11e CONGRÈS INTERNATIONAL DE GÉNIE INDUSTRIEL – CIGI2015
Québec, Canada
26-28 octobre 2015

Vers la maîtrise de l'évasion tarifaire dans un réseau de transport collectif

HUGO POURMONET[1,2], SAMUEL BASSETTO[1,2], MARTIN TRÉPANIER[1,3]

[1] ÉCOLE POLYTECHNIQUE DE MONTRÉAL, Département de mathématiques et génie industriel
C.P. 6079, station Centre-Ville, Montréal, Québec, Canada H3C 3A7

[2] CIMAR-LAB

[3] CENTRE INTERUNIVERSITAIRE DE RECHERCHE SUR LES RÉSEAUX D'ENTREPRISE, LA LOGISTIQUE ET LE TRANSPORT (CIRRELT)

hugo.pourmonet@polymtl.ca , samuel.bassetto@polymtl.ca , martin.trepanier@polymtl.ca

Résumé **-** L'objectif principal de cette étude est d'exploiter les données de carte à puce et de comptage à bord des bus du réseau de transport de la Société de transport de Montréal (STM) afin de caractériser l'évasion tarifaire et développer un outil de contrôle. Ce phénomène cause plusieurs millions de dollars de pertes par an aux agences de transport. L'étude commence par une analyse fine des disparités entre les données provenant du système de perception par carte à puce d'une part, et du système de comptage des passagers d'autre part. Ces résultats sont ensuite intégrés dans un outil d'allocation dynamique des contrôles en fonction d'observations d'opportunités d'améliorations. L'étude de faisabilité a été réalisée sur des données réelles issues du système d'information de la STM.

Abstract **-** The main objective of this study is to exploit the smart card data and counting on buses of the Société de transport de Montréal (STM) public transportation network to analyze the phenomenon of fare evasion and provide a dedicated inspection allocation tool. Several million dollars are lost per year for transport agencies due to fare evasion. This study starts by analyzing row data to highlight a potential fare evasion locations. The second steps introduces these results into a dynamic allocation dispatcher based on improvement opportunities. The feasability study has been performed over real data issued from STM information system.

Mots clés **– cartes à puce / transport collectif / comptage à bord / évasion tarifaire**
Keywords **– smart card / public transit / passenger count / fare evasion**

1 INTRODUCTION

L'évasion tarifaire, lorsque l'on parle de transport en commun, concerne le fait d'utiliser ce service sans posséder de droit de passage ou en utilisant des tickets contrefaits. Pour les entreprises gérant ces services, ce phénomène est préoccupant [Diebel 1981]. En effet, l'évasion tarifaire implique à la fois des pertes financières, mais elle impacte également la sécurité [Clarke et al. 2009] et l'image des transports collectifs.

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 demeurent très conséquents. Basée sur des données provenant de 31 compagnies de transports en commun de 19 pays différents, représentant 300 000 employés et 20 milliards de passagers par an, l'étude de Bonfanti, G., & Wagenknecht, T. (2010) montre que la perte pour ces compagnies due à la fraude s'élève à 1 milliard d'euros par an. À Montréal, la fraude représente moins de 1% du volume de transport [STM 2014], mais elle atteint jusqu'à 27,7% dans des villes telles que Santiago, au Chili [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 2011.

Afin de lutter contre ce phénomène et d'adopter des politiques pertinentes, plusieurs études ont été réalisées de manière à savoir quel était le profil de ces fraudeurs.

Lorsqu'elle est technique, la fraude est volontaire; cependant la difficulté pour falsifier un ticket est bien plus faible qu'une carte à puce [Mayes, 2009]. Lorsqu'elle est non technique, elle peut être souhaitée ou non. Le fraudeur peut être opportuniste (une 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. Reddy et al. (2011) ont 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. Les auteurs ont mis en évidence qu'une saisonnalité existait pour ce phénomène avec une augmentation du taux d'évasion tarifaire en été. Ils montrent également 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. Ils expliquent cela par le fait que le fraudeur doit être moins à l'aise de frauder devant le regard des autres usagers. Cet aspect est intéressant pour nos travaux, car il pointe la possibilité de mettre en œuvre une démarche de surveillance par les usagers.

Afin de maîtriser ce phénomène, les compagnies de transport installent des dispositifs d'inspections et de contrôle d'accès, et doivent embaucher des inspecteurs. Par exemple, la Société de Transport de Montréal a déboursé 174M$ en 2008 pour se doter d'une technologie de carte à puces : une 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 13M$ seraient sauvés chaque année, en plus de faciliter le travail des conducteurs [STM 2014]. Les entreprises doivent alors trouver un équilibre pour que les systèmes mis en place pour réduire l'évasion tarifaire soient les mieux adaptés.

Des travaux de recherche 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 Bonfanti [2010] le nombre de fraudeurs est inversement proportionnel au nombre d'inspecteurs. La présence d'un contrôleur dans un bus ou près des 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. Afin d'augmenter les capacités d'inspection et raccourcir les délais de traitement d'action, certains auteurs ont proposé une approche de partenariats entre plusieurs institutions (police, justice, éducation) [Bonfanti et Wagenknecht 2010, Dauby et Kovacs 2006]. L'optimisation de l'allocation repose sur un modèle comportemental des fraudes, mais également de l'achalandage, journalier, hebdomadaire et annuel et ce sur les plans temporel et géographique.

