University of Padova

Department of Physics and Astronomy "Galileo Galilei"

Master Thesis in Physics of Data

Effects of Topological Changes in the Individuals Decision-Making in relation to Transportation Methods in a City.

Supervisor
Prof. Samir Suweis
University of Padova

Co-supervisor
Dr. Andrea Guizzo
Dr. Riccardo Gallotti
Fondazione Bruno Kessler

Master Candidate
David Alejandro Altamirano-Coello

Student ID
2081255

Academic Year
2023-2024

— To my parents for their unconditional support all the days of my life

Abstract

Urban mobility plays a crucial role in the organization of a city for reducing the travel time, or in the use of sustainable transport methods to reduce the pollution. We focus here in the first aspect particularly, studying the impact of the topological effects in the individuals decision-making process of commuters using two transport modes within a city: bicycles and cars. To perform the study, we combine two models: the first is a network that simulates a city and it is characterized by the traffic flow theory in order to compute the commuting time that will serve as the fitness parameter in the replicator equation, which is our second model. This equation indicates the evolution of the modal share distribution along the time. We also compute the optimal and sub-optimal average commuting time when the system reaches the equilibrium, and the aim is to find a constraint that allows us to obtain the shortest possible time. We obtained that a network with triangular lattice, and random origins and destinations, gives the shortest time while being resistant to a greater number of commuters as well. Finally, we also found that add a small portion of parking spots decreases considerably the average commuting time.

Contents

Listing of figures

Listing of tables

Listing of acronyms

NE .... Nash Equilibrium

ESS .... Evolutionary Stable Strategies

OD .... Origin-Destination

MFD .... Macroscopic Fundamental Diagram

LoS .... Level of Services

HCM .... Highway Capacity Manual

SRI .... Speed Reduction Index

SPI .... Speed Performance Index

VoC .... Volume-over-Capacity

CT .... Commuting Time

SPT .... Shortest Path Tree

ASPL .... Average Shortest Path Length

BCT .... Baseline Commuting Time

1 Introduction

Once I have finished the bachelor on Physics I have two clear thoughts mind: the first one was I love physics but I would like to use in a field with more applications; and the second was I wish I had gotten into the study of complexity sooner. The Master in Physics of Data offered by the University of Padova matched exactly what I was looking for. My knowledge about machine learning and networks increased exponentially in one year, and I was willing to apply in research projects with impact in the society. In that sense, I found the Complex Human Behaviour Lab (CHuB) in Foundazione Bruno Kessler (FBK) that opened the doors for me to explore spatial networks describing urban infrastructure. The first time I contacted Riccardo Gallotti and Andrea Guizzo, they had a project in mind which consists on adding a degree of difficulty to a previous work related to evolutionary dynamics that explored the individuals decision-making process of commuters using various forms of transport within a city. The new degree of difficulty was using a spatial network to characterize the city and replicate the previous work. Basically, the network allows us to study how populations change when the topology and geometry of the city also changes. This has an impact on public policy-making regarding the growth of urban areas into cities and their sustainability in order to save the time that users take to reach their destinations. Definitively I wanted to work in this project.

Studying the urban mobility plays a crucial role in the organization of a city to reduce the travel time or in the use of sustainable transport methods to reduce the pollution. Cars are the transportation method in a city that pollute more in comparison with the number of beneficiaries. The pollution is not only related with the fuel consumption but also from their manufacture and the constant changes of their parts. For example, the average lifespan of tires is two years, which generates an inordinate amount of waste. Beyond the air pollution, cars also require expensive infrastructure and much space, produce congestion, noise and possibly crashes and road injuries [4]. Stimulating the use of active transport, such as walking or cycling, helps to address the previous issues [5, 6]. However, convincing drivers to switch into a different transportation method is a big challenge because the use of private cars give to them comfort and status. In fact, every time the use of cars in a city increases more and more, but instead of reducing the commuting times they increase due to the congestion [7]. Increasing the road capacity in response to the congestion can make the congestion even worse because more people will switch to using a car and the time lost increases [8].

The way how users decide about the transportation method is not only about the commuting time, but also based on other aspects such as the risk involved, accessibility, comfort of their own preferences [9]. However, the commuting time is still the key parameters [10] because users prefer to reach their destination as soon as possible. This leads to selfish behaviors where the social cost at equilibrium goes away the minimum social cost [11]. This decrease in the social efficiency is often referred to as the price of anarchy [12], which is understood as the measure between the worst possible equilibrium and the social optimum. The tragedy of the common [13] summarizes this kind of situations where the players of a population act motivated solely by their self-interest such as reducing their commuting time, but resulting in the unintended consequence of the deterioration of a wider social outcome such as the increase in the average commuting tie of the whole system. Furthermore, cooperation sometimes also appears in the cities [14].

The baseline work for this thesis developed by Prieto-Curiel [15] shows the tragedy of the commons using the population game theory in the daily urban transit of a city, where we understand the city as the game and the users or commuters as the players. Every players can choose a finite number of strategies that in this case are the commuting methods and the cost or reward of the game is expressed as the travel time. To describe the decision-making process among the strategies, the replicator equation is used benefiting the transport method that display a smaller commuting time in comparison with the average commuting time of the whole system. In other words, the smaller travel time the better. From a initial random distribution of the modal share, a stable and new modal share is reached corresponding to an equilibrium (Nash equilibrium) in the model which is the sub-optimal one. On the other hand, the optimal equilibrium is the ideal scenario where the average commuting time is the smallest possible. There is only one case where the optimal and sub-optimal equilibria are the same and it is when the costs related to the congestion and the number of users are equal for all transport modes. In general, the purpose of users when switching between different transport methods is to reduce their individual commuting time, however, sometimes it causes the contrary effect increasing the average commuting time of the system. A clear example are the cars because they are, in principle, the faster method, moreover provide the freedom to commute when and where the user want, but imagine an hypothetical case where all the users have decided to use a car. Then, the congestion increases making the commuting time higher and the method useless. This is described as the *tragedy of the commons*. The paper also explore the idea of introducing a new transport mode with a small baseline commuting time but has a high impact in the other users, thus the congestion level will increase and also the average commuting time. In conclusion, the preferred method will be the one who present less costs related to the congestion.

The aim of this work is to study the effect in the congestion costs when we modify the topology and geometry of the spatial network. Basically, the methodology to develop the task is building a model based on a spatial network that simulates a city. Next, we characterize the network with the traffic flow theory and extract physical parameters such as the commuting time. It will be the input in the replicator equation that study the dynamics. Once the model is built, we could modify the network and explore the effect of the topology in the decision-making process. We will use two networks which are differentiated by the degree, one network will have a triangular lattice while the other will have a quadratic lattice. We will also propose two algorithms to generate the origins and destinations within the city. The first approach is randomly and the second is considering the existence of busier areas known as hotspots. Finally, we are going to introduce the idea of parking spots and analyze how affects in the dynamics.

The following document is organized as follow. In Chapter 2 we will discuss the background required to understand the first concepts presented in this work: game theory, population games, evolutionary dynamics and Nash equilibrium in all the cases.

In Chapter 3 we will explain the second necessary concepts required to understand this work: network theory and spatial networks. Also the shortest paths to visit all the points in the network. Finally, we will also introduce the traffic flow theory to characterize the network and understand how the congestion can be introduced in the network.

The Chapter 4 is devoted the complete model where both the population game and the network were combined describing in detail the methodology. We describe in detail the work developed by Prieto-Curiel. And latter we explain the developed algorithms based in the network and the traffic flow theory for computing the commuting times. Finally, we show how combine both approaches.

Chapter 5 summarizes the results. We will start with the first result using a triangular-lattice network and random origins and destinations. Then, we will explore what happen if we increases the total number of users. Next we modify the topology of the network reducing the degree and using quadratic-lattice network. Then, we generate the paths assuming hotspots in the network which represent the busiest areas in the city. Finally, we introduce the idea of parking spots and measure the impact of reducing some of them.

Finally, in Chapter 6, we will discuss some conclusions about the results and what are the possible implications of this results. Future work related with the results obtained in this thesis will also be explained in this chapter.

