Étude des nouvelles formes de protestations

written by
Guillaume Moinard

supervised by
Matthieu Latapy and *Nicolas Maudet*

Internship report for the
Master 2 Physics of Complex Systems

Laboratoire d'informatique de Paris 6
Sorbonne Université
April 1 - June 25, 2022

Abstract

During the past decades, new kind of protests have emerged. They focus on maximizing the car traffic perturbation, mainly by blocking streets until they are dispersed. While a full simulation or analysis of this ecosystem is out of reach, many articles have already studied how traffic can respond to congestion.However, the literature mainly focuses on explaining self-induced traffic jams, without much interest for a possible underlying dynamic of an exogenous perturbation. In this study, we present how an agent-based model able to form aggregates can induce congestion on a city modelled by a directed graph. We start by gathering data about Paris, especially with the library OSMnx in order to model the road infrastructure as a graph. We also look for different classes of graph having the properties cities exhibit, such as quasi-planarity or specific degree distributions. We then explore different metrics to quantify the impact of activist strategies and to find weak points in the infrastructure. After those prerequisites, we design our agent-based model regarding strategies adopted by activists, and we discuss response of the model to different scenarios.

This report is the result of the first three months out of five of research around this topic. As this is only the beginning of our study, we have mainly focused on exploring the different aspect of our study and the associated state of the art. We will present the details of our analysis and we will propose some perspectives.

Keywords : Protests, OpenStreetMap, distributed algorithm, agent-based model, active particles, aggregates, graph theory, metrics, robustness, link streams

Contents

1 Introduction

I decided to do my internship at the *complex networks* team in the LIP6 for many reasons. The first was because the subject fascinated me, and this was a major factor of choice in order to be dedicated to the task. The second concerns the nature of the team. Besides being marvelous people, the team main focus is the development of a clear formalism to analyse dynamic graphs. Unlike static graphs, those graphs change over time. Nodes are created or destroyed, edges are added or disappear.
As my academic project is to keep on studying social oriented questions, I know what an important role graph theory plays, and we easily see how social interactions inevitably imply having a dynamic network. However, this dynamic is often put under the rug, and is only presented as a succession of static graphs. This is not the most efficient, nor the simplest way to deal with this object and I will present the formalism of **link streams** [1] that the team of *complex networks* has developed.

All over the world, new kinds and methods of protest arise. While the more "traditional" ways to protest, especially **demonstrations**, are more focused on showing to the politics and the public, via the press, that a large number of people gathered to show their reaction to a decision, those new actions are more focused on having an important impact on the underlying structure ; their goal is to disturb the traffic.
Examples of those actions come mainly from the last climate activists actions like *extinction rebellion*, which decided to block a bridge in London.[1] Or when a large number of citizens cross a city on bikes, creating a succession of congestion wherever they pass.[2]

In this report, we won't comment on the legitimacy of those actions. Nor will we detail the political and sociological reasons that push those people to act, at a point which can put themselves in danger, and which can go beyond the limits defined by the law. We are not sociologists or moral philosophers.

However, one important question arises when one wonders **what is the impact of those actions.** Do those activists achieve their primary goal ? And what **exactly** is their goal ? Is changing everyone average travel time by 1% enough, or wouldn't they prefer to slow down more efficiently a smaller number of travelers ?
And how do we measure the impact of those actions ? Couldn't we model the underlying structure they act upon in order to define measures to estimate the impact a certain perturbation has ? As those perturbations are defined by strategies chosen by activist, couldn't we model them with simple behaviors from which would emerge more complex dynamics ?

Those questions have rarely been tackled rigorously. Almost two centuries ago Auguste Bianchi, facing the failures of the 1848 riots, was one of the only person who tried to offer a methodology to analyse how to better protest in a city [2]. Today, it is still possible to find books by activists which try to analyse different types of protest[3]. But none of them do it from the scientific point of view, with the perspective of giving a quantitative analysis from real protest data and simulations.

Before starting, we have to specify what other aspect of protests have already been studied. Indeed, a few articles study this subject from a different point of view. For example, I had the pleasure to discuss with Jean-Pierre Nadal about his study on french 2005 riots in suburbs. [3] His study focuses on the spread of the riot as a contagion process. They look, in the data, at how people in given area started being active in those protests based on how their geographical neighbourhood behaved. When people from a certain suburb started rioting, the ones from the closest regions followed them, which led to a riot wave that could be analysed with an epidemic model.
This echoes another study [4], agent based modelled this time, which analysed how each agent start rioting by evaluating the power balance between the number of people already rioting and his estimated number of peace keepers.

This literature remains very limited, and with respect to it, our approach looks at a blind spot on the current state of the art. Therefore, the goal of my internship was to give a first approach on how

[1] BBC article about XR protest
[2] Website about veloruptions
[3] Website for the book *Hacker Protest*

to tackle those questions with the tools from physics and computational science. Indeed, while our questions are not explicitly dealt with in articles, by translating this study in the language of natural sciences, I will show we can link it to known problems and an already existing rich state of the art.

2 How to model a city

All along this study, we will be using the python library *networkx* to generate graphs and handle them[4], as it is easy to use, offers a large number of already implemented algorithms to create and make measurements on graphs, and is supported by a rich documentation and an active community.

First of all, we need to correctly define our playground. We wish to simulate some kind of protests and evaluate quantitatively their impact. Therefore we need to choose how to model the city, the network they act upon. Our context for this study is an urban city. Actually *any* urban city. Which means that we first want to be able to model a city from real data.

2.1 Real data models

A simple way to construct a model for a city is to represent it as a graph. Indeed, we can easily describe any intersection as a node, and the streets between those intersections as edges. Moreover, using a directed multigraph with attributes, we can include the fact street often have many lanes in two opposite directions. We can also add properties to each node and edge so our model is even richer and more realistic. Now our problem lies in finding the data and being able to easily convert this network into a graph. Hopefully, recent upgrades in geographic open source data, and the creation of easy to use libraries solved those problem. We used the python library OSMnx [5] to generate graphs from real city data. The data are extracted from the website OpenStreetMap, which is the open source version of Google Maps with extra features, thanks to a dedicated community.[5] With OSMnx it is also possible to give properties to each node or edge of the generated graph, based on informations available, such as the number of lanes, speed limits or lengths of the roads, only to cite a few of those that will be relevant to our study.

We focus on Paris 1a, as our personal knowledge of the city can help us detect inconsistencies in the data. But also as Paris is the French city the most subject to the kind of protests we study, and as big cities are more detailed on OSMnx. The first thing I did, after having great fun mapping graphs of all the places I know, was to plot on figure 1b a skeleton of Paris' road infrastructure. What I mean is that I highlighted, in yellow, streets which have an important number of lanes, while in purple we see streets with only one or two lanes. This allows us to see what are the main arteries of the network, as streets with an important number of lanes are the ones that can handle the biggest charge in terms of car traffic.