Cependant, la création d'un modèle statistiquement fiable pour un réseau donné demeure un défi de grande ampleur. On peut lui préférer la fourniture d'une information synthétique, représentative de la fraude, également dynamique temporellement et adaptée spatialement permettant de s'engager vers une allocation intelligente (faire intervenir aux bons moments, aux bons endroits les inspecteurs), mais non-optimale. Nous évacuons ici le débat concernant les questions légales et éthiques pouvant émaner de l'exploitations de données fines sur la clientèle (Thomopoulos et al. 2015), qui, du reste, dans le cadre de cette étude, sont complètement anonymisées et maintenues confidentielles à leur niveau individuel.

C'est dans ce cadre que s'inscrit notre article de recherche. Nous avons cherché à utiliser les données disponibles et à les fusionner avec des informations de terrain, remontées par des observations ponctuelles. À la lueur des nouvelles technologies mobiles de communication massivement répandues, il est possible de bâtir à frais réduit des dispositifs d'information permettant de relever les points chauds, par tout utilisateur et d'informer les équipes d'intervention en temps quasi-réel. Nous pensons que ces informations alliées à une analyse des données issue des systèmes de comptage et de paiement constituent un intermédiaire intéressant pour les groupes d'inspection. Ils peuvent ainsi disposer d'une information pertinente, en attendant de disposer d'un modèle comportemental complet du réseau.

La suite de cet article présente en section 2, une analyse des données issue des systèmes de comptage à bords et de validation de cartes à puces. La section 3 analyse les résultats de manière détaillée. La section 4 présente l'outil de visualisation et d'enregistrement pour mieux aider les inspecteurs à choisir leurs destinations d'intervention.

2 MÉTHODOLOGIE ET DESCRIPTION DES DONNÉES

Cette partie a pour objectif de présenter le processus de traitement des données. L'étude s'appuie sur deux ensembles de données fournies par la STM: le comptage à bord et les transactions de cartes à puces. Deux semaines distinctes de ces deux sources de données sont disponibles : du 7 au 13 avril 2014, et du 6 au 12 octobre 2014.

2.1 Le système de comptage

La Société de Transport de Montréal compte 312 bus équipés du système de comptage à bord (*automated passenger counting system*, APC), sur les 1600 autobus qu'elle opère. Ce système est basé sur une technologie de capteurs infrarouges actifs – permettant de détecter la coupure d'un rayon- et passifs – permettant de détecter une source de chaleur [Boyle 2008]. L'ordinateur de bord va pouvoir récupérer l'ensemble des données des capteurs, pour les transmettre à ce qu'on appelle l'ordinateur du garage (centralisé). L'ordinateur de bord est équipé d'un module AVL (*Automatic Vehicle Location*) permettant de suivre le parcours du véhicule à l'aide d'un lecteur GPS et d'un odomètre comptant le nombre de kilomètres parcourus. 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.

2.2 Le système de carte à puces

Les données de carte à puce peuvent être exploitées pour trois *principaux* usages selon Pelletier et al. (2011) :

- Niveau stratégique : Les données peuvent être utilisées pour mieux cerner le comportement des utilisateurs [Morency et al., 2007] et [Bagchi et White, 2005]

- Niveau opérationnel : Les données peuvent être utilisées pour générer des tableaux de bords de performance pour l'amélioration de la conformité des horaires [Trépanier et al., 2009] ou sur les erreurs de transaction [Chapleau et Ch, 2007]
- Niveau tactique : Les données de validations (Date, Heure, Lieu) peuvent servir à générer des matrices origine-destination [Munizaga et al., 2012] ou permettre d'adapter en temps réel la coordination entre plusieurs modes de transports (bus et métros par exemple) [Bagchi et White, 2005].

Nous utiliserons ici les transactions provenant du système de cartes à puce sous un angle tactique, en examinant de façon fine chaque véhicule et chaque ligne. Il est à noter ici que le système de perception par cartes à puces de la STM enregistre non seulement les transactions par cartes, mais également celles liées à des tickets et à des paiements comptants.

2.3 Traitement des données

Pour le mois d'avril, le fichier du comptage à bord possède 782 765 enregistrements, tandis que les données de cartes à puce comptent 10 502 628 transactions. Les nombres sont du même ordre de grandeur pour la semaine d'octobre; celle-ci a été analysée afin de vérifier si le phénomène connaît une certaine stabilité entre deux périodes distinctes.