2
Population Games Theory

This chapter will summarize the mathematical tools required to understand the concepts of game theory. We will start explaining the classical game theory, passing through fundamental concepts as Nash equilibrium (Section 2.1). Section 2.2 explains the game theory applied on a large amount of agents. Finally, Section 2.3 explores the idea of game theory with large population that evolves in time due to the Deterministic Evolutionary Dynamics.

2.1 Classical Game Theory

Game theory is the mathematical study of multiple players making-decision processes where the final output is not a direct consequence of one's own choices but is influenced by the choices made by other people [16].

The more common way to represent a game theory is by to the strategic also called normal form. For this, we need 3 key concepts: the *player*, $i \in N$, which is an individual or entity that participates in the game and have their own preferences, the *strategy*, $s_i \in S_i$, which is the set of choices that a player can take, and the *payoff* also called *utility*, $u_i = u_i(s_1, s_2, \ldots, s_N) \in \mathbf{R}$, which is the outcome or the reward that the player receives based on the strategies. So the normal form of the game is specified by

$$G = \{S_1, \ldots, S_N; u_1, \ldots, u_N\}, \tag{2.1}$$

and the main feature is that the players have to do simultaneous moves. It does not necessarily mean that it has to happen at the same time, but that players choose without knowing the choices of others. In this sense, the previous knowledge of every player is common, everyone know about all possible actions of all players, all possible outcomes resulting from every previous actions and that player are rational. A rational player means that they always maximize their utility and choose the strategy which provides the best reward, as well as that the other players will maximize their payoff.

The game representation of the normal form generate a N-dimensional space where all the possible game resolutions can be shown. Each player is represented as a dimension where each dimension is split in s_i sections. The intersection of these sections in the N-dimensional space creates a grid which summarize all the outcomes that the players can obtain. For instance, if we assume 2-dimensional space (2 players), the grid is understood as a bi-matrix known as the *payoff matrix*. For a better understanding we are going to explore an example and introducing gradually to the core of the present work.

Example 1. Consider the following problem: Ana and John are working in a project separately and need to download a large dataset from internet. They are sharing the network connection and have two options: limit its bandwidth or don't do it. If both limit their bandwidth the network will not have a high congestion and they can download the dataset in 2 hours. If both do not limit the congestion level of the network is higher and they will need 5 hours to download the dataset. If only one of them limit the network, the one with limited network will need 6 hours while the other just 1 hour to download the dataset. This game is represented in Table 2.1, The previous is an example of a symmetric game where each player faces the same set

		John	
		Limited	Unlimited
Ana	Limited	2h, 2h	6h, 1h
	Unlimited	1h, 6h	5h, 5h

Table 2.1: Example of a game in normal form

of strategies choices, and the resulting payoffs are the same regardless of the player's identity.

The payoff matrix allows to visualize in a easier way which is the **dominant strategy**. It is the strategy that yields the highest payoff for a player, regardless of the strategies chosen by the other players. On the other hand, a **dominated strategy** is a strategy that results in a lower payoff for a player compared to at least one other available strategy. In the previous example, since the lower time the better strategy, taking into account to player *Ana*, $u_A(L, L) = 2\text{h} < u_A(L, U) = 6\text{h}$

and $u_A(U, L) = 1\text{h} < u_A(L, L) = 5\text{h}$, we realize that the limited choice, $u_A(L, x)$, is the dominant strategy, while the unlimited choice, $u_A(U, x)$, is the dominated strategy. Analyzing to player *John* we also realize that car choice, $u_J(x, L)$, is the dominant strategy. Since it is a symmetric game, it makes sense that the dominant strategy is the same for both players. When all strictly dominated strategies are excluded and only one strategy pair remains, the strategy pair (limited, limited) is called **dominant strategy equilibrium**. However, to find a strictly dominant strategy pair as in the previous example is very rare in real life because there are strategies that dominate some and are dominated by others. To solve that issue, Nash Equilibrium is introduced.

2.1.1 Nash Equilibrium

Nash Equilibrium (NE) is the most important notion about game theory and the study of decision-making because describes the set of strategies that most benefits to the whole system. It makes use of the **best response** idea which refers to the strategy that yields the most favorable outcome for a player, considering other players' strategies as given.

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i}) \quad \forall s_i' \in S_i \,. \tag{2.2}$$

Then, **pure NE** is the point at which each player has chosen the best response to the strategies of the other players. Moreover could exist **mixed NE** points which are stochastic and consist on strategies that could be played with a given probability. As the probabilities are in a continuous space, there are an infinite number of mixed strategies.

For describing NE in the formal way, let's consider a game with N number of players, for each player i exist a set of S_i available strategies, and a payoff function u_i. A strategy is a NE $(s_1^*, ..., s_n^*)$ if, for all i, s_i^* is a best response to $s_{-i}^* = (s_1^*, ..., s_{i-1}^*, s_{i+1}^*, ..., s_N^*)$,

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \quad \forall\, s_i \in S_i \,, \tag{2.3}$$

or, equivalently,

$$s_i^* = \arg\max_{s_i \in S_i} u_i(s_i, s_{-i}^*) \,. \tag{2.4}$$

2.2 Population Games

The previous idea has been explored in many other fields, for example, studying large population of agents, where the Nash equilibrium can be also determined. Population games [17, 18] study strategic interactions among large number of agents, where each agent is identical, small and anonymous. It means that agents have same choices, make decisions independently, every choice has only a minor or not impact on the other agents' payoff, and the payoffs depend on the distribution of others' strategies. Typically, two additional restrictions are imposed on the agents to robust the focus: the number of strategies is finite and have identical payoff functions for each strategy. The simplicity of population games offer a powerful tool which can be applied in many fields, such as economics, biology, computer science, and transportation science.

Let's describe in a formal way how population games is defined. Let $\mathcal{P} = \{1, ..., p\}$ be a *society* consisting of $p \geq 1$ *populations* of *agents*. The available set of *strategies* in population p is denoted $S^p = \{1, ..., n^p\}$ and the total number of pure strategies in all populations is given by $n = \sum_{p \in \mathcal{P}} n^p$. When the game runs, each agent selects a (pure) strategy from S^p. The set of *population state* or *strategy distributions* for population p is $X^p = \{x^p \in \mathbb{R}^{n^p}_+ : \sum_{i \in S^p} x^p_i = m^p\}$, where m^p is the *population mass* and $x^p_i \in \mathbf{R}_+$ is the mass of players in population p choosing strategy $i \in S^p$.

Elements of $X = \prod_{p \in \mathcal{P}} X^p = \{x = (x^1, ..., x^p) \in \mathbf{R}^n_+ : x^p \in X^p\}$, the set of *social states*, describe behavior in all p populations at once. The elements of $X_v = \prod_{p \in \mathcal{P}} X^p_v$ are the vertices of X and are called the *pure social states*. When there is just one population ($p = 1$), the mass is 1 and the notation is simplified: the strategy set is $S = \{1, ..., n\}$, the state space is $X = \{x \in \mathbf{R}^n_+ : \sum_{i \in S} x_i = 1\}$, and the set of pure states $X_v = \{e_i : i \in S\}$ is the standard basis vectors in $\mathbf{R}^n$.

Taking the sets of populations and strategies fixed, the *payoff function* $F : X \to \mathbf{R}^n$ is a continuous map that assigns a vector of payoffs to each strategy in each population. The (*weighted*) *average payoff* obtained by members of population p at social state x is defined as,

$$\bar{F}(x) = \frac{1}{m^p} \sum_{i \in S^p} x^p_i F^p_i(x). \tag{2.5}$$

To describe the optimal behavior, as in the classical game theory, we should define how best response and, by consequence, Nash equilibrium is defined. Define *pure best response correspon-*

dence, $B^p : X \Longrightarrow S^p$, which specifies the optimal strategies S^p for each social state x,

$$B^p(x) = \arg\max_{i \in S^p} F_i^p(x), \tag{2.6}$$

and, the *mixed best response correspondence* for population p, $b^p : X \Longrightarrow \Delta^p$, is given by

$$b^p(x) = y^p \in \Delta^p : y_i^p > 0 \Longrightarrow i \in b^p(x). \tag{2.7}$$

Social state $x \in X$ is a *Nash equilibrium* of the game F if every strategy earns the maximal payoff in each population,

$$NE(F) = \{x \in X : [x_i^p > 0 \Longrightarrow F_i^p(x) \geq F_j^p(x)] \text{ for all } i, j \in S^p, p \in \mathcal{P}\}. \tag{2.8}$$

In population games always will exist at least one Nash equilibrium. In order to visualize better the concept of population games, we will a use an example described by Novak in [19].