(a) Google Earth picture of Paris

(b) Graph of Paris from OSM data

Figure 1: From Paris to its graph representation

[4] Networkx website
[5] OpenStreetMap website

On the other hand, we also see on the plot, and directly by sorting the data, which areas of an important path are vulnerable to perturbations. An example of that is the narrow region of the boulevard periphérique at the south of Paris, which has only two lanes in both directions, and has indeed been perturbed by activist protests.[6]

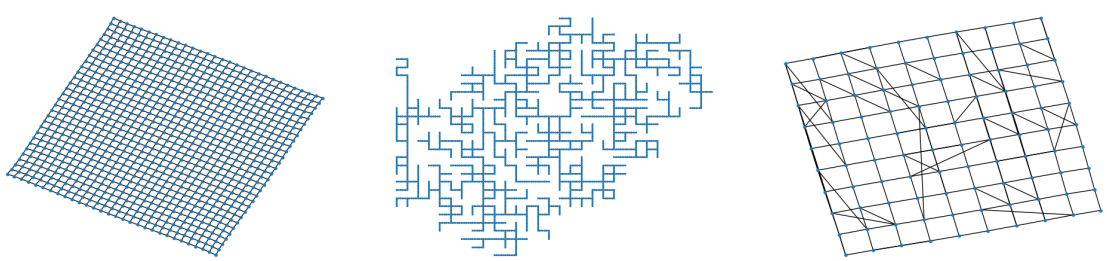
2.2 Theoretical models

Being able to generate such an accurate graph to describe our city is marvelous. But every city is different, and we also would like to know what general properties a model of a city should exhibit. Moreover our goal is to use an agent based model on a graph representing our city, in order to see what impact the agents will have on the network, and then to measure the efficiency of these actions. However, if we want to have any insight in the impact our agent based model, we first need to understand its behaviors on very simple city models. Moreover, we need those models to be "realistic enough" for us to be exploring the properties of our activist model in a meaningful way.

For example, a random network such as Erdős–Rényi [6] is a very bad choice. While it is a good starting point for studying a random graph (for example in social networks) its properties make no sense for a graph that should represent a geographical infrastructure as a city, as you could link by a road any intersections of your Erdős–Rényi city whatever the distance between them.

For those reasons, in this section we are going to explore different graph models that we decided to use as different "steps" in order to go from the simplest idea we can have, until a model that would resemble the most a graph generated with real world data.

So what classes of graph are fitted to model a city ? The first kind of graph we can use is a two dimensional grid as on figure 2a. This is probably the simplest graph we can think of in order to model a geographic system. First, this kind of graph is planar, which means you can project the graph on a surface without having any edge crossing each other. Second, if you choose edges with very short lengths, which means you discretize with a small unit length, you can get a space that looks almost continuous. Which explains why two dimensional grids are often used to simulate stochastic processes and active particles in motion.
In the section on agent based models we will start and focus mainly on a two dimensional grid, but for now couldn't we get closer to reality ?



(a) Two dimensional grid (b) Percolation on edges (c) Navigable small-world

Figure 2: Classes of graph used to model cities

One refinement we can start with is to make our two dimensional grid non planar. Here we have to justify ourselves, because this seem to be in contradiction with our previous statements. The thing is, city networks are only almost planar graphs, the truth is slightly more complicated. In a city, there can be bridges and tunnels, which implies that some links won't respect planarity. [7] Those roads are designed to be short paths from one point to another of the network. In a sense, those paths tend to add a small-world feature in the city graph. Therefore we are looking for a simple modification on our grid that would introduce this property. Such class of graph was first described by Kleinberg [8] in

[6] Website of Dernière Rénovation

order to understand how agents with only a local vision on a graph, could find the shortest path to a target which location they know. We shall only focus here on the properties of the graph he used. Starting from our two dimensional grid, we add new long-range connections from a node u until an endpoint v. Those links are chosen randomly with probability proportional to $d(u,v)^{-r}$ where d is the distance between two nodes and r a chosen constant. While important properties, in terms of shortest path lengths, arise from the value $r = 2$, it might be more accurate to choose higher values to get a model that won't create too many ridiculously long bridges and tunnels. After getting rid of the self-loops generated for some nodes, we end up with the graph 2c for $r = 3$.
Comparing the lengths of those new links to the links between intersections, and doing the same on graphs from real data, we should be able in a near future to fit this parameter r.

This stylized grid exhibits some of the properties we are looking for. However, to approach a realistic model for the city, we can't stay on a lattice without any obstacles forever. Indeed, buildings shape a city and we somehow need to introduce their presence in our model. Barthelemy [9] mentions the idea that a good, and still simple, candidate could be the giant component of a percolated two dimensional grid close to the phase transition. Performing a percolation means deleting each edge with a given probability. On a 2D grid, there is a critical percolation value $p_c = 0.5$. Below it, about all components are small and disconnected, and above it almost all nodes are connected to one giant component, while the remaining components are very small in comparison. This component, taken close but just below the phase transition, is our candidate as a city model. This produces the graph in 2b. Beside having a better shape, this network also has a degree distribution closer to real cities. Indeed almost all nodes have degree 3 or 4 on city graphs. This is a problem as we can't apply most of complex networks techniques based on the analysis of degree distribution, but it allow us to make a simple first test on our generated city graphs. Counting the number of nodes of degree 3 and of degree 4, we get a ratio of 0.360 for Paris and of 0.373 on average for our percolated 2D grid with $p_c = 0.4$ and roughly the same number of nodes.

We still are far from reality. No complex patterns such as places or dense regions are generated, and some part of this graph seem to be easy to disconnect. This never happens in city, except with bridges when a a river separates the city, but we haven't yet explored if river-cities present specific common characteristics and we still don't know what would they would be in terms of graph properties.

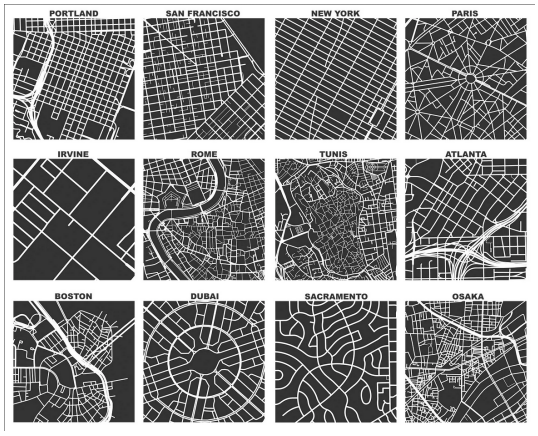
2.3 Refinements

Our theoretical cities can be greatly improved. We will start by merging those different simple classes of graph. Indeed, a real network exhibits properties from the percolated model and from the navigable small-world. Moreover, as OpenStreetMap provides us a very large source of data, we can fit the parameters controlling the generation of our graphs, like we did with the percolation threshold. Concerning the navigable small-world 2c, we will use metrics [7] evaluating how far from planarity is a city network. Combined with lengths and proportions of brides and tunnels among the streets, we should be able to fit our models to real data.

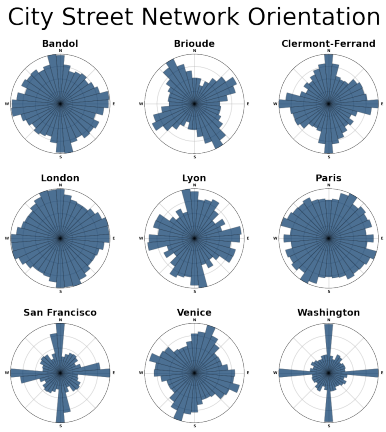
However, overfitting our model to one city could be dangerous. There actually might be different classes of city, with different common properties from one class to another. We can for example sort cities based on street orientations 3b, or similarity of topology : like all cities crossed by a river. Or degree distributions, etc. How to study and extract those informations have been done [10] and we will build on the already existing work. Assuming some cities fall under a same class, it wouldn't be surprising that they would have common weak points and then that the impact of a strategy in one city would be similar on the other cities of this class.