Le traitement des données consiste à relier les deux ensembles de données relativement volumineux (plusieurs dizaines de milliers d'enregistrements par jours) de manière à comparer le nombre d'usagers montants dans les bus et ceux qui valident. Cette comparaison est synthétisée dans un ratio. Une méthode similaire a déjà été testée sur un ensemble restreint par Lomone (2014). Le logiciel de traitement des données et d'analyses statistiques est (http://www.r-project.org). La jointure des tables issues du comptage à bord et des cartes à puce a été réalisée par véhicule et par ligne de bus, grâce à une synchronisation entre les dates des deux fichiers. Cette jointure a permis de calculer, pour chaque enregistrement du système de compte à bord (nombre de montants), le nombre de transactions de cartes à puce correspondant (nombre de validations), ce qui a permis de calculer le ratio validations/montants qui, idéalement, devrait être 1. Si le ratio est inférieur à 1, alors certains montants n'auront pas validé leur paiement (volontairement ou par action du chauffeur), et nous supposerons la présence du phénomène d'évasion tarifaire (fraude). Si le ratio est supérieur à 1 (ce qui est beaucoup plus rare), alors cela traduit un dysfonctionnement de l'un des deux système de comptage.

Après la jointure des deux sources de données, de nombreuses données aberrantes apparaissaient. Nous avons remarqué à plusieurs reprises un grand nombre de montants pour un arrêt, mais aucune validation à ce moment-là. Ces données aberrantes proviennent selon nous d'une défaillance du collecteur ou de sa désactivation par le conducteur, ce qui rendait 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 empêchent les analyses. Nous avons également enrichi les données en ajoutant pour chaque lieu de montée (arrêt) l'identifiant du quartier et la proximité à une station de métro, à l'aide du logiciel de système d'information géographique QGIS (http://www.qgis.org).

3 ANALYSES ET INTERPRÉTATIONS DES RÉSULTATS

Suite au prétraitement des deux fichiers, nous présentons dans cette partie les résultats et leurs analyses. Le lecteur excusera les auteurs d'avoir anonymisé les résultats, à la demande de la STM. Dans le texte, le constat est identifié par (C). L'analyse de ce constat est identifiée par (A). Les données sont analysées suivant les perspectives temporelles et géographiques. Pour chacune d'elle deux niveaux de granularité sont considérés. Au total, cela représente donc quatre perspectives d'analyse des données : sur les jours de la semaine, sur les heures du jour, sur les lignes de bus, sur les arrêts.

3.1 Jour de la semaine

La première analyse que nous allons aborder correspond au jour de la semaine. Nous avons analysé sur les deux semaines 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 montées à bord. La Figure 1 présente le ratio en fonction des différents jours de la semaine, pour les deux semaines étudiées.

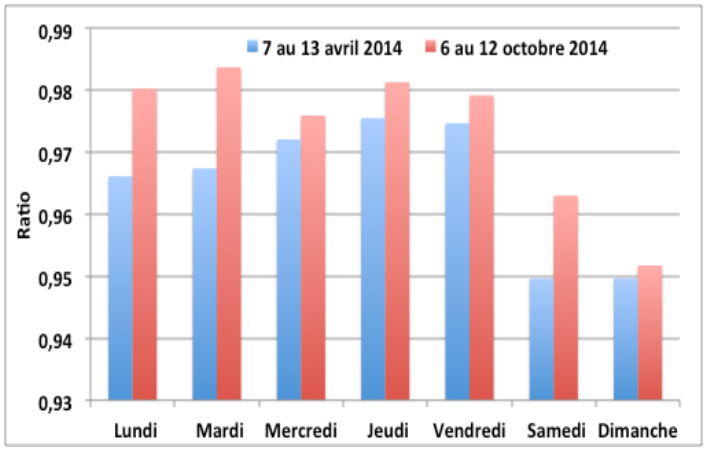

Figure 1 : Ratio comparatif Validations/Montants pour les deux semaines (7 au 13 avril 2014 et 6 au 12 avril 2014)

(C) On remarque que lors de la fin de semaine le ratio est plus faible que la semaine d'environ 1,5% (en points de pourcentage). 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. Il est également à 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.

(A) Il s'avère qu'à la STM, les enfants de moins de 12 ans ne paient pas les transports en commun la fin de semaine. Ce fait peut justifier à lui seul cet écart par rapport au reste de la semaine, et on ne peut donc pas vraiment tirer de conclusion sur le phénomène d'évasion tarifaire la fin de semaine. Pour la suite des analyses, nous éviterons donc de prendre en compte ces ratios des fins de semaine.

3.2 Heure de la journée

Les transports en commun sont soumis à des variations importantes de fréquentation durant la journée de semaine.

(C) En général, les heures de pointe correspondent au moment 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 de 15h à 18h. Le graphique de la Figure 2 représente le ratio validations/montants pour les jours de semaine (lundi au vendredi) d'avril et d'octobre

(7 au 11 avril 2014 et 6 au 10 octobre 2014). On remarque des ratios plus faibles aux heures creuses pour la semaine d'octobre, notamment le soir après 19h ou entre 13h et 14h.