Example 2. Imagine two phenotypes, A and B, where A can move while B cannot. A pays a certain cost for the ability to move, but also gains the associated advantage. Suppose the cost-benefit analysis leads to a fitness of 1.1 for A compared to a fitness of 1 for B. In this setting, fitness is constant, and A will certainly outcompete B. But imagine that the advantage of being able to move is larger when few others are on the road, but diminishes as the highways get blocked up. In this case, the fitness of A is not constant, but is a declining function of the frequency of A. A has a higher fitness than B when A is rare, but has a lower fitness than B when A is common.

2.3 Deterministic Evolutionary Dynamics

Once strategic interaction of large number of agents was modelling, one would like to forecast how these agents will behave during a repeated interaction. John Maynard Smith and George R. Price introduce the *Evolutionary Game Theory* [20] improving the classical game theory with the concepts of Darwinian natural selection, modelling an evolving population of competing strategies. This theory points out that the previous local stability analysis is just the first step and when we consider a large number of agents, the equilibrium seems most difficult to achieve.

In this context, there are two basic assumptions for reinforce the model. The first assumption, *inertia*, assumes that agents occasionally can switch their strategies in order to receive a better payoff and the second assumption, *myopia*, assumes the behavior of agents is specified by the information available at the current strategic environment to make the decisions. These two assumptions are summarized in the mechanism called *revision protocol* which describes when and how agents decide to switch strategies based on the available information at the moment. Revision protocols can be imitation, optimization, or any criterion that agents can employ to react to current conditions.

When we put all together, a population game and a revision protocol, a differential equation is generated. It describes the evolution of aggregate behavior when the game is played recurrently and the revision protocol is employed. The generated process is a Markov process on the population states, namely a stochastic process [21] describing the sequence of possible strategies where the probability of each strategy depends only on the previous pure strategy.

2.3.1 Replicator Dynamics

A revision protocol ρ, describes when and how agents update their strategies along time, it is defined as a map that assigns each population game F a function $\rho^F : X \to \mathbb{R}_+^{n\times n}$, which maps population states $x \in X$ to collections of *conditional switch rates*, $\rho_{ij}^F(x)$, from strategy $i \in S^p$ to strategy $j \in S^p$, $\forall p \in \mathcal{P}$. All together define a stochastic evolutionary process that runs on the discrete grid $\mathfrak{X}^N = X \cap \frac{1}{N}\mathbb{Z}^n = \{x \in X : Nx \in \mathbb{Z}^n\}$.

Let's follow the analysis showed by Ohtsuki and Novak [22] on different population games to understand how the revision protocol works. Consider a population game with finite n strategies, $(s_1, \ldots, s_n)$. The payoff matrix A, is a $n \times n$ matrix where each entry, $a_{ij} = \pi(s_i, s_j)$, is the payoff for the individual playing strategy i against strategy j. The relative abundance, also called *frequency*, of each strategy is given by x_i, where $\sum_{i=1}^n x_i = 1$. Imagine we have a population of two strategies: the first strategy V with a frequency x_V and the second strategy W with a frequency x_W. The population is fixed, so $x_V + x_W = 1$ and we can express on variable in terms of the other, $x_V = 1 - x_W$. For representing the current population we define the vector $\vec{x} = (x_V, x_W)$. In general, the fitness of each strategy i is given by

$$f_i(\vec{x}) = \sum_{j=1}^{n} x_j\, a_{ij}\,, \tag{2.9}$$

and the average fitness of the population is described as

$$\varphi(\vec{x}) = \sum_{i=1}^{n} x_i f_i\,. \tag{2.10}$$

Finally, we can express the differential equation which defines the changes in the frequency of every strategy in the next generation, known as the *replication equation* introduced by Taylor and Jonker [23]:

$$\dot{x}_i = x_i \left(f_i(\vec{x}) - \varphi(\vec{x}) \right) . \tag{2.11}$$

The meaning of equation (2.11) is that the fitter a strategy is at any given moment, the more likely it will be used in the future. The equation, as proved by Hofbauer [24], satisfies:

1. a stable rest point is a Nash Equilibrium,
2. a convergent trajectory in the interior of the strategy space evolves to a Nash Equilibrium,
3. a strict Nash Equilibrium is locally asymptotically stable.

The replicator dynamics was originally studied in symmetric games with finite number of strategies, then extended to other games and their dynamics summarized in [25].

Now we focus on understand some properties of evolutionary dynamics including the convergence to and stability of the *Evolutionary Stable Strategies* (*ESS*) and the relation with the Nash Equilibrium, exploring game with more than two strategies. Given the payoff $\pi(s_i, s_j)$ for a strategy s_i versus s_j, we have that a strategy s_k is a strict NE if and only if $\pi(s_k, s_k) > \pi(s_i, s_k) \ \ \forall i \neq k$ or, s_k is a Nash equilibrium if $\pi(s_k, s_k) \geq \pi(s_i, s_k) \ \ \forall i$. Furthermore, we define s_k as an evolutionary stable strategy *ESS* if satisfies these two conditions:

$$\pi(s_k, s_k) \geq \pi(s_i, s_k) \quad \text{(NE condition)} \,, \tag{2.12}$$

and

$$\pi(s_k, s_i) > \pi(s_i, s_i) \quad \text{if} \quad \pi(s_i, s_k) = \pi(s_k, s_k) \quad \text{(stability condition)} \,. \tag{2.13}$$

We can observe how all strategies that satisfies strict NE condition, also satisfies ESS condition. But not true for simple NE. That means that strict NE implies ESS that implies simple NE but not vice-versa.

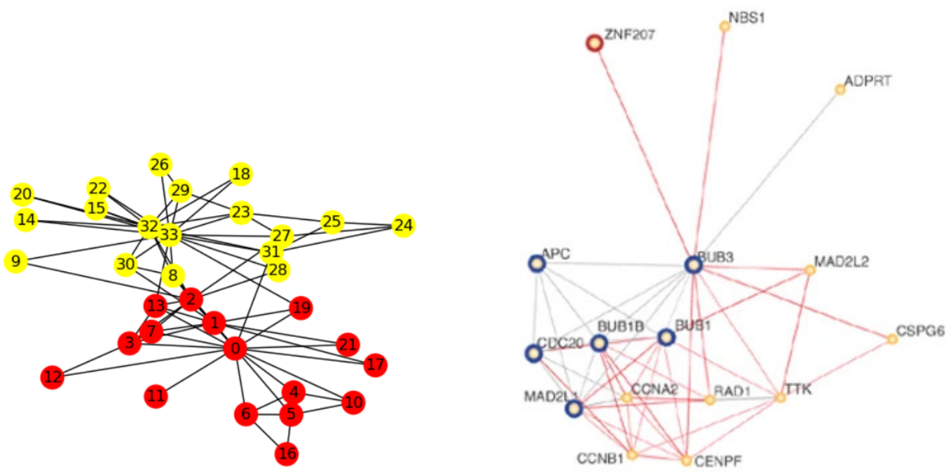
3

Networks Theory

In this chapter we introduce a basic concepts of networks and mathematical tools for network analysis. A network, in its basic definition, is a collection of nodes connected with other nodes by edges. A network analysis approach in widely used in different research field, such as statistical physics, social physics, web and many others. In this chapter we display a brief summary about networks, Section 3.1. We will do special emphasis on spatial networks in Section 3.2. Then, we describe the notion of connectivity through the shortest paths and its importance in Section 3.3. Finally, in Section 3.4, we will explore the traffic flow theory in order to give a sort of physical meaning to the paths in spatial networks and be able to study urban systems, being the cornerstone of the current thesis.