To illustrate our point, we see from figure 3a that a city like Portland looks like 2a, while New York might be a graph like 2b. And from figure 3b you see you can exploit more regular patterns on a city like Washington than on Paris.

We also started sequencing our edges, like actually shown on figure 2b. We choose a unit step, and we divide our edges in new paths, with new nodes, and with as many new unit edges in order to keep the length of the original edge. This artefact is only there in order to simplify the future activists



(a) Wide range of shapes between cities



(b) Comparison between French and worldwide cities

Figure 3: Cities exhibit different properties and shapes

motion we will generate on the graphs. It is necessary because on real graphs you get edges of very different lengths, depending on the number of intersections a given street has, and we want our agents to make moves of the same length at every step.
Moreover, this will generate a new graph with a locally tree-like structure. Therefore, this opens the door to the usage of mean field approaches.

On the other side, while the network generated by OpenStreetMap is good, it lacks a few realistic features we still need. Some of them are details, like the possibility to delete edges that appear as exiting on OSM maps, but which are temporally unavailable. This can be due to work sites and construction sites present on the road[7], or actually any other kind of event[8]. Some of them are more important, like the lack of information about the streets' width on OSM, that can be filled in part with other sources[9].

Traffic

While we have focused on building a realistic city, we still haven't talked about the main actor of this network. The question of how to model the traffic is tedious. Indeed, if we could successfully simulate car traffic on our graph, then the evaluation of the impact of activists action would be straight forward. If a bridge was to be shut down by a group of activists, it would induce a congestion at his entrance, the drivers would need to choose another path, which could induce a cascade of congestion, directly visible by the motion of cars on our graph. However, this is too complex. Besides requiring enormous numeric capacity and computing time, the state of the art in terms of traffic prediction doesn't allow us to evaluate how would really behave cars on a graph of a city. Especially with the massive usage of GPS while driving, driver behaviour is very different as it is dependent on the analysis of real time data from traffic. Indeed, car traffic can not be modelled as a fluid passing through a network with simple systematic decision rules. Drivers are agents, who take decisions based on their evaluation as in a game. This lead to complicated equilibriums with congestion, and cascade of congestion, which are different from the optimum. This is known as Braess's paradox [11]. This paradox is very important in our study. Indeed, it implies that deleting some roads, might lead to making the network more efficient at accomplishing its task. Therefore, making the impact negative.

Last but not least, we realized recently that Germany was the best mapped country on OpenStreetMap. This is mainly due to the behavior of the government who has forbidden Google Street View, based on their more rigorous definition of privacy. But also to a very active open source community in Germany. Therefore, analysing the city graph properties generated from german cities might

[7] Travaux perturbants la circulation / Chantiers à Paris
[8] Project listing every "event" occurring in the world. From music festival to riots
[9] Data about the width of roads in France

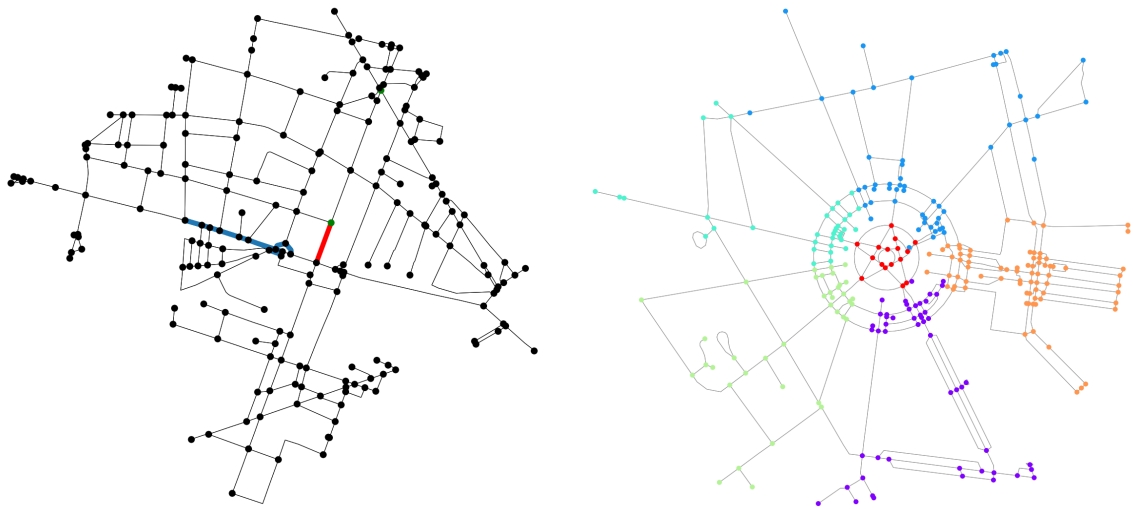
be more accurate, and we will try to compare if more details, like specific edges or nodes properties are better referenced on those cities.

3 How to measure a perturbation

Before putting to test our cities by perturbing them, we must define tools to evaluate the impact of that perturbation. Moreover, we want to be able to find weak points in the network, in order to tackle questions such as strategy design, but also to evaluate the robustness of the infrastructure.
In this section, we will start by exploring different sets of metrics to identify what edges or nodes should be targeted, and how the infrastructure is transformed if they are deleted. We then define how the activists and the network are coupled together, before getting in the section where we will model the activists.

3.1 Robustness and weakness in an infrastructure

Measures of the impact of perturbations on networks of infrastructures have already been done [12]. Those studies focus on uncorrelated attacks or random failures. As a consequence, perturbations on planar graphs don't seem to have been investigated much. However metrics and scores exist to evaluate the efficiency of a network at performing a given task.



(a) Network of Porte Saint-Denis

(b) Network of Place de la Nation

Figure 4: On the left 4a, in blue, the region blocked by Extinction Rebellion. Other colors are weak nodes are edges identified by different metrics. On the right 4b, the different colors are neighborhoods found by a community detection algorithm.

We have tested different classical metrics giving the importance of a node or edge in a graph, for example centrality measures. They tell how many shortest paths, between all pair of nodes in the network, pass through a given edge or node. Therefore parts of a network evaluated with a high coefficient are important, as their withdraw impact many shortest paths length, imposing them to grow. While centrality helps us find weak points, the change in the average length of shortest paths can be used to measure the impact of the perturbation. This is a basic combination illustrating how we want to proceed in order to characterize the robustness of the graph.

As an example, and without getting into the details of more metrics, we find on figure 4a that the most central edge is the red one. However if we delete it from the graph, the impact is in the end less important than the impact the group Extinction Rebellion had by blocking the blue area on the same plot. One important remark is that protests, like this one, usually impact a neighborhood instead of single isolated nodes. Therefore, it makes sense to identify clusters of nodes, that will be impacted if one of them is impacted, and therefore to measure the overall impact considering all of those nodes

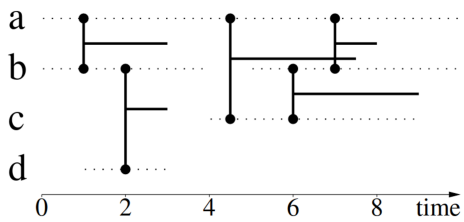
have been impacted. This means that instead of finding what node is the most important, maybe we can define group of nodes, and see what group will be lacking the most to the overall infrastructure if we delete them one at a time. This research is realised by applying community detection algorithms, which can even be specific to planar graphs [7]. On figure 4b, those algorithms are able to identify that nodes at the center of Place de la Nation, form a unique group, which indeed correspond in reality to the square at the center of this place. By finding dense regions among those communities, usually characterized by a high numbers of triangles and cliques, we will also be able to find robust sections of the graph.