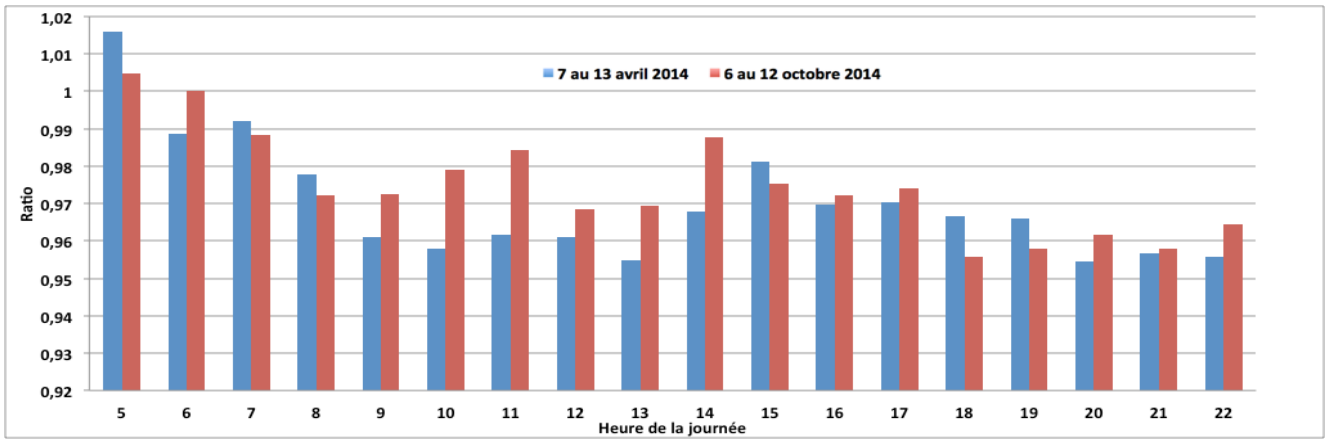

Figure 2 : Ratio Validations/Montants en fonction des heures de la journée pour les deux semaines

(A) Cette observation est confirmée par la littérature et notamment Reddy et al. (2011). Une explication avancée est la pression des utilisateurs et la pression plus grande à une conformation comportementale.

3.3 Lignes de bus

La fréquentation des lignes de bus est souvent bien différente en fonction de leur trajet et de leur niveau de service.

(C) Puisque l'évasion tarifaire peut être volontaire et quotidienne, nous caractérisons ce phénomène en fonction des différentes lignes qui composent le réseau. Nous recherchons à faire ressortir les lignes au faible ratio Validations/Montants et à leur récurrence dans le temps. Ce raisonnement permet de distinguer et de catégoriser les lignes de bus en fonction de leur risque potentiel d'évasion tarifaire. Le Tableau 1 nous montre les douze lignes ayant le plus faible ratio sur la semaine d'avril (les numéros de ligne ont été anonymisés pour fins de confidentialité).

Tableau 1 : Ratios les plus faibles pour la semaine d'avril (du 7 au 13 avril 2014)

Ligne	Ratio	Montants
Ligne régulière A	0,884	155
Navette métro A	0,884	2023
Ligne régulière B	0,889	415
Ligne touristique	0,901	1490
Ligne aéroport	0,905	3407
Ligne régulière C	0,911	360
Ligne régulière D	0,915	2047
Ligne régulière E	0,919	2727
Ligne nocturne A	0,921	671
Ligne nocturne B	0,921	114
Ligne régulière F	0,925	4995
Ligne régulière G	0,930	2005

(A) Pour pouvoir analyser ce tableau, il est nécessaire préalablement de 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. Les dix premiers ratios sont donc bien inférieurs à cette moyenne, et nous faisons l'hypothèse qu'au moins une partie de cet écart est due à l'évasion tarifaire. Nous ne pouvons pas décréter à l'aide de ces résultats que la ligne régulière A est la ligne la plus à risque de Montréal. Puisque seul 16% des bus sont équipés du système de comptage à bord sur l'île de Montréal, les affectations sont faites de façon régulière par les répartiteurs. Il s'avère que la ligne régulière A 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. À ce faible montant, les bas ratios peuvent être dûs aux montées des conducteurs en terminus. Il est donc nécessaire de 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 dans cette optique la que nous examinerons également les ratios pour les arrêts de bus.

Il y a par contre deux phénomènes qui ont pu être identifiés. Pour la ligne aéroport, il est vraisemblable que des passagers entrent et sortent du bus pour porter leurs bagages, d'où un possible double comptage et un ratio plus faible. Quant à la ligne touristique, le faible ratio peut être dû à l'utilisation de titres touristiques, qui sont plutôt rares, mais non enregistrés dans le système de perception par cartes à puce. L'analyse montre également que des lignes nocturnes sont sujettes à l'évasion tarifaire, possiblement à cause de la clientèle ou pour des raisons de sécurité. Somme toute, cette analyse permet de constater la sensibilité de la méthode et sur sa potentialité à détecter les non-validations.

