

Université de Montréal

**Réseaux sociaux et renseignement policier :
Facebook, quel apport pour les données policières ?**

par Marine Freyne

École de Criminologie
Faculté des Arts et des Sciences

Rapport de stage présenté
en vue de l'obtention du grade de Maîtrise
en Criminologie
option stage analyse

Avril, 2017

© Marine Freyne, 2017

Résumé

Les données disponibles sur les médias sociaux présentent une opportunité unique de pallier les limites des données officielles, notamment dans un contexte d'analyse de réseau et de renseignement. Les réseaux sociaux de dix individus, réputés appartenir à un groupe de motards criminalisés, ont été construits à partir à la fois des données policières et des liens d'amitié publics de Facebook. Une analyse comparative de la structure des réseaux, des indicateurs de centralité des individus ainsi que de leur rôle, est effectuée. L'influence que les caractéristiques individuelles telles que le lieu de résidence ou le secteur d'activité professionnelle, peuvent avoir sur l'appartenance de certains individus aux deux sources de données, ainsi que sur le rôle qu'ils occupent, est aussi analysée. Il en ressort que l'apport de Facebook est relativement faible, pouvant ajouter du bruit et disperser les résultats. L'information pertinente qui y est disponible est similaire à celle déjà présente dans les données policières. Les données policières sont donc une première source fiable et complète d'informations, bien que Facebook puisse être un complément s'il en est fait un usage ciblé.

Mots-clés : Facebook, médias sociaux, groupes de motards criminalisés, renseignement policier, analyse de réseaux

Abstract

Data available online on social networking sites represent a great opportunity to solve the limits of official data in a context of social network analysis and police intelligence. The social network of ten criminalized individuals, reputed to be members of an outlaw motorcycle club (OMC) have been built from public Facebook's friendships and official police data simultaneously. A comparative analysis of the network structure, the evolution of the centrality measures among the individuals depending of the network and their position and role, has been realized. The influence of individual characteristics like the residence place or the professional sector, on the belonging of some of them to both sources, and on their role and position in the different networks, has been analyzed too. It results in a low contribution of Facebook, with some noise and dispersion of the results. The relevant information available on Facebook is similar to that available in police data. Police data are therefore a complete and reliable source of information, even if Facebook could bring more when specifically used and targeted.

Keywords : Facebook, social networking sites, outlaw motorcycle clubs, police intelligence, social network analysis

Table des matières

Résumé.....	i
Abstract.....	ii
Table des matières.....	iii
Liste des tableaux.....	v
Liste des figures.....	vi
Liste des sigles.....	vii
Remerciements.....	viii
Introduction.....	1
1. Revue de littérature.....	3
1.1. Analyse de réseaux d’individus criminalisés et données policières.....	3
1.1.1. Pertinence et intérêt de l’analyse de réseau dans le milieu policier.....	3
1.1.2. Application aux réseaux de codélinquance.....	7
1.1.3. Le cas des groupes de motards criminalisés.....	9
1.1.4. Les données policières : source fiable, mais limitée.....	12
1.2. Le potentiel de l’usage des sources ouvertes : les médias sociaux.....	14
1.2.1. Caractéristiques.....	15
1.2.2. Police et médias sociaux.....	20
2. Problématique.....	26
3. Méthodologie.....	28
3.1. Données et opérationnalisation.....	28
3.1.1. Données issues du réseau social Facebook.....	29
3.1.2. Données issues des banques policières.....	33
3.1.3. Le réseau combinant simultanément les deux sources de données.....	35
3.1.4. Les réseaux des individus présents simultanément dans les deux sources de données.....	38
3.2. Analyses.....	43

3.2.1. Analyse comparative des différents réseaux.....	43
3.2.2. Analyse de l'impact des caractéristiques individuelles dans les différents réseaux	46
4. Analyse et interprétation des résultats	49
4.1. Analyse comparative des différents réseaux.....	49
4.1.1. Analyse comparative des réseaux comprenant l'ensemble des individus	50
4.1.2. Analyse comparative des différents réseaux comprenant uniquement les individus présents dans les deux sources de données.....	57
4.2. Impact des caractéristiques individuelles sur l'appartenance aux réseaux et le rôle occupé.....	64
4.2.1. Influence des caractéristiques individuelles sur l'appartenance aux différents réseaux	65
4.2.2. Influence des caractéristiques individuelles sur le rôle et la position occupés.....	67
Conclusion	71
Bibliographie.....	i
ANNEXE 1 : Tableau récapitulatif des variables utilisées au cours des différentes analyses	i
ANNEXE 2 : Classification des activités professionnelles par secteur d'emploi basée sur la classification des activités économiques du CIDREQ.....	ii
ANNEXE 3 : Représentation graphique du réseau total comprenant l'ensemble des individus (logiciel Gephi).....	iii
ANNEXE 4 : Statistiques descriptives des différents indicateurs de centralité selon les différents réseaux comprenant l'ensemble des individus	iv
ANNEXE 5 : Statistiques descriptives des différents indicateurs de centralité selon les différents réseaux comprenant uniquement les individus présents dans les deux sources de données	v

Liste des tableaux

Tableau I.	Classement des dix individus impliqués dans le plus d'évènements.....	34
Tableau II.	Portrait de l'échantillon du réseau total	36
Tableau III.	Portrait de l'échantillon du réseau total des individus présents dans les deux sources de données.....	41
Tableau IV.	Fréquence d'apparition dans les différents TOP 15 des réseaux comprenant l'ensemble des individus.....	54
Tableau V.	TOP 15 des acteurs centraux de chaque réseau comprenant l'ensemble des individus par indicateur de centralité.....	55
Tableau VI.	Fréquence d'apparition dans les différents TOP 15 des réseaux comprenant l'ensemble des individus présents dans les deux sources de données	61
Tableau VII.	TOP 15 des acteurs centraux de chaque réseau comprenant les individus présents dans les deux sources de données par indicateur de centralité.	62
Tableau VIII.	Matrice de corrélation entre les différents réseaux.....	64
Tableau IX.	Prédicteurs d'appartenance aux deux sources de données.....	65
Tableau X.	Prédicteurs d'une position influente au sein des réseaux (vecteur propre).....	68

Liste des figures

Figure 1.	Schéma récapitulatif de la démarche adoptée pour construire les différents réseaux	
		29
Figure 2.	Comparaison des indicateurs de centralité de chacun des réseaux comprenant l'ensemble des acteurs	51
Figure 3.	Comparaison des indicateurs de centralité des acteurs présents dans le TOP 15 des réseaux comprenant l'ensemble des acteurs	52
Figure 4.	Comparaison des indicateurs de centralité de chacun des réseaux comprenant les acteurs présents dans les deux sources de données.....	58
Figure 5.	Comparaison des indicateurs de centralité des acteurs présents dans le TOP 15 des réseaux des acteurs présents dans les deux sources de données	59

Liste des sigles

CIDREQ : Centre informatique du registre des entreprises du Québec

HAMC : Hells Angels Motorcycle Club

IRC : Internet Chat Relay

OMCs : Outlaw Motorcycle Clubs

RMR : Région Métropolitaine de Recensement

SPVM : Service de Police de la Ville de Montréal

Remerciements

Le présent projet n'aurait pu être réalisé sans le soutien de nombreuses personnes.

À commencer par mon directeur de recherche, Rémi Boivin. C'est avec beaucoup de plaisir que j'ai pu apprendre et évoluer à tes côtés depuis près de deux ans. Exigeant lorsqu'il le fallait pour que je m'améliore et me surpasse, tu as aussi su te montrer bienveillant et disponible dans les moments plus difficiles, notamment sur le plan personnel, et je ne saurais assez te remercier pour tout ce que tu as pu faire pour m'aider.

Je remercie également le SPVM, plus particulièrement la Section Recherche et Planification et Michelle Côté qui m'a permis d'y réaliser mon stage. Je remercie toute l'équipe, qui m'a donné une place à part entière, et aux côtés de laquelle j'ai réellement pu apprendre et m'épanouir. Outre les connaissances en recherche, vous avez, chacun à votre manière, contribué à la réussite de ce projet : le party de Noël chez Valérie, l'écoute attentive et les chocolats du vendredi de Karine, les pauses cafés de Mathieu et Maurizio, l'horaire précis de Gilbert permettant de se souvenir de l'heure du diner, le pop corn magique de Romilda, les lunchs riches en anecdotes, tranches de vie et recettes avec notamment Julie, Anne, ou encore Joséphine, sans oublier le soutien moral indéfectible et les conseils précieux de Stéfanie. Je remercie également les autres stagiaires, notamment J-P et David, avec qui j'ai partagé le stress, le doute mais aussi de supers moments. Mais c'est sans aucun doute envers Véronique Chadillon-Farinacci, ma superviseure, que je suis la plus reconnaissante. Aucun mot n'est suffisant pour décrire ce qu'on a partagé et vécu pendant ces quelques mois. Tu m'as tout appris. Tu as toujours été disponible, dans les meilleurs moments comme dans les plus difficiles, tant professionnellement que personnellement. Tu m'as permis de réaliser un projet passionnant et de progresser en me guidant au quotidien, tout en me permettant de m'appropriier le projet et d'aller à mon rythme. J'ai trouvé bien plus qu'une superviseure. Grâce à toi, on a réussi à réaliser ce projet, « step by step »...

Évidemment, je remercie ma famille et mes amis, en particulier ma mère pour les nombreuses heures passées à me soutenir, me rassurer et me relire, ainsi que mon père pour m'avoir transmis sa passion pour son métier de policier. Vous me permettez de réaliser mes rêves, même si cela implique qu'un océan nous sépare.

Introduction

L'application des techniques de l'analyse de réseau à des groupes d'individus criminalisés s'est développée au fil du temps. De plus en plus adoptée dans les études, sa pertinence et l'intérêt qu'elle représente pour le milieu policier ont été démontrés. Elles peuvent notamment pallier certaines limites inhérentes aux données policières, telles que leur caractère incomplet, ou guider l'interprétation qui doit en être faite. L'analyse de réseau permet en effet de révéler une structure particulière ou plus complexe, ainsi que certains acteurs. Si leurs rôles et positions peuvent apparaître aux premiers abords secondaires, ces acteurs se révèlent être importants et essentiels. Ces outils d'analyse sont particulièrement pertinents pour le renseignement policier. Ces méthodes ont alors été étendues à d'autres objets d'études, tels les réseaux de codélinquance ou, plus particulièrement, dans le cadre de groupes de motards criminalisés. Toutefois, les données policières présentant leurs limites, même lorsqu'il est fait usage de l'analyse de réseau, il devient nécessaire d'exploiter d'autres sources d'information. L'analyse de réseaux offre un cadre d'analyse suffisamment souple et ces méthodes peuvent s'appliquer à de nouveaux sujets et champs de recherche, notamment centrés sur l'utilisation d'Internet et l'existence de communautés virtuelles. Cela permet ainsi la production et l'analyse de jeux de données différents. Parmi les sources potentielles de renseignement figurent les médias sociaux, plus particulièrement Facebook qui compte plus d'un milliard d'utilisateurs. Ces plateformes connectent de nombreux individus mais génèrent également une quantité importante de données, tant individuelles que relationnelles, la plupart des utilisateurs ne les protégeant pas par des paramètres de sécurité. Elles constituent par conséquent une opportunité unique, tant vis-à-vis de leurs caractéristiques, que des possibilités d'utilisations s'offrant aux organisations policières. Il s'agit en effet d'outils librement accessibles, contenant de l'information de nature publique, utilisables pour la communication, le renseignement ou encore les enquêtes.

Toutefois, la question se pose de savoir si les données disponibles en ligne sur Facebook apportent effectivement une information complémentaire utile et nécessaire aux données policières. Ainsi, le défi du présent rapport est d'utiliser simultanément ces deux sources de données. Il s'agit de les utiliser dans la construction et l'analyse de réseau d'individus criminalisés réputés appartenir à un groupe de motards criminalisés, ainsi que de déterminer l'apport qu'une source telle que Facebook peut offrir dans une optique de renseignement. Il sera alors question de l'application de l'analyse de réseaux dans ces deux sphères différentes, l'une appartenant au monde réel, l'autre au monde virtuel. Plus particulièrement, il s'agira de déterminer s'il est pertinent d'utiliser Facebook en renseignement et d'en préciser l'utilisation, à savoir si cette plateforme et notamment les profils qu'elle contient doivent être consultés systématiquement, au-delà de toute considération de ressources et de temps, ou bien s'ils doivent être utilisés de manière plus spécifique et ciblée.

Il sera tout d'abord présenté l'utilisation de ces deux outils par les agences d'application de la loi, à savoir l'analyse de réseau d'une part, et les réseaux sociaux d'autre part, plus particulièrement dans une optique de renseignement, telle qu'établie dans de précédents travaux de recherches. Les données utilisées, les démarches suivies pour la construction des différents réseaux ainsi que les analyses effectuées seront ensuite présentées. Enfin, les résultats portant sur différents aspects et points de vue du réseau et des informations qu'il peut être tirées de Facebook seront analysés. Le présent rapport se terminera par une brève discussion autour des résultats qui méritent d'être retenus et des implications pratiques qu'ils engendrent, en termes de pratiques policières et de renseignement.

1. Revue de littérature

1.1. Analyse de réseaux d'individus criminalisés et données policières

L'application des techniques de l'analyse de réseau à des groupes d'individus criminalisés s'est développée au fil du temps. Sa pertinence et son intérêt pour le milieu policier ont été démontrés par plusieurs études, notamment quant au fait de pouvoir pallier certaines limites inhérentes aux données policières. Ces méthodes sont utilisées pour l'étude de phénomènes criminels, tels les réseaux de codélinquance ou, plus particulièrement, les groupes de motards criminalisés.

1.1.1. Pertinence et intérêt de l'analyse de réseau dans le milieu policier

L'application de l'analyse de réseau trouve une résonance particulière dans le milieu policier. Basée sur les relations unissant des individus, elle offre une autre vision du phénomène criminel. Il s'agit ainsi d'un outil utile aux organisations policières, notamment dans le cadre d'enquêtes.

a) Un outil offrant une perspective différente

Le point de départ de l'analyse de réseau implique que les individus ont des relations avec d'autres, illustrant leur interdépendance (McGloin et Kirk, 2010). Les réseaux sociaux sont définis comme un ensemble de nœuds ou membres, liés par un ou plusieurs types de relations (Wasserman et Faust, 1994). Le comportement social passe

principalement par les liens et relations sociales (Van der Hulst, 2008 ; Bhat et Abulaish, 2013), tels que des liens de familles ou d'amitié, d'affinités, ou de partage de ressources. Ils peuvent prendre la forme d'une co-adhésion, d'une communication ou d'un comportement partagé (Morselli, 2009 ; McGloin et Kirk, 2010 ; Bhat et Abulaish, 2013).

Les interactions fournissent tous les types d'opportunités et de ressources pour que les choses soient accomplies (Van der Hulst, 2008). Selon la théorie du capital social, les individus ayant un meilleur accès à des ressources sociales utiles, notamment par la mobilisation de leur réseau de relations, réussissent mieux dans ce qu'ils entreprennent (notamment Bourdieu, 1986 ; Coleman, 1990 ; Lin, 2001). Le capital social associé aux réseaux est l'une des premières façons de faciliter le passage à l'acte (Van der Hulst, 2008). La participation au crime requiert la capacité de réagir rapidement. Les réseaux sont des systèmes organisationnels permettant cela. Il peut s'agir de simples décisions de codélinquance pour saisir une opportunité criminelle disponible et attractive, ou bien de modèles sophistiqués pour monopoliser un marché ou un territoire géographique donné. La clé est d'être et de rester positionné de manière flexible (Morselli, 2009). Certains acteurs se démarqueront ainsi par leur visibilité au sein du réseau, caractérisée par une centralité de degré élevée. Ils ont une proportion importante de contacts au sein du réseau avec lesquels ils sont directement connectés (Freeman, 1979 ; Morselli, 2009). D'autres, se montreront plus discrets mais stratégiques, occupant la place d'intermédiaires ou de courtiers entre des individus non connectés. Cela procure un avantage important dans le contrôle de l'information ou du flux de communication (Freeman, 1977, 1979 ; Morselli, 2009). Enfin, d'autres encore préféreront être directement connectés à des individus ayant eux-mêmes de nombreux contacts directs, ayant ainsi une forte centralité de vecteur propre, signe d'influence (Bonacich, 1972 ; Morselli, 2009).

L'analyse systématique des réseaux criminels est donc une manière viable d'obtenir une meilleure compréhension du comportement criminel (Van der Hulst, 2008). Ainsi, dans le contexte de l'analyse de gangs de rue, l'avancement des connaissances à propos de la structure de groupes criminels a été aidé par l'utilisation des techniques de

l'analyse de réseau et de représentation graphique (McGloin et Kirk, 2010). Cela a permis notamment de démontrer que le crime organisé a une structure moins hiérarchique que ce qui pouvait être affirmé par le passé, mais beaucoup plus souple et caractérisée par un tissu logistique dans lequel plusieurs petits réseaux sont actifs (Klerks, 2001). La nature dynamique des relations parmi les membres de gang de rue est un élément définissant l'organisation et pouvant exercer une solide influence sur le comportement du groupe, collectivement et individuellement (McGloin et Kirk, 2010). Par exemple, le meurtre de gang serait mieux compris en examinant les réseaux sociaux de l'action et de la réaction qui le créent plutôt que ses déterminants individuels. Cela consiste en la manière dont ils sont construits et le placement des individus, ce qui génère un réseau institutionnalisé de conflit de groupe et qui se propage par un processus de contagion sociale (Papachristos, 2009). Ces analyses peuvent alors avoir un intérêt particulier pour le milieu policier.

b) Un outil utile pour les agences d'application de la loi

Martineau et Lecocq (2016) ont identifié plusieurs défis étroitement liés à la réalité du domaine du renseignement, notamment un environnement opérationnel dynamique, divers et complexe, la gestion de différentes sources et leur fiabilité, la crédibilité des informations, les données manquantes ou la surabondance d'informations.

La pertinence d'appliquer les techniques de l'analyse de réseau et ses outils dans le cadre des agences d'application de la loi et du renseignement a été démontrée depuis plusieurs années, partant du constat que ce type d'analyses pourrait permettre d'extraire de l'information pertinente et d'exploiter les nombreuses données récoltées (Sparrow, 1991 ; Van der Hulst, 2008 ; McGloin et Kirk, 2010). Considérée comme l'équivalent scientifique de l'analyse de liens (Harper et Harris, 1975), l'analyse de réseaux permet d'aller au-delà des représentations graphiques et de l'aspect visuel, et d'analyser des modèles relationnels d'acteurs et de relations basés sur des mesures arithmétiques (Van der Hulst, 2008 ; McGloin et Kirk, 2010). Ces dernières portent sur la structure du réseau

ou paramètres qui quantifient les caractéristiques de l'activité du réseau, les rôles, les positions et les mécanismes sociaux associés tels que le pouvoir, le contrôle des ressources et la dépendance (Wasserman et Faust, 1994). L'un des buts principaux est de détecter et interpréter les modèles des liens sociaux parmi les acteurs et d'identifier l'impact, c'est-à-dire les bénéfices ou les contraintes, de la structure sociale sur le fonctionnement des acteurs et des réseaux (Van der Hulst, 2008).

L'analyse de réseau constitue ainsi un outil d'enquête (Van der Hulst, 2008). Elle est aussi un outil de soutien pour des prises de décisions appropriées et efficaces, permettant de posséder une meilleure connaissance de la situation, en en saisissant la dynamique sous-jacente et en répondant à certaines lacunes en renseignement, voire en anticipant des actions adverses (Martineau et Lecocq, 2016). Parmi les utilisations possibles figurent la construction de scénarios, l'analyse de risque et l'évaluation de la menace, les tests d'hypothèses, la déstabilisation de réseaux, l'identification des alias, des rôles identiques ou substitutifs, ou encore un support décisionnel pour le déploiement de ressources (Sparrow, 1991 ; Koschade, 2006 ; Van der Hulst, 2008 ; Bouchard et Konarski, 2014). Utiliser les techniques automatisées d'analyse de réseau et de visualisation permettrait de révéler différentes configurations au sein d'un réseau, de détecter des sous-groupes et de découvrir leurs modèles d'interaction en identifiant les individus centraux (Xu et Chen, 2005). Les connaissances ainsi acquises aideraient les agences d'application de la loi à lutter contre le crime de manière proactive, notamment pour comprendre et perturber les réseaux criminels au cours d'enquêtes en identifiant les acteurs clés et déterminer quelles interventions seraient plus efficaces (Xu et Chen, 2005 ; Koschade, 2006). En effet, les réseaux sociaux sont des structures créant, facilitant et transformant les limites en opportunités. Les connexions aux autres personnes, ainsi que la structure du réseau dans son ensemble, fournissent un accès aux ressources. En ce sens, ils peuvent être considérés comme un des principaux blocs de construction du succès criminel (Van der Hulst, 2008). Ces considérations sont particulièrement importantes, lorsqu'il s'agit de sélectionner des complices pour réaliser une entreprise criminelle.

1.1.2. Application aux réseaux de codélinquance

Les réseaux de codélinquance forment et affectent le comportement criminel. Les éléments situationnels jouent aussi sur la tendance à rejoindre de tels réseaux et à y adopter certains rôles. Par conséquent, il est important de les étudier (McGloin et Nguyen, 2014). Il est rare qu'un délinquant agisse seul. La compréhension des modèles et des processus de codélinquance fournit des éléments importants pour les interventions des agences d'application de la loi (McGloin et Nguyen, 2014). En effet, la codélinquance affecte le développement de la carrière criminelle, qui lui-même affecte les modèles de codélinquance (Reiss, 1988 ; Carrington, 2009). Le crime en groupe peut augmenter la propension d'un individu à être introduit ou à s'insérer dans un style de vie criminogène, du fait notamment de l'influence des pairs (Andresen et Felson, 2012). Les individus tendraient par ailleurs à éviter des relations redondantes. Cela leur permet de s'engager dans une variété de crimes lorsqu'ils les commettent à plusieurs. Ils ont en effet accès à des habiletés et opportunités plus variées (Granovetter, 1973 ; Burt, 1992 ; McGloin et Piquero, 2010). La structure de réseau s'applique donc à l'étude des réseaux de codélinquance. Le but devient alors la prévention de l'existence du réseau, et pas tant son démantèlement (Warring, 2002 ; McGloin et Nguyen, 2014). Cependant, le fait que les complices diffèrent d'un événement à un autre pose difficulté (Sarnecki, 2001). La position d'instigateur semblerait alors être celle ayant le plus de pouvoir et méritant attention (McGloin et Kirk, 2014). Elle serait en effet à la source d'un nombre disproportionné d'affaires criminelles dans le système judiciaire. Ces individus ont un impact sur leurs complices, immédiat en les influençant à commettre un crime, et à long-terme en les amenant à poursuivre dans cette voie (Andresen et Felson, 2012).

Deux types de délinquants peuvent effectivement être distingués, selon leurs niveaux d'activité criminelle et le rôle que la délinquance joue dans leur carrière criminelle (Carrington, 2009). Une petite proportion de délinquants, dits persistants, ont un taux élevé d'activité, sont autonomes et ont le rôle de recruteurs, tandis qu'une majorité ayant un taux d'activité faible, moins expérimentée, se joint à eux (Reiss, 1988).

