

Analyse de Nouvelles Formes de Protestation

– projet de thèse –

– SCAI 2022 –

Matthieu Latapy

Matthieu.Latapy@lip6.fr

LIP6 (SU et CNRS)

Julien Randon-Furling

Julien.Randon-Furling@univ-paris1.fr

SAMM (Panthéon Sorbonne)

Contexte

Depuis quelques années, **les militants et activistes à travers le monde inventent de nouvelles façons de faire entendre leurs voix**, de protester^{1, 2}. Parmi les initiatives les plus récentes, on peut notamment relever deux mouvements emblématiques :

- Des activistes écologistes ont lancé les opérations *rebellion of one*³ : plusieurs individus s'assoient en pleine rue, chacun à un endroit différent d'une ville, seuls ou presque, à un même moment. Le but est de créer un effet psychologique et médiatique fort, et de perturber le fonctionnement de la ville.
- Les opposants aux restrictions sanitaires liées au Covid-19 ont organisé les *convois de la liberté*⁴ : des camions et autres véhicules convergent de façon décentralisée mais synchronisée vers de grandes villes (souvent des capitales) afin de perturber les transports dans le pays et le fonctionnement des villes.

L'impact réel ou potentiel de tels mouvements est mal connu, et très peu étudié. Par exemple, la perturbation du trafic est-elle potentiellement forte ? Avec combien de participants, sous quelles hypothèses, et quels lieux cibles ? Est-elle comparable à celle d'une manifestation classique ?

Ces questions sont très liées à des problématiques classiques en informatique, comme par exemple la robustesse des graphes en cas de suppressions de sommets et/ou d'arêtes⁵. Ici, le graphe représente le réseau de routes, et les activistes suppriment (temporairement) un sommet ou une arête du graphe. Ils peuvent être vus comme des agents ayant pour objectif de maximiser la perturbation tout en minimisant les risques, ou le nombre d'agents nécessaires. Des modélisations de dynamiques de graphes à base de flots de liens⁶ semblent alors particulièrement pertinentes.

Par ailleurs, **les données nécessaires pour explorer ces questions sont aujourd'hui disponibles**. En particulier, OpenStreetMap⁷ fournit librement des cartes extrêmement précises de réseaux de routes dans le monde entier. On dispose également de traces GPS de mobilité de nombreux individus.

Objectifs

L'objectif central de ce projet est d'**étudier l'impact potentiel d'actions comme *rebellion of one* ou les *convois de la liberté***, avec un haut niveau de réalisme obtenu grâce aux données de terrain et à une modélisation appropriée.

En reposant sur OpenStreetMap, nous voulons tout d'abord modéliser la ville (ou plus largement le réseau de transports) comme un graphe de rues et de carrefours^{8, 9}. Nous verrons ensuite les activistes

1. Article du Courrier International "Huit nouvelles formes de protestation pacifique".

2. Article du Journal du CNRS "Les nouvelles formes de la contestation".

3. Site officiel de la campagne *Rebellion of One* en France (Extinction Rebellion France).

4. Pages Wikipédia "Canada convoy protest" et "Convoi de la liberté"

5. *Impact of Random Failures and Attacks on Poisson and Power-Law Random Networks*, Clémence Magnien, Matthieu Latapy and Jean-Loup Guillaume, ACM Computing Surveys, 43 (3), 2011. Lien arxiv.

6. *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. Lien arxiv.

7. Site OpenStreetMap.

8. *A typology of street patterns*, Rémi Louf and Marc Barthelemy, Journal of The Royal Society Interface 11.101 (2014) : 20140924. Lien.

9. *Urban street networks, a comparative analysis of ten European cities* Emanuele Strano, Matheus Viana, Luciano

comme des agents susceptibles de supprimer ou surcharger, pour un temps (court) donné, des arêtes de ce graphe. L'objectif devient alors de mesurer l'impact (en termes de connexité, ou de durée des trajets, par exemple) d'une action. Cet impact dépend du placement et des déplacements des agents ; nous voulons donc étudier l'efficacité de plusieurs stratégies simples pour ces placements et déplacements.