3.4 Arrêts de bus

Dans cette partie, nous allons nous intéresser à la caractérisation de l'évasion tarifaire à travers les arrêts de bus. Pour cela, nous allons comparer le ratio Validations/Montants à court terme, c'est-à-dire entre les jours d'une même semaine et à long terme, entre les deux semaines d'avril et d'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 tous les véhicules ne sont pas équipés du système de comptage à bord, les lignes de faible affluence ne sont pas forcément parcourues par un bus équipé par jour. Pour certains arrêts, la comparaison avec les autres jours de la semaine est impossible puisqu'il manque d'observations. 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 à proximité des stations de métro. Nous avons choisi de regarder les 5 arrêts avec plus de 1500 montants ayant le plus faible ratio sur les 5 jours de la semaine et d'observer sur ce ratio était aussi faible sur l'ensemble des jours ou si c'était qu'un seul jour. Les résultats sont au Tableau 2 (les noms d'arrêts ont été anonymisés pour fins de confidentialité).

(C) L'arrêt situé près du centre commercial A présente un très faible ratio le lundi (0,8016) et un ratio proche de la moyenne (0,967) les autres jours.

(A) 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 (forte affluence qui empêche la validation ou panne d'équipement par exemple).

(C) À l'inverse, l'arrêt près du collège A présente pour sa part un ratio très faible 4 jours sur 5, tout comme l'arrêt situé près de la station de métro C.

(A) On peut donc faire l'hypothèse que l'évasion tarifaire a bien une composante récurrente, mise en évidence ici. Afin de tirer des analyses pertinentes, il conviendra de tenir compte du nombre d'observation et ainsi s'appuyer sur un nombre important d'enregistrement (1500 sur une semaine). En effet, par exemple nous n'avons enregistré aucune observation le mardi à l'arrêt situé près de la station de métro A.

Tableau 2 : Ratio Validations/Montants sur les 5 jours de la semaine d'avril

	ARRÊT COLLÈGE A	ARRÊT CENTRE COMMER-CIAL A	ARRÊT MÉTRO A	ARRÊT MÉTRO B	ARRÊT MÉTRO C
Lundi	0,949	0,802	0,858	0,876	0,890
	(118)	(620)	(436)	(614)	(499)
Mardi	0,912	0,989	X	0,895	0,893
	(479)	(180)	X	(929)	(261)
Mercredi	0,873	0,967	0,897	0,850	0,953
	(573)	(241)	(392)	(766)	(570)
Jeudi	0,871	0,952	0,987	0,924	0,936
	(851)	(125)	(508)	(1383)	(703)
Vendredi	0,864	0,961	0,871	0,984	0,865
	(213)	(411)	(294)	(543)	(920)
Lundi au Vendredi	0,884	0,902	0,903	0,905	0,906
	(2234)	(1577)	(1630)	(4235)	(2953)

L'autre partie de l'analyse de la vue – arrêts – porte sur la comparaison entre les faibles ratios entre 2 dates (avril et octobre). La présence d'un arrêt aux deux dates traduirait un indice de sensibilité de l'arrêt face à l'évasion tarifaire. Le tableau 3 présente les 20 arrêts ayant le plus faible ratio et plus de 1500 passagers pour chacune des semaines.

(C) Afin de construire ce tableau, 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 de Montréal (le territoire de la STM). En regardant ces 20 arrêts, on remarque très vite, grâce au code de couleur, que 7 se retrouvent à la fois en avril et en octobre, ce qui laisse supposer une stabilité longitudinale du phénomène à ces endroits.

(A) Tous ces arrêts ont des ratios relativement faibles comparés à l'ensemble du réseau. En outre, on retrouve dans le groupe des arrêts situés aux abords des mêmes stations de métro ou terminus intermodal. Cette petite analyse nous montre encore une fois que l'évasion tarifaire peut être mise en évidence ici de cette façon et qu'une forme de récurrence semble exister dans ce phénomène.

Tableau 3 : 20 arrêts ayant les plus faibles ratios pour les 2 semaines (entre parenthèses, le nombre de montants)

AVRIL		OCTOBRE	
Arrêt	**Ratio**	**Arrêt**	**Ratio**
Collège A	0,884	Autre arrêt près du terminus intermodal B	0,895
Centre commercial A	0,902	Métro C	0,899
Métro A	0,903	Arrêt A	0,912
Métro B	0,905	Métro K	0,916
Métro C	0,906	Métro D	0,916
Métro D	0,910	Métro O	0,923
Centre commercial B	0,910	Métro N	0,924
Métro E	0,918	Métro P	0,929
Métro F	0,920	Métro E	0,930
Autre arrêt métro F	0,924	Arrêt B	0,930
Terminus intermodal A	0,924	Collège A	0,932
Métro G	0,925	Métro Q	0,934
Métro H	0,927	Centre commercial A	0,935
Métro I	0,927	Métro F	0,938
Métro J	0,927	Autre arrêt près du métro O	0,938
Métro K	0,928	Métro N	0,939
Terminus intermodal B	0,931	Arrêt C	0,939
Métro L	0,931	Arrêt D	0,940
Métro M	0,932	Métro K	0,941
Métro N	0,933	Métro R	0,941

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, les bus ne sont pas tous équipés du système de comptage à bord ce qui rend l'analyse et l'interprétation plus complexe. L'analyse comparée nous a permis de souligner 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 là-dessus que s'appuie l'outil d'aide à la décision.

