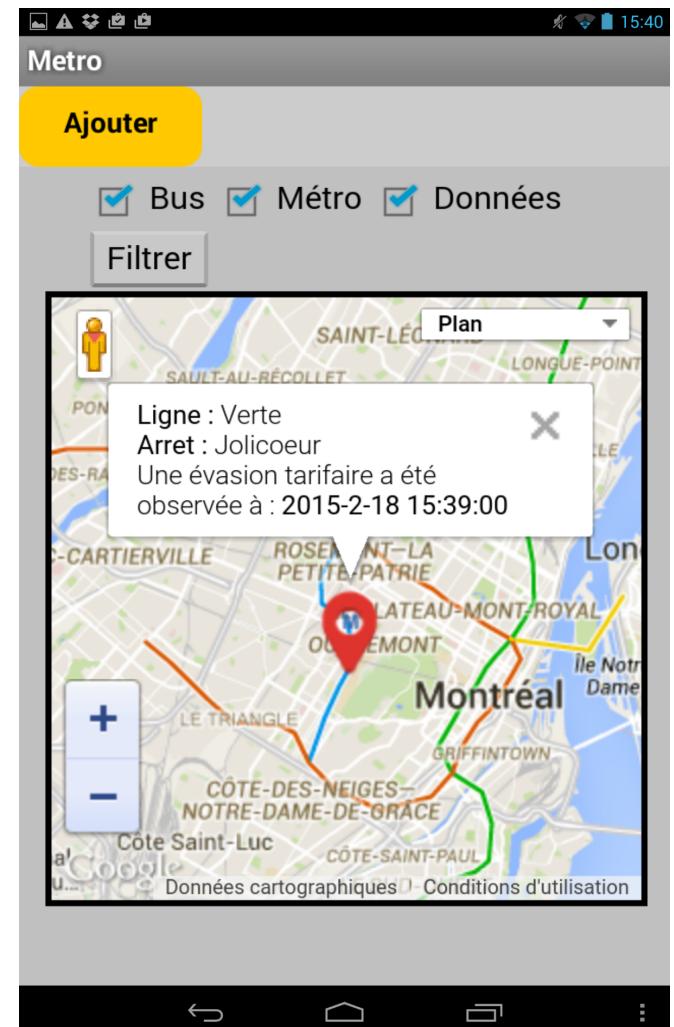
ANNEXE D – EXEMPLE FICHIER RÉSULTAT ANALYSE EXCEL

A	B	D	E	F	G	H	J	K	L	M	N	O
DATE_SORTIE	HEURE_DEP/ LAT_GPS	LONG_GPS	QUARTIERS	RTA	METRO	NUMERO_AF DESC_ARRET_CLIENT	NUMERO_LIGNE	NUMERO_VEHICULE	NOMBRE_VALIDATIONS	SOMME_MONTANTS		
2014-04-07	0,46483	45.519865	-73.665265	St. Laurent	H4N	50000 Gar. Stinson	105	22215	1	0		
2014-04-07	0,49856	45.473788	-73.604295	Cote des Neiges H4A	Station Vendome	126979 Station Vendome / Terminus	105	22215	18	19		
2014-04-07	0,50074	45.474146	-73.608166	Cote des Neiges H4A		110021 Decarie / Sherbrooke	105	22215	2	2		
2014-04-07	0,50166	45.47269	-73.611929	Cote des Neiges H4A		109602 Sherbrooke / Girouard	105	22215	2	2		
2014-04-07	0,50313	45.470373	-73.616199	Cote des Neiges H4A		109052 Sherbrooke / Wilson	105	22215	1	1		
2014-04-07	0,50644	45.462727	-73.630888	Cote des Neiges H4B		107102 Sherbrooke / Walkley	105	22215	2	2		
2014-04-07	0,51816	45.454502	-73.641368	Cote des Neiges H4B		105751 Gare Montreal-Ouest	105	22215	3	3		
2014-04-07	0,51843	45.45561	-73.640777	Cote des Neiges H4B		126934 Sherbrooke / Westmore	105	22215	1	1		
2014-04-07	0,51895	45.456422	-73.640193	Cote des Neiges H4B		105904 Sherbrooke / West Broadway	105	22215	1	1		
2014-04-07	0,51943	45.45747	-73.639004	Cote des Neiges H4B		106230 Sherbrooke / Campus Loyola	105	22215	3	2		
2014-04-07	0,51994	45.459412	-73.636752	Cote des Neiges H4B		106384 Sherbrooke / Coronation	105	22215	2	2		
2014-04-07	0,52095	45.460347	-73.635109	Cote des Neiges H4B		106594 Sherbrooke / Mariette	105	22215	5	5		
2014-04-07	0,52142	45.461561	-73.632889	Cote des Neiges H4B		106844 Sherbrooke / Park Row	105	22215	2	1		
2014-04-07	0,52189	45.462633	-73.630848	Cote des Neiges H4B		107104 Sherbrooke / Walkley	105	22215	2	2		
2014-04-07	0,52244	45.463445	-73.629278	Cote des Neiges H4B		107384 Sherbrooke / Cavendish	105	22215	3	1		
2014-04-07	0,52377	45.464783	-73.626775	Cote des Neiges H4B		107604 Sherbrooke / Benny	105	22215	4	6		
2014-04-07	0,52435	45.465816	-73.62478	Cote des Neiges H4B		107914 Sherbrooke / Madison	105	22215	4	4		
2014-04-07	0,52485	45.466515	-73.623386	Cote des Neiges H4B		108094 Sherbrooke / Grand Boulevard	105	22215	2	2		
2014-04-07	0,52563	45.467639	-73.62136	Cote des Neiges H4A		108344 Sherbrooke / Hingston	105	22215	5	6		
2014-04-07	0,52627	45.468863	-73.618839	Cote des Neiges H4A		108714 Sherbrooke / de Clifton	105	22215	4	4		
2014-04-07	0,52677	45.470206	-73.61635	Cote des Neiges H4A		109054 Sherbrooke / Wilson	105	22215	1	2		
2014-04-07	0,52726	45.471107	-73.614634	Cote des Neiges H4A		109224 Sherbrooke / Oxford	105	22215	2	2		
2014-04-07	0,53748	45.473788	-73.604295	Cote des Neiges H4A	Station Vendome	126979 Station Vendome / Terminus	105	22215	20	20		