Going back to the protest on figure 4a, did activists chose the right place ? Not so easy to say. Despite the visual impact of organizing a meeting below the gate at Porte Saint-Denis, we are still not taking here traffic into account. We have already seen that realistic simulations of traffic on such a scale are out of reach. Another way is to start by attributing specific weights to each edges. They could be associated to real traffic data from sensors[10]. Then, if one is able to only evaluate how a flux of car is roughly redistributed when an edge is closed, we can give an estimation of the impact of this action.

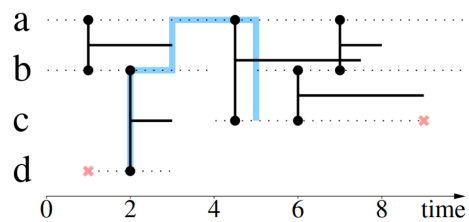
Another idea to measure the impact of an action is to use the city scores defined by Vittorio Loreto [13]. In his article he defines scores to evaluate how good is a city to connect its different areas and to transport people from one point to another. When I had the chance to talk with him, he confirmed me that it would be simple to use those scores in the reverse way. Which is to measure how less efficient the city is to transport people after a protest. This is a direct measure of its impact.

Whatever the angle we looked at it, we only thought the question from the static point of view. First, static in space. Indeed, on a very local space, the impact induced by the perturbation of a node is enormous. For the drivers blocked in front of the deleted node, time travel increases a lot. However, at the scale of the whole city, or at the department scale, it is probably as if nothing had happened. How does the transition of the impact behaves between those two scales ? Is there a sharp distance value at which the impact drops drastically, or do we have a smoother transition ? We still need to study this question. Second, our analysis was static in time. Indeed, a perturbation impacts the infrastructure over time. A huge perturbation, resolved in a very narrow window of time might have almost no impact. While a long perturbation, even if weak, could bring the entire system to another, and maybe worse, equilibrium state. We will try to extend our analysis, at least to the dynamic graphs, using the formalism of link streams.

This gives me the opportunity to present what are link streams. Without going into the details of a formalism I am still learning, the idea of link stream is to represent dynamic graphs in the same way as functions represent a series of numbers. Dynamical graphs, which are simply graphs evolving in time, are often represented as a succession of static graph. A link streams, on the other hand is a static object, exhibiting all the properties the dynamical graph had over time, without the need of being showed as a movie to show those properties. The figure 5a shows a link stream for a graph with four nodes : a, b, c, d and with links appearing and disappearing over time. When a link appears, a connection is drawn, and the perpendicular line shows until what time it persists.



(a) Example of a simple link stream



(b) Shortest path from d at $t = 1$ to c at $t = 9$

Of course, this formalism is not only for drawing dynamical graphs. The core idea is to use it to export static graph concept to dynamic graphs. We can think of shortest paths, as illustrated on

[10] Comptage routier - Données trafic issues des capteurs permanents

figure 5b. Indeed, having a list of a dozen pictures of graphs indexed by time, and being able to find the shortest path between two points of space-time is a complicated task, while it is almost straight forward on this link stream.

In our study, link streams can first be used to characterize the impact of a protest. Concepts such as an extension of the derivative are being studied on link streams in the complex networks team and could be used to tell how abrupt is a transformation in a dynamic graph from one time to the following. Applying this to the dilation of a node or edge because of a protest could be translated as a direct calculation of its impact on the network. In the next section we will also present how link streams could be used to study the copresence graph induced by the simultaneous presence in the field of vision of two agents.

3.2 Coupling between the charge and the structure

Coupling the charge and the structure can be straight forward. We can say that each node possess a certain capacity, given by the surface of the intersection the node represent. When enough protesters are gathered on the node and the capacity is reached, we consider it is deleted from the graph, from the point of view of the drivers, until enough people have left the node for the capacity not to be saturated anymore.

However, a few problems remain. First of all, evaluating the area of an intersections isn't a given data. Still, it will be an inevitable problem, and a possible approximation is to multiply the number of lanes (length) from the biggest road by the number of lanes of the second biggest road. If we are on a crossroad or an intersection, this is the exact answer. If crossing roads are more than two or don't perpendicularly intersect, this at least gives an approximation.

Let us remind we want to evaluate how much a perturbation on a given node impact traffic from the incoming edge, without modelling the overall traffic. We thought about taking the *TASEP* model [14], already used to model traffic jams, and that has already been applied on networks [15]. The reason is because this model can be a great way to couple the charge and the structure. Indeed, in this model, the behavior of the traffic is only defined by the transition rates of the exit node, β and of the entrance node, α. Depending on only those two parameters, and with a simple model of particles hopping on a lattice in between, the model provides an already rich phase diagram with mainly a fluent phase and an other one with traffic jams. So instead of deleting nodes above a threshold, the node charge could define those values α and β. Indeed, even if the node is not filled by protesters, their simple presence impacts how easily drivers can get out of an edge or get into a new one. Having values of α and β for each node, depending on the charge of those nodes, we therefore could deduce if the state of the in between edge is traffic jam or not.

4 How to implement the activists behavior

During the French Regional Conference on Complex Systems, I was lucky enough to talk with Alexandre Nicolas from the Institut Lumière Matière in Lyon, which main field of study is pedestrian dynamics. This was a great opportunity to get a clear idea of the state of the art. Pedestrian dynamics are described with many different formalism. Physicists got into this field using techniques from granular materials, but this is now rejected as it doesn't include the ability to anticipate collisions [16]. As presented by the authors of this study, nowadays methods can sometimes be based on mean field games as Denis Ullmo does. But they mainly focus on *Newtonian force* models. Those models describe pedestrians as particles attracted to a certain point, their destination, and feeling repulsive forces from objects and other pedestrians, while having an internal energy function leading to an equilibrium speed along their path. Complex behaviors emerging from those models can for example describe how people evacuate a specific building, and they are precise enough to be sold to companies. But they also reproduce stop and go waves in crowds, and queues formation in corridors with opposite incoming pedestrian flux [17].

However, we realised we are actually not interesting by the dynamic at this scale. Sure, in an ideal world it would be very useful to be able to model the microscopic dynamic of our system, between activists and officers, such as kettling [18]. But, having explicitly asked to Alexandre Nicolas, this clearly is out of reach giving the current state of the art. Moreover, the time needed to compute such interactions all other a city would be too important to derive any information from simulations.

Our solution to this problem is to propose a **mesoscopic** approach of the situation. We are somehow more interested by the general state of agents in different regions of the city, than the microscopic interactions between every infantryman. None of those two scales are the macroscopic scale, where we see the aggregates formed. That is why we can introduce and talk about a mesoscopic scale, between the microscopic and the macroscopic scale. In this section, we will present the agent-based model we have chosen as a first approach to simulate the mesoscopic behavior of the activists on a city. We will discuss and justify its details and explain how it behaves on our city graphs.