4 Outil d'aide à la décision pour mieux contrôler l'évasion tarifaire

Dans cette section, 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 nous baser sur les résultats issus de l'analyse précédente 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 mentionné précédemment, des 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 ne valident pas leur titre 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 ces irrégularités.

4.1 Présentation et fonctionnement de l'outil

Nous parlons volontairement d'outil dynamique, car le but est de s'appuyer sur le code R issu du traitement des données pour automatiser le processus. Une fois les données traitées, une extraction des arrêts à risque est fait dans un fichier permettant une consultation relativement simple. Ce fichier est envoyé sur un serveur et ensuite chargé dans une base de données sécurisée en ligne qui nous permet de représenter ces arrêts sur un fond de carte Google Maps (http://maps.google.com).

La Figure 3 résume le processus de fonctionnement de l'outil.

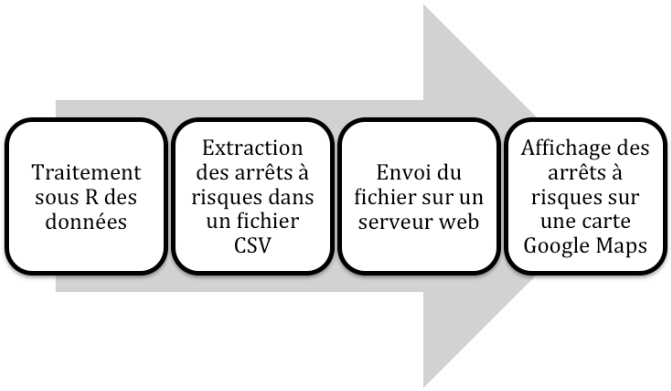

Figure 3 : Processus de fonctionnement de l'outil

Un processus d'automatisation via des commandes UNIX permet de rendre l'outil très facile à employer. La visualisation sur la carte se fait par le biais d'un navigateur web et peut donc être consultable en tout temps sur un ordinateur, une tablette ou un téléphone intelligent. À titre d'exemple, la carte de la Figure 4 représente des arrêts à risques de la semaine d'avril 2014.

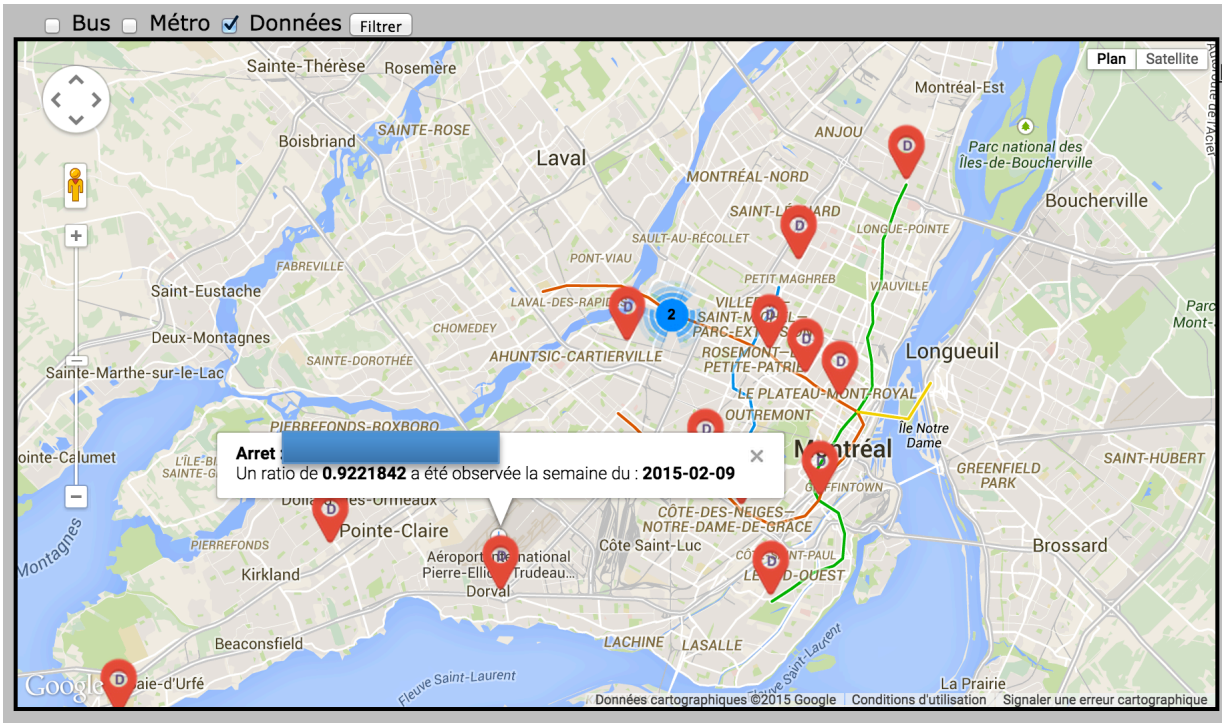

Figure 4 : Exemple de visualisation des arrêts à risque

La quantité d'information à remonter est paramétrable, mais doit être soigneusement dosée afin de ne pas surcharger d'information les inspecteurs. L'affichage dynamique de Google permet de concentrer le nombre d'observations selon le niveau d'échelle affiché.