Il peut également y être référé comme délinquants persistants à vie, plus probables d'être recruteurs, et délinquants limités à l'adolescence. Ces derniers commettent des crimes principalement ou exclusivement sous l'influence ou l'imitation de leurs pairs plus autonomes et peuvent ainsi être utiles aux leaders (Moffitt, 1993). Par participation criminelle est entendue l'implication criminelle d'une personne dans un événement (Frank et Carrington, 2007 ; van Mastrigt et Farrington, 2009 ; Carrington, 2009) et par activité, le nombre d'événements dans une carrière criminelle. La plupart des délinquants sont des hommes et ne comptent qu'un seul incident, avec une moyenne de 1,98 participation (Carrington, 2009). La proportion de codélinquance, importante et commune au début d'une carrière criminelle, diminue pour tous selon l'âge et varie selon le type d'infraction. Cependant, elle varie peu avec l'expérience criminelle pour ceux ayant un taux d'activité élevé. Ceux-ci continueraient donc de trouver une certaine valeur à leurs complices pour la continuité de leur activité criminelle, commettant à la fois des infractions seuls ou à plusieurs (Reiss, 1988 ; Moffitt, 1993 ; Carrington, 2009). Toutefois, les caractéristiques à la fois individuelles et situationnelles jouent un rôle dans la compréhension de l'incitation au crime de groupe (Warr, 1996 ; McGloin et Nguyen, 2012). Cela complique l'identification des individus ayant le profil d'instigateur ou de recruteur, qui ne servirait donc que partiellement (McGloin et Nguyen, 2014).

Les individus occupant la position de courtiers ont une position stratégique au sein du réseau (Morselli, 2009 ; McIver, 2016). Ils présentent certaines spécificités, telles le fait de connaître plus de personnes au sein du réseau, d'avoir débuté leur carrière criminelle plus tôt que les autres membres, d'avoir recours occasionnellement à la violence plutôt que fréquemment. Ils ne font également pas l'objet de plus de condamnations pour les infractions commises malgré une surveillance policière plus importante, puisqu'ils sont capables d'agir par contacts interposés (McIver, 2016). Ainsi, les courtiers se distinguent par une centralité d'intermédiation élevée et une centralité de degré plus faible, offrant un compromis entre sécurité et efficacité, puisqu'ayant davantage le contrôle de l'information ou du flux de communication (Freeman, 1977, 1979 ; Morselli, 2009). Les données de codélinquance ont par exemple été utilisées dans le cadre de l'analyse de la configuration des partenariats dans l'importation de drogues

illicites. Elles ont permis de montrer comment la vulnérabilité peut être partagée par d'autres. Bien que n'agissant pas dans les sections les plus risquées, ils s'exposent néanmoins lors de l'exécution complète de la chaîne de codélinquance et offrent de la flexibilité au réseau (Morselli, 2009). En effet, les acteurs, pris en tant que collectif, ont de la difficulté à échanger l'efficacité au profit de la sécurité, à l'inverse des réseaux s'engageant dans des tâches à plus long terme, tel un réseau terroriste (Morselli, 2009). L'addition d'autres acteurs au cœur du réseau, tels que des investisseurs, des blanchisseurs ou des contributeurs logistiques, contribue alors à étendre la périphérie du réseau criminel, en isoler le cœur, et apporter plus de sécurité (Morselli, 2009).

L'étude de l'évolution dans le temps d'un réseau de personnes co-arrêtées pour des crimes liés à la drogue révèle notamment que les délinquants établissent initialement des groupes soudés avec ceux ayant été arrêtés pour le même type de crime et que des relations entre les groupes sont faites que par la suite. Ainsi, les délinquants arrêtés pour des types de crimes similaires travaillent ensemble (Iwanski et Frank, 2014). Cela découle du principe d'homophilie, c'est-à-dire le fait que des liens se créent entre des personnes se ressemblant (Lazarsfeld et Merton, 1954). Ce phénomène structure les réseaux et les homogénéise (McPherson, Smith-Lovin et Cook, 2001).

L'application de l'analyse de réseau aux données de codélinquance permet donc d'approfondir la connaissance de certains phénomènes, notamment quant à la sélection des complices et à l'importance des individus occupant la place de courtiers ou d'instigateurs. Un exemple concret serait celui des groupes de motards criminalisés.

1.1.3. Le cas des groupes de motards criminalisés

Les groupes de motards criminalisés ou « Outlaw Motorcycle Clubs » (OMCs) permettent d'illustrer l'application des méthodes et concepts de l'analyse de réseau à un

groupe criminel et à des données de codélinquance. L'exemple le plus connu est le chapitre québécois des Nomads des Hells Angels qui essaya de monopoliser les marchés de la drogue à Montréal. De 1994 à 2001, la « guerre des motards » qui opposait les Rock Machine aux Nomads, a marqué les esprits par la violence des confrontations. Le nombre de meurtres était conséquent et a fait des victimes collatérales. De plus, ces événements étaient constants et regroupés autour d'un groupe spécifique sur une période de temps étendue (Morselli, 2009 ; Barker, 2014). Les Hells Angels sont le groupe de motards criminalisés le plus important (Barker, 2014). D'après leur site officiel¹, ils compteraient des chapitres dans plus d'une cinquantaine de pays, dont cinq au Québec, à savoir celui de Montréal qui est le plus ancien chapitre québécois (1977), ainsi que ceux de Sherbrooke (1984), de Québec (1988), de Trois-Rivières (1991) et du Sud (Montréal et les environs) (1997).

La structure et les règles de chaque club sont basées sur celles figurant dans le « guide original » des Hells Angels, à savoir des officiels élus, incluant généralement un président, un vice-président, un trésorier, un « road captain » et un « sergent at arms » (Lauchs, Bain et Bell, 2015). Ils fonctionnent sur un style militaire de leadership, tout en ayant un processus de décision démocratique (Lauchs, Bain et Bell, 2015). Les membres placent leurs collègues avant toute autre relation, votent démocratiquement sur les décisions majeures, supportent le club financièrement et à travers leur présence aux différentes rencontres (Lauchs, Bain et Bell, 2015). Une notion militaire d'obéissance et d'engagement envers le club est présente, impliquant une stricte conformité aux règles, dont la participation aux événements et l'adhésion au code vestimentaire (Lauchs, Bain et Bell, 2015). L'adhésion est basée sur la franchise locale du club, le chapitre, et la sélection des membres s'effectue avec grand soin. Les clubs soumettent le candidat à une période d'initiation avant de lui accorder possiblement à l'unanimité une adhésion complète, ainsi que le droit de porter le patch complet sur le dos de sa veste et le statut de membre titulaire (Lauchs, Bain et Bell, 2015). Les clubs de soutien ou clubs écoles sont

¹ D'après leur site Internet : <http://affa.hells-angels.com>

quant à eux des groupes de motards criminalisés affiliés à un autre groupe dominant. Ils obéissent à ses ordres, servent comme sources potentielles de recrutement ou lors de guerres entre clubs ou gangs et lui versent une partie des gains illicites (Barker, 2014). Le club dominant s'occupe de la distribution en gros des drogues, tandis que les clubs écoles conduisent les ventes dangereuses au détail, isolant ainsi les membres du club dominant des poursuites (Barker, 2014). À l'occasion, le club subordonné qui est pris en train de commettre un crime et qui démontre sa capacité à garder le silence et à endosser la responsabilité, peut être récompensé en devenant un « prospect » (Barker, 2014). Toutefois, l'application des méthodes d'analyse de réseau aux transcriptions des surveillances électroniques ayant servi à l'opération Printemps 2001 portant sur les Hells Angels, a permis d'établir que les rangs officiels apparaissent dans les modèles de communication analysés et reflètent la hiérarchie. Le réseau n'agirait cependant pas nécessairement comme tel (Morselli, 2009). En effet, la plupart des membres ayant un rang élevé, c'est-à-dire les membres les plus réputés mais aussi les plus âgés et expérimentés, ont une faible centralité de degré et de vecteur propre. Ils sont indirectement impliqués, avec un capital de courtage élevé. Cela caractérise une position stratégique dans le réseau, mais ne reflète pas nécessairement un système de commandement et d'ordre (Morselli, 2009).

Ce mode de fonctionnement, ainsi que leur nature variée, entre groupes de passionnés de moto pouvant compter quelques membres criminalisés et véritables organisations criminelles, font des groupes de motards criminalisés un phénomène complexe. Ils sont à l'origine de nombreuses réponses, tant législatives avec la loi antigang de 2001, que policières avec la mise en place d'escouades régionales mixtes (escouade Carcajou qui mena à l'opération « printemps 2001 » à l'encontre des Hells Angels par exemple). Un travail de renseignement de qualité est donc nécessaire pour établir la nature des clubs individuels et des chapitres, afin d'établir des réponses policières appropriées, tel le retrait sélectif de membres criminellement actifs du reste du club dans le but de renforcer les factions traditionnalistes et de réduire les activités criminelles sérieuses (Lauchs, Bain et Bell, 2015). Initialement perçus comme de simples bagarreurs de bars ivres, les motards sont désormais considérés comme des groupes de

criminels sophistiqués et organisés utilisant la violence et l'intimidation pour sécuriser leur place au sein de l'environnement criminel, et de sérieux acteurs dans le crime organisé international (Barker, 2014 ; Lauchs, Bain et Bell, 2015). Ils offrent alors un terrain propice à l'application des méthodes de l'analyse de réseau. En effet, l'exigence d'être accepté par chaque membre du chapitre pour en devenir membre titulaire implique que chaque membre se connaisse sur une base personnelle (Van der Hulst, 2008). Cela permet d'établir des bases de données relationnelles et de renseignement, à partir des informations collectées par les policiers de terrain, lors d'un évènement impliquant un membre d'un groupe de motards criminalisés ou un associé (Lauchs, Bain et Bell, 2015).

Cependant, si ces stratégies policières ont eu un certain succès dans le ralentissement ou l'arrêt de l'activité criminelle des membres de groupes de motards criminalisés, leur efficacité reste modeste face à l'entreprenariat démontré par certains groupes tels que les Hells Angels (Lauchs, Bain et Bell, 2015). Ils mettent notamment à contribution des individus provenant de sphères légales, ce qui a pour effet d'isoler le cœur du réseau (Morselli, 2009). De plus, les données de sources policières utilisées, bien que représentant une source fiable d'informations, ne sont pas exemptes de limites.

1.1.4. Les données policières : source fiable, mais limitée

Les données officielles révèlent des informations importantes sur les interactions d'un type spécifique, non couvertes par d'autres sources. Il s'agit de la codélinquance officielle (Bouchard et Konarski, 2014). Les réseaux générés à partir des données d'arrestation, pas complètement exacts, constitueraient tout de même une riche source d'information à propos des réseaux illicites (Xu et Chen, 2008). Ces données sont utiles dans l'identification des membres constituant le « cœur » du réseau d'un gang de jeunes, que la nature dynamique et diffuse rend difficile. Beaucoup d'individus non-ciblés étaient aussi centraux que ceux ciblés, et inversement (Bouchard et Konarski, 2014).

Plusieurs défis existent effectivement en analyse de réseau criminel dans le contexte des données policières. Les données d'arrestation comportent leur lot de limites (Iwanski et Frank, 2014 ; Xu et Chen, 2008). Le temps nécessaire pour établir de tels réseaux, l'accès aux données, leur validité et leur classement, la définition des limites du réseau, le changement au cours du temps, ou encore la gestion des données manquantes, sont des paramètres à prendre en compte. Un protocole précis doit être suivi, les erreurs pouvant avoir des répercussions importantes sur les résultats (Van der Hulst, 2008). Il s'agit de la problématique de spécification des liens. La difficulté réside dans les liens manquant plutôt que les nœuds manquants (Morselli, 2009). En effet, les délinquants ne font partie du réseau que s'ils sont arrêtés par la police. Inversement, de faux liens peuvent être établis si les délinquants sont arrêtés en commettant le même crime, malgré l'absence de relation réelle (Iwanski et Frank, 2014 ; Xu et Chen, 2008). En outre, certains acteurs peuvent être centraux parce qu'ils sont les cibles centrales d'une enquête. Cela a pour conséquence de n'offrir qu'une vision partielle du réseau (Morselli, 2009).

Plus particulièrement, la nature des données utilisées doit être prise en considération dans l'interprétation des résultats, différents designs de recherche pouvant mener à différentes conclusions quant aux frontières d'un gang (Fleisher, 2005 ; Bouchard et Konarski, 2014). Cela a également été illustré par l'étude de Morselli (2009) en utilisant des données d'écoute électronique, concernant le rôle d'un gang majeur dans la distribution de drogue à Montréal. L'analyse de réseau montre ses limites quant au fait de pouvoir fournir une image claire des acteurs clés d'un réseau, lorsqu'elle est utilisée seule. De l'information complémentaire sur la nature des interactions entre les acteurs du réseau, leurs fonctions et leurs rôles, est nécessaire pour avoir une représentation précise de l'importance de l'individu dans le réseau (Bouchard et Konarski, 2014). Devant le coût et la difficile accessibilité des écoutes téléphoniques, d'autres sources complémentaires ou alternatives peuvent être considérées. La participation aux réunions d'une organisation offre ainsi une autre perspective dans l'identification des leaders, en les détectant par la mesure d'intermédiarité, et de l'information à propos de la structure interne du réseau (Calderoni, 2016). Papachristos, Hureau et Braga (2013) ont par ailleurs adopté une approche intégrée pour analyser l'influence des propriétés du quartier sur la

violence de gang, considérant à la fois des dimensions géographiques et des mécanismes de réseau social. Ainsi, des processus spatiaux (territoires adjacents, conflits antérieurs) et de réseau (réciprocité, recherche de statut) médient les effets raciaux. Cela suggère la primauté de la place et du groupe dans la génération de la violence de gang. De futures recherches devraient ainsi continuer de produire des comparaisons systématiques des résultats d'analyse de réseaux réalisées à partir de deux ou plusieurs jeux de données (Bouchard et Konarski, 2014). Les sources ouvertes, c'est-à-dire des données librement accessibles en ligne, plus particulièrement les médias sociaux, pourraient être utilisées.

1.2. Le potentiel de l'usage des sources ouvertes : les médias sociaux

Les données policières présentent leurs limites, même lorsqu'il est fait usage des techniques et des outils de l'analyse de réseau. Elles sont souvent incomplètes, incohérentes ou même incorrectes (Sparrow, 1991 ; Xu et Chen, 2005 ; Morselli, 2009). Il est donc utile de pouvoir exploiter des données provenant d'autres sources d'information. L'analyse de réseaux offre un cadre d'analyse suffisamment souple pour appliquer ces méthodes à de nouveaux sujets et champs de recherche, notamment centrés sur l'utilisation d'Internet (e.g. l'existence de communautés virtuelles). Cela permet la production et l'analyse de jeux de données différents (Boivin et Morselli, 2016). Parmi les sources potentielles de renseignement figurent les médias sociaux, plus particulièrement Facebook. C'est l'une des plateformes les plus populaires. Elle compte désormais en moyenne plus d'un milliard d'utilisateurs actifs chaque jour, dont 934 millions à partir d'appareils mobiles (Facebook, 2016). Très populaires et générant une quantité importante de données, ces sites constituent une opportunité unique (Bhat et Abulaish, 2013), tant vis-à-vis de leurs caractéristiques que des manières dont ils peuvent être utilisés par les organisations policières.

1.2.1. Caractéristiques

Les médias sociaux présentent des caractéristiques particulières, notamment du fait de la définition dans laquelle ils s'inscrivent. Ils consolideraient des relations préexistantes et faciliteraient les relations sociales par une présence constante en ligne. Cela rend les informations personnelles de nombreux individus visibles et accessibles.

a) Définition

Les médias sociaux sont définis comme « des services Web qui permettent aux individus de (1) construire un profil public ou semi-public dans le cadre d'un système délimité, (2) d'articuler une liste d'autres utilisateurs avec lesquels ils partagent des relations et de (3) voir et de croiser leurs listes de relations et celles faites par d'autres à travers la plateforme » (Boyd et Ellison, 2007). Ils fournissent aux individus une manière de joindre un réseau et de devenir utilisateur. Ces derniers peuvent fournir de l'information sur eux ou leurs préférences grâce au profil, publier du contenu qu'ils aimeraient partager avec d'autres utilisateurs, et créer des liens (Bhat et Abulaish, 2013).

b) Consolidation de relations préexistantes

Si certaines de ces plateformes aident des inconnus à entrer en contact en se basant sur des intérêts partagés, des opinions publiques ou encore des activités, la plupart serviraient à maintenir des réseaux sociaux préexistants (Ellison, Steinfield, and Lampe, 2007 ; Lewis, Gonzaleza, et Kaufman, 2012). Elles permettent aux usagers de les articuler et de les rendre plus visibles qu'auparavant, à une large population (Boyd et Ellison, 2007 ; Hansen et al., 2010). Elles constituent généralement un portrait précis, contenant de l'information véridique sur leur identité hors ligne (Back et al., 2010).

Ainsi, l'étude du réseau Facebook d'une cohorte d'étudiants de niveau collégial a permis de démontrer que le processus de formation des liens serait influencé par d'autres mécanismes tels l'équilibre, la réciprocité d'amitié ou d'être amis avec l'ami d'un ami, la proximité, c'est-à-dire la fréquentation d'un même environnement physique, et l'homophilie selon des catégories non raciales telles que la classe sociale (Wimmer et Lewis, 2010). L'homophilie est un phénomène reconnu dans la formation et le maintien des liens. C'est une caractéristique importante dans les réseaux sociaux, faisant référence à la tendance à ce que des relations amicales se forment entre ceux se ressemblant à différents égards (Lazarsfeld et Merton, 1954). Ce principe structure tout type de réseau et les rend homogènes au regard de nombreuses caractéristiques sociodémographiques, comportementales et intrapersonnelles (McPherson, Smith-Lovin et Cook, 2001). L'homophilie sur le plan de la race et de l'ethnie crée les plus fortes divisions dans les environnements personnels, suivie de l'âge, la religion, l'éducation, l'emploi et le sexe. La proximité géographique, les liens familiaux, les structures telles que l'école et l'emploi ou l'occupation de positions similaires constituent des contextes propices au fait de se lier à des personnes semblables (McPherson, Smith-Lovin et Cook, 2001).

L'homophilie a pour conséquence de limiter les mondes sociaux des individus. Elle engendre d'importantes implications quant à l'information reçue, les attitudes adoptées, et les interactions vécues (McPherson, Smith-Lovin et Cook, 2001 ; Grund, 2014). En effet, des liens peuvent être établis entre les attributs d'un individu, c'est-à-dire ses caractéristiques individuelles, et la position occupée dans un réseau. Ces liens peuvent être le fait d'individus centraux dans un réseau qui adoptent certaines caractéristiques. Inversement, ils peuvent être dus au fait que le réseau évolue autour d'individus possédant certaines caractéristiques (Grund, 2014). Par exemple, un individu central du réseau est un motard criminalisé. Il est devenu motard parce qu'il était positionné de manière centrale dans le réseau et que plusieurs de ces relations ont pu l'introduire à ce milieu. À l'inverse, sa position centrale pourrait s'expliquer par le fait qu'il soit motard. En d'autres termes, il s'agit de l'énigme de l'œuf et de la poule. Par conséquent, les amis présents dans le réseau d'un individu sont non-représentatifs de l'ensemble de la population. La distribution des attributs est affectée par le paradoxe de taille du groupe ou

paradoxe de l'amitié, c'est-à-dire que la taille perçue des groupes est plus grande qu'elle l'est réellement (Feld et Grofman, 1977). Nos amis ont alors en moyenne plus d'amis que nous, les individus ayant beaucoup d'amis apparaissant simplement plus souvent dans des dyades de relations (Feld, 1991), seraient susceptibles d'être mieux informés (centralité de proximité), d'être de meilleurs intermédiaires (centralité d'intermédiation) et d'avoir plus de pouvoir (centralité de vecteur propre) (Grund, 2014). En outre, le paradoxe de l'amitié peut être crucial dans l'identification de sous-populations cachées, par exemple pour identifier les codélinquants importants d'un groupe de délinquants choisis au hasard. Ce dernier sera plus important dans le réseau pris dans son ensemble et présentera des attributs non-aléatoires et non-représentatifs de la population générale (Grund, 2014).

Toutefois, bien que les effets principaux de l'homophilie sur différentes dimensions soient positifs, il semblerait qu'avoir un nombre croissant d'attributs en commun atténue l'émergence et la persistance de liens amicaux. Les individus chercheraient le contact social avec des personnes similaires sur certaines dimensions mais pas toutes (Block et Grund, 2014). Cumuler tant de similarités auraient un effet relatif et redondant. Les relations sociales sont utiles lorsqu'elles donnent accès à un ensemble différent d'individus (Granovetter, 1973; Burt, 1992 ; Block et Grund, 2014). Un choix sera alors effectué selon les situations, certaines provoquant une préférence à la similarité et d'autres aux différences (Block et Grund, 2014). En outre, l'impact social du partage de certains goûts dépendrait du médium utilisé et du contenu. Les goûts partagés par tous sont trop généraux pour servir de marqueurs de différenciation sociale (Lewis, Gonzaleza, et Kaufman, 2012). L'objectif serait alors de renforcer des liens sociaux avec des individus nous ressemblant plutôt que de s'influencer mutuellement, bien que Facebook se base sur des liens faibles (Lewis, Gonzaleza, et Kaufman, 2012), réputés être un type de lien propice à la diffusion d'informations (Granovetter, 1973).

Le partage de ses goûts ou de ses intérêts avec d'autres individus ou « amis » engendre une visibilité des informations personnelles concernant un individu au reste de la communauté et, par conséquent une plus grande accessibilité à celles-ci.

c) Visibilité et accessibilité aux informations personnelles

Les médias sociaux consistent en des profils visibles qui exposent publiquement une liste articulée d'amis, également usagers du système. Ces profils en sont une des composantes essentielles. Il est notamment possible d'y trouver des indications descriptives telles que l'âge, le lieu, les intérêts, une section « à propos », ainsi qu'une photo de profil (Boyd et Ellison, 2007). Une importante quantité de données relationnelles est ainsi créée, c'est-à-dire de l'information au sujet de qui connaît, est ami ou parle avec qui, qui sort dans les mêmes lieux ou encore aime les mêmes choses (Hansen et al., 2010). En effet, la liste d'amis contient des liens vers chacun des amis du profil, permettant à ceux regardant le profil de parcourir le réseau (Boyd et Ellison, 2007). La visibilité d'un profil est à la discrétion de l'utilisateur. Toutefois, seul le quart des utilisateurs prendraient des dispositions quant à la protection de leurs données privées. Cela prévient notamment le fait que d'autres usagers que ceux de leur propre réseau accèdent à leur liste d'amis (Catanese et al. 2011). Par ailleurs, il suffirait de posséder les attributs disponibles sur Facebook de seulement 20% des utilisateurs d'un réseau pour déduire ceux de l'ensemble. Cela se base sur le principe d'homophilie, les individus qui partagent des attributs communs étant plus probables d'être amis et formant des communautés denses (Mislove, Viswanath, Gummadi, et Druschel, 2010).

C'est un élément important. Les compagnies fournissant des services de réseautage social telles que Facebook sont le plus souvent réticentes à partager leurs données à des fins de recherches. Le forage des données (« crawling ») de l'interface, à savoir collecter de l'information publique à propos d'utilisateurs via leur profil sans leur participation active, constitue ainsi une solution viable d'acquérir l'information (Catanese et al., 2011 ; Gjoka et al., 2011). Toutefois, des limites s'appliquent du fait des politiques restrictives de protection de la vie privée (Gjoka et al., 2011). Le forage des données est contraire aux règles d'utilisation de Facebook, qui exige une demande d'autorisation. Cela lui permet notamment d'assurer l'anonymat des utilisateurs lorsque de l'information est transmise à des tiers (Facebook, 2016).

Outre le fait de permettre la consolidation de relations préexistantes, et de rendre plus visibles et accessibles certaines informations personnelles, les médias sociaux faciliteraient les relations sociales, par une présence en ligne de plus en plus constante.

d) Un facilitateur des relations sociales et une présence constante en ligne

L'accessibilité grandissante d'Internet, à travers de nombreux médias, donne à la plupart des usagers une présence en ligne 24/7. Cela les encourage à construire une interconnexion solide de relations en ligne et en fait un terrain propice à l'analyse de réseau (Catanese et al., 2011). Les médias sociaux peuvent être représentés sous forme de graphiques, les nœuds représentant les utilisateurs et les liens, les relations.

