

# Floating Car Data Mining to Feature out Mobility Patterns: Individual-Centered and Place-Based Analyses

Danyang SUN

*Thesis directed by Fabien LEURENT and co-supervised by Xiaoyan XIE*

## ABSTRACT

The presence and movement of human beings in space and time constitute their mobility: it is a physical phenomenon and also a socioeconomic phenomenon, since people choose their locations and their trips between them to satisfy their needs and desires. The scientific knowledge of mobility as a physical phenomenon involves causalities and patterns. Owing to the recent surge in sensing technologies, trajectory data are now massively available, thereby empowering the observation of human mobility. Among various sources, Floating Car Data (FCD) pertain to vehicle-based mobility and yield discretized trajectories of positions in space and time for the “vehicle” entity. Most of the academic literature has concentrated on methods for characterizing traffic conditions from an engineering aspect. Only a few studies have shifted the focus of FCD data mining towards semantic-oriented excavation by exploring behavioral representations.

This thesis aims to explore and analyze mobility patterns by leveraging FCD to contribute a better understanding of vehicle-based movements. More specifically, mobility patterns are studied at two levels: the individual level of human behaviors from the trajectory “authors”, and the more global level of capturing spatial relations and structure on the basis of aggregated mobility features in space. Furthermore, another fold of the thesis objective is to build up methodological approaches for trajectory mining, contributing to broadening the way of using trajectory data in mobility analytics.

The first part of the thesis pertains to individual mobility with three research questions addressed. First, **Chapter 2** discovers vehicle usage patterns based on their daily mobility-making and constitutes a typology for the ways of vehicle usage. Then, **Chapter 3** investigates significant places to mobility makers, by identifying the “anchoring” geolocations and further extrapolating their meaningful representations such as homes, workplaces, and other secondary places. Next, **Chapter 4** proposes a novel approach for travel time estimation by building a stochastic model to exploit FCD materials.

The second part of the thesis pertains to places and spatial relations with another three research questions addressed. First, **Chapter 5** aims to reveal the functional occupation of urban areas by looking at related vehicle movements and further characterize the spatial land use divisions. Second, **Chapter 6** studies the spatial organization of a territory, with a particular emphasis on the jobs-housing spatial relations. It establishes a data-driven method to recognize employment core areas by density and identify corresponding residential catchment areas by core-periphery patterns. Third, **Chapter 7** investigates the spatial interaction between places. It deals with the estimation of the Origin-Destination matrix flows based on two kinds of data: vehicle trajectory data and local traffic counts, along with a Bayesian assignment framework to account for the heterogeneous sampling rate issues of such data.

Overall, this thesis expands the “mobility analytics” of “mobility patterns” from a data mining standing point. It contributes to overcoming traditional limitations on extensive mobility analysis in terms of inter-day variations and large-scale observations by employing massive digital trajectories and artificial intelligence. Through various applications, this thesis shows the feasibility of mining semantic context behind individual mobility at a micro-level and the possibility of capturing grouped phenomena reflected

in geographical spaces at a macro-level. However, this thesis pays particular attention to vehicle-related mobility based on FCD. Future work can bring with other modes of transportation to have a more complete investigation of the mobility system.

**Keywords:** Mobility pattern; Floating Car Data; Data mining; Machine Learning; Individual mobility behavior; Mobility-based spatial structure

---

## **Formes spatio-temporelles de la mobilité individuelle et configurations territoriales : caractérisation statistique à partir de données Floating Car Data**

### **RESUME**

La présence et le mouvement des êtres humains dans l'espace et le temps constituent leur mobilité : c'est un phénomène physique et aussi un phénomène socio-économique, puisque les gens choisissent leurs emplacements et leurs déplacements entre eux pour satisfaire leurs besoins et leurs envies. La connaissance scientifique de la mobilité en tant que phénomène physique implique des causalités et des formes. En raison de l'essor récent des technologies de détection, les données de trajectoire sont désormais massivement disponibles, permettant ainsi l'observation de la mobilité humaine, avec une large couverture spatiale et temporelle et éventuellement des taux de pénétration significatifs.

Cette thèse vise à explorer et analyser les formes (modèles, motifs) de mobilité en tirant parti de la Floating Car Data (FCD) pour contribuer à une meilleure compréhension des mouvements basés sur les véhicules. Plus précisément, les formes de mobilité sont étudiés à deux niveaux : le niveau individuel des comportements humains à partir de la trajectoire « auteurs », et le niveau plus global de capture des relations et de la structure spatiale à partir des caractéristiques de mobilité agrégées dans l'espace. Par ailleurs, un autre volet de l'objectif de la thèse est de construire des approches méthodologiques pour l'exploration de trajectoires, contribuant à élargir la manière d'utiliser les données de trajectoires dans l'analyse de la mobilité.

La première partie de la thèse porte sur la mobilité individuelle. Tout d'abord, le chapitre 2 découvre les modes d'utilisation des véhicules en fonction de leur mobilité quotidienne et constitue une typologie des modes d'utilisation des véhicules. Ensuite, le chapitre 3 étudie les lieux significatifs pour les décideurs de la mobilité, en identifiant les géolocalisations « d'ancrage » et en extrapolant davantage leurs représentations significatives telles que les maisons, les lieux de travail et d'autres lieux secondaires. Ensuite, le chapitre 4 propose une nouvelle approche pour l'estimation du temps de trajet en construisant un modèle stochastique pour exploiter les matériaux FCD.

La deuxième partie de la thèse porte sur les lieux et les relations spatiales. Premièrement, le chapitre 5 vise à révéler l'occupation fonctionnelle des zones urbaines en examinant les mouvements de véhicules associés. Deuxièmement, le chapitre 6 étudie l'organisation spatiale d'un territoire, avec un accent particulier sur les relations spatiales emploi-logement. Il établit une méthode basée sur les données pour reconnaître les zones d'emploi principales par densité et identifier les zones de chalandise résidentielles correspondantes par les modèles noyau-périphérie. Troisièmement, le chapitre 7 étudie l'interaction

spatiale entre les lieux. Il traite de l'estimation des flux matriciels Origine-Destination sur la base de deux types de données : les données de trajectoire des véhicules et les comptages de trafic local.

Dans l'ensemble, cette thèse étend les « analyses de la mobilité » des « formes de mobilité » du point de vue de l'exploration de données. Il contribue à surmonter les limitations traditionnelles de l'analyse approfondie de la mobilité en termes de variations inter-journalières et d'observations à grande échelle en utilisant des trajectoires numériques massives et l'intelligence artificielle. À travers diverses applications, cette thèse montre la faisabilité de l'exploration du contexte sémantique derrière la mobilité individuelle à un niveau micro et la possibilité de capturer des phénomènes groupés reflétés dans des espaces géographiques à un niveau macro. Cependant, cette thèse porte une attention particulière à la mobilité liée au véhicule basée sur la FCD. Les travaux futurs peuvent apporter avec d'autres modes de transport pour avoir une enquête plus complète du système de mobilité.

**Mots-clés:** Formes de mobilité; Données de voiture flottantes (Floating Car Data); Exploration de données; Apprentissage automatique; Comportement de mobilité individuelle; Structure spatiale basée sur la mobilité