

# PROPOSITION DE SUJET DE THESE

---



Campagne 2016

Laboratoire L3i

## Sujet de la thèse

Analyse et modélisation de trajectoires utilisateur dans des systèmes réels.

## Résumé du travail proposé

Les déplacements des personnes dans l'espace ont été étudiés en sciences sociales, en géographie et dans beaucoup d'autres domaines depuis de nombreuses années. Il a été montré que les individus ne se déplacent pas de manière aléatoire mais, au contraire, suivent des schémas plus ou moins prédictibles. Cependant, dans de nombreux contextes, les déplacements sont contraints et ne peuvent se produire qu'en suivant les connexions d'un réseau, connexions qui ne correspondent pas nécessairement à une notion de distance, comme par exemple le parcours d'un internaute qui navigue de page en page. Les trajectoires lorsqu'elles se produisent sur un réseau ont été assez peu analysées. Ainsi, l'objet de cette thèse est de décrire ces processus dynamiques, de les comprendre et également de les modéliser dans un contexte temporel : quelles trajectoires sont possibles en fonction de contraintes sur le réseau ? Est-il possible de proposer efficacement des trajectoires vérifiant des contraintes ? Etc.

## Mots clés

Graphes/réseaux, trajectoires, contraintes temps-réel

## Informations complémentaires

### Encadrant(s) :

- Jean-Loup Guillaume (directeur de thèse) [ 50% d'encadrement ],
- Armelle Prigent (encadrante scientifique) [ 50% d'encadrement ].

### Thématiques :

- X Ingénierie des connaissances
- X Analyse et gestion de contenus
- X Interactivité et dynamique des systèmes

### Domaine d'application :

- X Pertinence – contenu – interactions
- Environnement

**Co-financement possible ?** Non

**Cadre de coopération :**

**Date de début du contrat :** septembre 2016

**Durée du contrat :** 3 ans

## Contexte de l'étude

Les données numériques produites par et pour l'homme intègrent différents types de contenus très hétérogènes qui s'entremêlent. De plus, la dimension temporelle de ces contenus prend une place de plus en plus importante du fait qu'elle influence fortement leurs modalités de production, de traitement et d'utilisation. L'hétérogénéité, à laquelle s'ajoutent les aspects temporels, justifie la nécessité d'avoir des méthodes de fouille de données adaptées.

Un exemple de telles données sont les traces laissées par les utilisateurs : elles sont très riches et présentent un fort potentiel pour comprendre comment les utilisateurs interagissent avec leur environnement, mais également pour leur apporter des services plus performants et personnalisés.

Dans cette thèse, nous souhaitons étudier les trajectoires d'entités (qui seront souvent des personnes) lorsque leurs déplacements se produisent dans un réseau/graphe ou dans un espace pouvant être représenté par un réseau/graphe. Cela peut correspondre aux déplacements d'une personne dans une ville (modélisée par le réseau des rues ou des attractions touristiques), à la visite d'un site web par un internaute (qui navigue dans le réseau des pages du site), ou à l'apprentissage d'une activité en passant de tâche en tâche.

L'objectif est double : d'une part, comprendre ces trajectoires temporisées et les relations qu'elles ont avec la structure du réseau sous-jacent et, d'autre part, pouvoir proposer des trajectoires, éventuellement personnalisées, à des utilisateurs, en vérifiant que ces trajectoires sont pertinentes et vérifient des propriétés fixées *a priori* (qu'elles soient temporelles ou structurelles).

## Description du sujet

Deux méthodologies complémentaires seront étudiées durant cette thèse. Tout d'abord, les outils classiques issus de la science des réseaux seront utilisés et étendus pour s'adapter à la notion de trajectoire dans un contexte dynamique et temporisé [REL+14]. Ensuite, des algorithmes de model-checking [Da09] seront définis pour valider des trajectoires, prouver leur faisabilité avec des contraintes temporelles éventuellement paramétrées (une journée pour visiter La Rochelle pour un visiteur ayant une vitesse de déplacement à fixer) ou des contraintes de conflits entre trajectoires (une activité ne pouvant autoriser qu'un visiteur à la fois).