En un second temps, nous souhaitons développer une modélisation *adversariale* : les agents activistes suppriment des arêtes ou des sommets, mais d'autres agents les rétablissent peu après, typiquement en supprimant (arrêtant) ou en dispersant des activistes. Les questions autour de la dynamique des agents des deux camps, ainsi que des questions de budget (combien d'agents dans chaque camp) deviennent alors centrales.

La réalisation de ces objectifs produira des connaissances (empiriques et théoriques) nouvelles publiées sous forme d'**articles académiques**, mais également du **code libre** pour extraire et traiter les données OSM, du code de simulation de contestations, et des **ensembles de données** (notamment des graphes de villes). L'ensemble sera mis librement à disposition de la communauté scientifique.

Approche

Nous proposons une approche ancrée en sciences informatiques et mathématiques, mêlant étroitement des travaux empiriques et théoriques.

En effet, la pertinence des résultats reposera sur le réalisme de la modélisation des villes, autant que des agents. Ceci nous amène en premier lieu à exploiter au maximum les très riches données aujourd'hui disponibles : cartographiques (OpenStreetMap), mais également des traçages GPS, des mesures de trafic de divers modes de déplacement (piéton, motorisé, collectif, etc), des données INSEE, et autres. En second lieu, nous serons extrêmement vigilants quant au réalisme de nos modélisations : par exemple, les agents n'ont essentiellement qu'une connaissance locale de leur environnement, doivent prendre des décisions totalement décentralisées, et ils ne peuvent effectuer que des calculs très simples.

Réciproquement, le travail repose sur un riche état-de-l'art dans un ensemble des domaines scientifiques, notamment l'analyse de graphes massifs et dynamiques, les marches aléatoires en temps et/ou espace discrets et/ou continus, les modèles multi-agents, ou l'algorithmique distribuée, pour ne citer que les principaux. Cet état-de-l'art ne répond toutefois pas à tous nos besoins, et il s'agira donc de le faire avancer. Par exemple, nos marches aléatoires ne sont généralement pas en régime stationnaire, mais en régime transitoire, bien plus mal connu. Pire, ces marches ont lieu sur des graphes fortement dynamiques, ce qui implique que le régime transitoire est en fait la règle. De plus, les agents eux-mêmes modifient le réseau sur lequel ils évoluent. Modéliser et analyser ces processus nécessite le développement de travaux fondamentaux, comme par exemple des formalismes capturant toute la dynamique et l'étude de marches sur des graphes dynamiques⁶.

Cet aller-retour permanent entre théorie et pratique, qui se nourrissent l'un de l'autre, est caractéristique de notre approche. Une autre illustration en est donnée par l'analyse des graphes issus des données : elle met en évidence des propriétés, ensuite utilisées comme hypothèses pour l'obtention de théorèmes sur les processus qui s'y déroulent.

Un **stage de Master 2** a déjà été encadré sur le sujet du projet (6 mois, jusqu'à fin février 2022), et un autre est planifié sur les aspects agents (avril à août). Les travaux menés ont permis d'établir la **faisabilité** du projet (données disponibles et exploitables, premières modélisations et simulations, résultats prometteurs). Ils ont également permis d'**identifier les problématiques clés et les méthodes pertinentes** pour les attaquer (graphes fortement dynamiques, modélisation par agents simples, marches aléatoires, algorithmique distribuée). Le présent projet de thèse est issu de ces travaux préliminaires.

da Fontoura Costa, Alessio Cardillo, Sergio Porta and Vito Latora, Environment and Planning B : Planning and Design, 40(6), 1071-1086, 2013. Lien.