4.2 Extension pour l'évasion tarifaire observée

En plus de l'évasion tarifaire estimée, nous proposons d'étendre l'outil aux évasions observées. À cette fin, nous avons développé un prototype fonctionnel 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 avec un marqueur différent.

Le fonctionnement de l'application est simple pour une meilleure efficacité. L'utilisateur choisit d'abord s'il se trouve dans un bus ou dans le 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. Les données envoyées sont alors complétées par les

informations circonstancielles disponibles sur l'appareil (date, heure, lieu). Cela permet alors de mettre en vis-à-vis les évasions tarifaires réellement observées dans les bus et les informations remontées par le traitement des données. L'application a été développée sous Androïd, en utilisant le code AppInventor (http://mit.appinventor.edu). Le code est disponible sur demande aux auteurs.

4.3 Utilisation de l'outil et limitations

Nous voyons cet outil comme un complément pour les inspecteurs. Les informations récoltées pourront aider à la planification des inspections de la journée. De plus, puisque sa consultation est relativement facile à l'aide d'un ordinateur ou d'un téléphone intelligent, l'outil pourrait servir de sonnette d'alarme dans une utilisation temps réel. Grâce à un tel outil, la planification des inspections pourrait donc être facilitée, et leur réactivité sur le terrain améliorée.

Limitation 1 : Le profilage : Afin de se prémunir du phénomène de profilage que pourrait générer un tel outil, nous préconisons de séparer les ressources d'inspection en deux : une partie utilisatrice de ces services, qui ferait des interventions ciblées et l'autre non, qui effectuerait des patrouilles aléatoires. La partie utilisatrice pourrait employer cet outil de manière réactive face aux observations relevées ainsi que les faibles ratios.

Limitation 2 : le temps de déplacement des inspecteurs : pour les bus, nous pouvons imaginer que si une évasion est notifiée sur un arrêt quelconque sur la ligne, les inspecteurs devrait alors se déplacer vers l'arrêt où il y a le plus de descentes (souvent le terminus, proche d'un métro par exemple). Concernant les métros, la réactivité parait plus complexe pour arriver à intercepter le ou les fraudeur(s). 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.

La figure 5 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 mis 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.

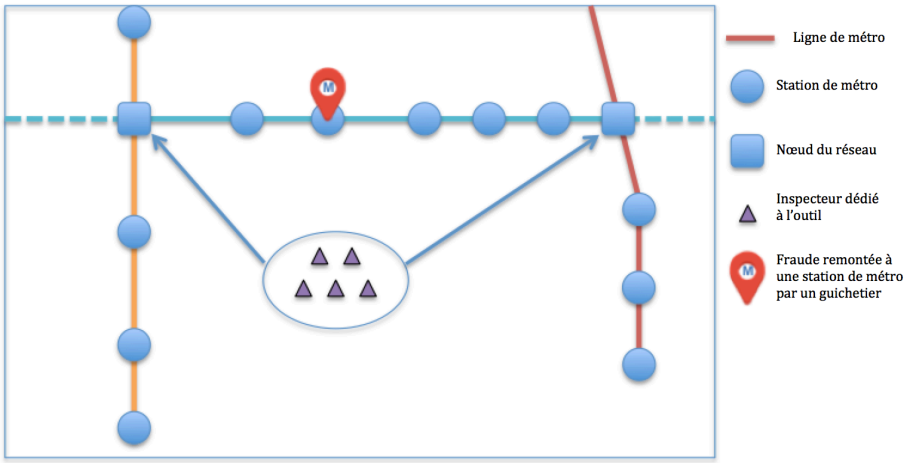

Figure 5 : Utilisation de l'outil dans le métro

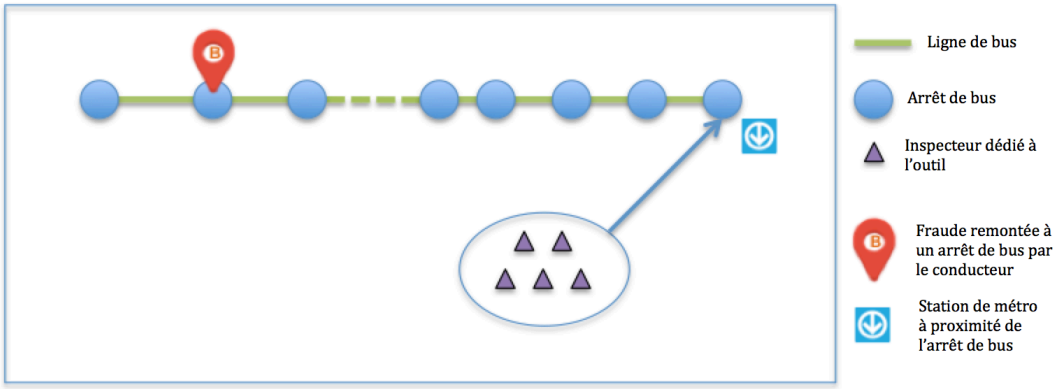

Figure 6 : Utilisation de l'outil dans les bus

La figure 6 schématise l'utilisation pour les bus. Le principe serait le même que pour les métros et vise à cibler les arrêts où le fraudeur est le plus à même de descendre.