Plus particulièrement, Facebook se caractérise par un schéma simple de relations amicales. Sa structure peut être représentée par un graphique non pondéré et non-dirigé, à l'inverse de Twitter permettant de savoir qui suit qui (Catanese et al., 2011 ; Hansen et al., 2010). Les nombreuses activités qui y sont réalisées laissent des données concrètes et observables. Il est donc utile de concevoir Facebook comme une base de données en continu et en temps réel de l'activité sociale (Wilson, Gosling et Graham, 2012).

Ainsi, les réseaux d'amis Facebook révèlent les mondes sociaux importants d'un individu donné. Ils tendent à se regrouper autour de différents domaines de la vie de la personne tels que les amis, les camarades de classe, les collègues de travail, ou la famille (Hansen et al., 2010). La plupart des réseaux personnels sur Facebook sont alors presque entièrement connectés. Cela inclut un certain nombre de groupes représentant les contextes sociaux aux côtés d'individus ayant une centralité de degré ou d'intermédialité élevée qui sont notables ou personnellement importants (Hansen et al., 2010). Plusieurs études ont établi que 92% des utilisateurs étaient en moyenne séparés par moins de quatre relations intermédiaires. Le nombre d'amis suit une distribution curvilinéaire, avec une moyenne de 214 amis aux États-Unis (Backstrom et al., 2012 ; Ugander et al., 2011). De

plus, en se basant sur la distinction entre liens forts et liens faibles (Granovetter, 1973), un utilisateur type de Facebook communique directement avec un petit groupe central d'amis en postant des commentaires ou des messages, indiquant des liens forts, et suit la majorité des amis de manière passive tel qu'en visionnant le fil d'actualité, indiquant alors des liens faibles (Burke, Marlow, et Lento, 2010).

Ainsi, la place qu'occupent les médias sociaux et la quantité d'informations qu'ils permettent de générer, tant d'un point de vue individuel que relationnel, en font une source non négligeable pouvant être utile aux organisations policières.

1.2.2. Police et médias sociaux

Notamment du fait de l'expansion des médias sociaux et de la place qu'ils prennent dans la vie quotidienne, les organisations policières exploitent les nouvelles opportunités qu'offrent les plateformes de réseautage social telles que Facebook (Frank, Cheng et Pun, 2011 ; Trottier, 2015). Selon un sondage de 2015 effectué par l'association internationale des chefs de police, 96,4% des agences participantes utilisent les médias sociaux, en particulier Facebook (94,2%) et Twitter (71,2%) (IACP, 2015). La motivation du personnel, l'implantation réussie des outils, leur utilisation simple et leur gratuité font une réussite de l'adoption des médias sociaux dans le contexte policier (Altunbas, 2013). Plusieurs types d'utilisation peuvent être observés. Il peut s'agir d'outils de communication, de renseignement ou d'enquête (Frank, Cheng et Pun, 2011 ; Altunbas, 2013 ; IACP, 2015).

a) Un outil de communication avec le public

L'une des utilisations principales des médias sociaux consiste à communiquer avec le public. Il s'agirait alors par exemple d'informer le public de problèmes relatifs à la criminalité (84,3%) ou de mener des activités de prévention (79,2%). Il peut aussi s'agir de fournir de l'information en cas d'urgence ou de crise (79,9%) (IACP, 2015). Le département de police de la ville de Boston a par exemple saisi cette opportunité lors des attaques à la bombe du marathon de 2013, en engageant la conversation avec le public et en l'informant via Twitter. Ces nouveaux outils technologiques sont donc compatibles avec les objectifs traditionnels de la police et ses considérations (Davis, Alves et Sklansky, 2015). Cela souligne aussi le pouvoir de communication par la transmission d'un message rapide, actuel et précis sans l'intermédiaire de la presse traditionnelle, ainsi que le pouvoir de conversation des médias sociaux. Un dialogue ouvert peut être engagé à grande échelle avec le public (Davis, Alves et Sklansky, 2015). En outre, les technologies peuvent avoir un impact sur la façon dont le public perçoit la police (Goldsmith, 2010). Cette dernière se tourne vers les médias sociaux pour gérer sa visibilité et son image (Schneider, 2015 ; Trottier, 2015). Un certain nombre d'agences gérant ainsi les relations publiques et la communication soutiennent que cela a amélioré les relations entre la police et la communauté dans leur juridiction (IACP, 2015).

Mais ce n'est pas la seule utilisation qui peut être faite des réseaux sociaux en ligne par les organisations policières. Ils peuvent aussi servir de source de renseignement et d'outil d'enquête.

b) Un outil d'enquête et de renseignement

Les agences d'application de la loi utilisent aussi les médias sociaux à des fins d'enquête et de renseignement. L'utilisation dans le cadre d'enquêtes est la plus fréquente (88,7%), les organisations policières considérant qu'il s'agit d'une aide pour la résolution

de crimes. Les médias sociaux sont aussi utilisés pour solliciter de l'information sur des crimes auprès du public (76,3%) ou en renseignement (75,0%) (IACP, 2015). En effet, le fait d'établir un lien de communication avec le public via de telles plateformes, permet aussi de collecter du renseignement. Cela consiste pour des personnes non affiliées aux agences d'application de la loi de documenter des événements criminels et de rassembler à l'attention de la police de l'information qui pourrait constituer d'éventuelles preuves à une vitesse sans précédent. Cela a été notamment observé lors des émeutes de 2011 à Vancouver (Schneider et Trottier, 2012).

Cela note également l'utilisation croissante et officielle par la police de sites de médias sociaux dans un but d'action de police et de surveillance. Il s'agit d'une approche proactive des médias sociaux et marque un changement profond dans le rassemblement de preuves pour les enquêtes (Schneider et Trottier, 2012). Cela indique un futur proche dans lequel la police et d'autres enquêteurs auront un accès important au contenu de médias sociaux, leur permettant de collecter plus d'informations qu'ils ne le pouvaient auparavant. Cela engendre de profondes implications pour la surveillance et le contrôle social, et une surveillance augmentée de la vie sociale (Schneider et Trottier, 2012). En effet, les policiers peuvent enquêter à travers les interfaces de la plateforme sociale, examiner son contenu et collecter des preuves, et ce sans nécessiter la coopération des utilisateurs (Trottier, 2012). Ils recourent ainsi à de faux profils ou à une identité cachée pour rassembler et gérer de l'information, en devenant par exemple l'ami d'un suspect, visionnent son profil et ses activités sur les médias sociaux, de même que ceux de victimes (Frank, Cheng et Pun, 2011 ; IACP, 2015). L'une des manières les plus faciles d'obtenir de l'information est toutefois d'imiter la manière des autres usagers d'accéder à l'information sur le site. Cela consiste à se servir des avantages des services de recherche de ces sites tels que la recherche d'identité, de groupes, d'événements ou d'autres mots clés (Trottier, 2012). En effet, de nombreuses informations générées par les utilisateurs (détails de contact, photos, listes d'amis, posts sur les murs) (Bhat et Abulaish, 2013) ne sont pas protégées et ainsi accessibles aux enquêteurs (Catanese et al., 2011 ; Trottier, 2012).

Les médias sociaux constituent alors une source riche de données. Ils permettent de réaliser de l'analyse de réseau et des tâches de data-mining, et ainsi apporter des résultats plus pertinents et réalistes (Bhat et Abulaish, 2013). La facilité d'utilisation et l'accessibilité à tant de contenu en font une première étape intéressante pour commencer une enquête ou chercher un individu précis. L'avantage est aussi que l'individu sous examen n'en a pas connaissance, ce qui n'est pas sans soulever son lot de problématiques et de critiques (Trottier, 2012, 2015 ; Décary-Héту, Dupont et Fortin, 2014). La visibilité des liens sociaux devient une source de preuve pour les enquêteurs mais d'insécurité pour les suspects d'actes criminels (Trottier, 2012).

En effet, les criminels utilisent Internet et ses outils de communication facilement accessibles à tous (Décary-Héту et Morselli, 2011 ; Décary-Héту, Dupont et Fortin, 2014). L'observation des communautés en ligne procure ainsi des jeux de données innovants qui améliorent la compréhension du phénomène criminel en ligne, par exemple dans le cadre d'« Internet Chat Relay » ou IRC (Décary-Héту, Dupont et Fortin, 2014) ou à l'aide des hyperliens reliés aux sites Internet de pédopornographie (Bouchard et Westlake, 2016). L'analyse de tels types d'interaction, différents de la délinquance traditionnelle, permet de connaître un ensemble de relations d'un délinquant ou d'une entité délinquante, c'est-à-dire un ensemble de codélinquants potentiels (Bouchard et Westlake, 2016). Du point de vue pratique, cela permet aux agences d'application de la loi d'être plus efficace dans l'identification et l'investigation de cybercriminels, en intégrant l'observation de ces sites à leur quotidien, puisqu'il s'agit d'outils simples, automatisés et librement accessibles, soit une information de première ligne exempte de biais (Décary-Héту, Dupont et Fortin, 2014). S'agissant plus particulièrement des médias sociaux, ils aident à relier les personnes ayant des intérêts communs, à faciliter un large éventail d'activités du secteur légal. Ils peuvent alors également faciliter les activités du secteur illicite, voire celles des organisations criminelles (Frank, Cheng et Pun, 2011).

Le renseignement en sources ouvertes, c'est-à-dire les données accessibles librement en ligne d'une variété de sources, joue ainsi un rôle croissant pour aider des responsables de la sécurité nationale à déterminer les caractéristiques, les motivations et les intentions de groupes adverses menaçant la stabilité de la société (Rhodes, 2011). Cela donne lieu à la gestion des médias sociaux ou « social media monitoring » (Trottier, 2015). En effet, de plus en plus d'aspects de la vie sociale figurent sur les réseaux sociaux tels que Facebook, Twitter ou encore YouTube. La probabilité que des événements criminels apparaissent ou soient rendus visibles sur ces sites est plus forte, pouvant ainsi constituer une preuve d'un crime commis hors ligne (Trottier, 2015). En outre, si les méthodes de l'analyse de réseau se sont montrées utiles dans le fait d'organiser et de représenter les organisations de réseaux cachés, il s'agit d'une technique gourmande en données, en temps et en ressources (Rhodes, 2011 ; Bhat et Abulaish, 2013). Le renseignement en sources ouvertes permet de collecter des données complètes et non-contradictoires, sur une échelle de temps plus rapide et donc plus pertinente vis-à-vis du tempo des opérations policières (Rhodes, 2011).

En effet, les groupes criminels peuvent aussi tirer profit des médias sociaux, pour recruter de nouveaux membres et diffuser des messages (Décary-Héту et Morselli, 2011), coordonner des activités criminelles ou encore trouver des victimes potentielles (Frank, Cheng et Pun, 2011). La présence sur ces sites est essentiellement reliée à la promotion d'une culture de gang ou de rue à travers des initiatives individuelles. Les médias sociaux les rendent cependant accessibles à un public plus large, leur permettent d'annoncer leurs activités, de diffuser leurs réputations à travers des médiums plus conventionnels et d'interagir avec des personnes qui n'y auraient probablement pas été exposées dans le monde réel (Décary-Héту et Morselli, 2011 ; Morselli et Décary-Héту, 2013). Toutefois, le contenu général de ces pages pose toutefois la question de l'identité des individus se trouvant à leur origine ou pourrait manifester la volonté de garder l'identité réelle secrète, les membres étant conscients que les images de membres actuels pourraient être utilisées par les agences d'application de la loi (Décary-Héту et Morselli, 2011). Ainsi, malgré leur importance, les Hells Angels existent en ligne en tant que chapitres ou groupes mais non en tant que membres ou individus spécifiques, et n'affichent pas leurs exploits criminels

ou violents, contrairement à d'autres groupes (Décary-Héту et Morselli, 2011). Cela pourrait être dû au fait que les personnes impliquées dans le crime organisé seraient d'un âge plus avancé que les personnes qui fréquentent les sites de médias sociaux, et donc moins susceptibles de les utiliser (Frank, Cheng et Pun, 2011). Cependant, malgré les difficultés de bruit et d'identification des membres réels, les réseaux sociaux permettraient d'avoir une image plus actuelle de la présence de gangs en ligne (Décary-Héту et Morselli, 2011).

Certains auteurs soulignent ainsi l'intérêt que représenterait la comparaison de ces comportements en ligne à ceux du monde réel (Décary-Héту et Morselli, 2011), ainsi que le besoin de cumuler plusieurs sources de données face aux limites des données policières (notamment Sarnecki, 2001 ; Fleischer, 2005 ; Morselli, 2009 ; Bouchard et Konarski, 2014). Toutefois, enquêter sur les médias sociaux présente aussi son lot d'obstacles et de difficultés. Il n'est jamais certain que le profil appartient réellement au suspect, surtout lorsque peu de détails sont disponibles. Des recherches longues et approfondies peuvent alors s'avérer vaines, menant à un nombre trop important de résultats ou s'il est fait usage de surnoms. Si l'individu ne se sert des médias sociaux qu'aux mêmes fins que tous, soit pour entretenir les relations avec son entourage, les recherches doivent se porter sur les contacts de son réseau pour espérer obtenir de l'information sur son activité criminelle. De telles recherches doivent donc être combinées à de l'information la plus précise possible (Frank, Cheng et Pun, 2011). Une comparaison des sources de données permettrait aussi de déterminer si les tendances observées dans le monde réel se traduisent dans le monde virtuel ou, inversement, si les tendances observées en ligne peuvent être considérées comme signes avant-coureurs de l'implication d'individus dans un incident futur. Cela peut également permettre de colliger du renseignement sur les organisations criminelles et d'analyser l'utilité des graphes sociaux, c'est-à-dire les contacts d'organisation criminelles connues, d'observer leurs membres et leur évolution dans le temps, d'en identifier les acteurs les plus visibles, et les relations non encore détectées (Décary-Héту et Morselli, 2011).

2. Problématique

Une perspective d'analyse de réseau a été progressivement adoptée et développée au cours des dernières années. L'application des méthodes d'analyse de réseau est pertinente dans le milieu policier, notamment comme outil d'enquête. Le terrain est alors propice à ce que de telles méthodes soient appliquées à des données officielles de codélinquance. Celles-ci apportent une autre vision de la criminalité. Elles permettent aussi de déterminer un ensemble de complices potentiels. Parmi eux, ceux qui occupent la position d'instigateurs ou de courtiers méritent une attention particulière (Morselli, 2009 ; McGloin et Piquero, 2010 ; Andresen et Felson, 2012 ; McGloin et Nguyen, 2014). Ces considérations sont importantes dans l'étude de groupes criminels, tels les groupes de motards criminalisés, qui sont un phénomène particulier. Il faut prendre en compte leur nature variée et complexe, certains groupes constituant de véritables organisations criminelles, mais aussi l'étendue de leurs activités dans différents secteurs, entremêlant sphères légales et illégales. Par conséquent, ces groupes méritent des réponses policières spécifiques (Barker, 2014 ; Lauchs, Bain et Bell, 2015).

Toutefois, les données policières représentent une source fiable mais limitée d'informations. Elles sont souvent incomplètes et procurent de l'information partielle (Sparrow, 1991). Cela est dû notamment à la méconnaissance de l'existence d'une relation criminelle ou de complicité pouvant exister entre deux individus. À cet égard, les recherches sur le sujet soulignent l'importance et le besoin d'exploiter simultanément différentes sources de données (Sarnecki, 2001 ; Fleischer, 2005 ; Morselli, 2009 ; Bouchard et Konarski, 2014). Par ailleurs, la popularité des médias sociaux et la quantité importante d'informations présentes sur de tels sites, offrent une opportunité grandissante d'étudier les sources ouvertes et de déterminer l'intérêt et l'apport qu'elles représentent (par exemple Décary-Héту et Morselli, 2011 ; Bhat et Abulaish, 2013 ; Boivin et Morselli, 2016). En effet, les informations personnelles de nombreux individus sont de plus en plus visibles et librement accessibles. Elles concernent différentes sphères de la

vie des individus, telles la famille, la scolarité et l'emploi, mais aussi la sphère criminelle (Hansen et al., 2010). Les organisations policières ont compris le potentiel de telles sources et y portent de plus en plus attention. Elles l'utilisent notamment pour communiquer avec le public, pour générer du renseignement ou lors d'enquêtes (Frank, Cheng et Pun, 2011 ; Altunbas, 2013 ; IACP, 2015 ; Trottier, 2015). Ces sites, solidifiant des relations préexistantes et facilitant les relations sociales ainsi qu'une présence constante en ligne, sont donc potentiellement un bon complément aux données officielles.

Le présent projet s'inscrit donc dans ce besoin d'exploiter simultanément plusieurs sources de données. En effet, bien que mentionné dans d'autres études, cet aspect semble avoir été négligé. Il s'agit de construire le réseau social de sujets criminalisés, à partir à la fois des événements, criminels ou non, figurant dans les banques de données policières et des liens d'amitié publics de Facebook. Les données concernent dix individus réputés appartenir à un groupe de motards criminalisés. Ces deux types de données présentent l'intérêt d'être de nature différente, les données policières étant des données officielles et appartenant au monde réel tandis que les données Facebook représentent une source de données publiques et relèvent du virtuel. Ce faisant, l'objectif est de déterminer l'apport des sources ouvertes en matière d'analyse de réseau, plus particulièrement en renseignement, comparativement aux données officielles. Il est nécessaire de déterminer si une certaine complémentarité existe réellement. En d'autres termes, il s'agit de déterminer si Facebook a la capacité d'apporter des relations nouvelles et intéressantes, ainsi que de l'information nouvelle ou complémentaire sur celles-ci et les individus concernés, qui ne figureraient pas déjà dans les données policières. Cela consistera en une comparaison des informations qu'apporte chacune des deux sources, plutôt que de les considérer indépendamment l'une de l'autre. L'objectif sous-jacent porte sur l'intérêt pour les organisations policières d'exploiter Facebook comme source de renseignement et comment optimiser son utilisation pour contrer le caractère partiel des données policières et bonifier l'information dont elles disposent à partir d'une source publique et librement accessible. Plus spécifiquement, cela questionne la nécessité d'investir de nombreuses ressources sur Facebook et s'il convient d'utiliser cette plateforme de manière systématique ou de manière plus spécifique.

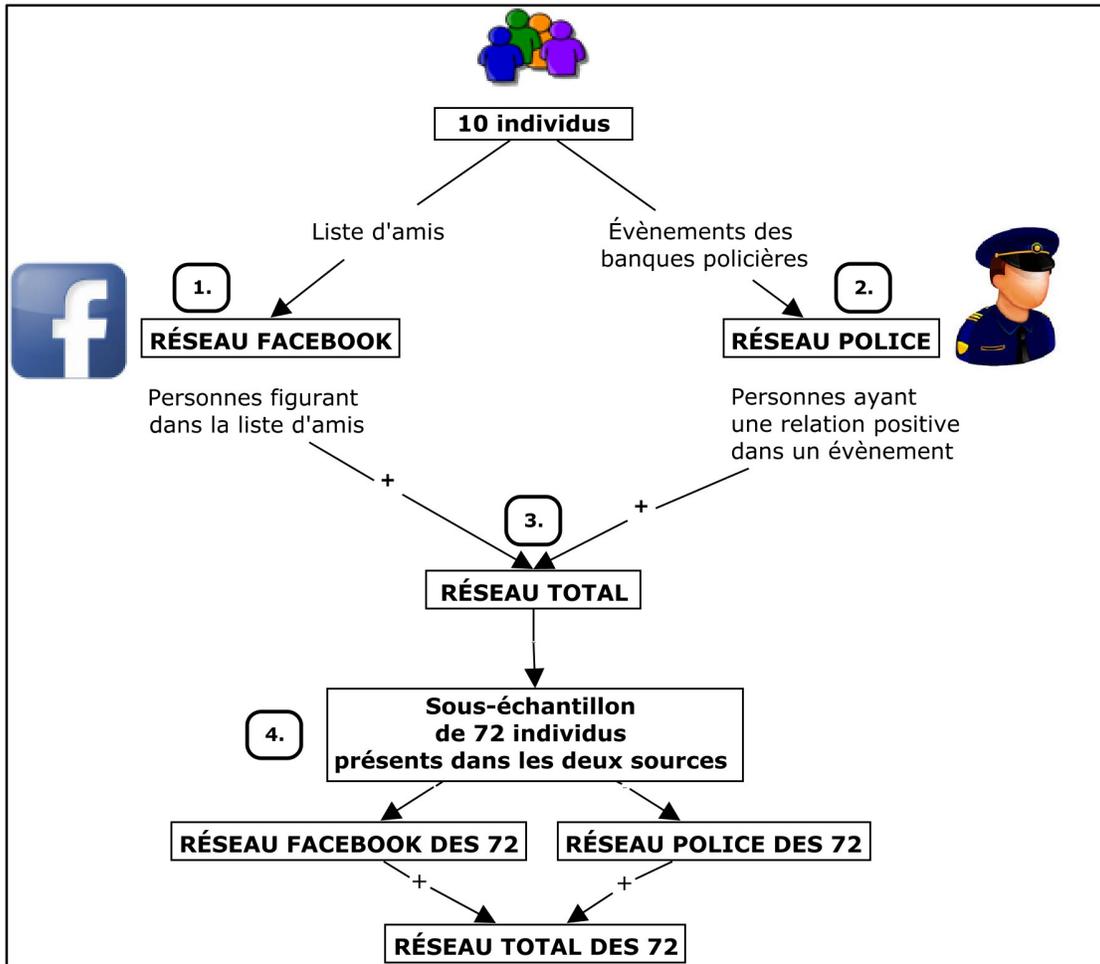
3. Méthodologie

3.1. Données et opérationnalisation

Le présent projet vise à construire le réseau social de dix individus réputés appartenir à un groupe de motards criminalisés, dont les noms ont été fournis par le Renseignement du Service de Police de la Ville de Montréal (SPVM). Il implique deux composantes, à savoir une matrice binaire, et un sociogramme représentant les relations directes et indirectes d'un ensemble d'acteurs (Morselli, 2009). Le logiciel Gephi (Bastian, Heymann et Jacomy, 2009) a été utilisé, puisqu'il s'agit d'un logiciel libre pour l'exploration visuelle et l'analyse de graphes et de réseaux (Akhtar, 2014).

Deux sources de données ont été exploitées afin de déterminer leur apport et leur possible complémentarité. Il est donc question de trois réseaux, soit le réseau construit à partir des données Facebook, celui construit à partir des données policières et celui construit en combinant ces deux sources de données. Cette démarche a été tout d'abord suivie avec l'ensemble des individus, puis uniquement avec les individus présents dans les deux sources de données (Figure 1). Chaque personne s'est vue attribuer un identifiant anonymisant unique (ID). De plus, les relations sont non-dirigées. On ne connaît pas le niveau de réciprocité des liens d'amitié. À l'inverse, Twitter permet de déterminer quel individu suit l'autre (Hansen et al., 2010 ; Catanese et al., 2011). Un tableau récapitulatif comprenant l'ensemble des variables est disponible en annexe (Annexe 1).

