

Proposition de communication pour les 2e Rencontres Francophones Transport Mobilité (RFTM) - Montréal, 11-13 juin 2019

Titre :

Apport de données de geotracking dans l'analyse des mobilités quotidienne en Ile-de-France.
Pistes de recherches à partir du projet FUI Geolytics

Auteur(s) :

Julie CHRETIEN, Docteure., Ecole Nationale des Ponts et Chaussées

Florent LE NECHET, Maître de Conférences, Université Paris-Est Marne-la-Vallée,
florent.lenechet@u-pem.fr

Laurent PROULHAC, Ingénieur d'études, IFSTTAR

Mots-clés :

Données de téléphonie mobile, mobilité quotidienne, données longitudinales, comparaison de sources

Résumé :

La communication proposée s'inscrit dans le champ du traitement de données de mobilité dites « nouvelles », et de leur croisement avec des données classiques (enquête ménage déplacement). Il s'agit d'une recherche exploratoire visant à évaluer le potentiel de données issues d'une application de geotracking pour mesurer des aspects de la mobilité urbaine ne pouvant être captés par les données classiques. En s'appuyant sur une revue de bibliographie extensive (Chrétien et al., 2018), nous avons souhaité explorer en particulier l'apport de ce type de données pour la connaissance des mobilités quotidiennes vues dans leur longitudinalité (Järv et al., 2014). La proposition de communication combine donc des aspects techniques (retraitement des données, inférence du motif des déplacements, des lieux de résidence et d'emploi) et des aspects applicatifs avec trois exemples.

Les données de téléphone mobile sont collectées de différentes manières, et sont générées par des « événements » contenant au moins un horodatage et une information de localisation. Les jeux de données obtenus diffèrent par la technologie de géolocalisation utilisée, la fréquence des événements et le fait que les événements soient générés automatiquement via le fonctionnement normal du périphérique (génération « passive ») ou impliquent une intervention active de l'utilisateur. Les données mobilisées dans cette recherche proviennent d'un programme « tracker » installé sur une application grand public en région parisienne, développé par l'entreprise IT4PME, partenaire du projet au sein du programme FUI Géolytics.

Les données de géolocalisation sont générées par le smartphone sur lequel l'application est installée, sur la base des signaux disponibles localement. Elle repose sur la meilleure information dont le smartphone dispose (en ordre de qualité décroissant : Wi-Fi, GPS, antenne...). Les données utilisées portent sur l'Ile-de-France (région de Paris), environ 3000

téléphones pour lesquels 1 million de points au total ont été recueillis sur une période d'un mois (Janvier-Février 2018). Ces points sont déjà le résultat d'un prétraitement effectué par IT4PME : des points anormaux ont été supprimés et chaque mode de transport a été attribué à chaque point, selon des algorithmes propriétaires qui sont de notre point de vue une « boîte noire ». Un travail important de traitement de données a été nécessaire pour pallier la grande hétérogénéité de disponibilité des points au fil de la journée.

Un premier travail consistait à reconstruire les déplacements à partir de la succession de points fournie. En réalité, la startup IT4PME fournissait déjà des déplacements reconstruits mais en comparant avec les données de l'enquête globale de transport, les bons ordres de grandeur ne se retrouvaient pas en termes de nombre de déplacement par jour par personne. Nous avons donc retravaillé les données en créant un index de déplacement, réunissant les points sur des critères de durée et de distance entre deux points. Enfin, nous avons reconstruit les programmes d'activités en inférant à la fois les lieux de résidence, de travail et les motifs de déplacement (retour domicile, vers travail ou « autre »). Pour ce faire, nous avons décidé d'utiliser une approche basée sur les délimitations des municipalités françaises, et à partir des heures et des jours de présence dans chaque municipalité pour en déduire la municipalité de résidence et de travail. L'algorithme retenu, après plusieurs essais, est inspiré de (Kung et al. 2014, Alexander et al. 2015, Çolak et al. 2015, Picornell et al. 2015, Toole et al. 2015, Phithakkitnukoon et al. 2017). Il a été calibré sur l'enquête nationale déplacement du Royaume-Uni (UKNTS) qui décrit plusieurs milliers de voyages d'individus sur une semaine.