4.1 Agent-based modelling

We are going to use agent-based modelling to generate the activist dynamic on the city graphs [19]. The idea behind the paradigm is to create objects, with their own internal rules, that will interact together. This is a kind of distributed algorithm used to solve a problem. It is the collective action that will lead to the emergence of complex patterns that we want to study. Agent-based models are more and more studied in fields such as economics, where each agent has rules on how to buy and sell certain quantities of product on a given market. The collective action of many agents on this market can lead to the emergence of real life economic behavior such as periodic crisis [20], and so give us insights on how endegeneous information can transform drastically the equilibrium of the system. To implement our agents, we will use the python library *Mesa*[11] designed specifically for simple implementation of agent based modelling.

But what exactly do we want our agents to do ? As we said, our goal is no longer to describe the microscopic interactions. And we don't aim for a realistic global behavior to begin with. We would like to start by implementing a model that catches some of the features exhibited by protesters, in order to measure what impact those features will have the networks we previously generated. We identified two main properties, which are the capacity to form **aggregates**, even after dispersion, and the capacity to evaluate their numbers as well as the power balance with the officers. The first one falls under the class of *rendezvous* problems, while the second is a question of consensus among players. We define behaviors such as a vision, a dynamic, and the ability to communicate to get a model able to reproduce those features.

Vision

The first characteristic to give to our agents is a vision. Defining a vision will be very important in order to allow them to count other agents, and to choose the one they will communicate with. For the moment we simply consider every agent sees all the other agents on its node. However we could be pushed to give them a topological vision in stead of a geographical one, as agents in swarms seem to take decisions based only on a restricted number of other agents instead of as many as they can stare at [21]. A restricted number seems also more credible in terms of people I can communicate to. On the other side, it seems more realistic that I see further than one node, and that my dynamic is not defined only by the number of people in my node, but also in nearest nodes, so that I can be attracted by populated nodes if mine is empty.

Dynamic

Most agent-based model we will explore are able to describe animals or particles motion, but they might not be correct to describe pedestrians in a street. But as we said, our goal is to look at another scale, at which we are only interested in specific already aggregated properties.

[11] Mesa website

As our first point of interest was to generate aggregates of agents, I spontaneously thought about *boids* [22]. They are a general class of agent-based model for swarms, defined by three properties. Alignment, the capacity to mimic the direction of motion of neighbours. Separation, a reaction in order to avoid collision, but in the same time cohesion, the will to get go closer to groups. This model might not be true for general pedestrians, but could be interesting for a group which have a strategy of similar to moving as a swarms. But while we should keep their principles in mind, boids might be too much of a microscopic model for us to use.

I then reminded Julien Tailleur course about **active brownian particles** (ABPs). Active matter describe interacting particles which are self-propelled [23]. ABPs translations are subject to fluctuations as well as their direction of motion. It can also be subject to external forces. This leads to the set of equations 1-4.

$$\dot{\mathbf{r}} = v_0\hat{\mathbf{n}} + \nabla V(\mathbf{r}) + \sqrt{2D}\,\boldsymbol{\eta}_{\text{trans}}(t) \tag{1}$$

$$\dot{\theta} = \sqrt{2D_r}\,\eta_{\text{rot}}(t). \tag{2}$$

where we have a gaussian white noise :

$$\langle \eta_{\text{rot}}(t)\rangle = 0; \qquad \langle \eta_{\text{rot}}(t)\eta_{\text{rot}}(t')\rangle = \delta(t-t') \tag{3}$$

$$\langle \boldsymbol{\eta}_{\text{trans}}(t)\rangle = \mathbf{0}; \qquad \langle \boldsymbol{\eta}_{\text{trans}}(t)\boldsymbol{\eta}_{\text{trans}}(t')\rangle = \mathbf{I}\delta(t-t') \tag{4}$$

Actually we need no potential for now. Moreover, in order to form aggregates, we need the velocity to depend on the local density ρ. We get :

$$\dot{\mathbf{r}} = \frac{v_0}{\rho}\hat{\mathbf{n}} + \sqrt{2D}\,\boldsymbol{\eta}_{\text{trans}}(t) \tag{5}$$

Those stochastic differential equations, and the phase separation in equation 5 between a dense and dilute phase have been well studied [23]. They are already used in simulations on continuous space to simulate the motion of ABPs. However, in our case we want them to evolve on a discrete space. While this is already well studied on 2D lattice space, with simplified equations [24], the road network is a non regular space, with no symmetry and unambiguous notion of left/right, up/down. Therefore we propose the following algorithm 1 to extend ABPs motion to our city graphs. We simply simulate a classic ABPs motion, and we pick among nearest node neighbours the closest to the position we would have gotten to in continuous space. This leads to the formation of aggregates on lattice 6a as already known, but also on other kind of graphs such as figure 6b. While only at the beginning at this stage, an advanced goal will be to find under what graph conditions those aggregates form.

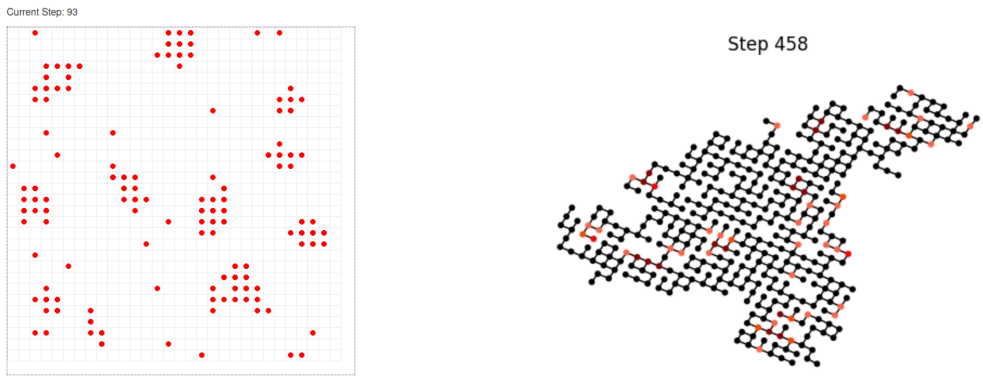
Algorithm 1 Discrete ABM on a geographic network

Require:
- N players $i \in [1, N]$ on a graph ▷ The nodes must have positions
- D, D_r, v_0 given
- each player i has x_i, y_i, θ_i
- **for** $j \in N$ **do** ▷ Find closest node to "real" ABP motion
  - $\rho_j \leftarrow$ number agents on my node
  - $dx_j \leftarrow \frac{v_0}{\rho_j} \times \cos(\theta_j) + \sqrt{2D} \times random.normal()$
  - $dy_j \leftarrow \frac{v_0}{\rho_j} \times \sin(\theta_j) + \sqrt{2D} \times random.normal()$
  - $x' \leftarrow x_j + dx_j$
  - $y' \leftarrow y_j + dy_j$
  - $(x_j, y_j) \leftarrow$ $node \in$ neighbours(x_j, y_j) **closest** to (x', y')
  - $\theta_i \leftarrow \theta_i + \sqrt{2D_r} \times random.normal()$
- **end for**

We realised that on city networks, the largest distance are relatively short with respect to the average length between nodes. Therefore, agents had a tendency to aggregate only on the borders of the map. To correct this behavior, we introduced an elastic rebound. Every time the agent would

point outside the graph, it would get pushed in the symmetric direction with respect to the boundary, as if it was an **elastic collision** with a wall.