Si l'analyse des trajectoires (détection de trajectoires type, calcul de similarité) est un sujet de recherche actif actuellement, certains enjeux primordiaux restent à résoudre. En effet, dans un contexte de données massives, les algorithmes de traitement des représentations temporelles nécessitent d'être optimisés. Ainsi, par adaptation des algorithmes d'analyse d'accessibilité temporisés et utilisation de structures de données optimales dédiées, notre objectif est de permettre un traitement efficace du calcul de ces trajectoires temporisées. L'une des questions principales liées à l'analyse des trajectoires est de pouvoir les partitionner afin de les modéliser plus finement.

Ensuite, les trajectoires ont surtout été étudiées dans un espace physique (type 2D), notamment pour comprendre la mobilité humaine. Or, dans notre cas, la notion de trajectoires sur réseaux est moins contrainte et pose de nouvelles questions. Dans certains cas enfin, on dispose de réseaux mais qui correspondent à un espace physique (réseau de rues, par exemple) et il convient d'étudier l'impact de la géographie sur les trajectoires observées ou possibles dans le réseau.

De plus, l'un des points de la thèse étant de permettre la construction de trajectoires selon un but donné à partir des trajectoires existantes, nous pouvons envisager l'insertion de contraintes temporelles paramétrées dans les graphes étudiés. Ainsi, il sera possible de définir des déplacements entre les nœuds à des vitesses indéterminées (temps  $p$  pour une activité) ou des durées de trajectoires globales à définir (visiter La Rochelle en  $n$  heures). Les modèles temporisés paramétrés restent difficile à analyser sur de grandes masses de données car ils impliquent un indéterminisme structurel. La thèse aura pour objectif de proposer des algorithmes et structures de données optimales pour ce type de représentation.

Enfin, dans certains contextes, le réseau lui-même peut être dynamique et offrir de nouvelles possibilités de trajectoires (ou en supprimer) au cours du temps. Cela peut être régulier, comme par exemple une activité ouverte seulement pendant quelques heures par jour, une rue interdite à la circulation certains jours de la semaine... ou imprévu.

Enfin, cette thèse est fortement interdisciplinaire et mènera à des contacts avec des géographes ou des juristes, notamment concernant l'utilisation de traces tout en respectant l'anonymat ou du moins le caractère personnel des données.

## Prérequis et contraintes particulières

La thèse nécessitera des compétences en graphes et en modélisation et vérification formelle de systèmes. Des compétences dans l'un, l'autre ou les deux domaines seraient un plus.

Par ailleurs, ce travail amènera des collaborations avec des spécialistes en marketing, en géographie et en droit : un attrait pour l'interdisciplinarité sera donc un élément positif.

## Références bibliographiques

- [REL+14] *Memory in network flows and its effects on spreading dynamics and community detection*. Martin Rosvall, Alcides V. Esquivel, Andrea Lancichinetti, Jevin D. West, Renaud Lambiotte. *Nature Communications* 5, 4630 (2014).
- [EDB+15] *Trajectory Box Plot: a new pattern to summarize movements*. Laurent Etienne, Thomas Devogele, Maïke Buckin, Gavin Mcardle. *International Journal of Geographical Information Science (IJGIS)*, Taylor & Francis (2015).
- [AI03] *Timed Automata*. Rajeev Alur. In *Computer Aided Verification*, Volume 1633 of the series *Lecture Notes in Computer Science* pp 8-22, 2003.
- [Da09] *Tools for Model-Checking Timed Systems*. Alexandre David, Gerd Behrmann, Peter Bulychev, Joakim Byg, Thomas Chatain, Kim G. Larsen, Paul Pettersson, Jacob Illum Rasmussen, Jiri Srba, Wang Yi, Kenneth Y. Joergensen, Didier Lime, Morgan Magnin, Olivier H. Roux, and Louis-Marie Traonouez. In *Communicating Embedded Systems - Software and Design* (2009).

## Contacts

Email(s) : [jean-loup.guillaume@univ-lr.fr](mailto:jean-loup.guillaume@univ-lr.fr), [aprigent@univ-lr.fr](mailto:aprigent@univ-lr.fr)