2014-04-07	0,53917	45.474146	-73.608166	Cote des Neiges H4A		110021 Decarie / Sherbrooke	105	22215	1	1		
2014-04-07	0,54073	45.47269	-73.611929	Cote des Neiges H4A		109602 Sherbrooke / Girouard	105	22215	2	2		
2014-04-07	0,54296	45.470373	-73.616199	Cote des Neiges H4A		109052 Sherbrooke / Wilson	105	22215	6	6		
2014-04-07	0,54427	45.467872	-73.621413	Cote des Neiges H4A		108342 Sherbrooke / Hingston	105	22215	1	1		
2014-04-07	0,54487	45.466799	-73.623268	Cote des Neiges H4B		108092 Sherbrooke / Grand Boulevard	105	22215	0	1		
2014-04-07	0,54598	45.465023	-73.6267	Cote des Neiges H4B		107602 Sherbrooke / Benny	105	22215	2	2		
2014-04-07	0,54652	45.463884	-73.628882	Cote des Neiges H4B		107382 Sherbrooke / Cavendish	105	22215	1	1		
2014-04-07	0,54709	45.462727	-73.630888	Cote des Neiges H4B		107102 Sherbrooke / Walkley	105	22215	1	1		
2014-04-07	0,55274	45.45561	-73.640777	Cote des Neiges H4B		126934 Sherbrooke / Westmore	105	22215	2	2		
2014-04-07	0,55412	45.459412	-73.636752	Cote des Neiges H4B		106384 Sherbrooke / Coronation	105	22215	1	1		
2014-04-07	0,55449	45.460347	-73.635109	Cote des Neiges H4B		106594 Sherbrooke / Mariette	105	22215	2	2		

ANNEXE E – FONCTIONNEMENT DE L’OUTIL



ANNEXE F – IMPRESSIONS ECRAN OUTIL



ANNEXE G – RÉSUMÉ DES LANGAGES ET OUTILS UTILISÉS POUR LA RÉALISATION DE L'OUTIL DYNAMIQUE

Etapes	Langages	Version	Commentaires
Traitement	R	3.1.1	Logiciel de traitement de données et d'analyse statistique (Open Source)
Envoi des fichiers sur serveur	Terminal		Commande Unix pour envoyer des fichiers sur un serveur (fonction scp). Nécessite le téléchargement et l'installation de package.
Envoi des fichiers du serveur vers la base de données	Terminal		Commande unique pour se connecter au serveur (sshpass) et mysql. Nécessite le téléchargement et l'installation de package.
Création de l'interface web affichant la carte Google Maps et les marqueurs	HTML	5	Langage de balisage permettant de créer des pages web
	CSS	3	Langage de balisage permettant de donner du style aux pages
	Javascript, Jquery Ajax		Langage de programmation de scripts pour les pages web interactives (téléchargement de bibliothèques pour afficher carte Google Maps, les marqueurs, les clusters). Lien entre la carte google maps et la base de données

			via l'API Jquery ajax.
Création de la base de données	MYSQL		Système de gestion de base de données interrelationnelles (Open Source)
Communication avec la base de données	PHP	5.4.27	Langage de programmation côté serveur pour les communications entre les différentes pages et la transmission de variables ainsi que certaines liaisons avec la base de données.
Application mobile	App inventor		Module de prototypage fonctionnel d'application mobile Android (Open Source)