However, defining such an abrupt objective end of the map is not realistic. We recently realised that we could correct this behavior by reintroducing the $\nabla V(\mathbf{r})$ term. Indeed, with $\mathbf{r}$ being the position from the center of the map, and d being the average distance from the center to a boundary, $\left(\frac{\mathbf{r}}{d}\right)^p$ gives a potential $V(\mathbf{r})$ that forbids agents to go far from their region of action, and bias them into going to the center of this area, if p is a positive even number.
But we can go further by fine tuning this potential. While this $\left(\frac{\mathbf{r}}{d}\right)^p$ is a good start, we could also modify it locally depending on the density of the regions in the graph. Many clustering metrics allow us to evaluate density on a graph. We could for example decide that places with an important number of triangles are more attractive than others, as they represent squares in cities.



(a) On a 2D grid

(b) On a percolated lattice

Figure 6: Agents, initially uniformly distributed, gather in aggregates

While the previous dynamic is the one we started using for simulations, here is an example of a model of human mobility, already tuned for the mesoscale we are looking at. It is inspired by the Schelling model [25] proposed to generate a dynamic of segregation in a city in which all agents are separated in two groups and are *tolerant*. This is defined in the sense that the agents utility function, without giving the details, tells us they are the happier when there is as much agents from their own community as agents from the other community. Still, when the agents will be given the opportunity to relocate, this will lead in on overall segregation phenomenon, with people moving to neighborhoods with members of their groups.

Back to our cities, here is briefly how Barthélemy changed it [26]. We are no longer dealing with two groups of people that all would like to be in a mixed society, but on group of people that want to be in half-filled neighborhoods. With enough people, bur not too much. The probability of transition from one neighbor to an other is then given by equation 6.

$$P(\{\mathcal{C}\} \to \{\mathcal{C}'\}) = \frac{1}{1 + \exp\left(-G(\Delta U)/T\right)} \tag{6}$$

Without presenting the analysis on the mechanics, and on the function G depending on the difference between the utilities in the neighbors $\mathcal{C}$ and $\mathcal{C}'$, I will just say that this new model leads to the formation of over and under populated neighborhoods, instead of an uniformly distributed population which would be the best configuration. We started trying this dynamic for the agents, as it seems to lead to the formation of aggregates. We simply have to specify the agents can only move to nearest nodes instead of relocating to any node of the city. Moreover it can more easily be modelled by a Markov chain, which could help us to find if the system converges to an equilibrium state.

Communication

Leaving the agents without any communication wouldn't produce a very rich dynamic, as it would mostly look like an active particles bath with a few obstacles defining the city. Moreover, in reality, activists are able to communicate. We choose to start by including in our model the least they are able to communicate. One important factor for them to keep on acting is the evaluation of their own number.

Nicolas Maudet advised us to use a simple communication protocol used in agent based modelling called *push-sum* [27]. The idea is the following, we have a list of values, that will be distributed among all the agents. Those values, of course, have all together an average. The goal of our protocol is that, by exchanging information about their knowledge, all the values the agents initially had must evolve and converge toward the average. This can be summarised by the following algorithm 2.

Algorithm 2 Push-sum protocol

Require:
N players $i \in [1, N]$ on a social network ▷ The graph must be "regular"
At time $t = 0$, value $s_{0,i} \leftarrow x_i$ and weight $w_{0,i} \leftarrow 1$
$y = \frac{1}{N} \sum x_i$
Ensure: $s_{\infty,i}/w_{\infty,i} = y \ \forall \ i$ ▷ Every s_i reaches y
while $t <$ Stop **do** ▷ Takes $\mathcal{O}(N \log N)$ to converge
  for $x \in N$ **do**
    y ← random.choice(neighbours(x))
    $s_{t,y} \leftarrow s_{t,y} + s_{t,x}/2$, $w_{t,y} \leftarrow w_{t,y} + w_{t,x}/2$
    $s_{t,x} \leftarrow s_{t,x}/2$, $w_{t,x} \leftarrow w_{t,x}/2$
  end for
end while

What does that means in our study ? If we declare that all the agent start at a certain position, and evaluate their surroundings, they then initiate their estimation of the number of agent in their area, based on the density of agents they see and on the surface of the area they act upon. If initially, the agents are uniformly distributed on the network, then the more agent you have and the greater the initial average converge to the real number of agents there are. However, at the beginning, most agents have in incorrect evaluation of that number. Still, during the simulations, while agents move, meet each others and execute the push sum protocol, we see that their evaluation of their number converge correctly.

One of our goal will be to prove why the protocol still converges even if the agents are moving and therefore are not linked by a static and regular graph. More importantly, we will try to identify, under what conditions on the dynamic and on the shape of the underlying graph, our protocol converges and we will try to investigate if the complexity of the algorithm is modified.

4.2 Refinements

This model is only a very first approach to simulate the behavior of protesters. As explained, the goal here was only to reproduce some important features. But we can easily improve this model, even by using already existing models. We are thinking about using the activation behavior agents have in the riot model proposed by epstein [4]. In this model, agents move freely, and start rioting if enough people around them have decided to riot, and if they judge positive their force balance with the repressive forces putting them out of the game. This behavior of *activation* if we estimate we are enough to do something, could be added to our agents, leading to introducing a more game theorist approach of the behavior our agents will have, as they will adapt depending on the situation they face.

Another aspect of the model will be to introduce a adversarial dynamic. This is the dispersion, arrests and kettling of the protesters. Once again, we will base our approach on real world behaviors [18]. Moreover, this opponent doesn't behave in a decentralized way. Simulating it with an

agent-based model might not have any use. Also, we could be able to implement the opponent capacities from real data, such as a vision based on city camera[12].

One last important thing is to notice we can understand our complex system as a bipartite graph. A graph is bipartite when it has 2 different kind of nodes and when the nodes from one kind are linked to the nodes of the other kind. This is our case if we treat the intersections and the agents as two different kind of nodes instead of considering only the intersections as nodes. Using this kind of graph, in order to model complex systems, has been studied by the team of *complex systems.* [28]

First, all the geographic nodes linked on the city graph are linked together on the bipartite graph. Then, all agents which have a geographic node in sight, are linked to this geographic node with an edge on the bipartite graph. Now we can also do a projection of our bipartite graph in a *copresence graph*, where agent nodes are linked if they see each other. This can give a nice dynamical representation of our evolving copresence graph while protesters are moving on the city. However, applying this transformation is problematic. While it seems useful to visualize how agents are in sight of each other, and while we might think we simplify our problem by reducing the information, we are actually losing information and increasing the numbers of edges from d to $d(d-1)/2$. Instead of doing a film of the dynamic projected graph, we can keep the bipartite graph and once again think of it as a link stream. In this paradigm, we hope to we'll be able to tackle the question of how the aggregates form and under what conditions, depending on the road network properties.

