**Prévision à court terme des flux de voyageurs :  
une approche par les réseaux bayésiens**

Résumé

Dans ces travaux de thèse, nous proposons un modèle de prévision à court terme des flux de voyageurs basé sur les réseaux bayésiens. Appliqué au réseau de transport public de la région Île-de-France, ce modèle est destiné à répondre à des besoins opérationnels divers, liés à l'information voyageurs, la régulation des flux ou encore la planification de l'offre de transport. Conçu pour s'adapter à tout type de configuration spatiale, il permet de combiner des sources de données hétérogènes et fournit une représentation intuitive des relations de causalité entre ces dernières. Sa capacité à gérer les données manquantes lui permet de réaliser des prédictions en temps réel même en cas de défaillances techniques ou d'absences de systèmes de collecte.

Après un état de l'art des méthodes de prévision des flux à court terme, nous introduisons les réseaux bayésiens et leur extension temporelle (les réseaux bayésiens dynamiques), ainsi que les algorithmes d'apprentissage et d'inférence qui leur sont associés. Nous nous intéressons plus particulièrement aux réseaux bayésiens à mélanges gaussiens, dont la flexibilité permet d'approximer une grande variété de distributions et de représenter ainsi le comportement non linéaire des flux. Dans ce contexte, nous développons notamment un algorithme d'apprentissage des paramètres permettant d'optimiser automatiquement le nombre de composantes des modèles de mélanges gaussiens.

Collectées sur le réseau ferré, les données utilisées dans notre étude sont issues de trois sources différentes : les validations des titres de transport, les comptages par pesée des voyageurs à bord des trains et l'offre de transport. Ces données sont mises en relation par le biais d'un référentiel spatial commun défini à partir d'une description topologique du réseau de transport public. La structure de notre modèle dérive alors directement de ces relations spatiales. Issue de la connaissance experte du fonctionnement du réseau, cette dernière repose sur les dépendances causales entre les flux de voyageurs et leur voisinage spatio-temporel. La prise en compte de l'offre de transport, par l'intermédiaire des intervalles de départ des trains, constitue l'un des aspects novateurs de notre méthode de prédiction.

En appliquant notre approche par les réseaux bayésiens à la ligne 2 du métro de Paris, les résultats obtenus surpassent ceux fournis par d'autres méthodes. Ils témoignent entre autres de la supériorité des réseaux bayésiens à mélanges gaussiens par rapport à l'utilisation de simples distributions gaussiennes. Ils illustrent également l'intérêt d'exploiter le voisinage spatio-temporel des flux, mais aussi et surtout le rôle fondamental de l'offre de transport dans la prédiction.