Figure 1. Schéma récapitulatif de la démarche adoptée pour construire les différents réseaux



3.1.1. Données issues du réseau social Facebook

La première source de données comprend l'information disponible dans les listes d'amis des profils Facebook des dix individus de départ. Celle-ci varie selon le niveau de confidentialité déterminé par l'individu dans les paramètres de sécurité. Il choisit quel renseignement rendre public ou non. Les données ont été collectées uniquement à partir des listes d'amis fournies. Cette étape s'est faite manuellement. Aucun algorithme n'a été utilisé. Les données ont ensuite été fusionnées afin de construire le réseau global comprenant les individus de départ et leurs relations respectives.

a) Variables individuelles

Les variables individuelles collectées comprennent le sexe, le lieu de résidence, les études réalisées, l'emploi occupé, ainsi que le degré de relations auquel il appartient, c'est-à-dire s'il s'agit d'un ami direct d'égo (degré 1) ou d'un ami d'un ami d'égo (degré 2). Les individus appartenant au second degré ont également pu être identifiés grâce aux suggestions d'amis communs dans les listes d'amis des individus concernés. Certaines variables comme l'âge étaient manquantes pour la plupart sinon la totalité des acteurs. Cela peut être dû au fait que l'accessibilité à cette information est restreinte par les paramètres de sécurité ou que l'individu n'a pas renseigné cette information. L'utilisation d'alias a été notée pour lier plusieurs comptes d'un même individu, à l'aide du lien URL du profil contenant souvent le nom réel de la personne.

- Lieu de résidence :

Il a été déterminé à l'aide des classifications établies par Statistique Canada² si le lieu de résidence se trouve à Montréal, c'est-à-dire sur l'Île de Montréal, dans la province du Québec, au Canada ou aux États-Unis. Si le lieu se trouve au Québec, il a été précisé s'il se rattache à l'une des six régions métropolitaines de recensement (RMR) de la province (Montréal, Trois-Rivières, Québec, Saguenay, Sherbrooke, Ottawa-Gatineau). Chaque possibilité constitue une variable binaire. En outre, la distance en véhicule (km) séparant le lieu de résidence de Montréal, selon le trajet le plus court préconisé par Google Maps, a été relevée.

² Statistique Canada (lieu de résidence) : http://www23.statcan.gc.ca/imdb/p3VD_f.pl?Function=getVD&TVD=117369&CVD=116982&CPV=24&CST=01012011&CLV=2&MLV=5

- Niveau de scolarité :

Les études réalisées répertorient les niveaux de scolarité établis par Statistique Canada³. Ils incluent le niveau secondaire, le niveau d'une école technique ou de métiers, d'un collège communautaire ou du cégep et le niveau universitaire. Ces derniers sont mutuellement exclusifs. Le niveau le plus élevé déclaré a été retenu. Chaque niveau constitue une variable binaire. De même, une cote de fiabilité indique si l'individu a écrit une information peu plausible comme « Harvard ».

- Emploi :

Les emplois ont été regroupés par secteurs, en se basant sur la classification des activités économiques du CIDREQ (Centre Informatique Du Registre des Entreprises du Québec)⁴, ainsi que sur internet pour les entreprises ne figurant pas au CIDREQ. Au vu des données collectées, les catégories esthétique et soins personnels, hôtellerie et restauration, bar, construction et chantiers, loisirs et divertissement, transports, habillement et mode, services sociaux et médicaux, public, alimentation, finances immobilier et assurances, enfance et enseignement, ont été retenues, auxquelles s'ajoutent les catégories « autres » et « motard » pour ceux ayant fait une référence explicite au fait d'être rattaché à un groupe de motards criminalisés (Annexe 2). Par ailleurs, comme pour les études, une cote de fiabilité a été créée. Toutes font l'objet d'une variable binaire.

³ Statistique Canada (établissement de fréquentation scolaire) : <http://www.statcan.gc.ca/fra/concepts/definitions/educationclass01b>

⁴ CIDREQ (classification des activités économiques) : http://www.registreentreprises.gouv.qc.ca/fr/demarrer/immatriculer/classification_activite_economique.asp

b) Variables relationnelles

La date de référence est le 1^{er} octobre 2016. De là, le temps depuis lequel deux individus sont amis sur Facebook en mois a été codifié. Il a été possible d'accéder à un nombre substantiel d'amis Facebook de 4 des 10 individus de départ.

c) Portrait du réseau obtenu via Facebook

Le réseau Facebook des dix individus de départ comporte un total de 2038 acteurs, dont 90,8% (n=1850) se trouvent au premier degré de relations. Cela pourrait s'expliquer par le fait que le réseau est incomplet en termes d'individus de second degré de relations.

Le réseau est composé de 54,5% d'hommes et de 45,5% de femmes pour ceux dont le sexe a pu être identifié (n=1796 ; 88,1%). Parmi ceux ayant renseigné un lieu de résidence (n=376 ; 18,4%), 93,6% résident au Canada, dont 88,3% au Québec. Ces derniers habitent principalement la RMR de Montréal (31,1%), de Sherbrooke (14,1%) ou de Québec (9,8%), habitant en moyenne à 267,23 km de Montréal ($\mu=267,23$; Md=107,00). En outre, un peu plus de la moitié des acteurs ayant renseigné leurs études (n=885 ; 43,4%) ont un niveau correspondant au secondaire (53,4%) tandis que 34,4% fréquentent un établissement de niveau postsecondaire, et 12,1% n'ont pas fourni d'information fiable. Plus de la moitié des individus composant le réseau n'ont fourni aucune information concernant leur parcours académique (56,6%). S'agissant de l'emploi occupé, les secteurs d'activités qui ressortent le plus sont la construction et les chantiers (14,5%), les transports (10,5%) et les bars (8,6%), alors que 5,3% ont fait une allusion explicite au fait d'être lié à un groupe de motards criminalisés et que 17,5% n'ont pas donné une information fiable. Toutefois, 59,4% n'ont pas renseigné d'informations (n=1210). Le présent réseau comprend un total de 3812 relations. Les acteurs du réseau sont en moyenne amis sur Facebook depuis moins d'un an ($\mu=16,81$; Md=9,00 en mois).

3.1.2. Données issues des banques policières

La seconde source de données comprend les crimes et les contrôles d'identité ayant eu lieu à Montréal de 2011 à 2016 où neuf des dix individus de départ sont présents (n=107), l'un d'eux n'y figurant pas (ID 10). Les événements ont été tirés de banques de données policières. Aucun n'est antérieur à 2007, date de création de Facebook. La présence d'autres individus permet alors d'établir les relations.

a) Variables individuelles

Les variables individuelles incluent le sexe et le statut attribué pour l'évènement. Pour un même évènement, un individu peut s'être vu attribué plusieurs statuts. Chaque statut fait l'objet d'une variable binaire. Le nombre d'évènements dans lesquels apparaît un individu, a aussi été considéré.

b) Variables relationnelles

La construction de la banque relationnelle issue des données policières s'est effectuée en plusieurs étapes, à savoir générer les relations à partir des données de chaque évènement, puis sélectionner les relations pertinentes dans le cadre du présent projet.

La question se posait de savoir si toutes les relations méritaient d'être retenues, parmi les 17 468 établies initialement. Seules celles ayant une signification d'un point de vue relationnel ont été conservées. Par exemple, deux personnes citées dans un évènement ne sont pas nécessairement présentes physiquement. On ne peut pas assumer qu'elles se connaissent. L'accent s'est alors porté sur les dyades des statuts susceptibles d'avoir une relation positive, soit un lien où la présence de deux individus est certaine au

même endroit et au même moment. De plus, il s’agit de relations non conflictuelles, plus susceptibles de trouver une correspondance dans le réseau Facebook. Les relations positives criminelles (e.g. co-arrêtés) et les relations positives non criminelles (e.g. co-interpelés) ont été distinguées. Le fait de garder uniquement les relations positives, qu’elles soient criminelles ou non, entraîne la perte de 5842 relations (33,44%) et de 114 individus parmi les 679 composant initialement le réseau. Ces derniers, dont l’un des dix individus de départ (ID 9), ne possèdent aucune relation positive.

c) Portrait du réseau obtenu via les données policières

Le réseau issu des données policières comporte 565 acteurs, une fois les seules relations positives sélectionnées ($n_{rel}=11626$). Il est majoritairement composé d’hommes (95,1%), étant impliqués en moyenne dans près de deux événements ($\mu=1,92$; $Md=1,00$). Il diffère du réseau Facebook, car la proportion de femmes est moins grande. Dans de précédentes recherches, la plupart des délinquants sont des hommes et ne sont présents que dans un seul incident criminel (Carrington, 2009). Toutefois, il mérite d’être noté qu’au vu des 107 événements ayant servi à construire ce réseau, certains individus se démarquent par leur implication dans un grand nombre d’évènements (Tableau I).

Tableau I. Classement des dix individus impliqués dans le plus d’évènements

ID	Rang	Nombre d’évènements	Présence dans le réseau Facebook	Individu de départ
ID5	1	55	oui	oui
ID2	2	36	oui	oui
ID73	3	29	oui	non
ID3053	4	17	non	non
ID3	5	15	oui	oui
ID8	6	10	oui	oui
ID47	7	10	oui	non
ID3023	8	10	non	non
ID3468	9	10	non	non
ID3374	10	10	non	non

Ainsi, six des dix individus ayant été impliqués dans le plus d'évènements sont également présents sur Facebook, tandis que quatre ne sont présents que dans le réseau des données policières. Cela pourrait tendre vers une complémentarité des deux types de source. Parmi ceux présents dans les deux sources de données, quatre figurent parmi les individus de départ (ID5, ID2, ID3, et ID8). Les autres individus de départ, bien que ne figurant pas dans ce TOP10, démontrent une participation supérieure à la moyenne, avec une implication dans 5 à 8 évènements chacun. En outre, le statut qu'un individu tend à obtenir le plus souvent au cours d'un évènement est celui de personne interpellée ($\mu=1,71$; $Md=1,00$). Concernant les liens les unissant, il s'agit essentiellement de relations positives non criminelles ($n=11548$; 99,3%). Il s'agirait le plus souvent d'une relation unique ($\mu=1,10$; $Md = 1,00$). Dans notre échantillon, à l'exception d'une relation, les relations positives criminelles et non criminelles sont mutuellement exclusives.

Il s'agit alors de savoir si le réseau construit à partir des deux sources de données ressemble plus à celui issu de Facebook ou à celui issu des données policières, ou bien s'il regroupe les caractéristiques observées pour ces deux derniers.

3.1.3. Le réseau combinant simultanément les deux sources de données

L'objectif du présent rapport est de déterminer l'apport de chaque type de sources de données. Aux variables décrites dans les sections précédentes, s'ajoutent celles portant sur la présence ou l'absence d'un individu dans le réseau Facebook, ainsi que sa présence simultanée dans ces deux réseaux. Le réseau ainsi obtenu comprend un total de 2531 individus et 15384 relations. Parmi eux, 77,7% ($n=1966$) sont exclusivement sur Facebook, 19,5% ($n=493$) exclusivement dans les données policières, et 2,8% ($n=72$) appartiennent aux deux réseaux (Tableau II).

Tableau II. Portrait de l'échantillon du réseau total

	N	%	% valide	Moyenne	Médiane	MIN	MAX	Écart-type
CARACTERISTIQUES INDIVIDUELLES (N=2531)								
Informations obtenues à l'aide de Facebook								
Degré								
Degré 1	1850	73,1	90,8					
Degré 2	188	7,4	9,2					
Manquant	493	19,5						
Total	2531	100,0	100,0					
Lieu de résidence								
Lieu de résidence	376	14,9	100,0					
Aux USA	8	0,3	2,1					
Au Canada	352	13,9	93,6					
Au Québec	332	13,1	88,3					
<i>RMR Montréal</i>	117	4,6	31,1					
À Montréal	71	2,8	18,9					
<i>RMR Québec</i>	37	1,5	9,8					
<i>RMR Saguenay</i>	1	0,0	0,3					
<i>RMR Sherbrooke</i>	53	2,1	14,1					
<i>RMR Trois-Rivières</i>	3	0,1	0,8					
<i>RMR Ottawa-Gatineau</i>	1	0,0	0,23					
Manquant	2155	85,1						
Total	2531	100,0	100,0					
Distance en véhicule de Montréal (km)	361	14,3	100,0	267,27	107,00	0,00	4862,00	721,11
Manquant	2170	85,7						
Total	2531	100,0	100,0					
Études								
Secondaire	473	18,7	53,4					
École technique ou de métiers, collège communautaire ou Cégep	196	7,7	22,1					
Université	109	4,3	12,3					
Pas fiable	107	4,2	12,1					
Manquant	1646	65,0						
Total	2531	100,0	100,0					
Emploi								
Esthétique et soins personnels	52	2,1	6,3					
Hôtellerie et restauration	60	2,4	7,3					
Bar	71	2,8	8,6					
Construction et chantiers	120	4,7	14,5					
Loisirs et divertissement	36	1,4	4,4					
Transports	87	3,4	10,5					
Habillement et mode	22	0,9	2,7					
Services sociaux et médicaux	53	2,1	6,4					
Public	14	0,6	1,7					
Alimentation	26	1,0	3,2					
Finances, immobilier et assurances	30	1,2	3,6					
Enfance et enseignement	27	1,1	3,3					
Autres	39	1,5	4,7					
Motard	44	1,7	5,3					

Pas fiable	144	5,7	17,5					
Manquant	1706	67,4						
Total	2531	100,0	100,0					
Informations obtenues à l'aide des données policières								
Évènements								
Participation aux évènements	575	22,7	100,0	1,90	1,00	1,00	55,00	3,32
Manquant	1956	77,3						
Total	2531	100,0	100,0					
Informations obtenues via la combinaison des deux sources de données								
Sexe								
Sexe								
Hommes	1448	57,2	63,3					
Femmes	839	33,1	36,7					
Manquant	244	9,6						
Total	2531	100,0	100,0					
Présence dans les réseaux								
Facebook	2038	80,5	80,5					
Exclusivement Facebook	1966	77,7	77,7					
Police	575	22,7	22,7					
Exclusivement réseau police	493	19,5	19,5					
Dans les deux réseaux	72	2,8	2,8					
Total	2531	100,0	100,0					
CARACTERISTIQUES RELATIONNELLES (N_REL=15384)								
Informations obtenues à l'aide de Facebook								
Temps amis Facebook								
En mois	2068	13,4	100,0	16,81	9,00	1,00	105,00	21,63
Manquant	13316	86,6						
Total	15384	100,0	100,0					
Informations obtenues via les données policières								
Nature et fréquence des relations								
Nombre	11626	75,6	100,0	1,10	1,00	1,00	26,00	0,45
Criminelles	78	0,5	0,7	1,00	1,00	1,00	1,00	0,00
Non criminelles	11549	75,1	99,3	1,10	1,00	1,00	26,00	0,45
Manquant	3758	24,4						
Total	15384	100,0	100,0					

Le réseau comprend donc un plus grand nombre d'acteurs que les réseaux provenant indépendamment de chaque source. En combinant les informations provenant des deux sources de données, ce réseau compte 63,3% d'hommes (n=1448) et 36,7% de femmes (n=839), l'information étant manquante pour seulement 9,6% des individus (n=244). La proportion d'hommes est donc plus importante que pour le réseau Facebook et plus faible que pour le réseau issu des données policières. Cela amènerait à déduire que le réseau total combine les tendances obtenues via chacune des deux sources de données.

Les trois réseaux ainsi décrits concernent l'ensemble des individus composant le réseau des dix individus de départ, présents dans les différentes sources de données. La question se pose de savoir si les mêmes tendances peuvent être observées pour les réseaux comprenant uniquement les individus présents dans les deux sources de données.

3.1.4. Les réseaux des individus présents simultanément dans les deux sources de données

La même démarche méthodologique est appliquée pour les individus présents dans les deux sources de données. En effet, 72 individus sont à la fois présents dans les données Facebook et dans les données policières. Leur réseau Facebook, leur réseau issu des données policières et leur réseau total sont successivement présentés.

a) Le réseau Facebook des individus présents dans les deux sources de données

Le réseau Facebook des individus présents dans les deux sources de données (n=72), se compose en grande majorité d'hommes (n=59 ; 90,8% d'hommes et n=6 ; 9,2% de femmes). Il n'est pas surprenant que cela suive la même tendance que celle du réseau issu des données policières comprenant l'ensemble des individus plutôt que celui de Facebook. Les individus dont il est question dans cette section sont en effet présents dans les deux sources de données, donc également dans les données policières.

De même que pour le réseau Facebook comprenant l'ensemble des individus, la plupart se trouve au premier degré de relations (n=62 ; 86,1%). S'agissant du lieu de résidence, seul 15,3% d'entre eux ont renseigné cette information (n=11). Tous habitent au Québec, plus précisément la RMR de Montréal dans 72,7% des cas (n=8), dont 5 à Montréal même (45,5%). En effet, ils résident à une distance médiane de 20,00 km de

Montréal en véhicule ($\mu=89,18$; $Md=20,00$). Comme pour le sexe, cela n'est pas surprenant, ces individus étant présents également dans les données policières montréalaises, donc sur le territoire montréalais.

S'agissant des sphères académiques et professionnelles, les tendances générales sont similaires à celles observées pour le réseau Facebook comprenant l'ensemble des individus. Ainsi, 43,4% ($n=30$) ont indiqué de l'information sur les études suivies, dont la plupart ont un niveau correspondant au secondaire ($n=14$; 46,7%). Notons que 20% ($n=6$) n'ont pas fourni d'information fiable. En outre, seul 36,1% ($n=26$) ont renseigné leur emploi ($n=26$). Parmi les secteurs d'activité professionnelle les plus présents au sein du réseau, figurent la construction ($n=5$; 19,2%), les bars ($n=4$; 15,4%), les loisirs ($n=3$; 11,5%) et ceux se définissant comme motards ($n=3$; 11,5%). Notons que le secteur des transports, l'un des secteurs les plus représentés au sein du réseau Facebook comprenant tous les individus, ne figure pas au sein du présent réseau. Par ailleurs, de même que pour la scolarité, certains individus ont renseigné de l'information mais cette dernière n'est pas fiable ($n=5$; 19,2%). Aussi, 58,3% ($n=42$) n'ont fourni aucune information sur leur parcours académique et 63,9% ($n=46$) sur leur carrière professionnelle. Ces résultats peuvent révéler une certaine volonté de discrétion, soit par un accès restreint à ce type d'information ou par l'absence volontaire de toute indication. En effet, un individu figurant dans les données policières et sachant que les réseaux sociaux sont une source d'information pour les agences d'application de la loi, peut avoir intérêt à se montrer le plus discret possible sur ces plateformes (Décary-Héту et Morselli, 2011). Cela ne semble toutefois pas une précaution prise par tous, ni spécifique à ces derniers, puisqu'il s'agit d'une tendance également observée dans le réseau Facebook de l'ensemble des individus.

Les 128 relations composant le réseau Facebook des individus présents dans les deux sources de données sont relativement récentes. Les individus sont amis sur Facebook en moyenne depuis 14 mois ($\mu=13,83$; $Md=10,00$).

b) Le réseau issu des données policières des individus présents dans les deux sources de données

Le réseau issu des données policières ne comprenant que les individus présents dans les deux sources de données est majoritairement composé d'hommes ($n=68$; 94,4%), soit une proportion proche de celle observée pour celui comprenant l'ensemble des individus. Ces derniers sont en moyenne impliqués dans plus de quatre événements, criminels ou pas ($\mu=4,22$; $Md=2,00$), soit plus que l'ensemble des individus. Cela peut paraître étonnant, puisqu'il pouvait être attendu que les individus présents dans les deux sources de données soient moins actifs, ceux l'étant plus pouvant avoir intérêt à faire preuve de plus de discrétion et ne pas posséder de compte Facebook. Cela semble être le cas des acteurs 3053, 3023, 3468 et 3374, qui figuraient parmi les dix individus comptant le plus d'événements au sein du réseau comprenant l'ensemble des individus mais qui ne figurent pas dans Facebook. En outre, le statut qu'un individu tend à obtenir le plus souvent au cours d'un événement est le même que pour le réseau comprenant l'ensemble des individus, à savoir celui de personne interpellée ($\mu=3,39$; $Md=1,50$).

Ainsi, excepté le nombre d'événements dans lesquels ces individus sont impliqués, les mêmes tendances s'observent pour les réseaux issus des données policières, qu'ils comprennent l'ensemble des individus ou seulement ceux présents dans les deux sources de données. Il s'agit de savoir s'il en est de même pour le réseau total.

c) Le réseau total des individus présents dans les deux sources de données

Le réseau total des individus présents dans les deux sources de données comprend donc les mêmes 72 individus, regroupant à la fois les caractéristiques obtenues à l'aide de Facebook, celles obtenues à l'aide des données policières et celles obtenues en combinant ces deux sources de données (Tableau III).

Tableau III. Portrait de l'échantillon du réseau total des individus présents dans les deux sources de données

	N	%	% valide	Moyenne	Médiane	MIN	MAX	Écart-type
CARACTERISTIQUES INDIVIDUELLES (N=72)								
Informations obtenues à l'aide de Facebook								
Degré								
Degré 1	62	86,1	86,1					
Degré 2	10	13,9	13,9					
Manquant	0	0						
Total	72	100,0	100,0					
Lieu de résidence								
Lieu de résidence	11	15,3	100,0					
Aux USA	0	0	0					
Au Canada	11	15,3	100,0					
Au Québec	11	15,3	100,0					
<i>RMR Montreal</i>	8	11,1	72,7					
A Montreal	5	6,9	45,5					
<i>RMR Québec</i>	1	1,4	9,1					
<i>RMR Saguenay</i>	1	1,4	9,1					
<i>RMR Sherbrooke</i>	0	0	0					
<i>RMR Trois-Rivières</i>	0	0	0					
<i>RMR Ottawa-Gatineau</i>	0	0	0					
Manquant	61	84,7						
Total	72	100,0	100,0					
Distance en véhicule de Montréal (km)	11	15,3	100,0	89,18	20,00	0,00	464,00	149,40
Manquant	61	84,7						
Total	72	100,0	100,0					
Études								
Secondaire	14	19,4	46,7					
École technique ou de métiers, collège communautaire ou Cégep	5	6,9	16,7					
Université	5	6,9	16,7					
Pas fiable	6	8,3	20,0					
Manquant	42	58,3						
Total	72	100,0	100					
Emploi								
Esthétique et soins personnels	2	2,8	7,7					
Hôtellerie et restauration	1	1,4	3,8					
Bar	4	5,6	15,4					
Construction et chantiers	5	6,9	19,2					
Loisirs et divertissement	3	4,2	11,5					
Transports	0	0	0					
Habillement et mode	0	0	0					
Services sociaux et médicaux	0	0	0					
Public	0	0	0					
Alimentation	1	1,4	3,8					
Finances, immobilier et assurances	2	2,8	7,7					
Enfance et enseignement	0	0	0					
Autres	0	0	0					

Motard	3	4,2	11,5					
Pas fiable	5	6,9	19,2					
Manquant	46	63,9						
Total	72	100	100					
Informations obtenues à l'aide des données policières								
Évènements								
Participation aux évènements	72	100,0	100,0	4,22	2,00	8,25	1,00	55,00
Manquant	0	0						
Total	72	100,0	100,0					
Informations obtenues via la combinaison des deux sources de données								
Sexe								
Sexe								
Hommes	68	94,4	94,4					
Femmes	4	5,6	5,6					
Manquant	0	0						
Total	72	100,0	100,0					
CARACTERISTIQUES RELATIONNELLES (N_REL=552)								
Informations obtenues à l'aide de Facebook								
Temps amis Facebook								
En mois	101	18,3	100,0	13,83	10,00	1,00	80,00	13,22
Manquant	350	63,4						
Total	552	100,0	100,0					
Informations obtenues via les données policières								
Nature et fréquence des relations (Police)								
Nombre	478	86,6	100,0	1,43	1,00	1,43	1,00	26,00
Criminelles	1	0,2	100,0	0,00	0,00	0,46	0,00	1,00
Non criminelles	477	86,4	100,0	1,43	1,00	1,43	1,00	26,00
Manquant	74	13,4						
Total	552	100,0	100,0					

La combinaison des deux sources permet d'établir que le réseau compte 94,4% d'hommes (n=68) et 5,6% de femmes (n=4), permettant alors de déterminer le sexe des individus pour lesquels cette information ne pouvait être déduite du profil Facebook. Ce réseau comprend un nombre plus important de relations, avec un total de 552 relations.