5 Conclusion

5.1 Overview

Along this report, we have presented our first ideas on how tackle the question of the evaluation of the impact of new forms of protests.
We presented the three basic ingredients needed to start our survey. We introduced different classes of graph that can be used together in order to model cities, but we also explained how to use real date to perfect them. We defined metrics that we use on our graph in order to evaluate the impact of a given perturbation. An important challenge will be to extend those metrics in time. This is to export them from graph to link streams. Finally we detailed the characteristics of our agent-based model, showed what behavior emerges from it, and explained how we are going to use it to implement activists strategies.

During this internship, I learned a lot about complex networks, and dove into the literature on agent-based modelling of crowds. Also, this survey was a great opportunity to discover the unlimited pleasures of real data analysis. Among many funny examples, I struggled a little time to extract and use some attributes such as the number of lanes ... before discovering it was given as a string while, naively, I thought it would be an integer.
In a sense, the most important thing I achieved along those three months was to obtain a good appreciation of the state of the art regarding my main topics : city networks and pedestrian simulations. Through reading the literature, but primarily thanks to my supervisors and also with important meetings such as the France Regional Conference on Complex Systems.

5.2 Complements

As this report is also meant to be a snapshot of what I have done, I found important to detail the collaborations and meetings I have been involved in the past three months, as long as those had an impact on the ideas and decisions relative to this work or my future career.
In this section, I will briefly describe major events, mainly related to complex systems, in which I have taken part. In a sense, I will use this section to describe my feelings and appreciation of my internship.

[12] Map of video surveillance in Paris from data.gouv

On interdisciplinarity

Nous ne sommes pas des sociologues et nous ne prétendons pas l'être. Il est nécessaire de souligner cet écueil, car les outils mathématiques et numériques que nous sommes habitués à utiliser peuvent paraitre supérieurs aux approches de collectes statistiques voir qualitatives qui sont de coutumes au sein des sciences sociales.

Cependant la simple discussion avec des professionnels du domaine (sociologues, anthropologues, économistes) permet de garder les pieds sur terre. Si leurs champs de recherche se sont développés de la sorte, c'est que les analyses qui prétendent comprendre de manière analytique les comportements humains restent peu fructueuses et nécessitent des modèles très complexes pour être analysées. L'approche des sciences sociales est plutôt celle de la compréhension des données à travers des théories, non pas formelles, mais qui permettent d'interpréter les données de terrain, elles même souvent recueillies par des entretiens longs, et qui décrivent des intéractions de manière inévitablement plus riches et complexes que des données binaires [29].

Les modèles multi-agents produisent des résultats exploitables, et les approches de théorie des graphes permettent de modéliser des dynamiques aux seins de communautés sociales. Il ne faut cependant surtout pas croire que ces phénomènes étaient invisibles, pour la plupart, aux yeux des sciences sociales, et elles avaient déjà produit des analyses avec leurs outils traditionnels. Dans ce cadre, il est important de se rappeler notre rôle : celui d'une approche complémentaire, qui exploitent ses outils pour traiter des quantités importantes de données, voir de les produire à travers des modèles dynamiques, et donc de proposer des intuitions complémentaires sur les dynamiques sociales aux études en sciences humaines. Une approche qui permet notamment de faire le pont entre les interactions humaines microscopiques et les phénomènes sociétaux macroscopiques [30]. Mais il ne faut jamais céder à l'orgueil et se targuer de produire de meilleurs modèles, ou pire, d'être le renouveau de ces disciplines.

Partnerships

Avec Matthieu Latapy nous cherchons à développer des liens forts avec d'autres chercheurs pour répondre à notre question initiale.
Cela se traduit à travers des collaborations "traditionnelles", comme les discussions avec Julien Randon-Furling, riche de son expertise sur les processus stochastiques, que nous souhaitons exploiter pour modéliser les dynamiques de nos piétons.

Mais nous nous entretenons aussi avec une équipe de géographes. En effet, bien que les données géographiques soient désormais libres et riches, notamment avec OpenStreetMaps, certaines zones de flou subsistent et nécessitent une modélisation. Je pense notamment au cas des largeur de rues, qui à notre surprise sont très dures à trouver. En effet, elles nécessitent de répondre à la question d'où commence un côté et fini un autre de la rue. Ce qui est non univoque : est-ce l'espace publique ? l'espace accessible par les piétons ? par les automobilistes ? Et une fois tranchée, les données ne sont pas souvent disponibles, et il est nécessaire de faire des inférences, potentiellement basées sur le nombre de voies, de trottoirs, de place de parking de droite et de gauche, ainsi que les tailles réglementaires de ces différentes infrastuctures. Tailles "réglementaires" qui ont, bien sûr, la facheuse tendance de différer d'une rue à l'autre.

Enfin nous sommes en train de contacter différents chercheurs en philosophie et sociologie qui auraient l'habitude de travailler sur les dynamiques des communautés militantes.

Conferences

One other important aspect of my internship was to discover as much as possible the community of complex systems. Luckily, the months of May and June have been rich in meetings of those kinds, and I was lucky enough to assist to many conferences related to my subjects such as the 75th *anniversary of informatics* at LIP6 or the conference *Repenser la justice sociale en contexte d'IA* at the ENS. This matters a lot to me regarding the fact I intend to become a researcher. Besides being pleasant moments of learning new exiting things, those gatherings are a perfect opportunity to get to know

who does what research among the community, and to have a glimpse at the state of the art. On last thing I didn't expect, as people almost always ask each other what they do, I ended up having to explain in very few sentences what I was doing. Which of course was a good exercise that I am not used to.

On the 1st and 2nd of June, Jean-Philippe Bouchaud organized a cycle of conference at the Collège de France, entitled *More is different*. The still huge influence from the eponymous article by Phil Anderson [31], led to the development of the study of complex systems by physicists. Behind this approach, there is the idea, reinforced by the success of coarse grain methods such as renormalization groups, that the micro details don't all matter in the emergence of the overall complex and macroscopic behavior. This implies that simple (agent-based) model could reproduce real life complexity, even if our agents don't capture all the single details of the real life agents, as long as they reproduce their main microscopic characteristics.
Speaking of renormalization group, and among all the incredible presentations, one very impressive talk was given by Andrea Cavagna on a pre-print article [32] in which he uses this method on an agent-based model. Along with experimental data, he was able to find the critical exponents characterising the structure of a swarm of insects.
Concerning my study, applying this methodology to my agent-based model might be an efficient way to characterise very important properties such as the size or robustness of the aggregates emerging. However, calculating Feynman diagrams to find the exponents is a very time-consuming task. But this could be an opportunity to do a future partnership with class colleagues going to study RG methods.

From June 20 to 22, David Chavalarias and his team invited a large panel of our community to the French Regional Conference on Complex Systems, to discuss an enormous variety of subjects from Covid issues with Vittoria Collizza, or ecosystems with Ana Altieri, or to dynamic graphs and SocioPatterns with Alain Barrat.
Among those works, one interested me in particular. Fernando Peruani presented a synthesis of his studies on *"Imitation, democratic leadership and collective intelligence in sheep"* [33] [34]. Of course, daily human behavior has nothing to do with sheep. Still the way they behave is an excellent example of strategy solving a dilemma between moving efficiently in herd to prevent attacks on isolated targets, and exploiting efficiently their resources. Their pattern is defined by the succession of two steps. First they are steady, are eat grass, then they move led by one sheep. Both patterns emerge from three individual states. The sheep can be eating, head down. He can alert, head up, and finally moving. The last step happens when enough sheep see another fellow sheep heads up, then one initiate the move, and the others follow if they have the same evaluation that induced the leader to move.