La représentativité de la base de données (qui n'est pas un échantillon puisque l'application étant grand public, nous n'avons pas de contrôle sur l'installation) a été testée par comparaison avec l'Enquête Globale de Transport (2010) en Ile-de-France. Nous observons une différence relative de moins de 4 pts dans les trois couronnes de la métropole parisienne lorsqu'on filtre dans l'EGT les actifs ayant un travail dans une commune différente de leur commune de résidence. Au niveau des flux regroupés par Origine-Destination, nous observons une surestimation des flux radiaux vers Paris et une sous-estimation des flux internes aux couronnes ; les temps et distances moyens de déplacement sont correctement estimés. Pour le mode de transport, l'estimation des déplacements en train et en voiture est correcte, sauf pour certains origines-destinations, et semble très peu fiable pour les trajets en métro et à pieds.

La recherche a pu déboucher sur l'exploration de trois aspects où nous avons pu montrer l'apport de ce type de données pour la compréhension des mobilités vues dans leur aspect longitudinal : la mobilité de loisir, le soir (Donnat, 2009) ; la mobilité le week-end, dans le contexte des études sur l'« effet barbecue » (Munafoz, 2017) ; la variabilité des trajets domicile-travail (heure de départ, mode, trajet) (Shen, 2012). Pour les trois applications proposées, les résultats s'appuient sur des observations classiques (mobilité de loisir par jour de semaine dans l'EGT), Enquête Nationale de Transport (2008) pour les mobilités le week-end, origine-destination Versailles-Paris dans l'EGT (OD choisie car il s'agit des données les plus nombreuses dans la base Géolytics).

Le potentiel de ce type de données et la possible reproduction de certains chiffres clés de mobilité justifie pour nous l'intérêt de la démarche exploratoire adoptée : les algorithmes développés dans le cadre de cette recherche sont génériques et peuvent être utilisés de la même manière avec une base de données beaucoup plus importante, ce qui ne dépend que de l'accord des acteurs développant les applications grand public. Les questions éthiques liées à l'utilisation de ces données fines au niveau individuel devraient néanmoins être abordée dans un cadre bien plus systématique que pour cette étude qui ne représente que peu d'individus sur une durée relativement courte (un mois).

Bibliographie

- Alexander L., Jiang S., Murga M., and González M.-C. (2015). "Origin–destination Trips by Purpose and Time of Day Inferred from Mobile Phone Data." *Transportation Research Part C: Emerging Technologies* 58 (September):240–50. <https://doi.org/10.1016/j.trc.2015.02.018>.
- Chrétien, J., Le Néchet, F., Leurent F., Yin B. (2018). Using mobile phone to observe and understand mobility behaviour, territories and transport usage. In, "Urban mobility and the smartphone, transportation, travel behaviour and public policy" by Aguilera, A. and Boutueil, V. (dir), Elsevier, 224 p.
- Çolak, S., Alexander, L. P., Alvim, B. G., Mehndiratta, S. R., González, M. C. (2015). Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. *Transportation Research Record: Journal of the Transportation Research Board*, (2526), 126-135.
- Donnat, Olivier. 2009. *Les pratiques culturelles des Français à l'ère numérique : enquête 2008*. Paris: Découverte ; Ministère de la culture et de la communication.
- Järv, Olle, Rein Ahas, et Frank Witlox. 2014. « Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records ». *Transportation Research Part C: Emerging Technologies* 38 (janvier): 122-35. <https://doi.org/10.1016/j.trc.2013.11.003>.
- Kung, K. S., Greco, K., Sobolevsky, S., & Ratti, C. (2014). Exploring universal patterns in human home-work commuting from mobile phone data. *PloS one*, 9(6), e96180.
- Munafò, Sébastien. 2017. « Forme urbaine et mobilités de loisirs : l'« effet barbecue » sur le grill ». *Cybergeographie : European Journal of Geography*, octobre. <https://doi.org/10.4000/cybergeographie.28634>.
- Picornell, Miguel, Tomás Ruiz, Maxime Lenormand, José J. Ramasco, Thibaut Dubernet, and Enrique Frías-Martínez. 2015. "Exploring the Potential of Phone Call Data to Characterize the Relationship between Social Network and Travel Behavior." *Transportation* 42 (4):647–68. <https://doi.org/10.1007/s11116-015-9594-1>.
- Phithakkitnukoon, S., Sukhvibul, T., Demissie, M., Smoreda, Z., Natwichai, J., & Bento, C. (2017). Inferring social influence in transport mode choice using mobile phone data. *EPJ Data Science*, 6(1), 11
- Shen, Y., & Chai, Y. W. (2012). Study on commuting flexibility of residents based on GPS data: A case study of suburban mega-communities in Beijing. *Acta Geographica Sinica*, 67(6), 733-744.
- Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58, 162-177.
- Wang, Z., He, S. Y., & Leung, Y. (2018). Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society*, 11, 141-155.