3.1 Classical Networks

The idea behind networks is simple but very powerful: we can represent any kind of relation due to graphs. A graph, $G(V, E)$, is a compilation of vertices which represent entities, and edges which represent the relation between those entities. For representing a graph we use an adjacency matrix A which has the form $N \times N$, where N is the number of nodes. For unweighted networks, the elements, a_{ij}, of the matrix has value 1 if there is a link between the vertices i and j, and $a_{ij} = 0$ otherwise. In Figure 3.1 a couple of examples are displayed. The first network represents the human friendship ties in a Karate club between 34 members studied by Zachary [1] at a USA university in 1977. The network shows a division in the club when the



(a) Zachary club Network

(b) Protein-protein interaction Network

Figure 3.1: Two example of real networks. The first one, (a), is a representation of the social ties in a karate club [1]. The second one, (b), represent a zoomed view of the human protein-protein interaction network [2].

president and the instructor decided to split the business in two. Some of the students went with the instructor and some other with the president. The second network is a small part of the human protein-protein interaction network which has $14,709$ proteins acting as the nodes and $117,268$ interactions which are the edges. The amount of nodes makes impossible to visualize in an understandable way the complete network, this piece was taken from the study of Rhodes [2]. Understanding the most probable interactions allow biologists to tackle problems as mutations, or drug development.

Now we introduce some notions and descriptors. Firstly, we can set the nature of the edge. They can be **directed** or **undirected** depending if there are a directionality in the relation between connected nodes. Take, for instance, social networks: Facebook could be represented by an undirected network because all the people you follow, also follows you. On the other hand, Instagram shows a broken symmetry and can be represented by an directed network because the people you follow can follow you back or simply do not. In directed graphs we can differentiate about outgoing or incoming edges depending on the direction. Another way to classify edges is between weighted and unweigthed. The previous examples can be easily understood as unweighted networks because you follow or you don't. However, the network of flights will need a weight in each edge representing the distance between geographical locations.

It can be extended to weighted networks with $a_{ij} = w_{ij}$ where the weight w_{ij} denotes the flux of any quantity on the link (i, j). Note that, in the case of directed networks, the matrix A is not strictly symmetric, on the other hand for undirected network, the adjacency matrix will be symmetric.

To describe the topology of the networks we have several metrics; the most used one is the mean degree. In an undirected network, the degree of a node is defined as the number od connections they have. Using the adjacency matrix the degree of a node i could be computed as,

$$k_i = \sum_{j=1}^{N} A_{ij}. \tag{3.1}$$

Then, the mean degree of the network is given by,

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i. \tag{3.2}$$

That measure is a characterizing factor of the topology and indicates the sparsity of the network. However, it does not provide any information about how the edges are distributed among the nodes.

In the case of directed network, we can define the *in-degree* of a node as the number of edges incoming, and *out-degree* is the number of edges outgoing from the node. Then, the degree of a node is given by the sum of the in-degree and out-degree of the studied node. If the node has in-degree 0 (i.e., no incoming edges), we refer to that node as a *source*, and if a node has out-degree 0 (i.e., no outgoing edges) we refer to it as a *sink*.

3.1.1 Random networks

In the decade of 1950s Paul Erdös and Alfred Rényi [26] proposed a way to generate networks where each node can be connected uniformly with any other node. Thus, the average number of edges in that network will be $\binom{n}{2}p$ and the probability that a given node has a given degree k is binomial,

$$P(X = k) = \binom{n-1}{k} p^k \,(1-p)\,(n-1-k). \tag{3.3}$$

If the number of nodes is large enough and the factor np is constant, the above equation could be approximated using a Poisson distribution,

$$P(X = k) = \propto \frac{(np)^k \, e^{-np}}{k!}, \tag{3.4}$$

which give us some versatility when we try to model complex scenarios. A interesting feature of this kind of network is that there no have any degree correlation, namely the degree of a node is independent of the degree of his neighbours.

3.1.2 Scale-free networks

To complete the brief overview of networks we must talk about one of the degree distribution models, the scale-free graphs proposed by Barabási and Albert in 1999 [27]. In this case the degree k, of the nodes has large fluctuations and has to be drawn from a power law distribution,

$$P(X = k) \propto k^{-\gamma}. \tag{3.5}$$

It has been observed in many real-world networks such as: the internet, collaborations, metabolic reactions, etc. The particularity of this kind of networks is that $\gamma \in (2, 3]$, producing that the degree distribution has a long tail which mean that the distribution is very heterogeneous: many nodes with few connections and a few nodes with very high degrees, known as *hubs*. To explain how this kind of networks growths, the widely accepted procedure is the *preferential attachment*. At every time step, some nodes are introduced to the systems will be attached preferentially to those nodes that have large degrees already. That is known as the "rich-gets-richer" phenomenon and the probability to attach a link to an existing node x is,

$$\Pi_x = \frac{k_x}{\sum_{i=1} k_i}, \tag{3.6}$$

from where we can appreciate how there exist a correlation between the degree of one node and the degree of the neighbours. That fact has a huge impact in the dynamics that run over the network.

3.2 Spatial Networks

The study of spatial networks was developed precisely for geographers who studied the structure and evolution of transportation systems [28, 29, 30]. Along the time it also have been used in other fields such as telecommunication [31], epidemics [32], power grids [33], botanic [34]. Essentially, the necessity to study spatial networks arises from the fact that understanding only the topological features was not enough to explain the whole behavior of networks, it is necessary to introduce the geometry as well [35]. In summary, a spatial network is a network embedded in space allowing to understand both topological and spatial aspects of large networks.

A spatial network is defined as a graph $G = (V, L)$ with the combination of a set V of N nodes and a set L of E links connecting the nodes. If we understand a city as a spatial network, the nodes will be the intersections and the links represent streets between these intersections. For introducing the spatial features, the network can be embedded in a plane called *embedding* which should be specified, usually described with the position of the nodes by a list $X = \{x_i\}$. These two elements allow us to characterize the network and extract coarse-grained information. One property of spatial networks is *planarity* namely, it can be represented in two-dimensional plane without any edge crossing. To prove if a graph is planar we can start with Euler's formula who showed that a finite connected planar graph satisfies the following formula

$$N - E + F = 2\,, \tag{3.7}$$

where N is the number of nodes, E the number of edges, and F the number of faces. Euler's formula allows to define a bound for the average degree $\langle k\rangle$ for any finite connected planar graph

$$\langle k\rangle = \frac{2E}{N} \leq 6 - \frac{12}{N}\,, \tag{3.8}$$

which is therefore always smaller than 6. So, a graph cannot be considered planar if has an average degree $\langle k\rangle > 6$.

The concept of directed networks, also called *digraph*, plays a key role since streets could be one-way or two-ways, namely, there is a flow indicating a preferred direction. The case of two-ways streets can be understood as an undirected edge because is the same goes from node i to node j than goes to node j to node i. When a network have directed and undirected edges is called mixed network. For digraphs, it is necessary to define an incoming degree $k_{in}(i)$ of a node i as the number of incoming links or arcs, as well as an outgoing degree $k_{out}(i)$ which is the number of outgoing links. We have the following trivial relation between the average quantities

$$\langle k_{in}\rangle = \langle k_{out}\rangle. \tag{3.9}$$

3.3 Shortest Paths in Networks

The notion of path $\mathcal{P}_{i,j}$ allows us to move in the network, going from a node i to another node j following the available connections given by the arcs. A graph is called *connected* if there exists a path connecting any two nodes in the graph. Usually cities are a connected network. Formally, in an undirected graph, a path is defined as a sequence of vertices $\mathcal{P} = (v_1, v_2, ..., v_n) \in V \times V \times ... \times V$ of length $n - 1$ such that v_i is adjacent to v_{i+1} for $1 \leq i < n$. The connecting edge between node v_i and v_j is defined as $e_{i,j}$ and the set of edges along the path is defined as $K\{e_{i,j}\}$. For the purpose of the present work we should explore the concept of shortest paths. Basically, a shortest path is the path with the least weight. Given a real-valued weight function $f : K \to \mathbb{R}$, the shortest path from v_1 to v_n is the path $\mathcal{P} = (v_1, v_2, ..., v_n)$ that over all possible n minimizes the sum,

$$\mathcal{P} = \min \sum_{i=1}^{n-1} f(e_{i,i+1}) \,. \tag{3.10}$$

If the network is unweighted or with the same weight in each edge, the shortest path is the one with fewest edges. In this sense, there may be more than one shortest path. Hillebrand et. al. [36] investigated the effect of a small perturbation (flip only one edge) in the shortest path on a directed two-dimensional square lattice, showing a power-law behaviour for the probability distribution of the differences in shortest path lengths. In the second half of the last century, some scientist turned their research to, proposing some algorithms to solve the problem [37], of which the Dijkstra's algorithm stands out.