Let's go back to our activists for a second. We see that they start by occupying a certain area. But maybe one interesting strategy to model would be that if they collectively "feel" it, they move until a new area, under the direction of one leader that would choose the path. Moreover, the idea that this leader could be chosen at random and change at every steps, sticks very well to the horizontally promoted in those groups. This corresponds in a striking way to our sheep motion.
I will try to implement this dynamic in order to see if this strategy works against the risk of dispersion that activists try to avoid in order to maximize the impact of their action.

5.3 Opening

To conclude this report, I will briefly talk about a project that could see the day in three years, as I am going to continue my work as a PhD student with Matthieu Latapy.

After having modeled our city, found good metrics to evaluate the impact of protests, and built a realistic agent based model for activists, we still need to compare the impact of different strategies. Of course, we can straight forwardly test known and already applied strategies. But this doesn't explore the full parameter space our model has access to. Of course, navigating through that space usually takes time, and one first need to reduce its dimensionality by finding what parameters matter most.
But there is another way. What about creating an interface through which people could test different strategies that would be evaluated by our metrics ?

This would be a perfect opportunity to include a dream of mine as a young researcher. I have always wanted to contribute to the link between the public and scientists, especially through citizen science. A great way to do that is with serious games which, if done in an entertaining way, can be a source of massively generated data. Famous examples, such as Foldit [35], proved the efficiency of this method. Moreover, our framework as the advantage of looking like a very appreciated type of games called tower-defense games. This would facilitate the creation of an entertaining interface that would lead to many people playing, not only for the sake of science, but also simply because it would be fun.

References

[1] Matthieu Latapy, Tiphaine Viard, and Clémence Magnien. Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, 8(1):1–29, 2018.

[2] Auguste Blanqui. *Instructions pour une prise d'armes: L'éternité par les astres, hypothèse astronomique, et autres textes.* Société encyclopédique française, 1973.

[3] Laurent Bonnasse-Gahot, Henri Berestycki, Marie-Aude Depuiset, Mirta B Gordon, Sebastian Roché, Nancy Rodriguez, and Jean-Pierre Nadal. Epidemiological modelling of the 2005 french riots: a spreading wave and the role of contagion. *Scientific reports*, 8(1):1–20, 2018.

[4] Joshua M Epstein. Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(suppl_3):7243–7250, 2002.

[5] Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

[6] Paul Erdős, Alfréd Rényi, et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[7] Marc Barthelemy. *Spatial Networks: A Complete Introduction: From Graph Theory and Statistical Physics to Real-World Applications.* Springer Nature, 2022.

[8] Jon Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 163–170, 2000.

[9] Marc Barthelemy. *The structure and dynamics of cities.* Cambridge University Press, 2016.

[10] Geoff Boeing. Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science*, 4(1):1–19, 2019.

[11] David Easley and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge university press, 2010.

[12] Clémence Magnien, Matthieu Latapy, and Jean-Loup Guillaume. Impact of random failures and attacks on poisson and power-law random networks. *ACM Computing Surveys (CSUR)*, 43(3):1–31, 2011.

[13] Indaco Biazzo, Bernardo Monechi, and Vittorio Loreto. General scores for accessibility and inequality measures in urban areas. *Royal Society open science*, 6(8):190979, 2019.

[14] Bernard Derrida, Steven A Janowsky, Joel L Lebowitz, and Eugene R Speer. Exact solution of the totally asymmetric simple exclusion process: shock profiles. *Journal of statistical physics*, 73(5):813–842, 1993.

[15] Izaak Neri, Norbert Kern, and Andrea Parmeggiani. Totally asymmetric simple exclusion process on networks. *Physical review letters*, 107(6):068702, 2011.

[16] Thibault Bonnemain, Matteo Butano, Théophile Bonnet, Iñaki Echeverría-Huarte, Antoine Seguin, Alexandre Nicolas, Cécile Appert-Rolland, and Denis Ullmo. Pedestrians in static crowds are not grains, but game players. *arXiv preprint arXiv:2201.08592*, 2022.

[17] Kardi Teknomo. Application of microscopic pedestrian simulation model. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9(1):15–27, 2006.

[18] Paul Rocher. *Gazer, mutiler, soumettre: Politique de l'arme non létale.* La fabrique éditions, 2020.

[19] Steven F Railsback and Volker Grimm. *Agent-based and individual-based modeling: a practical introduction.* Princeton university press, 2019.

[20] Stanislao Gualdi, Marco Tarzia, Francesco Zamponi, and Jean-Philippe Bouchaud. Tipping points in macroeconomic agent-based models. *Journal of Economic Dynamics and Control*, 50:29–61, 2015.

[21] Michele Ballerini, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Vivien Lecomte, Alberto Orlandi, Giorgio Parisi, Andrea Procaccini, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the national academy of sciences*, 105(4):1232–1237, 2008.

[22] Craig W Reynolds. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 25–34, 1987.

[23] Michael E Cates and Julien Tailleur. Motility-induced phase separation. *arXiv preprint arXiv:1406.3533*, 2014.

[24] Alexandre P Solon and Julien Tailleur. Flocking with discrete symmetry: The two-dimensional active ising model. *Physical Review E*, 92(4):042119, 2015.

[25] Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.

[26] Marc Barthelemy. The statistical physics of cities. *Nature Reviews Physics*, 1(6):406–415, 2019.

[27] David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-based computation of aggregate information. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 482–491. IEEE, 2003.

[28] Jean-Loup Guillaume and Matthieu Latapy. Bipartite graphs as models of complex networks. *Physica A: Statistical Mechanics and its Applications*, 371(2):795–813, 2006.

[29] Mathieu Génois, Maria Zens, Marcos Oliveira, Clemens Lechner, Johann Schaible, and Markus Strohmaier. Combining sensors and surveys to study social contexts: Case of scientific conferences. *arXiv preprint arXiv:2206.05201*, 2022.

[30] Rosaria Conte, Nigel Gilbert, Giulia Bonelli, Claudio Cioffi-Revilla, Guillaume Deffuant, Janos Kertesz, Vittorio Loreto, Suzy Moat, J-P Nadal, Anxo Sanchez, et al. Manifesto of computational social science. *The European Physical Journal Special Topics*, 214(1):325–346, 2012.

[31] Philip W Anderson. More is different: broken symmetry and the nature of the hierarchical structure of science. *Science*, 177(4047):393–396, 1972.

[32] Andrea Cavagna, Luca Di Carlo, Irene Giardina, Tomas S Grigera, Stefania Melillo, Leonardo Parisi, Giulia Pisegna, and Mattia Scandolo. Natural swarms in 3.99 dimensions. *arXiv preprint arXiv:2107.04432*, 2021.

[33] Sylvain Toulet, Jacques Gautrais, Richard Bon, and Fernando Peruani. Imitation combined with a characteristic stimulus duration results in robust collective decision-making. *PloS one*, 10(10):e0140188, 2015.

[34] Francesco Ginelli, Fernando Peruani, Marie-Helène Pillot, Hugues Chaté, Guy Theraulaz, and Richard Bon. Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proceedings of the National Academy of Sciences*, 112(41):12729–12734, 2015.

[35] Vickie Curtis. Motivation to participate in an online citizen science game: A study of foldit. *Science Communication*, 37(6):723–746, 2015.