



*4^e rencontres francophones transport mobilité (RFTM)
Esch-sur-Alzette, Luxembourg
8-9-10 juin 2022*

Proposition de session thématique

Titre de la session :

Données massives sur la mobilité : validation et valorisation des données

Identification du ou des organisateurs-trices : nom, position professionnelle, affiliation, email, pages web :

Patrick BONNEL, Responsable du Département Transports ENTPE, chercheur au LAET, patrick.bonnel@entpe.fr

Caroline BAYART, Maitre de conférences, Laboratoire de Sciences Actuarielle et Financière, Université de Lyon, Université Lyon 1, caroline.bayart@univ-lyon1.fr

Léa FABRE, doctorante, école doctorale sciences économiques et gestion, ENTPE, Lyon1, explain

Benjamin COTTREAU, doctorant, école doctorale sciences économiques et gestion, ENTPE, Keolis Lyon

Numéro spécial de revue projeté (O/N) ; si oui, nom de la revue

possible selon papiers proposés

Texte d'appel à communications (environ 300 mots)



De plus en plus de sources de données massives sont potentiellement disponibles pour appréhender les comportements de mobilité. Dans cette session, nous souhaitons nous focaliser sur des données collectées de manière passive. Cela signifie que les données sont collectées sans solliciter les individus concernés, à la différence des données d'enquêtes qui requièrent une action plus ou moins importante des répondants. Nous excluons ainsi les enquêtes traditionnellement menées à l'aide d'un questionnaire administré par mail, téléphone et/ou en face-à-face, mais aussi les collectes via GPS ou applications installées de manière active par les enquêtés sur leur téléphone. En revanche, sont concernées les données issues d'applications mobiles, dès lors que les individus ne sont pas sélectionnés pour participer à leur collecte, les données de téléphonie mobile, de billettique, mais également les données collectées via des capteurs WiFi ou bluetooth....

Les données collectées de manière passive ne sont généralement pas issues d'un plan d'échantillonnage permettant de contrôler l'échantillon et ne contiennent aucune information socio-démographique, pour des raisons de protection de la vie privée. Malgré leur volumétrie importante, elles sont donc rarement représentatives. Par ailleurs, elles contiennent le plus souvent des informations très partielles sur la mobilité, voire n'en contiennent pas. Leur validation reste donc problématique, dès lors que l'on souhaite les utiliser pour représenter la mobilité de l'ensemble de la population d'un territoire.

Toutefois, malgré ces nombreux biais, les données collectées de manière passive sont potentiellement très riches du fait de leur volumétrie, mais aussi de leur production en continue. Elles possèdent également des coordonnées spatiales (GPS, couverture téléphonique, entrée dans le réseau de transport collectif...) et temporelles permettant de proposer des méthodes pour les transformer en données de mobilité.

La session se focalise ainsi sur l'ensemble des travaux de recherche conduits dans une perspective de « validation » de ces données, notamment à travers la confrontation à d'autres sources de données considérées comme « ground truth », qu'il s'agisse d'enquêtes de terrain ou de données billettiques. Dans cette session, nous proposons d'aborder également les questions liées au traitement et à la valorisation de ces données. Les travaux recherchés pourront toucher les questions relatives à la fusion de données provenant de différentes sources.

5 mots clés

données massives, billettique, GPS, données de téléphonie, bluetooth, WiFi, systèmes de localisation et de comptage, validation de données, fusion de données



Document à retourner à : rftm2022@liser.lu