3.3.1 Origin-Destination Matrix

An origin-destination (OD) matrix [38] $\mathbf{T}$, is two dimensional array of M rows (origins) and N columns (destinations), where each element t_{ij} represents the number of trips from location i to location j in whatever temporal resolution needed. Usually the number of origins and destination are identical, $M = N$, giving a square matrix. The diagonal represents people or goods that do not commute large distances and it will depend on the zone definition. Then, the matrix is defined as $\mathbf{T} := (T_{ij})_{1 \leq i,j \leq L}$, where L is the number of origins (or destinations), i.e,

$$\mathbf{T} = \begin{pmatrix} 0 & T_{12} & T_{13} & \cdots & T_{1L} \\ T_{21} & 0 & T_{23} & \cdots & T_{2L} \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ T_{L1} & T_{L2} & T_{L3} & \cdots & 0 \end{pmatrix} , \tag{3.11}$$

In general, OD matrix plays a key role when we analyze the motion of people along networks such as sidewalk networks [39], or goods across spatial considerations [40].

3.4 Traffic Flow Theory

The last theoretical aspect and the most relevant for the model, explained in the next section, is to understand how the traffic flows in a city. We start with the Macroscopic Fundamental Diagram (MFD) of traffic flow which provide the relation between road traffic flux (vehicles/hour), the traffic density (vehicles/km), and the instantaneous speed (km/h). It is a macroscopic approach given by the follow relation,

$$q = kv\,, \tag{3.12}$$

where q represents the flow rate or volume, k is the traffic density, and v is the mean speed. This relation will depend on different features of the road (number and/or width of the lanes, type/material), the composition of the flow (number and type of vehicles, experienced drivers, number of users), external conditions (weather), political regulations, etc. Figure 3.2 displays the graphical relation between the components with the simplest model proposed by Greenshields proposed in 1947 [41]. Some remarks about the MFD model are: the higher density (more vehicles) the slower velocity. To prevent congestion, the number of vehicles entering in the studied zone during a time slot should be smaller or equal to the number of vehicles leaving the same zone. There are special quantities such as the $free - flow\ speed$, v_0, which is the speed of a vehicle without any other vehicle in the road, namely when $k = 0$ and $q = 0$. Also the critical quantities q_c, k_c and v_c where the flow changes from stable to unstable. Last important quantity is the *jam density*, k_j, where the system collapses and vehicles in the road cannot move, namely $v = 0$ and $q = 0$. These quantities are not easy to measure in a real city but with a big effort can be done, Geroliminis [42] collected the traffic flow data using 500 fixed sensors and 140 mobiles ones in Yokohama, Japan, and revealed that on a large urban area the macroscopic fundamental diagram exist. The importance of the fundamental diagram lies in the fact that allow us to compute the road capacity, the travel time, the *Level of Service* (LoS) and other parameters for understanding the congestion in the city. Afrin and Yodo [3] developed an survey summarizing the currently available congestion measures, a discussion about the advantages and the importance of understanding the traffic congestion.

The model proposed by Greenshields for the traffic flow theory is based on the simple assumption that speed decreases linearly with respect to the density and it is an uninterrupted traffic (without traffic signals). As the density of vehicles increases the speed decreases till density reaches the maximum referred as the jam density, k_j, where the speed becomes zero and vehicles are stuck. The flow-density diagram, Figure 3.2.a, is used to determine the traffic state

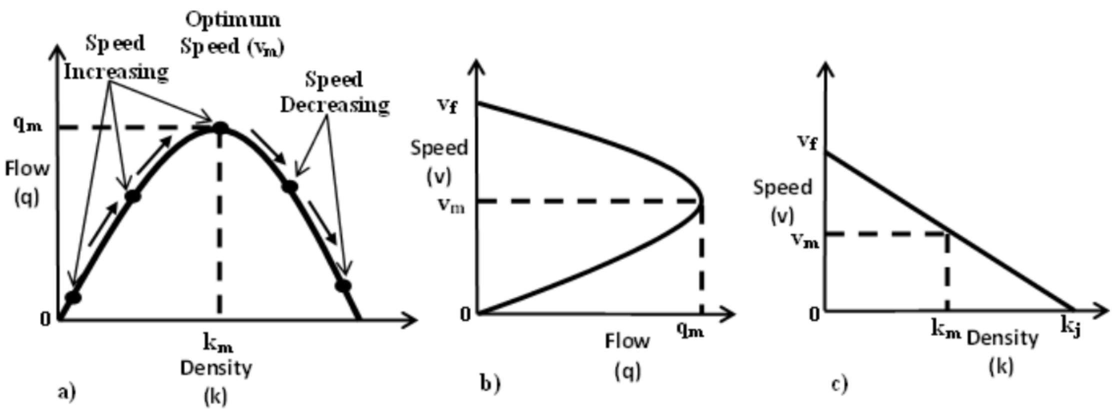


Figure 3.2: Greenshield's Fundamental Diagrams. a) Flow vs Density, b) Speed vs Flow, c) Speed vs Density.

where initially, the flow increases linearly with the density as well as the speed until reach an optimal point, k_m and q_m, after that a congested branch appears and implies that the higher the density the lower the flow. Therefore, even though there are more cars on the road, the number of cars passing a single point is less than if there were fewer cars on the road. Zaidi et. al. [43] uses the model and real data to present the Co-operative Detection And Correction (C-DAC) scheme to communicate each other of their position and speed to give them a sense of traffic around them. The Greenshields' model is given as,

$$v(k) = v_0 \left(1 - \frac{k}{k_j}\right) , \tag{3.13}$$

where v_0 is the free flow speed and k_j is the jam density. From Equation 3.12 and Equation 3.13 the relationship between flow and density can be found to be:

$$q(k) = kv_0 \left(1 - \frac{k}{k_j}\right) . \tag{3.14}$$

Doing an analysis about the stability we can compute the critical values for q, k and v from the condition that at the capacity point $\frac{dq}{dk} = 0$:

$$\frac{dq}{dk} = v_0 - \frac{2k^*}{k_j} v_0 = 0 \Longrightarrow k^* = \frac{k_j}{2} , \tag{3.15}$$

and thus, substituting into Equation 3.13 we obtain:

$$v^* = v_0 \left(1 - \frac{k^*}{k_j}\right) = \frac{v_0}{2}\,, \tag{3.16}$$

and finally, the critical flow is,

$$q^* = k^* v^* = \frac{1}{4} v_0 k_j\,. \tag{3.17}$$

The problem is widely studied in literature at mesoscopic and microscopic scales by Helbing [44].

3.4.1 Congestion

For the present work, describing the collective vehicle dynamics in terms of the spatial density per lane and the average velocity due to the MFD is a good start, and leads us to characterize the congestion in the city. Along the time, since the 1950 or even a bit earlier, many researchers have turned to understand, characterize and generalize the traffic intra and inter cities. For example, the Highway Capacity Manual (HCM) [45] provides a collection of techniques for estimating the capacity and determining the LoS for transportation facilities for transit, bicycles, and pedestrians. The LoS is a quality measure describing operational conditions within a traffic stream, generally in terms of such service measures as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience. On the other hand, capacity estimates the maximum number of persons or vehicles that can be accommodated with reasonable safety in a specified space region during a specified time period. Another definition for capacity is given by Othayoth et al. [46] and it is understood as the number of persons or vehicles that can be expected to traverse a point or a uniform section of a lane during a given time period under prevailing roadway, traffic, and control conditions. As we can appreciate in the Cost Benefit Analysis Manual [47], the capacity depends on the road material, the number of lanes in the road, the interval time, etc.

