

Modélisation et Analyse de Traces de Mobilités

– sujet de stage –

Matthieu Latapy – matthieu.latapy@lip6.fr

LIP6 – CNRS et Sorbonne Université – Paris

1 Contexte

De nombreux travaux de recherche reposent sur des traces de mobilités d’entités identifiées. Les exemples sont innombrables : bateaux naviguant de port en port, individus se déplaçant de ville en ville, vélos en libre service se déplaçant de borne en borne, appareils mobiles (téléphones ou voitures par exemple) se déplaçant sur un territoire, bétail se déplaçant d’élevage en élevage, etc.

Toutes ces données consistent en des séries de mouvements élémentaires où on sait qu’une entité identifiée (un bateau, un individu, un vélo, un téléphone, un animal, etc) s’est déplacée d’un lieu donné à un autre à un instant donné. Ces mouvements sont assimilables à des liens entre des nœuds représentant les lieux, donc à des graphes. Leur temporalité peut être représentée par des flots de liens (*link streams*)¹. L’identité des entités peut être associée à chaque lien.

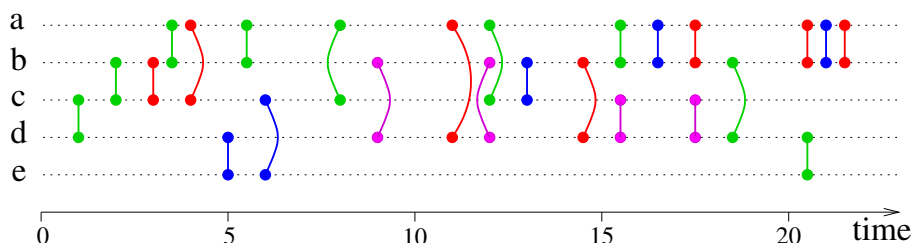


FIGURE 1 – **Un exemple de trace de mobilité modélisée comme flot de liens.** L’axe horizontal représente le temps (mesuré par exemple en heures), et les lettres *a*, *b*, *c*, *d*, et *e* représentent des lieux (par exemple, des bornes de vélos en libre service). On observe les mouvements de quatre entités (par exemple des vélos) représentées par les couleurs des liens : les entités bleue, verte, rouge et violette. Chaque lien représente un déplacement d’une entité d’un lieu à un autre à un moment donné. Par exemple, l’entité verte passe du lieu *d* au lieu *c* à l’instant 1, puis au lieu *b* à l’instant 2, au lieu *a* à l’instant 3, 5, et ainsi de suite.

Ces données contiennent de l’information précieuse dans de nombreux contextes allant de l’analyse des mouvements sociaux à la conception de réseaux mobiles, l’épidémiologie, la sécurité routière ou l’optimisation d’infrastructures pour n’en citer que quelques uns. Il est donc crucial de savoir les modéliser et les analyser de façon appropriée, ce qui permet

1. *Stream Graphs and Link Streams for the Modeling of Interactions over Time*, Matthieu Latapy, Tiphaine Viard and Clémence Magnien, *Social Networks Analysis and Mining*, 8 : 61, 2018. <https://arxiv.org/abs/1710.04073>

de les caractériser et d'utiliser les propriétés mises en évidence pour diverses applications (optimisation de protocoles, vaccination, etc).

La modélisation et l'analyse de ces données soulèvent toutefois de nombreuses problématiques scientifiques.

2 Objectifs

Le premier objectif de ce stage est de concevoir une modélisation appropriée pour les traces de mobilités. Par exemple, les déplacements sont orientés (on va d'un lieu à un autre) et ils ont en général une durée (le temps de trajet). Par ailleurs, les lieux et les relations entre eux sont parfois mal définis : il peut s'agir de coordonnées géographiques imprécises et/ou continues, reliées par des routes susceptibles d'avoir leur propre dynamique.

Cette modélisation se fera à partir des formalismes existants, mais aussi à partir de plusieurs cas concrets de traces de mobilité, afin de la confronter aux situations de terrain. L'équipe dispose de nombreux ensembles de données pour ce faire, qui constitueront un support pratique aux développements formels.

Il s'agira ensuite de définir des structures de données et des processus de traitement adaptés au cas de traces de mobilités massives ; notamment, il faudra traiter les données en *streaming*, c'est-à-dire au fur et à mesure de leur arrivée au cours du temps, sans les stocker entièrement en mémoire. Ces traitements conditionnent et limitent les calculs envisageables, et il s'agira aussi d'identifier les périmètres de ces restrictions.

En parallèle de ce travail algorithmique, il faudra identifier les métriques existantes pour décrire ce type de données. Par exemple, elles contiennent l'information du trafic (quantité de déplacements) entre les lieux au cours du temps, qui est un objet d'étude classique. Elles reflètent également le graphe des connexions (les routes, par exemple) entre les lieux, avec potentiellement des connexions de plus grande importance que d'autres, un autre objet d'étude classique. De nombreuses métriques disparates et redondantes existent dans la littérature de domaines variés² ; il s'agira de les identifier, de les unifier et de les compléter par de nouvelles métriques, si nécessaire.

Ce travail sur les métriques soulèvera des questions sur leur calcul à grande échelle, qu'il faudra résoudre (algorithmes et implémentations). Elles seront calculées sur plusieurs données types, afin d'en évaluer le coût et la pertinence descriptive. Ces mises en pratique fourniront des éclairages permettant en retour d'améliorer les métriques, formant ainsi une boucle vertueuse entre la conception de métriques et leur utilisation.

Enfin, les traces de mobilités contiennent aussi des traces d'événements *anormaux* comme des accidents, des travaux, des congestions, ou des blocages par des mouvements écologistes ou sociaux, par exemple. Mieux comprendre ces événements et leurs impacts est crucial pour de nombreuses raisons, et nous souhaitons utiliser les modélisations et métriques ci-dessus pour apporter un éclairage nouveau. On pourra considérer l'impact de grandes manifestations ou grèves sur les mobilités de taxis, les mobilités de téléphones, ou sur les mobilités maritimes (données disponibles).

Bien sûr, la priorité sera accordée à l'un ou l'autre de ses aspects, suivant les préférences de la personne retenue pour le stage, ses compétences, et les priorités de l'équipe.

2. voir par exemple le mémoire de thèse de géographie de Meihan Jin intitulé *A conceptual and semantic modelling approach for the representation and exploration of human trajectories*, Université de Bretagne occidentale – Brest, 2017.