En somme, six réseaux ont été construits. Trois d'entre eux concernent l'ensemble des individus, c'est-à-dire le réseau Facebook des dix individus de départ, ainsi que leur réseau issu des données policières et leur réseau total en combinant ces deux derniers réseaux. Trois autres (Facebook, données policières, total) concernent uniquement les individus présents simultanément dans les deux sources de données utilisées.

3.2. Analyses

L'analyse de réseau offre un certain nombre d'outils pertinents aux organisations policières. Dans un contexte de renseignement, elle permet par exemple l'analyse de l'affiliation et de la complicité, la déstabilisation de réseaux, en détectant l'existence de sous-groupes et leurs modèles d'interactions via l'identification d'individus centraux (Sparrow, 1991 ; Xu et Chen, 2005 ; Koschade, 2006 Van der Hulst, 2008 ; Bouchard et Konarski, 2014 ; Martineau et Lecocq, 2016). Différents aspects seront donc étudiés. Tout d'abord, une analyse comparative de la structure du réseau, des indicateurs de centralité des individus selon le réseau considéré ainsi que de leur rôle et position, est effectuée. L'influence que les caractéristiques individuelles peuvent avoir sur l'appartenance de certains individus aux deux sources de données, ainsi que sur le rôle et la position qu'ils occupent au sein des réseaux, est par la suite analysée.

3.2.1. Analyse comparative des différents réseaux

Une première partie des analyses est consacrée à l'étude comparative des différents réseaux construits, afin de définir l'apport de chacune des deux sources de données. D'une part, il s'agit de déterminer à l'aide des indicateurs de centralité quels sont les acteurs principaux de chaque réseau et s'ils diffèrent ou non d'un réseau à l'autre. D'autre part, il s'agit de déterminer à quel point les différents réseaux correspondent les uns aux autres. Il est donc à la fois question de la structure des différents réseaux, ainsi que des rôles et positions que les acteurs occupent au sein de ces derniers.

a) Rôles et positions au sein du réseau : indicateurs de centralité et acteurs principaux

La centralité d'un individu permet de connaître son positionnement au sein d'un réseau (Morselli, 2009). Le TOP 15 des acteurs les plus importants parmi les différents réseaux a été déterminé. La limite a été établie à quinze individus. Cela permet un juste milieu entre le ciblage d'un trop grand nombre d'individus, qui ne serait pas pertinent d'un point de vue opérationnel, et un nombre trop minime d'acteurs. Cela permet aussi de s'assurer que les dix premiers rangs ne sont pas occupés par les seuls dix individus de départ, ce qui indiquerait un biais dans les données. Les différents TOP 15 ainsi établis permettront de déterminer s'il s'agit des mêmes acteurs d'un réseau à l'autre.

Trois indicateurs ont été générés à l'aide du logiciel Gephi, à savoir la centralité de degré, d'intermédiarité et de vecteur propre. Le cumul de plusieurs indicateurs de centralité permet d'aborder l'importance des acteurs et leur rôle, sous différents aspects. En effet, la centralité de degré mesure le nombre de contacts directs entourant un nœud (Freeman, 1979 ; Morselli, 2009). Ainsi, un nœud ayant une centralité de degré élevée est un nœud attirant une forte concentration de connectivité directe au sein d'un réseau (Morselli, 2009). L'intermédiarité apporte une nuance, mettant en avant l'importance de la qualité des relations, en incluant le nombre de contacts indirects entourant un nœud (Freeman, 1977, 1979 ; Morselli, 2009). Elle mesure jusqu'à quel point un nœud médie les relations entre les autres nœuds par sa position parmi les chemins les plus courts au sein du réseau, c'est-à-dire parmi le nombre de degrés connectant une dyade. Elle représente la capacité de certains nœuds de contrôler le flux de connectivité ou de communication de manière indirecte au sein d'un réseau, devenant un intermédiaire ou courtier (Morselli, 2009). Enfin, la centralité de vecteur propre, développée par Bonacich (1972), apporte un aspect qualitatif à la centralité de degré. Certains nœuds seraient centraux parce qu'ils ont un degré élevé de contacts directs, ayant eux-mêmes une centralité de degré élevée dans le réseau (Morselli, 2009).

Ces différentes mesures permettent donc de déterminer la position et le rôle d'un individu au sein du réseau, ce qui peut être influencé par la structure du réseau mais également influencer celle-ci. Il est donc également nécessaire de comparer la structure des différents réseaux pour savoir si ces derniers correspondent les uns aux autres.

b) Correspondance des différents réseaux : analyse de corrélation

Afin de déterminer dans quelle mesure les différents réseaux correspondent les uns aux autres, une corrélation QAP (Quadratic Assignment Procedure) a été effectuée à l'aide du logiciel UCINET (Borgatti, Everett et Freeman, 2002). En effet, bien que Gephi ait été utilisé pour construire les différents réseaux et générer les indicateurs de centralité, UCINET, logiciel réputé pour l'analyse de réseaux et offrant un grand nombre d'analyses (Huisman et van Duijn, 2011), est l'un des rares logiciels à proposer cette mesure.

La corrélation QAP permet de générer des mesures de similarités entre différentes entrées de matrices carrées. Elle génère un coefficient de corrélation de Pearson. Elle effectue l'analyse plusieurs fois (500 itérations par défaut), afin de déterminer la fréquence qu'une mesure aléatoire soit plus grande ou égale à la mesure observée calculée. Une faible proportion ($p < 0,05$) suggère une forte relation entre les matrices qui soit peu probable d'être due à la chance. Elle est principalement utilisée pour tester l'association entre les réseaux, par exemple entre un réseau observé et un réseau modèle ou attendu (Hanneman et Riddle, 2005).

Cette mesure implique de comparer des réseaux comprenant le même nombre d'acteurs (Hanneman et Riddle, 2005). Par conséquent, elle sera générée parmi les réseaux des acteurs présents dans les deux sources de données ($n=72$). Ainsi pourront être constatées les correspondances entre le réseau Facebook et celui des données policières, entre le réseau Facebook et le réseau total, ainsi qu'entre le réseau des données policières et le réseau total.

La corrélation QAP permet donc de s'intéresser à la structure des différents réseaux. Celle-ci peut toutefois être également influencée par les caractéristiques individuelles des individus composant le réseau, ces dernières représentant une mise à disposition de ressources et un capital social (Van der Hulst, 2008).

3.2.2. Analyse de l'impact des caractéristiques individuelles dans les différents réseaux

Une seconde partie des analyses se concentre sur l'impact des caractéristiques individuelles. Le niveau de discrétion des individus selon leur présence dans les banques de données policières ou encore l'appartenance au réseau Facebook plutôt qu'à celui issu des données policières, n'ont pu être analysés. Trop peu de variables communes aux deux sources de données ou construites pour une telle analyse sans risquer un raisonnement tautologique, étaient disponibles. Les analyses visent alors à déterminer le type d'individu le plus susceptible de se retrouver dans un réseau précis, et les caractéristiques lui permettant d'occuper une position centrale.

a) Impact des caractéristiques individuelles sur l'appartenance aux différents réseaux

Il s'agit de déterminer les caractéristiques individuelles pouvant influencer la présence d'un individu dans un réseau. Plus précisément, il s'agit de savoir quels individus du réseau Facebook sont également présents dans celui issu des données policières, et inversement. La variable dépendante porte ainsi sur la présence dans les différentes sources de données de l'ensemble des individus (n=2531). Une régression logistique sera réalisée à l'aide du logiciel SPSS.

Les variables indépendantes varient selon la disponibilité de l'information. Excepté le sexe, certaines variables sont propres aux individus de Facebook et d'autres à ceux des données policières. Le sexe consiste en une variable binaire, obtenue en combinant les informations des deux sources de données.

Ainsi, les variables propres à Facebook sont utilisées pour prédire l'appartenance d'un individu présent dans le réseau Facebook (n=2038) au réseau total comprenant les individus présents dans les deux sources de données (n=72). Seront ainsi considérés, outre le sexe, si le lieu de résidence se situe à Montréal, le niveau de scolarité, et le secteur d'activité professionnelle. Le niveau de scolarité de l'individu consiste en quatre variables binaires, selon qu'il est au niveau secondaire, au cégep ou en formation professionnelle, au niveau universitaire, ou que l'information est inconnue, c'est-à-dire indisponible ou pas fiable. L'emploi est divisé par secteurs d'activités professionnelles les plus fréquents, à savoir s'il s'agit de personnes se déterminant comme motards, travaillant dans les loisirs, la construction, l'esthétique ou les bars. Chaque secteur fait l'objet d'une variable binaire. Deux autres variables binaires indiquant si l'information est inconnue ou concerne un autre domaine, ont été créées.

Inversement, les variables propres aux données policières sont utilisées pour prédire l'appartenance d'un individu présent dans le réseau issu des données policières (n=565) au réseau total comprenant les individus présents dans les deux sources de données (n=72). Dans le cadre de cette analyse, seront considérés le sexe et le nombre d'évènements dans lesquels un individu apparaît dans les données policières. Ce dernier consiste en deux variables binaires distinctes, à savoir si l'individu a participé à un évènement ou à plusieurs. En effet, les portraits d'échantillons établissent que la plupart des individus ont participé en moyenne à un évènement.

Toutefois, outre l'appartenance à un réseau, les caractéristiques individuelles pourraient également impacter la position et le rôle d'un acteur au sein du réseau.

b) Impact des caractéristiques individuelles sur le rôle central d'un individu au sein des différents réseaux

La première partie des analyses porte sur la fréquence d'apparition d'un acteur précis dans les différents réseaux et sur son rôle selon la position occupée. Les portraits d'échantillon indiquent que certaines caractéristiques individuelles, telles qu'un niveau de scolarité ou un secteur d'activité professionnelle précis, sont plus présentes que d'autres. Il s'agit de déterminer lesquelles permettraient à un individu d'occuper une place centrale au sein du réseau Facebook ou du réseau issu des données policières. Les mêmes variables indépendantes que précédemment seront considérées dans ces analyses. Des analyses de régression linéaire multiple seront effectuées.

S'agissant des variables dépendantes, c'est-à-dire les indicateurs de centralité, seule la centralité de vecteur propre sera analysée. En effet, les différents indicateurs de centralité étant corrélés entre eux, des résultats similaires sont attendus. Afin de respecter les postulats de multicollinéarité nécessaires à la réalisation d'une régression linéaire multiple, la centralité de vecteur propre a été préférée. Elle permet de révéler l'influence que peut avoir un acteur au sein du réseau (Bonacich, 1972 ; Morselli, 2009) et différentes recherches ont évoqué l'influence que certaines caractéristiques individuelles pouvaient procurer à un acteur au sein d'un réseau (Van der Hulst, 2008 ; Morselli, 2009). Cela mérite attention, notamment dans un objectif de renseignement policier. Cette variable a été loguée, afin de respecter le postulat de normalité de distribution.

Ces analyses seront réalisées d'une part avec l'ensemble des individus présents dans le réseau Facebook (n=2038) et d'autre part avec l'ensemble des individus présents dans le réseau issu des données policières (n=565). Cela permettra de déterminer si l'impact des caractéristiques individuelles est le même selon le réseau considéré, et lesquelles ont le plus d'influence sur le rôle ou la position d'un individu.

4. Analyse et interprétation des résultats

Les résultats consistent, d'une part, en une analyse comparative des différents réseaux, en termes de centralité, d'acteurs principaux et de concordance. D'autre part, l'impact des caractéristiques individuelles des acteurs sur leur présence dans un réseau plutôt qu'un autre, ainsi que sur leur position en termes de centralité, est analysé.

4.1. Analyse comparative des différents réseaux

Étant donné la complexité de la structure du réseau total comprenant l'ensemble des acteurs, des analyses plus approfondies basées sur des mesures d'analyse de réseau, telles la centralité ou la concordance entre plusieurs réseaux, étaient nécessaires. En effet, l'analyse de réseau va au-delà de la production de graphes et de l'aspect visuel. Elle permet d'analyser les acteurs et relations d'un réseau sur la base de mesures arithmétiques (Van der Hulst, 2008 ; McGloin et Kirk, 2010). Toutefois, un tel graphe est disponible en annexe (Annexe 3). Les résultats des analyses portant sur les réseaux comprenant l'ensemble des individus sont tout d'abord présentés, suivis de ceux concernant les individus présents dans les deux sources de données uniquement. Il est question des tendances des différents indicateurs de centralité selon le réseau considéré. Les individus les plus fréquemment présents dans les différents TOP 15 établis sur la base de chaque indicateur, ainsi que la position qu'ils occupent, sont ensuite analysés. Enfin, la concordance entre les différents réseaux est analysée afin de déterminer précisément l'apport réciproque de chaque source.

4.1.1. Analyse comparative des réseaux comprenant l'ensemble des individus

Dans la présente section, les résultats concernant les tendances centrales des différents indicateurs de centralité sont présentés, puis les individus présents le plus souvent dans les différents TOP 15 et leur position, sont analysés.

a) Tendances centrales des différents indicateurs de centralité selon les réseaux

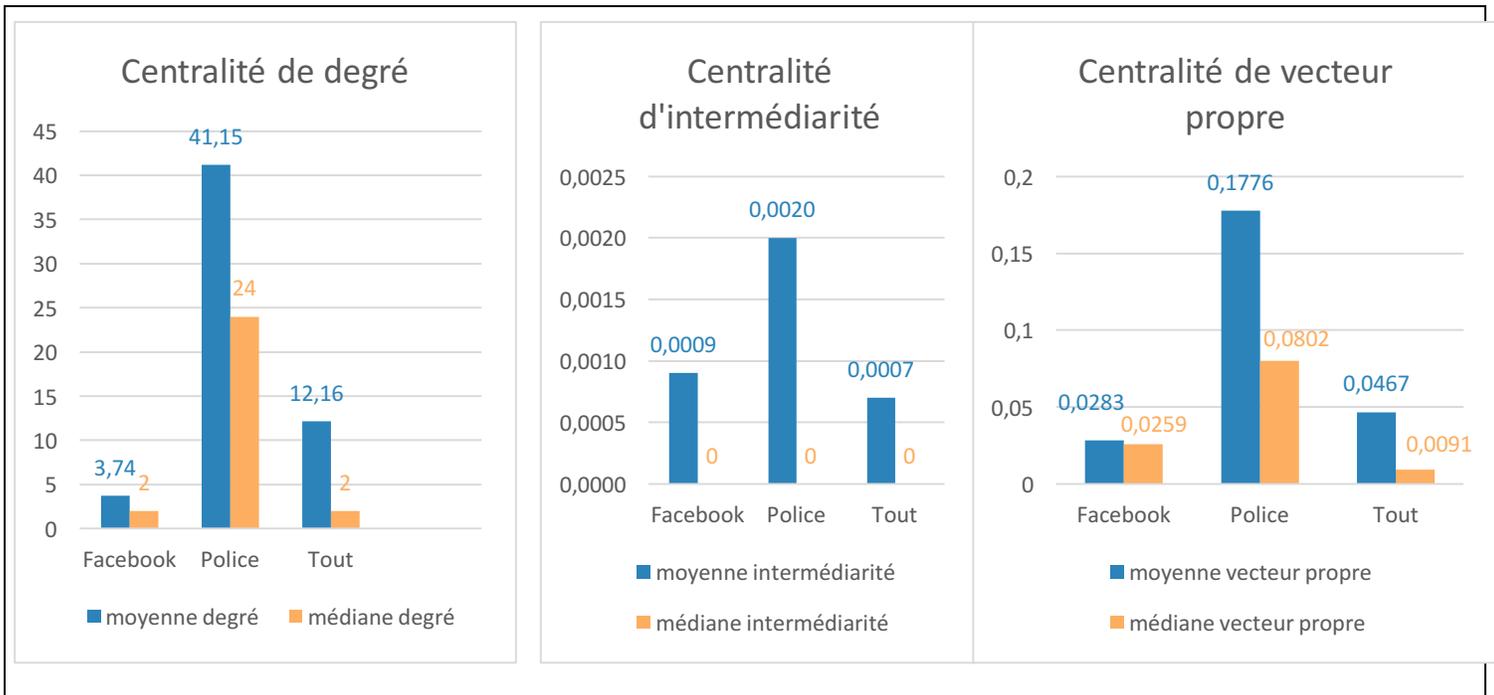
Les tendances centrales de chaque indicateur de centralité, dépendamment du réseau considéré, permettent d'avoir une vision générale de l'apport des différentes sources de données utilisées. En calculant les moyennes et médianes de chaque indicateur parmi les acteurs des différents réseaux, il a été possible de les comparer selon qu'est considéré le réseau Facebook, le réseau issu des données policières, ou le réseau total, avec l'ensemble des individus ou uniquement ceux présents dans les TOP 15. Les statistiques descriptives de ces analyses sont disponibles en annexe (Annexe 4).

- Comparaison des réseaux comprenant l'ensemble des individus

L'analyse comparative des réseaux comprenant l'ensemble des individus indique que le réseau issu des données policières a une plus forte centralité de degré ($\mu=41,15$; $Md=24$ pour les données policières contre $\mu=3,74$; $Md=2$ pour Facebook), d'intermédiation ($\mu=0,0020$; $Md=0$ contre $\mu=0,0009$; $Md=0$) et de vecteur propre ($\mu=0,1776$; $Md=0,0802$ contre $\mu=0,0283$; $Md=0,0259$) que le réseau Facebook (Figure 2). Les individus composant le réseau Facebook entretiendraient moins de contacts directs avec leur entourage et occuperaient des positions moins stratégiques et influentes que ceux composant le réseau issu des données policières. Ces résultats pourraient

s'expliquer par le nombre plus important de relations dans le réseau issu des données policières. Cela influence les mesures, particulièrement la centralité de degré.

Figure 2. Comparaison des indicateurs de centralité de chacun des réseaux comprenant l'ensemble des acteurs



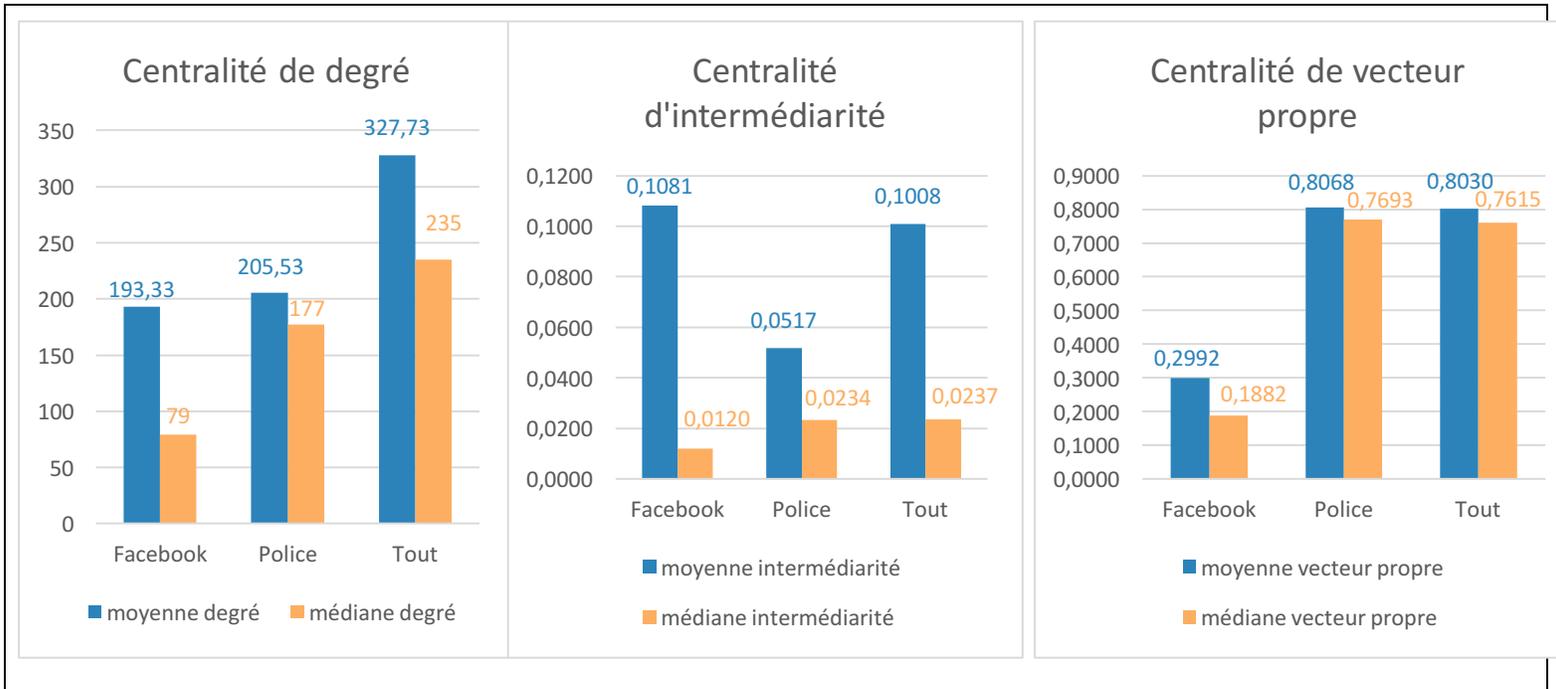
Toutefois, cet écart semble disparaître lorsque le réseau Facebook et le réseau issu des données policières sont fusionnés, les valeurs obtenues devenant alors proches de celles du réseau Facebook ($\mu=12,16$ et $Md=2$ pour la centralité de degré ; $\mu=0,0007$ et $Md=0$ pour l'intermédiarité ; $\mu=0,0467$ et $Md=0,0091$ pour la centralité de vecteur propre). Le nombre d'acteurs plus important dans le réseau Facebook influencerait les mesures du réseau total, ajoutant alors du bruit et pouvant induire en erreur.

Ces résultats tendraient vers le fait que l'analyse de Facebook n'apporterait pas une grande plus-value, mais au contraire disperserait les mesures obtenues à partir des données policières. Il s'agit de savoir s'il en est de même lorsqu'est considéré un groupe plus homogène d'acteurs au sein des réseaux, à savoir ceux présents dans les TOP 15.

- **Comparaison des réseaux avec les individus présents dans les TOP 15**

Les indicateurs de centralité de cette analyse présentent des valeurs plus élevées que précédemment. Ils portent sur les individus présents dans les TOP 15, c'est-à-dire les quinze individus ayant les valeurs les plus élevées en termes de centralité. Leurs tendances diffèrent de celles observées précédemment (Figure 3).

Figure 3. Comparaison des indicateurs de centralité des acteurs présents dans le TOP 15 des réseaux comprenant l'ensemble des acteurs



Les indicateurs de centralité présentent une valeur plus élevée dans les réseaux issus des données policières que dans Facebook, que ce soit pour la centralité de degré ($\mu=205,53$; $Md=177$ contre $\mu=193,33$; $Md=79$), d'intermédiation si l'on considère la médiane ($\mu=0,0517$; $Md=0,0234$ contre $\mu=0,1081$; $Md=0,0120$) ou de vecteur propre ($\mu=0,8068$; $Md=0,7693$ contre $\mu=0,2992$; $Md=0,1882$). Toutefois, les valeurs du réseau total sont plus proches de celles du réseau issu des données policières. Cela semble indiquer que les données policières ont un apport significatif quant au réseau total.

Il est alors nécessaire d'approfondir l'analyse de ces acteurs principaux en termes de centralité, ces derniers semblant présenter un profil différent des autres acteurs.

b) Acteurs principaux en termes de centralité selon les différents réseaux

Les quinze acteurs principaux de chaque réseau, en ce qu'ils en sont les acteurs les plus centraux, ont pu être identifiés pour chaque indicateur de centralité (n=38). Bien que les acteurs et leur nombre diffèrent d'un réseau à l'autre, certains sont plus fréquemment présents au sein de ces classements (Tableau IV).

