



## Avis de Soutenance

Monsieur Etienne THUILLIER

Présentera ses travaux en soutenance

Soutenance prévue le **lundi 11 décembre 2017** à 14h00

Lieu : IMI 62 Boulevard Sébastopol 75003 PARIS  
salle à définir

Titre des travaux : Fusion de données hétérogènes pour l'extraction d'informations de mobilité:  
un problème multi-source, multi-échelle, et multi-modal

Ecole doctorale : SPIM - Sciences Physiques pour l'Ingénieur et Microtechniques

Section CNU : 27

Unité de recherche : Optimisation et réseaux

Directeur de thèse : Alexandre CAMINADA

Codirecteur de thèse : Sid LAMROUS  HDR  NON HDR

Soutenance :  Publique  A huis clos

Membres du jury :

<u>Nom</u>	<u>Qualité</u>	<u>Etablissement</u>	<u>Rôle</u>
M. Alexandre CAMINADA	Professeur des Universités	Polytech Nice Sophia	Directeur de these
M. Sid LAMROUS	Maître de Conférences	Université de Technologie de Belfort-Montbéliard	CoDirecteur de these
M. Laurent MOALIC	Maître de Conférences	Université Haute-Alsace	CoDirecteur de these
M. Patrick SIARRY	Professeur des Universités	Université Paris-Est Créteil Val-de-Marne	Examineur
M. Vincent AGUILERA	Adjoint au chef d'arrondissement de gestion et d'exploitation Est, DRIEA/DIRIF	DRIEA - DIRIF	Rapporteur
Mme Latifa OUKHELLOU	Directeur de Recherche	IFSTTAR - GRETTIA	Rapporteur

## Résumé de la thèse (en français) :

Aujourd'hui c'est un fait, nous vivons dans un monde où les enjeux écologiques, économiques et sociétaux sont de plus en plus pressants. Au croisement des différentes lignes directrices envisagées pour répondre à ces problèmes, une vision plus précise de la mobilité humaine est un axe central et majeur, qui a des répercussions sur tous les domaines associés tels que le transport, les sciences sociales, l'urbanisme, les politiques d'aménagement, l'écologie, etc. C'est par ailleurs dans un contexte de contraintes budgétaires fortes que les principaux acteurs de la mobilité sur les territoires cherchent à rationaliser les services de transport, et les déplacements des individus. La mobilité humaine est donc un enjeu stratégique aussi bien pour les collectivités locales que pour les usagers, qu'il faut savoir observer, comprendre, et anticiper. Cette étude de la mobilité passe avant tout par une observation précise des déplacements des usagers sur les territoires. Aujourd'hui les acteurs de la mobilité se tournent principalement vers l'utilisation massive des données utilisateurs. L'utilisation simultanée de données multi-sources, multi-modales, et multi-échelles permet d'entrevoir de nombreuses possibilités, mais cette dernière présente des défis technologiques et scientifiques majeurs. Les modèles de mobilité présentés dans la littérature sont ainsi trop souvent axés sur des zones d'expérimentation limitées, en utilisant des données calibrées, etc. et leur application dans des contextes réels, et à plus large échelle est donc discutable. Nous identifions ainsi deux problématiques majeures qui permettent de répondre à ce besoin d'une meilleure connaissance de la mobilité humaine, mais également à une meilleure application de cette connaissance. La première problématique concerne l'extraction d'informations de mobilité à partir de la fusion de données hétérogènes. La seconde problématique concerne la pertinence de cette fusion dans un contexte réel, et à plus large échelle. Nous apportons différents éléments de réponses à ces problématiques dans cette thèse. Tout d'abord en présentant deux modèles de fusion de données, qui permettent une extraction d'informations pertinentes. Puis, en analysant l'application de ces deux modèles au sein du projet ANR Norm-Atis. Dans cette thèse, nous suivons finalement le développement de toute une chaîne de processus. En commençant par une étude de la mobilité humaine, puis des modèles de mobilité, nous présentons deux modèles de fusion de données, et nous analysons leur pertinence dans un cas concret. Le premier modèle que nous proposons permet d'extraire 12 comportements types de mobilité. Il est basé sur un apprentissage non-supervisé de données issues de la téléphonie mobile. Nous validons nos résultats en utilisant des données officielles de l'INSEE, et nous déduisons de nos résultats, des comportements dynamiques qui ne peuvent pas être observés par les données de mobilité traditionnelles. Ce qui est une forte valeur-ajoutée de notre modèle. Le second modèle que nous proposons permet une désagrégation des flux de mobilité en six motifs de mobilité. Il se base sur un apprentissage supervisé des données issues d'enquêtes de déplacements ainsi que des données statiques de description du sursol. Ce modèle est appliqué par la suite aux données agrégés au sein du projet Norm-Atis. Les temps de calculs sont suffisamment performants pour permettre une application de ce modèle dans un contexte temps-réel.

## Abstract (in English)

Today it is a fact that we live in a world where ecological, economic and societal issues are increasingly pressing. At the crossroads of the various guidelines envisaged to address these problems, a more accurate vision of human mobility is a central and major axis, which has repercussions on all related fields such as transport, social sciences, urban planning,

management policies, ecology, etc. It is also in the context of strong budgetary constraints that the main actors of mobility on the territories seek to rationalize the transport services and the movements of individuals. Human mobility is therefore a strategic challenge both for local communities and for users, which must be observed, understood and anticipated. This study of mobility is based above all on a precise observation of the movements of users on the territories. Nowadays mobility operators are mainly focusing on the massive use of user data. The simultaneous use of multi-source, multi-modal, and multi-scale data opens many possibilities, but the latter presents major technological and scientific challenges. The mobility models presented in the literature are too often focused on limited experimental areas, using calibrated data, etc., and their application in real contexts and on a larger scale is therefore questionable. We thus identify two major issues that enable us to meet this need for a better knowledge of human mobility, but also to a better application of this knowledge. The first issue concerns the extraction of mobility information from heterogeneous data fusion. The second problem concerns the relevance of this fusion in a real context, and on a larger scale. These issues are addressed in this dissertation: the first, through two data fusion models that allow the extraction of mobility information, the second through the application of these fusion models within the ANR Norm-Atis project. In this thesis, we finally follow the development of a whole chain of processes. Starting with a study of human mobility, and then mobility models, we present two data fusion models, and we analyze their relevance in a concrete case. The first model we propose allows to extract 12 types of mobility behaviors. It is based on an unsupervised learning of mobile phone data. We validate our results using official data from the INSEE, and we infer from our results, dynamic behaviors that can not be observed through traditional mobility data. This is a strong added-value of our model. The second model operates a mobility flows decomposition into six mobility purposes. It is based on a supervised learning of mobility surveys data and static data from the land use. This model is then applied to the aggregated data within the Norm-Atis project. The computing times are sufficiently powerful to allow an application of this model in a real-time context.