Let's focus back on the congestion to understand how these values are related. Numerous measures have been developed considering different performance criteria, categorizing the congestion measures into five categories (see Figure 3.3):

1. speed: there is congestion if the free-flow speed decreases,
2. travel time: there is congestion if the travel time increases,
3. delay: related to the rate of time loss,

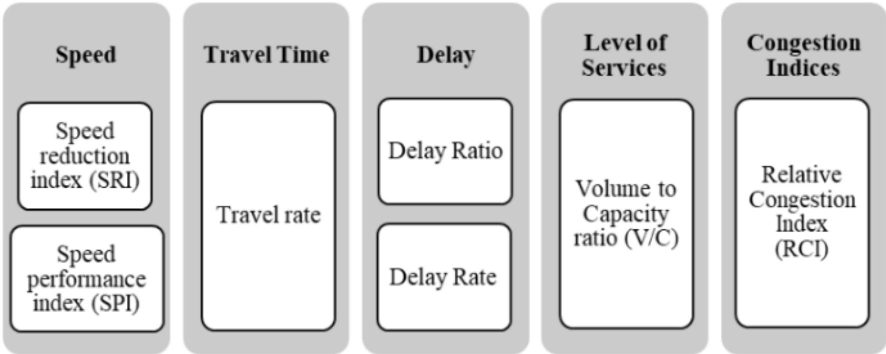


Figure 3.3: Congestion measures in different categories. [3]

4. level of services: use the volume-over-capacity ratio and classify the congestion in 6 classes,

5. congestion indices: related to the delay time and the free-flow travel time.

We must understand that some of these measurements are, in a some sense, equivalent and next we are going to explore them as a whole. Let's start with the **speed** measure may consists of *speed reduction index* (SRI) and *speed performance index* (SPI) [48]. SRI is the ratio of the relative speed change between the congested and the free-flow speed, given by,

$$SRI = \left(1 - \frac{v_{ac}}{v_0}\right) \times 10\,, \tag{3.18}$$

where v_{ac} is the actual travel speed and v_0 is the free-flow speed. The factor 10 is due to keep the SRI value in the range $(0, 10)$. Congestion occurs when the value exceeds 4 to 5. On the other hand, SPI (ranging from 0 to 100) can be defined by the ratio between vehicle speed and the maximum permissible speed and defined as,

$$SPI = \left(\frac{v_{avg}}{v_{max}}\right) \times 100\,, \tag{3.19}$$

where v_{avg} indicates the average travel speed and v_{max} denotes the maximum permissible road speed. Using this definition, the congestion is classified with three threshold values(25, 50, and 75). If $0 < SPI < 25$ the traffic state is considered as heavy congested. If $25 < SPI < 50$ the traffic state is considered as mild congested. If $50 < SPI < 75$ the traffic state is considered as smooth, and finally, if $75 < SPI < 100$ the traffic state is considered as very smooth.

Another and very interesting measure due to the simplicity is the **LoS** which can be deter-

LoS Class	Traffic State and Condition	VoC Ratio
A	Free flow	0.00 - 0.60
B	Stable flow with unaffected speed	0.61 - 0.70
C	Stable flow but speed is affected	0.71 - 0.80
D	High-density but the stable flow	0.81 - 0.90
E	Traffic volume near or at capacity level with low speed	0.91 - 1.00
F	Breakdown flow / jam traffic	> 1.00

Table 3.1: Level of services based on the corresponding VoC ratio and operating conditions

mined from various traffic quantities. We are going to use the capacity and the *volume-over-capacity ratio* (VoC) to determine the LoS of a road which is defined as,

$$VoC = \frac{N_v}{C}\,, \tag{3.20}$$

where N_v is the spatial volume in a time slot and C is the capacity or the maximum number of vehicles that a segment is able to contain during a time period. It can be further quantified as

$$C = \left(\frac{L_s}{L_v}\right) \times N_l \times \Delta t\,, \tag{3.21}$$

where L_s is the spatial segment length, L_v is the average vehicle length occupancy plus safety distance, N_l is the number of lanes, and Δt is the time interval on we are working. The relation with the congestion can be determined by a scale intervals with respect to the VoC ratio, as shown in Table 3.1. This definition had been used to determine the effect of congestion on **travel times**. To express the travel or commuting time, $t(v)$, (or cost) on a road link as a function of the traffic volume on the same road link, some *volume-delay functions* was proposed [49, 50]. The comparison of these methods using collected data was reported by David and Xiong in [51]. For example, the Bureau of Public Roads (BPR) function is given by,

$$t(v) = t_0\left(1 + \alpha\left(\frac{v}{c}\right)^{\beta}\right)\,, \tag{3.22}$$

where t_0 is the free-flow travel time, and the parameters α and β was fitted from real data on uncongested freeways as $\alpha = 0.15$ and $\beta = 4$. In that sense, when the volume in a link is very low, the predicted travel time is approximately equal to free-flow travel time. In turn, this predicted travel time allow to compute the **delays**, for example the delay rate is understood as the rate of time loss during congestion for a specific roadway segment, and can be calculated by the difference between the normal travel rate and the actual travel rate as

$$D = Tr_{ac} - Tr_{ap}\,, \tag{3.23}$$

where Tr_{ac} is the actual travel rate and Tr_{ap} is the normal travel rate. In summary, the commuting time savings are generally considered to be the most important component of transport projects designed to improve transport route and network efficiency. Reduction in congestion and lower travel times therefore represent the majority of benefits and goals in urban planning studies.

4

Network-Based Analysis of Urban Mobility Dynamics

This section will show the core of the present work which will be described in two sections. Section 4.1 shows a population game for studying individual decision-making processes of commuters with respect to transport mode choice in a city with the travel time as the key variable. The model was studied by Prieto-Curiel et al. in [15], and in this section we are going to summarize and replicate the concerning results of the modal share dynamics using a Python script. In Section 4.2 we display the network model which represents a city where some paths are generated to compute the average commuting time that will be used in the population game. The main goal of the present thesis is analyze the topology and the geometry of the network and study how changes will affect in the population game. In Section 4.3 we explain how to combine the two approaches.

4.1 The dynamics of modal choice

This part of the model makes use of the replicator equation to study the individual decision-making process about the transport mode in a city. The starting point of the current thesis was understand and replicate the work of Prieto-Curiel et al. In this section we make a summary and show the results that we replicate using pythosn scripts.

The paper explores the individual decision-making processes of commuters within a city regarding the transportation modes available, forming a modal share which will change dynamically using the commuting time as the key variable. A crucial assumption made is that some individuals change their commuting method to decrease their travel time, and the modal share changes governed by a population game. This kind of games can be considered as non-cooperative and iterative games with a finite set of strategies. For the present work, urban systems are understood as the game where each user (players) have to choose between different transport modes (strategies) to move inside the city from one point to another experiencing a travel time (cost or reward). Since every day thousands or millions of people have to choose the transport mode which interacts between them generating an impact to the other modes, and by consequence the commuting time may increase. For example, if there are many people using bikes or just walking, cars and buses have to stop at every crosswalk to give way increasing their commuting time. Also we have to consider the case of cars and buses which share the urban space. Prieto-Curiel identified that the whole process leads the model to reach an equilibrium which is always sub-optimal, unless both the direct and induced factors, as well as the number of users for all the transport mode are equal. Otherwise the tragedy of commons occurs, namely when individuals have a selfishness behaviour and choose a faster commuting time for themselves but creating a slower urban mobility within the city as a collective result, producing longer average commuting times.