En effet, l'acteur 11 est présent dans chacun des neuf TOP 15 générés. Il est suivi des acteurs 2 et 5, présents à huit reprises, de l'acteur 3 présent sept fois, ainsi que des acteurs 47 et 73, présents six fois. Cela signifie qu'ils ressortent parmi les acteurs les plus importants en termes de centralité, peu importe l'indicateur considéré. En outre, huit des dix individus de départ sont présents et ce, fréquemment, au sein de ces classements (ID 1 à 10). Cela pourrait cependant être le fait que les réseaux ont été générés à partir de leurs relations. Toutefois, la présence des acteurs 2, 3, 4, 5 et 6 dans les TOP 15 des différents réseaux, indique que leur présence dans le TOP 15 des réseaux issus des données policières ne serait pas seulement la conséquence d'une plus grande attention policière les concernant.

Tableau IV. Fréquence d'apparition dans les différents TOP 15 des réseaux comprenant l'ensemble des individus

ID	Fréquence (sur 9)	Rang selon fréquence
11	9	1
2	8	2
5	8	3
3	7	4
47	6	5
73	6	6
4	5	7
6	5	8
10	5	9
200	5	10
615	5	11
16	4	12
17	4	13
23	4	14
26	4	15
250	4	16
3179	4	17
3551	4	18
1	3	19
723	3	20
3039	3	21
3043	3	22
3129	3	23
3403	3	24
3468	3	25
7	2	26
12	2	27
3245	2	28
3509	2	29
22	1	30
30	1	31
31	1	32
37	1	33
44	1	34
3001	1	35
3023	1	36
3188	1	37
3301	1	38

Il s'agit alors de déterminer quelle position précise ces différents acteurs principaux occupent au sein du réseau. Le classement des quinze principaux individus, selon chaque indicateur et le réseau considéré, indique que les acteurs les plus récurrents, n'occupent pas nécessairement les premiers rangs, à l'image de l'acteur 11 qui n'apparaît que dans les rangs 8 à 15 (Tableau V).

Tableau V. TOP 15 des acteurs centraux de chaque réseau comprenant l'ensemble des individus par indicateur de centralité.

RANG	Facebook			Police			Total		
	Degré	Intermédiation	Vecteur propre	Degré	Intermédiation	Vecteur propre	Degré	Intermédiation	Vecteur propre
1	6	6	6	5	5	5	2	6	5
2	2	2	2	73	73	73	6	2	73
3	4	4	4	3179	2	3179	4	4	3129
4	10	10	10	47	3	615	5	10	2
5	200	200	200	2	3468	250	10	200	615
6	3	3	3	615	47	47	73	5	47
7	17	17	16	3468	1	3403	3	3	250
8	16	23	17	3551-250	3039	3551	200	73	11
9	23	16	23	3551-250	3301	11	11	17	3403
10	11	26	12	3129	3188	3043	3179	47	3551
11	22	11	26	11	3179	3509	47	26	3
12	44	12	5	3403	11	723	615	11	3043
13	26-31	5	11	3039	3129	3245	3468	1	723
14	26-31	37	7	723	3023	3039	250	16	3509
15	30	7	1	3043	615	3001	3551	23	3245

Toutefois, les autres acteurs les plus fréquents dans les différents TOP 15 sont également présents dans les premiers rangs. Ainsi, l'acteur 2 apparaît être le plus important en terme de centralité de degré dans le réseau total, et le deuxième plus important dans le réseau Facebook pour chaque indicateur, ainsi qu'en termes d'intermédiation dans le réseau total. Similairement, l'acteur 5 occupe le premier rang pour chaque indicateur dans le réseau issu des données policières, de même que dans le réseau total s'agissant de la centralité de vecteur propre. L'acteur 73, quant à lui, occupe la deuxième place dans le réseau issu des données policières pour chaque indicateur, ainsi qu'en termes de centralité de vecteur propre dans le réseau total. Enfin, les acteurs 3 et 47 apparaissent au 4^e rang dans le réseau issu des données policières, respectivement en termes de centralité d'intermédiation et de degré. Le classement quant à la fréquence d'apparition dans les différents TOP 15 ne reflète donc pas nécessairement les positions occupées, lorsque chaque indicateur est considéré individuellement.

En outre, le classement des acteurs principaux dans le réseau Facebook ou issu des données policières n'engendrerait pas forcément le classement du réseau total pour un indicateur donné. Un acteur qui se retrouve le plus central en termes de centralité de degré pour le réseau Facebook ou pour le réseau des données policières, ne le sera pas forcément pour le réseau total. En d'autres termes, l'individu qui a le plus de contacts directs sur Facebook ou dans la sphère criminelle, n'en aura pas forcément le plus dans le réseau total. Ainsi, l'acteur 6 qui apparaît être le plus central dans le réseau Facebook, peu importe l'indicateur de centralité considéré, n'occupe ce rang qu'en termes de centralité d'intermédiation dans le réseau total, le deuxième rang en termes de centralité de degré et n'apparaît pas dans le classement en termes de centralité de vecteur propre. Il s'agit toutefois d'une position stratégique dans le réseau total, permettant à la fois l'efficacité, la sécurité, et le contrôle du flux d'information et de communication (Morselli, 2009). Similairement, l'acteur 5 est le plus central dans le réseau issu des données policières mais n'occupe ce rang qu'en termes de centralité de vecteur propre dans le réseau total, ce qui est néanmoins signe d'influence (Morselli, 2009).

Cela tendrait donc vers une complémentarité des deux sources utilisées, lorsque la position de chaque acteur est analysée séparément. En effet, Facebook semble apporter plus que les données policières pour ce qui est de l'acteur 6. Toutefois, ce serait l'inverse pour ce qui est de l'acteur 5, d'autant plus qu'il s'agit d'un des individus de départ pour lequel nous ne possédions pas l'ensemble de sa liste d'amis Facebook. Le fait qu'un individu tel l'acteur 6 soit important dans le réseau Facebook et dans le réseau total mais absent des TOP 15 du réseau issu des données policières, pourrait signifier que son importance est réelle mais pas captée par les données policières. Ou bien, cela pourrait indiquer que l'importance dans le réseau Facebook est relative comparativement à la place occupée dans le réseau issu des données policières. Dans ce cas, l'individu ne devrait pas faire l'objet d'une attention policière particulière. Ces hypothèses signifieraient que les relations entretenues dans l'un des réseaux ont plus d'impact et d'importance que celles entretenues dans l'autre, sans pour autant déterminer desquelles il s'agit. Afin d'avoir un meilleur comparatif, les prochains résultats concernent donc les réseaux des acteurs présents dans les deux sources de données uniquement.

4.1.2. Analyse comparative des différents réseaux comprenant uniquement les individus présents dans les deux sources de données

Dans la présente section, la même démarche d'analyse que précédemment est appliquée, à savoir la comparaison des tendances centrales des différents indicateurs de centralité selon le réseau considéré, l'analyse des individus présents le plus souvent dans les différents TOP 15 et de leur position, ainsi que la concordance entre les réseaux.

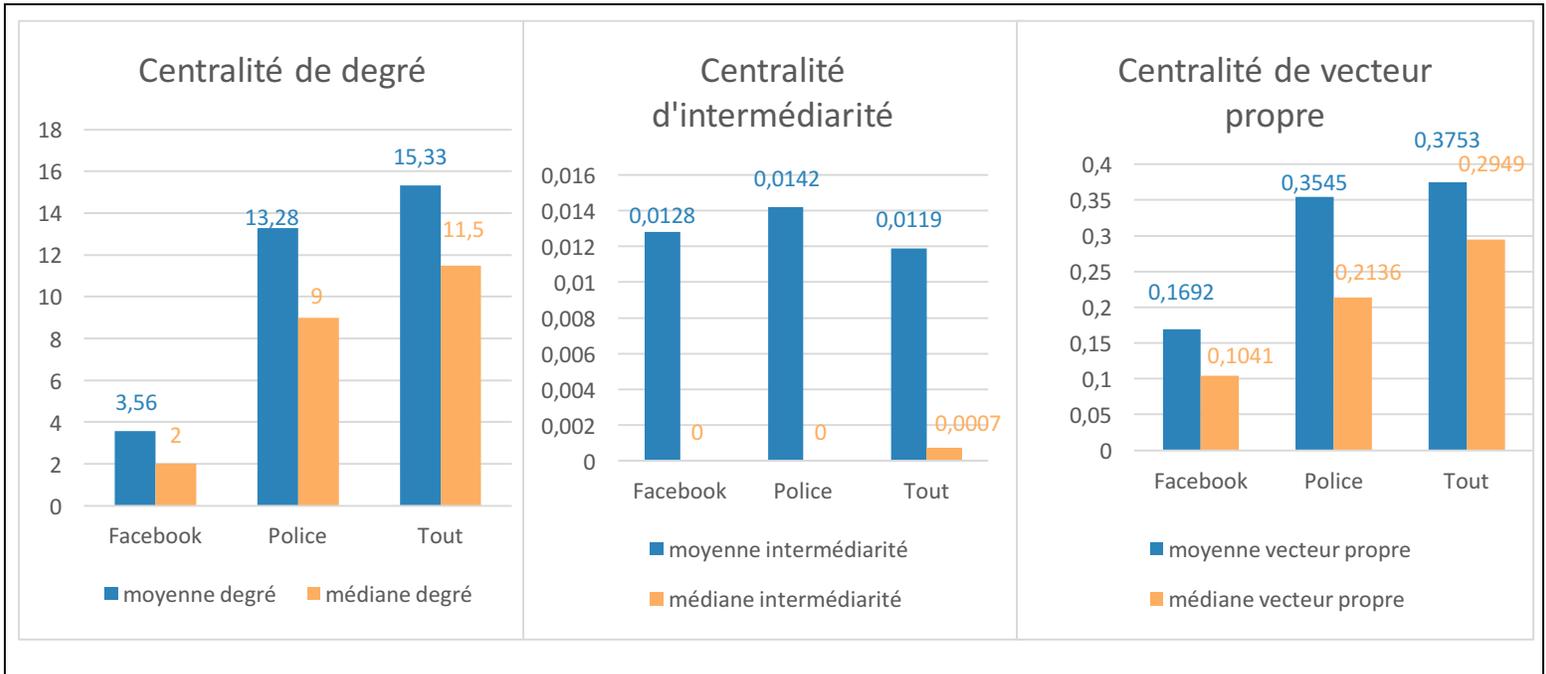
a) Tendances centrales des différents indicateurs de centralité selon les réseaux

Les tendances centrales de chaque indicateur de centralité selon le réseau considéré sont analysées, avec l'ensemble des individus présents dans les deux sources de données, puis ceux uniquement présents dans les TOP 15. Les statistiques descriptives de ces analyses sont disponibles en annexe (Annexe 5).

- Comparaison des réseaux comprenant les individus présents dans les deux sources de données

L'analyse comparative des réseaux comprenant l'ensemble des individus présents dans les deux sources de données, amène aux mêmes constats et considérations que lors de l'analyse comparative des réseaux comprenant l'ensemble des individus (Figure 4). Le réseau issu des données policières présente des valeurs plus élevées que le réseau Facebook. Toutefois, les valeurs de centralité de degré sont plus faibles que celles obtenues pour l'ensemble des individus, à la fois dans le réseau Facebook ($\mu=3,56$; $Md=2$) et dans celui issu des données policières ($\mu=13,28$; $Md=9$), à l'inverse des valeurs d'intermédiarité ($\mu=0,0128$; $Md=0$ pour Facebook et $\mu=0,0142$; $Md=0$ pour les données policières). Ces individus auraient donc tendance à entretenir des relations moins directes, mais à occuper une place plus stratégique au sein des différents réseaux.

Figure 4. Comparaison des indicateurs de centralité de chacun des réseaux comprenant les acteurs présents dans les deux sources de données



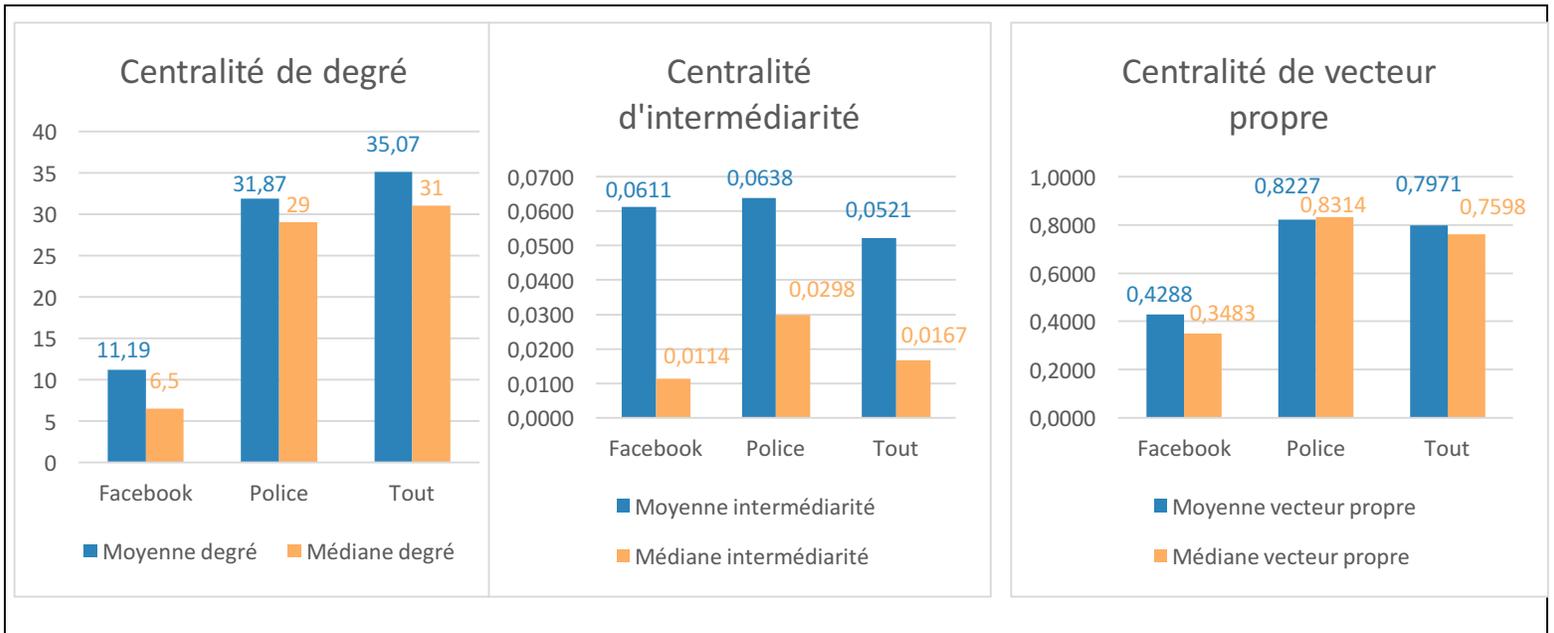
Ces acteurs occuperaient une place légèrement moins visible sur Facebook mais plus stratégique et moins vulnérable dans le réseau issu des données policières. Toutefois, la fusion des deux réseaux entraînerait une complémentarité des sources de données, les indicateurs de centralité pour le réseau total ayant une valeur plus élevée que ceux obtenus pour le réseau issu des données policières et le réseau Facebook ($\mu=15,33$ et $Md=11,5$; $\mu=0,0119$ et $Md=0,0007$; $\mu=0,3753$ et $Md=0,2949$, réciproquement pour la centralité de degré, d'intermédierité et de vecteur propre). Fusionner les deux réseaux engendrerait une plus grande centralité de degré et de vecteur propre et une légère dispersion de la centralité d'intermédierité. Cela signifierait que Facebook apporte plus d'informations pour les individus présents dans les deux sources de données. Cet apport semble toutefois limité, comparativement aux données policières.

Il s'agit de déterminer si les mêmes tendances s'observent lorsque seuls les individus présents dans les TOP 15 de ces réseaux sont considérés.

- **Comparaison des réseaux avec les individus présents dans les TOP 15**

Les tendances des indicateurs de centralité des acteurs présents dans les TOP 15 des réseaux des acteurs présents dans les deux sources de données sont similaires à celles observées pour ceux des réseaux comprenant l'ensemble des individus (Figure 5). Ils sont aussi plus proches des tendances observées pour les réseaux des acteurs présents dans les deux sources de données que pour ceux comprenant l'ensemble des individus.

Figure 5. Comparaison des indicateurs de centralité des acteurs présents dans le TOP 15 des réseaux des acteurs présents dans les deux sources de données



Les valeurs obtenues pour le réseau issu des données policières sont plus élevées que pour le réseau Facebook, aussi bien pour la centralité de degré ($\mu=31,87$; $Md=29$ contre $\mu=11,19$; $Md=6,5$), d'intermédiation ($\mu=0,0638$; $Md=0,0298$ contre $\mu=0,0611$; $Md=0,0114$) que de vecteur propre ($\mu=0,8227$; $Md=0,8314$ contre $\mu=0,4288$; $Md=0,3483$). Facebook ajouterait du bruit ou disperserait les résultats obtenus via les données policières s'agissant du réseau total, les valeurs étant plus faibles que pour le réseau issu des données policières, mis à part la centralité de degré.

Ces résultats suggèrent à nouveau que les deux sources de données sont complémentaires. L'apport de chaque source diffèrerait selon le type d'individus ciblé, à savoir les plus visibles (centralité de degré), les courtiers (centralité d'intermédiation) ou les plus influents (centralité de vecteur propre). Cependant, l'apport des données policières au réseau total semble plus important. Les valeurs de ces deux réseaux sont très similaires, notamment en termes de centralité de degré et de vecteur propre. Des analyses concernant plus particulièrement les différents acteurs sont donc nécessaires.

b) Acteurs principaux en termes de centralité selon les différents réseaux

De même que pour les réseaux comprenant l'ensemble des individus, les quinze acteurs les plus importants de chaque réseau ont été identifiés à l'aide des différents indicateurs de centralité (n=31).

Comme constaté précédemment, certains acteurs sont plus fréquemment présents que d'autres au sein des différents classements (Tableau VI). Les quatre acteurs dont la présence est la plus fréquente dans les différents TOP 15 générés sont les mêmes que pour ceux générés à partir des réseaux comprenant l'ensemble des individus. Il s'agit des acteurs 11, 2, 5, et 3, présents dans chacun des neuf TOP 15 générés. Les acteurs 2, 5 et 3 apparaissent donc plus souvent que lors de la comparaison des réseaux comprenant l'ensemble des individus. Ils sont suivis par l'acteur 7 en 5^e position, présent à huit reprises sur les neuf possibles, alors qu'il n'apparaissait qu'à deux reprises et était classé 26^e lorsque tous les individus étaient considérés. Les acteurs 73 et 47, qui figuraient parmi les plus fréquents lorsque l'ensemble des individus étaient considérés, se distinguent moins des autres et ne sont présents qu'à six reprises. De même, huit des dix individus de départ figurent dans les différents TOP 15.

Tableau VI. Fréquence d'apparition dans les différents TOP 15 des réseaux comprenant l'ensemble des individus présents dans les deux sources de données

ID	Fréquence (/9)	Rang selon fréquence
11	9	1
2	9	2
5	9	3
3	9	4
7	8	5
48	7	6
73	6	7
47	6	8
723	6	9
250	6	10
615	6	11
24	6	12
6	5	13
125	5	14
455	4	15
4	3	16
13	3	17
23	3	18
26	3	19
528	3	20
12	3	21
60	3	22
68	2	23
8	2	24
75	2	25
1	2	26
61	2	27
210	1	28
51	1	29
259	1	30
287	1	31

Il s'agit de déterminer la position précise de ces acteurs principaux au sein du réseau et s'il s'agit de la même que celle occupée dans les réseaux comprenant l'ensemble des individus. Les acteurs qui se démarquaient par une présence fréquente au sein des différents classements, apparaissent dès les premiers rangs (Tableau VII). Citons les exemples des acteurs 2 et 5, dont la présence à huit reprises dans les classements se perçoit dans les quatre premiers rangs. Il s'agit donc de deux acteurs visibles, importants et influents dans les différents réseaux.

Tableau VII. TOP 15 des acteurs centraux de chaque réseau comprenant les individus présents dans les deux sources de données par indicateur de centralité.

RANG	Facebook			Police			Total		
	Degré	Intermédiarité	Vecteur propre	Degré	Intermédiarité	Vecteur propre	Degré	Intermédiarité	Vecteur propre
1	2	2	2	5	5	5	2	2	2
2	6	6	6	73	2	73	5	5	5
3	3	3	4	2	73	615	73	3	73
4	4	5	3	615	7	47	3	11	3
5	11	4	12	11-47	11	2	11	6	615
6	5	11	5	11-47	3	723	615	73	47
7	12	210	11	3	615	455	47	615	11
8	26	23	26	723-250	48	250	6	48	723
9	23	12	7	723-250	47	11	723-250	75	250
10	1-7-13-60	60	1	48-455	75	3	723-250	7	455
11	1-7-13-60	26	13	48-455	259	528	455-48-7	47	24
12	1-7-13-60	13	61	24-125	24	48	455-48-7	723	125
13	1-7-13-60	287	8	24-125	125	125	455-48-7	24	7
14	8-61-68	48	23	7	250	24	24	250	528
15	8-61-68	51	68	528	723	7	125	60	48

Cela n'est toutefois pas le cas de tous. L'acteur 11 est présent dans tous les TOP 15 générés et donc l'acteur dont la présence est la plus fréquente au sein de ces classements, mais n'apparaît qu'à partir du 4^e rang dans le réseau total en termes d'intermédiarité, au 5^e rang en termes de centralité de degré pour le réseau Facebook, pour le réseau issu des données policières où il est ex-aequo avec l'acteur 47, et pour le réseau total. Similairement, l'acteur 7, présent dans huit des neuf TOP 15, n'apparaît qu'à partir du 4^e rang en termes de centralité d'intermédiarité dans le réseau issu des données policières, puis dans les rangs 9 à 15. Cette analyse mène donc aux mêmes constats que lorsque l'ensemble des individus sont considérés. De même, bien que les rangs diffèrent quelques peu de ceux constatés pour les réseaux comprenant l'ensemble des individus, le classement des acteurs principaux dans le réseau Facebook ou issu des données policières n'engendre pas nécessairement le classement du réseau total pour un indicateur donné. La conclusion serait donc la même, à savoir la complémentarité des deux sources lorsque la position de chaque acteur est analysée séparément.

Cependant, il peut être frappant de constater qu'un même acteur puisse occuper le même rang pour chacun des indicateurs de centralité, y compris pour les acteurs présents dans les deux sources de données. Il pourrait être attendu qu'un acteur occupant un certain rang pour la centralité d'intermédiation, ait une présence plus discrète au niveau de la centralité de degré, certains acteurs occupant une position plus stratégique et d'autres une position plus visible et vulnérable (Morselli, 2009). De plus, pour ces derniers, il semblerait qu'occuper une place importante en termes de centralité dans le réseau issu des données policières, n'empêche pas qu'il en soit de même pour le réseau Facebook, à l'image des acteurs 2 et 5. De tels individus n'adopteraient pas une plus grande prudence en ligne et y mettent potentiellement de l'information les concernant.

Ainsi, l'analyse comparative des différents réseaux amène au constat que l'apport de Facebook semble relatif, comparativement aux données policières. Ces dernières restent une première source fiable et complète, comme cela ressort des différents graphiques. Les résultats sont similaires entre le réseau issu des données policières et celui combinant les deux sources de données, bien qu'il faille garder à l'esprit que les données policières ne concernent que la région de Montréal. Cependant, Facebook peut apporter d'autres types de données, telles des informations personnelles ou des relations ne figurant pas dans les données policières, pouvant être pertinentes dans le cadre du renseignement. Cela pourrait constituer une étape qui interviendrait dans un second temps et de manière ciblée. Afin de confirmer ces observations, il est nécessaire d'analyser à quel point les différents réseaux correspondent les uns aux autres.

c) Correspondance entre les différents réseaux obtenus

Une corrélation QAP a été effectuée entre les réseaux comprenant les acteurs présents dans les deux sources de données (n=72), afin de déterminer leur correspondance et lequel du réseau Facebook ou issu des données policières contribue le plus au réseau total (Tableau VIII).