Limite 3 : Seuil d'intervention. Ces cas d'usage portent sur des cas isolés. Cependant, la capacité d'intervention n'étant pas infinie, la réaction ne peut pas être basée sur chaque observation. Cependant, étant donné que toutes les observations sont sauvegardées dans la base de données, cela permettrait dans une future étude de définir un seuil de risque pour une zone (arrêt, ligne, point d'intérêt) intégrant faibles ratios et observations de fraudes. Cela permettrait également d'identifier des patrons possibles d'évasion tarifaire.

5 Conclusion

Cette étude 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. Nous avons pu étudier les différents paramètres disponibles dans les données afin de mieux comprendre ce phénomène et de voir s'il était caractérisable. Il s'avère que l'on arrive à remonter des ratios validations/montants bien plus faible que la moyenne sur certains arrêts ou lignes de bus. Les jours ainsi que les heures de la journée sont aussi soumis à des variations de ce ratio.

Ce travail comporte cependant des limites. Comme nous l'avons dit auparavant, nous ne pouvons pas assurer que ce l'écart entre les validations et les montants est essentiellement dû à l'évasion tarifaire, mais nous pensons tout de même que cela en fait un bon indicateur. De plus, notre étude a été réalisée sur un réseau qui n'équipe pour le moment que 16% de ces bus avec le système de comptage à bord. Cela génère cependant assez de données pour que les résultats soient pertinents, mais il serait souhaitable d'atteindre 100% des observations, comme dans le cas du système de cartes à puce. Enfin, grâce à ces résultats, nous avons jugé utile d'étendre la recherche à la création d'un outil dynamique d'aide à la décision pour les inspecteurs des réseaux de transports en communs.

Parmi les perspectives de recherche, citons la possibilité de croiser ces données avec des informations exogènes sur les contraventions et les rapports d'inspection. Nous pensons que l'avenir réside dans les outils qui pourront utiliser et mettre en commun l'ensemble des données disponibles chez l'opérateur.

6 Remerciements

Nous tenons à remercier particulièrement Jocelyn Grondines et Jean-François Cantin de la Société de Transport de Montréal pour leur étroite collaboration. Nous soulignons également le soutien financier du groupe Thalès et du Conseil de recherche en sciences et en génie du Canada (CRSNG, projet RDCPJ 446107-12).

7 Références

Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. Transport Policy, 12(5), 464–474.

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.

Bonfanti, G., & Wagenknecht, T. (2010). Human Factors Reduce Aggression and Fare Evasion. Public Transport International, 59(1).

Boyle, D. K. (2008). Passenger Counting Systems. Transportation Research Board.

Casey, R. F., John A. Volpe National Transportation Systems Center (U. S.). Office of System and Economic Assessment, & United States. Federal Transit Administration. Office of Mobility Innovation, Washington, DC, (1999).

Chapleau, R., K. K. Chu. (2007). Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach. 11th World Conference on Transportation Research.

Clarke, R. V., Contre, S., & Petrossian, G. (2009). Deterrence and fare evasion: Results of a natural experiment. Security Journal, 23(1), 5–17.

Dauby, L., & Kovacs, Z. (2006). Fare Evasion in Light Rail Systems : A World of Applications and Opportunities. Joint International Light Rail Conference, 230–246.

Diebel, L. E. (1981). The introduction of European fare collection techniques in United States. Journal of Advanced Transportation, pp. 55–60.

Lomone, A. (2014, February). Exploration et traitement multidonnées appliqués à des corridors d'autobus (masters). École Polytechnique de Montréal.

Lu, H. K. (2007). Network Smart Card Review and Analysis. Comput. Netw., 51(9), 2234–2248.

Mayes, K. E., Markantonakis, K., & Hancke, G. (2009). Transport ticketing security and fraud controls. Information Security Technical Report, 14(2), 87–95.

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.

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.

Reddy, A. V., Kuhls, J., & Lu, A. (2011). Measuring and Controlling Subway Fare Evasion: Improving Safety and Security at New York City Transit Authority. Transportation Research Record, (2216), 85–99.

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

Société de transport de Montréal (2014). Rapport annuel.

Thomopoulos, N., Givoni, M., Rietveld, P. (2015). ICT for Transport: Opportunities and Threats, Edward Elgar Publishing, 29 mai 2015 - 336 pages.

Thorlacius, P., Clausen, J., Thorlacius, P., & Clausen, J. (2010). Scheduling of inspectors for ticket spot checking in urban rail transportation. Trafikdage På Aalborg Universitet 2008.

Torres Montoya, M. (2013). Tackling fare evasion in Transantiago: an integrated approach.

Trépanier M., Morency C., Agard B. (2009), Calculation of transit performance measures using smartcard data, Journal of Public Transportation, 12(1), 79-96.