For a better understanding of the problem, let's explore an example. Consider a city made of vertices which represent the street intersections and links connecting the vertices which represent the streets. Along the city would exist a population of agents that depending on the hour of the day will be in a residential or work zones. Each agent needs to commute from one zone (where he lives) to another (where he works) choosing a path to connect both points. The obtained payoff is the negation of the delay on the path he takes. The delay on the path is the sum of the delays on its constituent links, and the delay on a link is a function of the number of agents who use that link. Formally speaking, we consider a city with N individuals who, at time t, choose a transportation mode from κ choices. Let $x_i(t)$ be the number of users of mode i at time t, with $i = 1, 2, ..., \kappa$. We assume the effectiveness for transporting users inside the city can be measured in travel time and the costs for each transport mode can be expressed in minutes. The commuting time is divided into three components: 1) a baseline commuting time, which represents the free-flow travel time, and two variable costs representing the extra time due to congestion, that depend on 2) the number of users in the same mode of transportation, and 3) the number of users in the other modes. If the modal share, i.e., the distribution of users over different modes at time t, is given by the vector $\vec{x}(t) = (x_1(t), x_2(t), ..., x_\kappa(t))$, then the cost of using mode i can be expressed in minutes as

$$C_i(\vec{x}) = B_i + D_i(\vec{x}) + I_i(\vec{x}), \tag{4.1}$$

where B_i is a constant **baseline** commuting time for mode i, D_i is the **direct cost** that users of i impose to users on the same mode i, and I_i is the **indirect cost** imposed to mode i by the users of other transport modes. In general $D_i \geq 0, I_i \geq 0$ and $D_i \geq I_i$. Both the direct and indirect costs are not linear from a microscopic point of view, i.e. the city studied in detail, however from a macroscopic point of view (the city as a whole) the model can be simplified and the costs can be assumed as linear functions. The commuting time then for all transport modes can be expressed as

$$\vec{C}(\vec{x}) = \vec{B} + \mathcal{M}\vec{x}, \tag{4.2}$$

where $\vec{C}(\vec{x})$ is the commuting time vector (a function), $\vec{B}$ are the fixed baseline commuting times and $\mathcal{M}$ is a matrix related to the congestion (with the direct costs on the diagonal). The increasing number of, for example, car users imply a higher commuting time for driving, and, at the same time, the walkers and cyclists will commute faster. Similarly will occur for other transport modes.

We assume that there is some modal share for all transport modes for which the commuting time is lower than other transport modes, however no commuting mode is dominant because usually the transport modes compete between them. It implies that if everyone uses one transport method, it will be the slowest option. Also, there exists an internal point $\vec{x}^o$, where the number of users of all transport modes is greater than zero, $x_i > 0 \, \forall \, i$, such that $C_i(\vec{x}^o) = C_j(\vec{x}^o)$, for all pairs $i \neq j$, with $i, j = 1, 2, ..., \kappa$, meaning that there is some internal point associated with the Nash equilibrium for which all commuting methods take the same travel time.

For each modal share the average commuting time, i.e., the commuting time of each mode weighted by its share of users, can be computed by,

$$\mu(\vec{x}) = \frac{1}{N}[\vec{x}^T\vec{B} + \vec{x}^T\mathcal{M}\vec{x}] \, , \tag{4.3}$$

which is a paraboloid function. There is some modal share, $\vec{x}^\star$, for which the average commuting time is minimum. That is, $\mu(\vec{x})$ has some global minimum, $\mu^\star = \mu(\vec{x}^\star)$ for modal share $\vec{x}^\star$.

To study the dynamics of the modal share, we have to understand why the user choose one transport mode over the other, and a key factor is the travel time. For example, essentially cars provide the faster method but if everyone uses it the congestion increases considerably, then users would tend to change to a different transportation mode in order to reduce their time of the journey. We can realize that the game is evolutionary, dynamic and non-cooperative, where at each iteration the users switch between the distinct transport alternatives. The widely used way for modeling these games is the *replicator equation*, modified to urban systems as

$$\dot{\vec{x}}_i = \rho \vec{x}_i(\mu(\vec{x}) - \vec{C}_i) \tag{4.4}$$

where $\rho > 0$ is a speed parameter about how people switch between commuting methods. The overdot means differentiation with respect to time t. Here, $\mu(\vec{x}) - \vec{C}_i$ represents the difference between the average commuting time and the commuting time for mode i, remarking that the commuting method i becomes more popular when it is faster than the average.

Equation 4.4 defines a dynamical behaviour used to study the adaptation and co-evolution of large populations. After some iterations and from an arbitrary start, the dynamics will converge into an internal fixed point $\vec{x}^o$. This unique fixed point is described by the Nash equilibrium and exists if there is a point in X_N such that $C_i(\vec{x}) = C_j(\vec{x})$ for every pair $i \neq j$. This equilibrium means no user wants to change their commuting mode, because doing so their commuting time is longer. Figure 4.1 shows plots for three transportation modes given some baseline commuting time B, a cost matrix $\mathcal{M}$, and some initial condition $x(0)$. If $x(0)$ is changed but B and $\mathcal{M}$ are the same, the commuting time of all transport modes converge to the same value.

Eventually, $\vec{x}^o$ has the same commuting time for all the transport modes, regardless of the initial distribution. As we can appreciate in Figure 4.2, all trajectories converge to the same stable equilibrium distribution. With small value of $\rho > 0$ the process of convergence take a longer time. However, $\vec{x}^o$ is a suboptimal equilibrium of the system because there may be an internal equilibrium, $\vec{x}^\star$, if there is a modal share for which each commuting method has the lowest commuting time, defined as the optimal point of the system. In general $\vec{x}^o \neq \vec{x}^\star$ and the system converges into a non-optimum distribution which is always higher than the optimum. The only scenario where the equilibrium is optimum is when the baseline, direct and indirect commuting time are equal for all methods.

The replicator equation lies on the principle that faster commuting methods become more popular and users switch to reduce their commuting time. The modal share changes accordingly producing a trajectory in the space of transport modes X_N. Therefore, a trajectory is de-

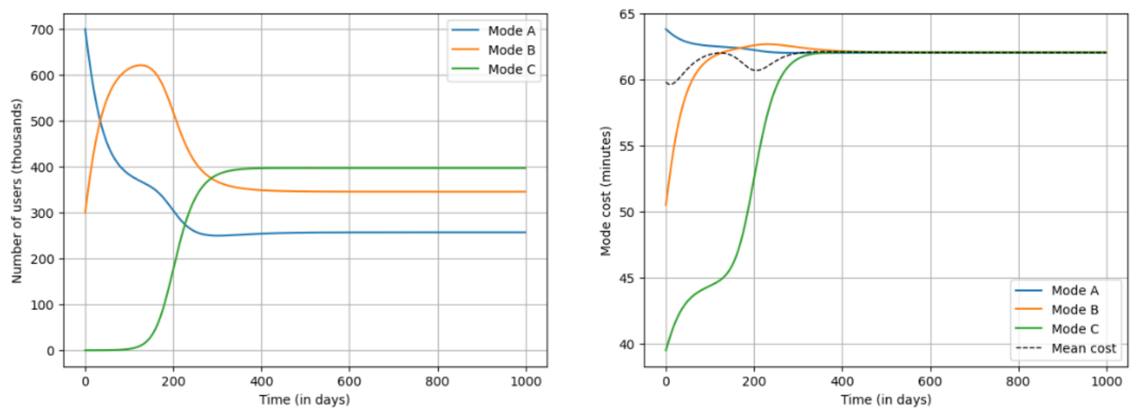


Figure 4.1: Modal share distribution versus time measured in 'days' scale for N = 1000000 individuals, considering three transport modes. Similar plot showed by Prieto-Curiel et al. in his work. The left figure indicates the number of users and their frequency. The right figure shows the commuting time in minutes for each transport mode and the mean commuting time experienced by all users which converge to some equilibrium $\vec{x}^{o}$.