Tableau VIII. Matrice de corrélation entre les différents réseaux

	Réseau Facebook	Réseau Police	Réseau Total
Réseau Facebook	1,00		
Réseau Police	0,13*	1,00	
Réseau Total	0,58***	0,61***	1,00

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Le réseau Facebook et le réseau issu des données policières sont peu corrélés. La probabilité qu'une relation présente dans l'un apparaisse dans l'autre est relativement faible ($r=0,13$; $p=0,016 < 0,01$). En revanche, la corrélation entre le réseau Facebook et le réseau total est plus élevée ($r=0,58$; $p=0,000 < 0,001$), de même qu'entre le réseau issu des données policières et le réseau total ($r=0,61$; $p=0,000 < 0,001$). Ces dernières sont similaires. Facebook n'apporterait pas plus que les données policières, comme l'analyse comparative des indicateurs de centralité et des acteurs les plus centraux le suggérerait.

Ainsi, Facebook aurait un apport plutôt relatif au réseau total comparativement aux données policières, tant du point de vue de la structure du réseau, que des rôles et positions occupés par les individus. Il peut apporter des nuances quant à la place d'un individu et à sa position centrale au sein du réseau lorsqu'un rang précis ou une position particulière sont envisagés. Les données policières semblent toutefois rester une première source fiable et complète, permettant d'identifier les mêmes individus. Il s'agit désormais de déterminer si les caractéristiques individuelles que Facebook permet de considérer, apportent des renseignements supplémentaires. Ces dernières pourraient influencer l'appartenance aux réseaux ainsi que le rôle occupé.

4.2. Impact des caractéristiques individuelles sur l'appartenance aux réseaux et le rôle occupé

Cette section aborde la question de l'impact des caractéristiques individuelles sur l'appartenance aux différents réseaux, ainsi que sur le rôle et la position occupés.

4.2.1. Influence des caractéristiques individuelles sur l'appartenance aux différents réseaux

Certains individus seraient plus susceptibles d'appartenir au réseau total des individus présents dans les deux sources de données, plutôt qu'exclusivement au réseau Facebook ou à celui des données policières. 72 individus du réseau Facebook (n=2038) sont aussi présents dans celui issu des données policières (n=565) et inversement. Certaines caractéristiques individuelles seraient liées au fait d'apparaître dans les deux réseaux, et expliqueraient 14,4% de la variance au sein du réseau Facebook ($\chi^2 = 75,70$; $p < 0,001$ et $R^2 = 0,144$) et 6.8% de la variance dans les données policières ($\chi^2 = 20,73$; $p < 0,001$ et $R^2 = 0,068$) (Tableau IX).

Tableau IX. Prédicteurs d'appartenance aux deux sources de données

	Facebook (n=2038)		Police (n=565)	
	B	Exp(B)	B	Exp(B)
Sexe	2,79***	16,27***	-0,61	0,55
Résidence à Montréal	0,99	2,69	-	-
Niveau de scolarité				
<i>Secondaire</i>	-	-	-	-
<i>Cégep</i>	2,39	1,27	-	-
<i>Université</i>	0,65	1,91	-	-
<i>Inconnu</i>	0,32	1,38	-	-
Emploi				
<i>Esthétique</i>	1,71	5,52	-	-
<i>Bar</i>	2,01**	7,45**	-	-
<i>Construction</i>	0,85	2,33	-	-
<i>Loisirs</i>	1,76*	5,78*	-	-
<i>Motards</i>	1,36	3,9	-	-
<i>Autres</i>	-	-	-	-
<i>Inconnu</i>	0,94	2,56	-	-
N évènements police				
<i>Présence une fois</i>	-	-	-	-
<i>Présence plusieurs fois</i>	-	-	1,20***	3,31***
Chi²	75,70***		20,73***	
R² (Nagelkerke)	0,144		0,068	

* $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$

Plus précisément, être un homme serait le meilleur prédicteur pour un individu du réseau Facebook d'appartenir aussi au réseau issu des données policières ($\text{Exp(B)}=16,27$, $p<0,001$). Toutefois, le raisonnement inverse ne donne pas de relation significative. Ainsi, un homme présent dans les données policières n'est pas plus susceptible qu'une femme d'apparaître sur Facebook. Cela est cohérent avec la considération selon laquelle Facebook compte des membres des deux sexes, comme cela pouvait aussi ressortir de notre échantillon. En outre, les individus exerçant un emploi dans le secteur des bars ($\text{Exp(B)}=7,45$; $p<0,01$) ou des loisirs, comme celui de coach sportif par exemple ($\text{Exp(B)}=5,78$; $p<0,05$), seraient plus probables d'être aussi présents dans les données policières. Toutefois, cela pourrait être la conséquence d'un ciblage plus important de ces secteurs d'activité par les agences d'application de la loi, en particulier pour les bars.

S'agissant des individus présents dans les données policières ($n=565$), la présence d'un individu dans plus d'un événement ($\text{Exp(B)}=3,31$; $p<0,001$) prédirait l'appartenance aux deux réseaux. Cela peut paraître surprenant. Les individus les plus présents dans les données policières font possiblement l'objet d'une attention plus élevée. Ils auraient pu être absents de Facebook afin d'éviter toute prise de risque quant à la divulgation d'informations supplémentaires aux organisations policières. Bien que certains fassent usage des paramètres de sécurité, ce ne semble pas être la règle pour tous, comme cela ressort des portraits d'échantillons. La majorité des relations étant de nature non criminelle, l'individu perçoit possiblement un danger moindre d'être présent en ligne.

Ainsi, les caractéristiques individuelles présentes dans les données policières auraient une moins grande valeur prédictive de l'appartenance simultanée au deux réseaux que celles de Facebook. Cela pourrait s'expliquer par le fait que la présence sur Facebook représente un phénomène plus large que la présence dans les données policières, et donc plus difficilement prédictible. De plus, peu de caractéristiques étaient disponibles, ce qui est une limite importante. Plusieurs caractéristiques individuelles de Facebook étaient aussi indisponibles pour de nombreux individus. Face à un nombre important de données manquantes, la valeur prédictive des caractéristiques individuelles

de Facebook quant à l'appartenance aux deux réseaux est relative. Elle doit d'autant plus être mise en perspective, qu'elle met en évidence des tendances reconnues au sein des données policières, telles le fait d'être un homme ou d'exercer un métier d'un secteur précis pouvant faire l'objet d'une attention policière particulière.

La question se pose de savoir si ces caractéristiques ont aussi une influence sur le rôle et la position occupés par un individu au sein des réseaux.

4.2.2. Influence des caractéristiques individuelles sur le rôle et la position occupés

Les caractéristiques individuelles d'un acteur au sein du réseau pourraient avoir une influence sur le rôle et la position qu'il y occupe. Plus particulièrement, un réseau consistant dans le partage de ressources (Van der Hulst, 2008), certains individus peuvent en posséder que d'autres n'ont pas, ce qui leur procure une certaine influence. Ces considérations sont particulièrement importantes lors de la sélection des complices pour réaliser une entreprise criminelle, la spécialisation de certains individus dans une tâche spécifique pouvant être mise à profit par le réseau dans la réalisation de ses objectifs (Van der Hulst, 2008 ; Morselli, 2009). Cette position influente est révélée par une forte centralité de vecteur propre (Bonacich, 1972 ; Morselli, 2009). Toutefois, les résultats suivants pourraient être également appliqués à la centralité de degré et d'intermédiarité, les trois indices de centralité étant fortement corrélés les uns aux autres. Cette similarité ressortait notamment lors de l'étude comparative des différents réseaux (cf. graphes). Les caractéristiques individuelles de Facebook prédiraient moins bien une position influente que celles des données policières ($R^2=0,039$ et $F=6,69$; $p<0,001$ pour Facebook contre $R^2=0,092$ et $F=28,07$; $p<0,001$ pour les données policières).

S'agissant du réseau Facebook, le fait de ne pas dévoiler son emploi diminuerait la probabilité de détenir une position influente au sein du réseau (Beta= -0,08 ; p<0,01) (Tableau X). Cela peut être dû à un manque de confiance ou de fiabilité. Les individus tendraient à se tourner vers ceux divulguant de l'information à leur sujet et faire de ces derniers des acteurs influents. En effet, le type d'individu dont il s'agit et le type de ressources potentiellement accessibles sont clairement établis. Cela peut aussi être le fait que certaines personnes utilisent Facebook dans un objectif professionnel, par exemple une esthéticienne qui recruterait des clients via cette plateforme sociale. À l'inverse, être un homme (Beta=0,06 ; p<0,05) ou exercer un métier dans le domaine de l'esthétique (Beta=0,05 ; p<0,05) augmenterait quelque peu cette probabilité.

Tableau X. Prédicteurs d'une position influente au sein des réseaux (vecteur propre)

	Facebook (n=2038)		Police (n=565)	
	b	Beta	b	Beta
Sexe	0,08*	0,06*	0,31	0,05
Résidence à Montréal	0,35***	0,09***	-	-
Niveau de scolarité				
<i>Secondaire</i>	-	-	-	-
<i>Cégep</i>	0,02	0,01	-	-
<i>Université</i>	0,20**	0,07**	-	-
<i>Inconnu</i>	-0,05	-0,04	-	-
Emploi				
<i>Esthétique</i>	0,20*	0,05*	-	-
<i>Bar</i>	0,07	0,02	-	-
<i>Construction</i>	0,08	0,03	-	-
<i>Loisirs</i>	0,09	0,02	-	-
<i>Motards</i>	0,36**	0,07**	-	-
<i>Autres</i>	-	-	-	-
<i>Inconnu</i>	-0,12**	-0,08**	-	-
N évènements police				
<i>Présence une fois</i>	-	-	-	-
<i>Présence plusieurs fois</i>	-	-	0,80***	0,29***
F		6,69***		28,07***
R²		0,039		0,092

*p<0,05 **p<0,01 ***p<0,001

Le fait d'être à l'université plutôt qu'au secondaire (Beta=0,07 ; p<0,01) et d'être qualifié de motards (Beta=0,07 ; p<0,01) seraient les meilleurs prédicteurs d'une position influente au sein du réseau Facebook, outre le fait de résider à Montréal (Beta=0,09 ;

$p < 0,001$). Cela pourrait être le fait que le réseau a été construit à partir d'individus fréquentant probablement le territoire montréalais. Or, la proximité géographique est un élément souvent présent lorsqu'il est question d'homophilie (McPherson, Smith-Lovin et Cook, 2001). En outre, cela rend ces personnes disponibles et proches dans l'espace si leurs ressources étaient nécessitées. S'agissant du niveau scolaire, il pourrait être affirmé qu'un plus grand nombre d'universitaires est présent sur Facebook ou alors qu'un individu à l'université est mieux outillé et plus expérimenté qu'un étudiant du secondaire. Toutefois, certaines caractéristiques ressortant comme influentes peuvent être le fait d'une distinction parmi les acteurs du réseau, les caractéristiques partagées (e.g. emploi fréquent dans le réseau) devenant trop générales pour se démarquer socialement (Lewis, Gonzaleza, et Kaufman, 2012).

Cependant, l'influence due au fait d'être motards est cohérente avec la littérature. Les motards criminalisés sont réputés être un groupe influent et maintiennent une certaine réputation (Barker, 2014 ; Lauchs, Bain et Bell, 2015). Il n'est toutefois pas anodin que de tels individus soient présents sur Internet et occupent une position influente au sein du réseau. Cela va à l'encontre des résultats de précédentes recherches affirmant que des groupes tels que les Hells Angels n'existeraient en ligne qu'en tant que chapitres ou groupes (Décary-Héту et Morselli, 2011 ; Morselli et Décary-Héту, 2013). Ils seraient désormais présents en tant qu'individus sur les médias sociaux. Leur position influente sur Facebook mérite attention. Le bassin large de délinquants potentiels (Bouchard et Westlake, 2016) permettrait à une telle organisation de procéder en ligne de la même manière que dans le monde réel, c'est-à-dire d'être connectée à des individus non criminalisés ou des entreprises légales pour mieux isoler le cœur du réseau (Morselli, 2009). Devant cette prise de risque de s'afficher en ligne, la question se pose de savoir si une position influente se reflète dans le monde réel. Il n'en reste pas moins que cela représente une source d'information pertinente en matière de renseignement policier.

Concernant le réseau issu des données policières, le sexe ne semble pas être un prédicteur d'influence. En revanche, un individu présent dans plusieurs événements serait

plus susceptible d'occuper une position influente au sein du réseau (Beta=0,29 ; $p < 0,001$). Cela pourrait être lié à l'accumulation d'expérience au fil des événements, permettant d'établir un statut de codélinquant fiable. Cela aurait pu toutefois être perçu comme un signe de faiblesse, c'est-à-dire le fait par exemple de s'être fait observer avec d'autres ou interpeler à plusieurs reprises par les agences d'application de la loi.

En somme, les différentes analyses semblent indiquer que la contribution de Facebook aux données policières est relativement minimale. Qu'il s'agisse de la structure globale du réseau, des positions et rôles occupés par les individus le composant, ou qu'il s'agisse des caractéristiques individuelles de ces derniers permettant de prédire leur appartenance et leur place au sein du réseau, les données collectées sur Facebook ne semblent pas avoir un apport significatif. Plus précisément, l'apport serait faible lorsque l'approche est générale. Les données policières constitueraient alors une première source d'information fiable et relativement complète. Disposer d'autres opportunités et d'informations, si elles doivent être étudiées, ne doit pas mener à sous-estimer la valeur et le potentiel des données policières. Ces dernières révèlent des informations importantes sur les interactions d'un type spécifique, à savoir la codélinquance officielle, non couvertes par d'autres sources (Xu et Chen, 2008 ; Bouchard et Konarski, 2014). Cela s'illustre notamment par de plus faibles valeurs prédictives des données Facebook que des données policières au cours des différentes analyses sur les caractéristiques individuelles. Toutefois, Facebook offrirait une information et un point de vue complémentaires aux données policières lorsqu'il s'agit de cibler ou identifier un individu ou un rôle précis. Cela ressort de l'analyse des différents TOP 15 des acteurs principaux en termes de centralité. Facebook serait ainsi une source complémentaire d'information dans un second temps, qui interviendrait de manière spécifique et ciblée. Il permettrait par exemple de collecter des informations personnelles sur un individu, qui ne seraient pas disponibles dans les données policières mais pertinentes en renseignement policier.

Conclusion

Le présent projet visait à exploiter plusieurs sources de données simultanément, ainsi qu'à déterminer l'apport des sources ouvertes en matière d'analyse de réseau, plus particulièrement en renseignement, comparativement aux données officielles. En effet, si les méthodes d'analyse de réseau se sont développées, de précédentes recherches soulignaient l'importance de l'étendre à d'autres types de sources, notamment pour l'étude de réseaux de codélinquance ou de médias sociaux. De plus, le besoin et l'importance d'utiliser simultanément plusieurs sources de données pour pallier les limites des données officielles étaient soulevés. Il s'agissait alors de construire le réseau social de sujets criminalisés, soit dix individus réputés appartenir à un groupe de motards criminalisés, à partir à la fois des événements, criminels ou non, figurant dans les données policières et des liens d'amitié publics de Facebook. L'objectif final étant de déterminer l'apport de chaque source et s'il est nécessaire d'investir beaucoup de ressources sur Facebook, les médias sociaux étant de plus en plus utilisés au quotidien par les agences d'application de la loi dans l'accomplissement de leurs tâches.

La popularité de Facebook et la quantité importante d'informations présentes sur cette plateforme en font une opportunité intéressante, notamment comme outil d'enquêtes ou de renseignement. Il donne accès librement à de nombreuses données relationnelles et personnelles. Toutefois, les différentes analyses effectuées semblent indiquer que sa contribution aux données policières est en réalité relativement minimale. Qu'il s'agisse de la structure globale du réseau, des positions et rôles occupés par les individus le composant, ou des caractéristiques individuelles de ces derniers permettant de prédire leur appartenance et leur place au sein du réseau, les données Facebook ne semblent pas avoir un apport significatif. Plus précisément, l'apport serait faible lorsque l'approche est générale, apporterait du bruit et pourrait induire en erreur. Les données policières constitueraient alors une première source d'information fiable et relativement complète. Leur potentiel ne doit pas être sous-estimé.

Cela ne signifie pas que dans un cas précis Facebook ne puisse pas bonifier le renseignement déjà disponible. Cela ne signifie pas non plus que dans le cadre d'autres types d'utilisation, telle la communication avec le public, ou en réponse à des contextes de crise particuliers, comme lors des émeutes de Vancouver en 2011 ou du marathon de Boston en 2013, une utilisation plus systématique de Facebook ne soit pas pertinente. Facebook offrirait de l'information et un point de vue complémentaires aux données policières, si une approche plus ciblée est adoptée. C'est par exemple le cas lors de l'identification d'un individu ou d'un rôle précis au sein du réseau, tel qu'illustré lors de l'analyse des différents TOP 15 des acteurs principaux en termes de centralité. Facebook serait ainsi une source complémentaire d'information, qui interviendrait de manière spécifique et ciblée dans un second temps. Il peut s'agir par exemple de collecter des informations personnelles sur un individu qui ne seraient pas disponibles dans les données policières mais pertinentes en renseignement policier.

Toutefois, le projet n'est pas exempt de limites. L'accès à des données complètes comporte ses interrogations, aussi bien avec les données officielles qu'avec les données obtenues en sources ouvertes. En effet, le réseau Facebook doit être considéré comme partiel. L'accès à l'information peut être impossible ou restreinte du fait des paramètres de sécurité de l'individu ou du fait de l'inaccessibilité à la liste d'amis des amis. Cela a pour conséquence de limiter la détection d'individus au deuxième degré de relations. L'information peut également être disponible mais non exploitable, en étant falsifiée ou non fiable ou encore du fait de surnoms. Ce n'est pas propre aux données en sources ouvertes, de fausses identités pouvant aussi paraître dans les données officielles. Cependant, le fait de disposer d'une liste partielle d'amis Facebook ne paraît pas empêcher que les individus concernés ressortent parmi les acteurs les plus importants. Ce serait alors un obstacle sur le plan informationnel, ne permettant pas d'obtenir de l'information sur les caractéristiques individuelles. De même, les données policières utilisées ne comprennent que les données de Montréal. D'autres individus que ceux détectés pourraient être présents dans les deux sources de données. Ils seraient en contact avec des individus montréalais mais pas actifs à Montréal. En effet, une partie de l'échantillon réside dans les RMR de Sherbrooke, de Québec ou de Trois-Rivières.

De plus, l'étude comporte des limites quant à sa généralisation. Les données concernent dix individus de départ, dont les noms et informations sur les profils Facebook ont été renseignés par un tiers. Des biais sont donc possibles, d'autant plus qu'il s'agit d'individus réputés reliés aux motards criminalisés. La structure et le fonctionnement des réseaux ainsi construits pourraient être influencés par cette caractéristique, bien que le réseau Facebook permette un bassin plus large d'individus. Par ailleurs, seul un type d'interactions de Facebook, à savoir les listes d'amis, a été pris en compte. D'autres types d'interactions, notamment si cette source ouverte est utilisée de manière ciblée comme cela est suggéré dans le présent rapport, pourraient être prises en considération. En effet, plus que de simplement regarder le contenu des profils et des posts Facebook, les « pages d'amitié », c'est-à-dire des pages rapportant l'historique de la relation entre deux individus, peuvent aussi permettre de connaître le nombre d'évènements auxquels ces deux individus ont participé, le nombre de statuts qu'ils ont en communs, le nombre de photos dans lesquelles ils sont tous deux identifiés, ainsi que leurs intérêts communs via le nombre de pages ou de groupes auxquels ils appartiennent tous les deux. Cela permettrait alors de quantifier une relation. Un nombre commun d'informations indiquerait un lien fort entre ces deux individus. Cela permettrait ultimement de cibler les relations pertinentes, comme cela a pu être fait dans les données policières en ne conservant que les relations positives. Enfin, un seul type de sources ouvertes, plus particulièrement un seul média social, Facebook, a été exploité. Ce dernier paraissait le plus approprié en termes de renseignement et d'analyses de réseau. Outre Twitter, d'autres sources ouvertes pourraient permettre de compléter les données. Par exemple, l'emploi renseigné sur Facebook, notamment lorsque le nom d'une entreprise est mentionné, permettrait de déterminer parmi les informations librement accessibles sur le CIDREQ, les noms et prénoms réels de la personne, d'établir d'autres liens, les noms des propriétaires et gérants apparaissant également, de même que l'adresse du siège social. Ces informations peuvent être intéressantes du point de vue du renseignement, puisqu'il est établi que certains groupes du crime organisé entretiennent des liens avec des entreprises légitimes. D'autres analyses, considérant plus de sources ou, plus précisément, d'autres types d'interactions mériteraient d'être menées.

Toutefois, comparer simultanément deux sources de données, ainsi que la structure d'un même réseau à la fois dans les mondes réel et virtuel, présente un aspect nouveau. Cela permet d'apporter une nuance et une orientation dans les pratiques, basées sur des données probantes. En effet, les données policières ne doivent pas être sous-estimées. Ces dernières présentent une première source fiable et relativement complète d'information, notamment pour les individus présents dans les deux sources de données. Il pourrait alors s'agir de bonifier les données policières, en incluant de l'information supplémentaire. En outre, afin que Facebook apporte une réelle complémentarité, l'utilisation de ce média social doit être repensée et ciblée. Elle doit intervenir dans un second temps, après que toute l'information disponible dans les données policières ait été exploitée, notamment à l'aide des méthodes d'analyse de réseaux. D'autres analyses sont nécessaires pour déterminer si un type d'interaction mériterait plus d'attention qu'un autre et serait susceptible d'apporter plus d'informations.

Bibliographie

- Akhtar, N. (2014). Social network analysis tools. In *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on* (p. 388-392). IEEE.
- Althaus, C., Bridgman, P., & Davis, G. (2007). *The Australian policy handbook*. Crows Nest, New South Wales: Allen & Unwin.
- Altunbas, F. (2013). *Social media in policing: A study of Dallas-Fort Worth area city police departments* (Thèse de doctorat). University of North Texas, Ann Arbor, United States.
Consulté à l'adresse
<http://search.proquest.com/docview/1667406095/abstract/35C53E19287B4C62PQ/1>
- Andresen, M. A., & Felson, M. (2012). Co-offending and the diversification of crime types. *International journal of offender therapy and comparative criminology*, 56(5), 811-829.
- Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook Profiles Reflect Actual Personality, Not Self-Idealization. *Psychological Science*. <https://doi.org/10.1177/0956797609360756>
- Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012). Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference* (p. 33-42). ACM.
- Barker, T. (2014). *Outlaw motorcycle gangs as organized crime groups*. Cham: Springer.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8, 361-362.
- Bhat, S. Y., & Abulaish. (2013). Analysis and Mining of Online Social Networks - Emerging Trends and Challenges. *WIREs Data Mining and Knowledge Discovery*, 3(6), 408-444.
- Block, P., & Grund, T. (2014). Multidimensional homophily in friendship networks. *Network Science*, 2(2), 189-212.
- Boivin, R., & Morselli, C. (2016). *Les réseaux criminels*. Montréal, Québec: Les Presses de l'Université de Montréal.

- Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1), 113-120.
<https://doi.org/10.1080/0022250X.1972.9989806>
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- Bouchard, M., & Konarski, R. (2014). Assessing the Core Membership of a Youth Gang from its Co-offending Network. In *Crime and networks* (p. 81-93). New York: Routledge/Taylor & Francis Group.
- Bouchard, M., & Westlake, B. (2016). La détection de communautés dans les réseaux criminels en ligne et la recherche en codélinquance. In *Les réseaux criminels* (p. 199-220). Montréal, Québec: Les Presses de l'Université de Montréal.
- Bourdieu, P. (1986). The Forms of Capital. In *Handbook of theory and research for the sociology of education* (Richardson J.). New York: Greenwood.
- Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.
<https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Burke, M., Marlow, C., & Lento, T. (2010). Social network activity and social well-being. In *Proceedings of the SIGCHI conference on human factors in computing systems* (p. 1909-1912). ACM.
- Burt, R. S. (1992). *Structural holes : the social structure of competition*. Cambridge, Mass.: Cambridge, Mass. : Harvard University Press.
- Calderoni, F. (2016). L'ARS pour l'identification des rôles de leadership dans les organisations criminelles. In *Les réseaux criminels* (p. 103-128). Montréal, Québec: Les Presses de l'Université de Montréal.
- Carrington, P. J. (2009). Co-offending and the development of the dlinquent career. *Criminology*, 47(4), 1295-1329. <https://doi.org/10.1111/j.1745-9125.2009.00176.x>
- Catanese, S. A., De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2011). Crawling Facebook for Social Network Analysis Purposes. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics* (p. 52:1–52:8). New York, NY, USA: ACM. <https://doi.org/10.1145/1988688.1988749>

- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, Ma.: Belknap Press of Harvard University Press.
- Combe, D., LARGERON, C., EGYED-ZSIGMOND, E., & GÉRY, M. (2010). A comparative study of social network analysis tools. In *Web Intelligence & Virtual Enterprises*.
- Davis, E. F., ALVES, A. A., & SKLANSKY, D. A. (2015). Social Media and Police Leadership: Lessons from Boston. *Journal of Current Issues in Crime, Law & Law Enforcement*, 8(2), 289-304.
- Décary-Héту, D., Dupont, B., & Fortin, F. (2014). Policing the hackers by hacking them: Studying online deviants in IRC chat rooms. In *Networks and network analysis for defence and security* (p. 63-82). Springer.
- Décary-Héту, D., & Morselli, C. (2011). Gang Presence in Social Network Sites. *International Journal of Cyber Criminology*, 5(2), 876–890.
- Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook « Friends: » Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication*, 12(4), 1143-1168. <https://doi.org/10.1111/j.1083-6101.2007.00367.x>
- Facebook. (2016). Company Info | Facebook Newsroom. Consulté 22 février 2016, à l'adresse <https://newsroom.fb.com/company-info/>
- Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, 96(6), 1464-1477.
- Feld, S. L., & Grofman, B. (1977). Variation in class size, the class size paradox, and some consequences for students. *Research in Higher Education*, 6(3), 215-222.
- Fleisher, M. S. (2005). Fieldwork research and social network analysis: Different methods creating complementary perspectives. *Journal of Contemporary Criminal Justice*, 21(2), 120-134.
- Frank, O., & Carrington, P. J. (2007). Estimation of offending and co-offending using available data with model support. *Journal of Mathematical Sociology*, 31(1), 1-46.
- Frank, R., Cheng, C., & Pun, V. (2011). *Sites de médias sociaux: nouveaux forums pour les occasions de crimes, de communications et d'enquêtes*. Sécurité publique Canada. Consulté à l'adresse http://publications.gc.ca/collections/collection_2012/sp-ps/PS14-5-2011-fra.pdf

- Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35-41. <https://doi.org/10.2307/3033543>
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215-239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Gjoka, M., Kurant, M., Butts, C. T., & Markopoulou, A. (2011). Practical Recommendations on Crawling Online Social Networks. *IEEE Journal on Selected Areas in Communications*, 29(9), 1872-1892. <https://doi.org/10.1109/JSAC.2011.1111011>
- Goldsmith, A. J. (2010). Policing's new visibility. *The British Journal of Criminology*, 50(5), 914-934.
- Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360-1380.
- Grund, T. (2014). Why your friends are more important and special than you think. *Sociological Science*, 1, 128-140.
- Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California Riverside.
- Hansen, D. L., Schneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL insights from a connected world*. Amsterdam; Boston: M. Kaufmann. Consulté à l'adresse <http://www.sciencedirect.com/science/book/9780123822291>
- Harper, W. R., & Harris, D. H. (1975). The application of link analysis to police intelligence. *Human Factors*, 17(2), 157-164.
- Huisman, M., & van Duijn, M. A. (2011). A reader's guide to SNA software. *The SAGE handbook of social network analysis*, 578-600.
- IACP. (2015). *Social Media Survey Results [PDF]*. Consulté à l'adresse <http://www.iacpsocialmedia.org/wp-content/uploads/2017/01/FULL-2015-Social-Media-Survey-Results.compressed.pdf>
- Iwanski, N., & Frank, R. (2014). The Evolution of a Drug Co-arrest Network. In *Crime and networks* (p. 52-80). New York: Routledge/Taylor & Francis Group.
- Klerks, P. (2001). The network paradigm applied to criminal organizations. *Connections*, 24(3), 53-65.

- Koschade, S. (2006). A Social Network Analysis of Jemaah Islamiyah: The Applications to Counterterrorism and Intelligence. *Studies in Conflict & Terrorism*, 29(6), 559-575. <https://doi.org/10.1080/10576100600798418>
- Krebs, V. E. (2002). Mapping networks of terrorist cells. *Connections*, 24(3), 43-52.
- Lauchs, M., Bain, A., & Bell, P. (2015). *Outlaw Motorcycle Gangs*. London: Palgrave Macmillan UK. <https://doi.org/10.1057/9781137456298>
- Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1), 18-66.
- Lewis, K., Gonzalez, M., & Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, 109(1), 68-72.
- Lin, N. (2001). *Social Capital: A Theory of Social Structure and Action*. New York: Cambridge University Press.
- Martineau, É., & Lecocq, R. (2016). L'orientation des recherches et les efforts en ARS pour la fonction du renseignement. In *Les réseaux criminels* (p. 61-82). Montréal, Québec: Les Presses de l'Université de Montréal.
- McGloin, J. M., & Kirk, D. S. (2010). An Overview of Social Network Analysis. *Journal of Criminal Justice Education*, 21(2), 169-181. <https://doi.org/10.1080/10511251003693694>
- McGloin, J. M., & Nguyen, H. (2012). It was my idea: Considering the instigation of co-offending. *Criminology*, 50(2), 463-494.
- McGloin, J. M., & Nguyen, H. (2014). The Importance of Studying Co-offending Networks for Criminological Theory and Policy. In *Crime and networks* (p. 13-27). New York: Routledge/Taylor & Francis Group.
- McGloin, J. M., & Piquero, A. R. (2010). On the Relationship between Co-Offending Network Redundancy and Offending Versatility. *Journal of Research in Crime and Delinquency*, 47(1), 63-90. <https://doi.org/10.1177/0022427809348905>
- McIver, A. (2016). Les courtiers dans le réseau criminel des gangs de rue. In *Les réseaux criminels* (p. 153-171). Montréal, Québec: Les Presses de l'Université de Montréal.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1), 415-444.

- Mislove, A., Viswanath, B., Gummadi, K. P., & Druschel, P. (2010). You are who you know: inferring user profiles in online social networks. In *Proceedings of the third ACM international conference on Web search and data mining* (p. 251-260). ACM.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological review*, *100*(4), 674.
- Morselli, C. (2009). *Inside criminal networks*. New York, NY, New York: Springer. Consulté à l'adresse <http://dx.doi.org/10.1007/978-0-387-09526-4>
- Morselli, C. (2014). *Crime and networks*. New York: Routledge/Taylor & Francis Group.
- Morselli, C., & Boivin, R. (2016). *Les réseaux criminels*. Montréal, Québec: Les Presses de l'Université de Montréal.
- Morselli, C., & Décary-Héту, D. (2013). Crime facilitation purposes of social networking sites: A review and analysis of the 'cyberbanging' phenomenon. *Small Wars & Insurgencies*, *24*(1), 152-170.
- Papachristos, A. V. (2009). Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *American Journal of Sociology*, *115*(1), 74-128. <https://doi.org/10.1086/597791>
- Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence. *American Sociological Review*, 3122413486800. <https://doi.org/10.1177/0003122413486800>
- Reiss. (1988). Co-Offending and Criminal Careers. *Crime and Justice*, *10*, 117-170. <https://doi.org/10.1086/449145>
- Rhodes, C. J. (2011). The Use of Open Source Intelligence in the Construction of Covert Social Networks. In U. K. Wiil (Éd.), *Counterterrorism and Open Source Intelligence* (p. 159-170). Springer Vienna. https://doi.org/10.1007/978-3-7091-0388-3_9
- Sarnecki, J. (2001). *Delinquent networks: youth co-offending in Stockholm*. Cambridge, UK; New York: Cambridge University Press.
- Schneider, C. J. (2015). Police image work in an era of social media: YouTube and the 2007 Montebello Summit Protest. *Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in the Age of Facebook, Twitter and YouTube*. New York, NY: Routledge, 227-246.

- Schneider, C. J., & Trottier, D. (2012). The 2011 Vancouver Riot and the Role of Facebook in Crowd-Sourced Policing. *BC Studies*, (175), 57-72-158.
- Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. London ; Thousand Oaks, Calif: SAGE.
- Sparrow, M. K. (1991). The application of network analysis to criminal intelligence: An assessment of the prospects. *Social Networks*, 13(3), 251-274.
[https://doi.org/10.1016/0378-8733\(91\)90008-H](https://doi.org/10.1016/0378-8733(91)90008-H)
- Trottier, D. (2012). Policing Social Media. *Canadian Review of Sociology/Revue Canadienne de Sociologie*, 49(4), 411-425. <https://doi.org/10.1111/j.1755-618X.2012.01302.x>
- Trottier, D. (2015). Coming to terms with social media monitoring: Uptake and early assessment. *Crime, Media, Culture*, 11(3), 317-333.
<https://doi.org/10.1177/1741659015593390>
- Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The Anatomy of the Facebook Social Graph. Consulté à l'adresse <http://arxiv.org/abs/1111.4503>
- Van der Hulst, R. C. (2008). Introduction to Social Network Analysis (SNA) as an investigative tool. *Trends in Organized Crime*, 12(2), 101-121.
<https://doi.org/10.1007/s12117-008-9057-6>
- Van Mastrigt, S. B., & Farrington, D. P. (2009). Co-offending, age, gender and crime type: Implications for criminal justice policy. *British Journal of Criminology*, azp021.
- Waring, E. (2002). Conceptualizing co-offending: A network form of organization. *Crime and Social Organization: Advances in Criminological Theory*. Transaction Publishers, Rutgers, NJ.
- Warr, M. (1996). Organization and instigation in delinquent groups. *Criminology*, 34(1), 11-37.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge; New York: Cambridge University Press.
- Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, 7(3), 203–220.
<https://doi.org/10.1177/1745691612442904>

- Wimmer, A., & Lewis, K. (2010). Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook 1. *American Journal of Sociology*, *116*(2), 583-642.
- Xu, J., & Chen, H. (2005). Criminal Network Analysis and Visualization. *Commun. ACM*, *48*(6), 100–107. <https://doi.org/10.1145/1064830.1064834>
- Xu, J., & Chen, H. (2008). The topology of dark networks. *Communications of the ACM*, *51*(10), 58-65.

ANNEXE 1 : Tableau récapitulatif des variables utilisées au cours des différentes analyses

Variables	Description	Facebook	Police
Sexe	Sexe de l'individu : hommes ; femmes	x	x
Degré	Degré de relations de l'individu dans le réseau Facebook : degré 1 ; degré 2	x	
Résidence	Lieu de résidence indiqué : USA ; Canada ; Québec ; RMR Montréal ; Montréal ; RMR Québec ; RMR Saguenay ; RMR Sherbrooke ; RMR Trois-Rivières ; RMR Ottawa-Gatineau	x	
Distance en véhicule de Montréal	Distance en km séparant le lieu de résidence de Montréal avec un véhicule	x	
Études	Niveau de scolarité de l'établissement fréquenté : secondaire ; école technique ou de métiers, collège communautaire ou cégep ; université ; pas fiable ; inconnu	x	
Emploi	Secteur d'activité professionnelle correspondant à l'emploi exercé : esthétique et soins personnels ; hôtellerie et restauration ; bar ; construction et chantiers ; loisirs et divertissement ; transports ; habillement et mode ; services sociaux et médicaux ; public ; alimentation ; finances, immobilier et assurances ; enfance et enseignement ; autres, motard ; pas fiable ; inconnu	x	
Évènements	Participation à un évènement figurant dans les banques policières (nombre)		x
Présence dans les réseaux	Présence dans un réseau : Facebook ; exclusivement sur Facebook ; Police ; exclusivement réseau police ; dans les deux réseaux	x	x
Indicateurs de centralité	Mesures de la centralité d'un individu dans un réseau : centralité de degré ; centralité d'intermédiarité ; centralité de vecteur propre	x	x
Temps amis Facebook	Temps en mois depuis lequel deux individus sont amis	x	
Nature et fréquence des relations	Nature et fréquence d'une relation entre deux individus : relations positives ; relation positives criminelles ; relations positives non criminelles		x

ANNEXE 2 : Classification des activités professionnelles par secteur d'emploi basée sur la classification des activités économiques du CIDREQ⁵

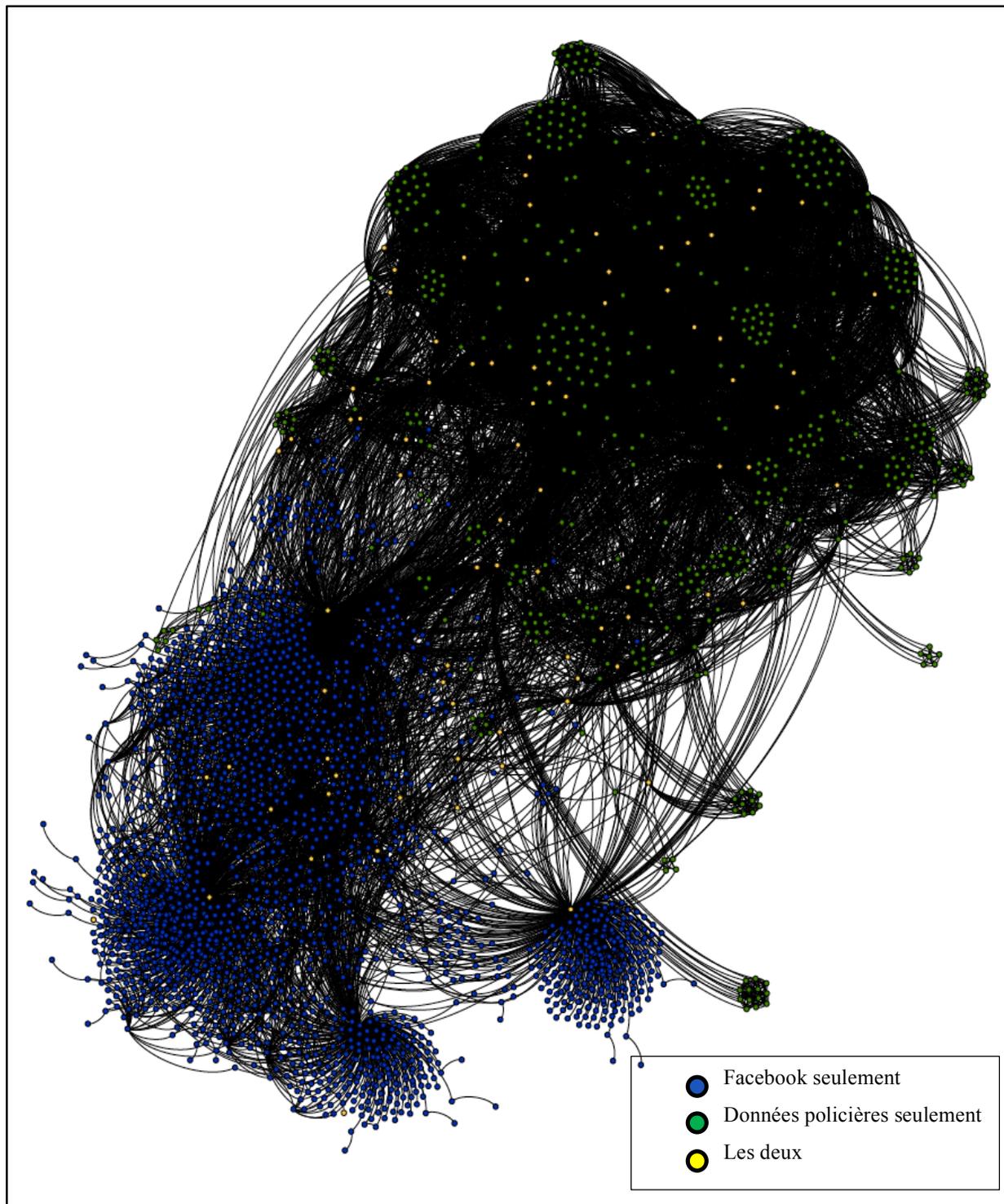
SECTEUR D'EMPLOI	EXEMPLES D'ACTIVITES PROFESSIONNELLES
Esthétique et soins personnels	Salons de coiffure et de beauté, tattoo, bronzage, spa massage
Bar	Tavernes, bar, club, stripclub, club de danseuses
Hôtellerie et restauration	Cafés, grandes chaînes (Tim Hortons, etc.), hôtellerie
Transports	Commerces et concessionnaires automobiles ou motos, pièces de détail, réparation, carrosserie, mécanique, chauffeurs
Loisirs et divertissement	Festival, arts, production télévision, course automobile, chanteuse comédienne, organisation d'évènements, jeux, entraînement (coach sportif, fitness, yoga...)
Construction et chantiers	Machinerie lourde, électricien, démolition, excavation, meubles préconçus, entretien asphalte, arpenteurs et géomètres, rénovation, peinture, usinage, réparation, construction bâtiment, maçonnerie
Habillement et mode	Vente de vêtements, chaussures, friperie, mannequin...
Services sociaux et médicaux	Santé, centre de soins hébergement personnes âgées, hôpital, infirmière, massothérapeute, entretien ménager, intégration personnes handicapées, préposé aux bénéficiaires
Enfance et enseignement	Garderie, magasin de jeux pour jeunes, école, commission scolaire, enseignement/université
Public	Villes ou municipalités, poste canada, gouvernement, grande bibliothèque, conseiller ville
Alimentation	SAQ, marché, épicerie/dépanneur, boulangerie, bar laitier distribution...
Finances, immobilier et assurances	Banque, courtier hypothécaire, comptabilité, conseiller en gestion financière, assurances, conseillers en gestion d'entreprise, gestion immobilier, société portefeuille (holding), courtage immobilier
Motard	Référence au fait d'être un motard, par exemple motard un jour motard toujours, support 81, lucky riders...
Autres	Technicien audio, protection sécurité, bijouterie, service client vidéotron, gérant en droit, mines, cinéma télévision, avocats, service informatique, agriculture, antiquité, maître-chien, imprimerie, érotique (pornographie, sexshop...)
Pas fiable	Peu plausible, introuvable, entreprise n'existant plus, travailleurs autonomes sans précision (« Boss à Moi-même », « à son compte », « rien à dire à c'est pas de tes affaires »...)

⁵

CIDREQ (classification des activités économiques) :

http://www.registreentreprises.gouv.qc.ca/fr/demarrer/immatriculer/classification_activite_economique.aspx

ANNEXE 3 : Représentation graphique du réseau total comprenant l'ensemble des individus (logiciel Gephi)



ANNEXE 4 : Statistiques descriptives des différents indicateurs de centralité selon les différents réseaux comprenant l'ensemble des individus

Tableau 1. Statistiques descriptives des indicateurs de centralité des réseaux comprenant l'ensemble des acteurs

	N	Moyenne	Médiane	Écart-type	Minimum	Maximum
Réseau Facebook	2038					
<i>Degré</i>		3,74	2	24,57	1	731
<i>Intermédiarité</i>		0,0009	0	0,0164	0	0,4910
<i>Vecteur propre</i>		0,0283	0,0259	0,0349	0,0007	1
Réseau Police	565					
<i>Degré</i>		41,15	24	42,29	1	385
<i>Intermédiarité</i>		0,0020	0	0,0128	0	0,2304
<i>Vecteur propre</i>		0,1776	0,0802	0,2006	0,0030	1
Réseau total	2531					
<i>Degré</i>		12,16	2	35,67	1	770
<i>Intermédiarité</i>		0,0007	0	0,0126	0	0,4015
<i>Vecteur propre</i>		0,0467	0,0091	0,1154	0,0002	1

Tableau 2. Statistiques descriptives des indicateurs de centralité des acteurs présents dans le TOP 15 des réseaux comprenant l'ensemble des acteurs

	N	Moyenne	Médiane	Écart-type	Minimum	Maximum
Réseau Facebook	15					
<i>Degré</i>		193,33	79	218,98	40	731
<i>Intermédiarité</i>		0,1081	0,0120	0,1629	0,0020	0,4910
<i>Vecteur propre</i>		0,2992	0,1882	0,2646	0,1032	1
Réseau Police	15					
<i>Degré</i>		205,53	177	66,57	153	385
<i>Intermédiarité</i>		0,1008	0,0237	0,1331	0,0084	0,4015
<i>Vecteur propre</i>		0,8068	0,7693	0,1006	0,6883	1
Réseau total	15					
<i>Degré</i>		327,73	235	191,64	177	770
<i>Intermédiarité</i>		0,1008	0,0237	0,1331	0,0084	0,4015
<i>Vecteur propre</i>		0,8030	0,7615	0,0952	0,7001	1

ANNEXE 5 : Statistiques descriptives des différents indicateurs de centralité selon les différents réseaux comprenant uniquement les individus présents dans les deux sources de données

Tableau 1. Statistiques descriptives des indicateurs de centralité des réseaux comprenant les acteurs présents dans les deux sources de données

	N	Moyenne	Médiane	Écart-type	Minimum	Maximum
Réseau Facebook	72					
<i>Degré</i>		3,56	2	6,44	0	42
<i>Intermédiarité</i>		0,0128	0	0,0560	0	0,4131
<i>Vecteur propre</i>		0,1692	0,1041	0,1759	0	1,0000
Réseau Police	72					
<i>Degré</i>		13,28	9	11,40	1	53
<i>Intermédiarité</i>		0,0142	0	0,0432	0	0,3078
<i>Vecteur propre</i>		0,3545	0,2136	0,2959	0,0324	1
Réseau total	72					
<i>Degré</i>		15,33	11,5	12,28	2	57
<i>Intermédiarité</i>		0,0119	0,0007	0,0348	0	0,2185
<i>Vecteur propre</i>		0,3753	0,2949	0,2654	0,0618	1

Tableau 2. Statistiques descriptives des indicateurs de centralité des acteurs dans le TOP 15 des réseaux comprenant les acteurs présents dans les deux sources de données

	N	Moyenne	Médiane	Écart-type	Minimum	Maximum
Réseau Facebook	15					
<i>Degré</i>		11,19	6,5	10,63	4	42
<i>Intermédiarité</i>		0,0611	0,0114	0,1128	0,0005	0,4131
<i>Vecteur propre</i>		0,4288	0,3483	0,2135	0,2475	1
Réseau Police	15					
<i>Degré</i>		31,87	29	7,63	24	53
<i>Intermédiarité</i>		0,0638	0,0298	0,0781	0,0152	0,3078
<i>Vecteur propre</i>		0,8227	0,8314	0,0871	0,6824	1
Réseau total	15					
<i>Degré</i>		35,07	31	9,72	26	57
<i>Intermédiarité</i>		0,0521	0,0167	0,0628	0,0099	0,2185
<i>Vecteur propre</i>		0,7971	0,7598	0,1043	0,6854	1