

# Understanding Social and Community Dynamics from Large-Scale Heterogeneous Urban Open Data and Their Application in Network Optimization

## Mots clés :

- **Directeur de thèse** : Thi-Mai-Trang Nguyen
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire d'informatique de Paris 6
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

## Résumé du projet de recherche (Langue 1)

The rapid progress of urbanization has transformed the urban landscape in several cities. Understanding the social and community dynamics in these cities has become a key challenge in urban management, given the complex and rapidly evolving changes of the city environment. Recently, with the phenomenal evolution of urban sensing infrastructures, social network services, and mobile crowd-sensing platforms, the large-scale digital traces left by individuals and crowds in cities are accumulating unprecedentedly. Such urban big data implies rich information about the social and community dynamics in big cities, providing us with new opportunities to address the challenges in urban management. The objective of this doctoral thesis is to design and implement a framework to understand the social and community dynamics in big cities by leveraging the large-scale, heterogeneous, urban big data. The framework will consist of four layers, i.e., urban big data acquisition, urban big data management, urban big data analytics, and urban big data service provision. (1) In the data acquisition layer, three sensing diagrams (i.e., infrastructural sensing, social network sensing, and mobile crowd-sensing) will be implemented to acquire the urban big data. (2) In the data management layer, various data structures, including tensors and graphs, will be employed to manage the highly versatile urban big data. (3) In the data analytics layer, we will employ state-of-the-art data mining and machine learning techniques, addressing the challenges in data variety, velocity, and volume, etc. (4) In the service provision layer, we will leverage the knowledge mined to enable applications in areas such as urban planning, transportation management, location recommendation and network optimization. Especially, in the context of this thesis, we will apply the results of Big Data analysis in 5G mobile networks. Urban big data analytics can provide context information, user profile and user mobility for a better network optimization. Big Data analysis is a very promising solution to an efficient network optimization. The real-time prediction of user mobility and user profile can improve a lot the performances of resource allocation and network planning.

## Résumé du projet de recherche (Langue 2)

The expected challenges and contributions of the thesis are three-fold. (1) Data selection: there are various kinds of urban data available, and each kind of data might contain large volume of information. Therefore, we aim to propose a semi-supervised method by combining human experience and correlation analysis results in the data selection procedure. (2) Data fusion: since the urban big data comes in different sources, format, and quality, we need to propose an effective data fusion mechanism to combine the information from the heterogeneous data and deal with data inconsistency and missing values. (3) Data streaming: the urban data is accumulating continuously, therefore we need to design an efficient data streaming pipeline to process the incoming data while updating the existing knowledge for online scenarios.

## Informations complémentaires (Langue 1)

This research is in a collaboration between University Pierre and Marie Curie, France, Telecom-SudParis, France, Zhejiang University, China, and Peking University, China.

## Informations complémentaires (Langue 2)

Relevant publications: [1] D. Zhang, B. Guo, and Z. Yu, "The emergence of social and community intelligence," *Computer*, vol. 44, no. 7, pp. 21–28, 2011. [2] P. S. Castro, D. Zhang, C. Chen, S. Li, and G. Pan, "From Taxi GPS Traces to Social and Community Dynamics: A Survey," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 17:1–17:34, Dec. 2013. [3] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban Computing: concepts, methodologies, and applications," *ACM Transaction on Intelligent Systems and Technology (ACM TIST)*, 2014. [4] O. Semiari, W. Saad, S. Valentin, M. Bennis, and V. Poor, "Context-Aware Small Cell Networks: How Social Metrics Improve Wireless Resource Allocation," *IEEE Trans. on Wireless Commun.*, Vol. PP(99), Jul. 2015.