Adéquation à l'appel

Ce projet est de nature particulièrement **interdisciplinaire** : il repose sur des méthodes et outils relevant de l'**informatique** (intelligence artificielle, graphes, agents, données, algorithmique distribuée, simulation, optimisation), des **mathématiques** (marches aléatoires, modélisations discrètes et continues, régimes transitoires), et de la **physique statistique** (processus stochastiques, transitions de phase, champ moyen). En outre, les liens avec les **sciences humaines et sociales (SHS)** sont essentiels, notamment dans la modélisation de la ville (géographie) et la modélisation des comportements (psychologie sociale). Le réseau de contacts en SHS des encadrants, ainsi que leur expérience en collaborations interdisciplinaires avec ces domaines, seront alors très utiles.

Ce projet est également caractérisé par une étroite combinaison de **considérations pratiques** (collecte de données, implémentation de simulations) **et théoriques** (modélisation du problème, conception et analyse formelle de stratégies).

Enfin, **la thématique est particulièrement originale** et, à notre connaissance, jamais abordée sous cet angle algorithmique/simulation. Le projet a vocation à répondre à des **questions sociétales importantes**, auxquelles nous pensons que la recherche en informatique et mathématiques peut contribuer de façon significative.

Ces caractéristiques nous semblent correspondre aux priorités mises en avant par l'appel à projets, notamment dans le cadre de **SCAI** (pour les méthodes d'intelligence artificielle mises en jeu et pour l'utilisation de données réelles massives). L'**ISCD** est également concerné, *via* le besoin de simulations intensives et de visualisation des résultats (cartes dynamiques).

Encadrement

L'encadrement de cette thèse réunit **un chercheur en informatique et un chercheur en mathématiques**, diplômé d'un doctorat en physique.

Tous deux ont une **expérience approfondie de l'interdisciplinarité**, entre sciences expérimentales et formelles aussi bien qu'avec les sciences humaines et sociales. Ils ont également une expertise sur l'analyse et l'exploitation de données réelles, dans le cadre de modélisations et simulations réalistes.

Ils **travaillent ensemble** depuis plusieurs mois sur une généralisation de l'algorithme *PageRank* au formalisme des *stream graphs*⁶, qui repose sur des marches aléatoires sur des graphes fortement dynamiques. Ils ont donc une compétence directe sur les sujets au centre du projet, et les contacts nécessaires pour les compétences complémentaires.

L'**encadrant principal, Matthieu Latapy**, a déjà encadré 21 thèses (donc 15 à 50% ou plus). Il n'est actuellement impliqué dans aucun encadrement doctoral et est donc pleinement disponible pour ce projet. Il apporte une compétence en analyse et modélisation de réseaux massifs, issus de données pratiques, ainsi qu'en algorithmique. Les publications des notes de bas de page 5 et 6 illustrent ses contributions directement liées au projet.

Le **co-encadrant, Julien Randon-Furling**, apporte des compétences complémentaires en marches aléatoires et processus stochastiques. Une délégation CNRS dans l'équipe de Matthieu Latapy a été demandée pour un an (2022-23, examen en cours). Il a encadré (à 100%) une thèse qui sera soutenue ce semestre, sur des approches probabilistes et de physique statistique appliquées au modèle multi-agents de Schelling. Il a animé en 2017-2018 l'Atelier Campus Condorcet "Faces of Paris" consacré aux dynamiques socio-spatiales dans l'agglomération parisienne.

Profil recherché

Le projet nécessite clairement une **forte ouverture interdisciplinaire**, entre informatique et mathématiques notamment, mais également avec la physique et les SHS. De façon similaire, une ouverture à une combinaison de **travaux à la fois théoriques et pratiques** est essentielle.

En termes de compétences, une formation sur les **réseaux complexes** et/ou les **processus stochastiques** sera nécessaire (préliminaire ou à acquérir en début de thèse). De même, des compétences en **programmation** et en **observation/analyse des résultats** (statistiques) empiriques seront cruciales.