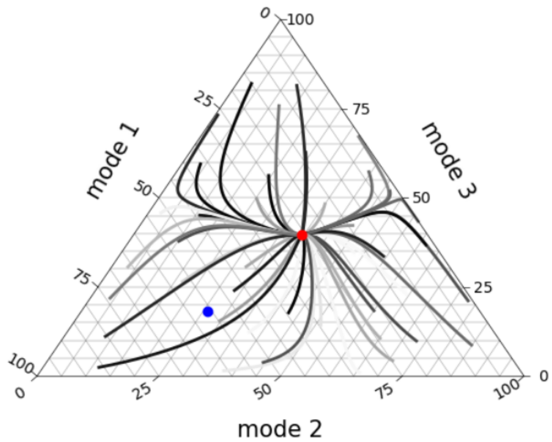


Figure 4.2: Similar simplex obtained by Prieto-Curiel. Each point in the triangle represents the distribution of users among the three transport modes which sum the total population. For some point inside the triangle, the amount of users of mode 1 can be identified following a diagonal line to the lower-right, for mode 2, a diagonal line to the upper-right, and for mode 3, a horizontal line from the right axis. The corners represent modal shares where only one transport system is used. There are different trajectories of the modal-share dynamics with random initial distribution and plotted with different colors across the simplex. All the curves converge to the same point (red) which is the attractor node and represent the Nash equilibrium of the dynamics, $\mathbf{x}^0$. From all modal shares, there is one configuration which has the lowest average commuting times (blue point) being the optimal modal share of the system, $\mathbf{x}^\star$.

fined for $\vec{x}(t)$ as t increases. With time passes and users switching the transport modes, the corresponding commuting time for each transport mode and the mean commuting time also change. The evolution of the mean commuting time with respect to time is given by combining Equation 4.4 and 4.3,

$$\frac{d\mu}{dt} = \frac{\rho}{N}\left[\vec{B}^T + \vec{x}^T(\mathcal{M} + \mathcal{M}^T)\right]\dot{\vec{x}}. \tag{4.5}$$

The goal of each user is to reach their destination as soon as possible which leads them to switch the transportation method in order to reduce their commuting time. However, the result of that switching is a higher average commuting time of the system. For example, if a person decides to drive instead of walking, their time might decreases, but creates more traffic for those who already drive. The city's transport evolution is represented by the trajectories $\vec{x}(t)$ of people trying to decrease their commuting time, although the collective result does not always succeed. To describe this, the *tragic* region is introduced where individual actions of switching between methods for a faster commute results in a slower collective system, increasing the average commuting time. With the replicator dynamics, $\dot{\mu} > 0$ means that the average commuting time increases. For any modal share $\vec{x}$, we say that it is *tragic* if

$$\left.\frac{d\mu}{dt}\right|_{\vec{x}} > 0, \tag{4.6}$$

and we call $\vec{x}$ to be *rewarding* if the derivative is not positive. A trajectory might transit between being tragic and rewarding along t in X_N. To conclude this section we want to remark that given a random initial modal share the system will evolve to a sub-optimal equilibrium, and when an individuals want to reduce their commuting time, the city will result in a collective tragedy increasing the average commuting time. Finally, an interesting and counter-intuitive result is that if a new mode of transportation with a small baseline commuting time but with a high impact to the other modes is introduced, the level of congestion will increase.

4.1.1 Beyond the model

Once we have replicated the model of Pietro-Curiel, we want to know what happen if we use a different number of transport modes. Specifically we are going to explore with two transportation modes which will help to understand and compare with our work. In Figure 4.3 we see the dynamics given some baseline commuting time B, a cost matrix $\mathcal{M}$, and some initial condition $x(0)$. As before, if we keep the same B and $\mathcal{M}$, although we change the $x(0)$ the commuting time of all transport modes will converge to the same value. We notice that the

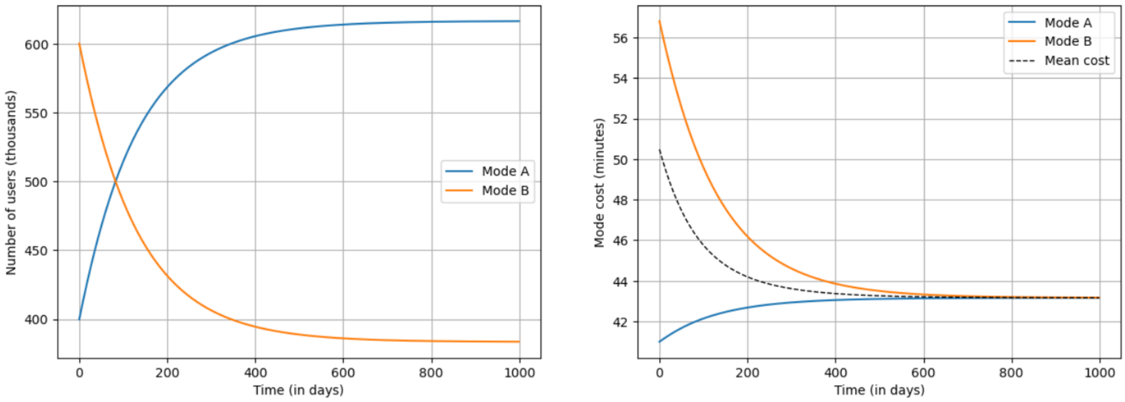


Figure 4.3: Modal share distribution versus time measured in 'days' scale for N = 1000000 individuals, considering two transport modes. The left figure indicates the number of users and their frequency. The right figure shows the commuting time in minutes for each transport mode and the mean commuting time experienced by all users which converge to some equilibrium $\vec{x}^o$.

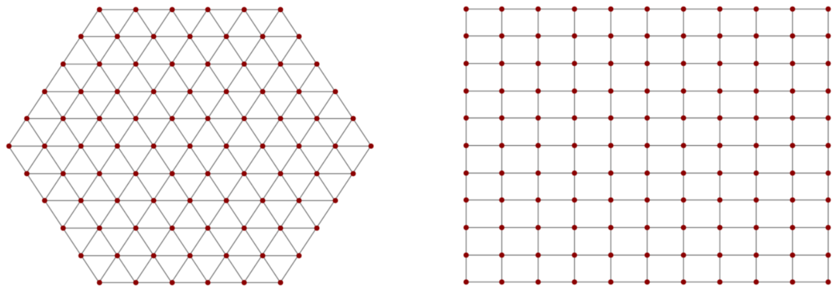
mean commuting time decreases in comparison than when three modes are used.

4.2 Network model

For the present section, the idea is to use a spatial network with edges representing the streets and nodes representing the intersections between the streets, all as a whole simulating a city. Then, N number of random nodes will be selected which will represent the origins and destinations (N is associated with the number of users using any transport mode). Using the Dijkstra algorithm we are going to compute the shortest paths between the selected points. Finally, using some assumptions and the traffic flow theory, we are going to be able to compute the baseline, direct and indirect commuting times. Now, we are going to show in detail every step.

Let's start defining the network and its embedding space. Consider an undirected graph with isotropic and centric properties, for preserving the rotational invariance. We consider two different graphs shown in Figure 4.4. The first is a hexagonal network with triangular lattice where the degree is 6. The second is a quadratic network with quadratic lattice and the average degree is 4. Both graphs are defined with the *radius* parameter, r, which controls the size of the graph and describes the number of links from the center to the border of the network, then the diameter (or largest shortest-path length) of the hexagonal lattice is given by

$$d_{hexa} = 2 \times r, \tag{4.7}$$



(a) Triangular lattice **(b)** Quadratic lattice

Figure 4.4: Example of graphs. Both are rotational invariant with $radius = 5$. **(a)** has 91 nodes and 240 edges. **(b)** has 121 nodes and 220 edges.

and for the quadratic lattice is,

$$d_{quad} = 2 \times (2 \times r). \tag{4.8}$$

4.2.1 Dijkstra's algorithm

Once the network is built, we have to generate the paths that users will cover to go from the origin to their destination. We consider the shortest path because, in the reality, most of the people want to arrive their destination as fast as possible. In 1959, the computer scientist Edsger W. Dijkstra [52] published his work proposing an algorithm to find the shortest paths given a origin node to every other node. The algorithm can also work for finding the shortest path between two known nodes. The idea behind Dijkstra's algorithm is construct a tree of vertices with the less distance between them. In order to do it, the algorithm maintains two main sets, one set is the *Shortest Path Tree* set (SPT) that keeps track of vertices whose show a minimum distance from the source. Initially, this set is empty. The other set, Q, contains all the unvisited vertices. The idea is to traverse all vertices of the graph starting with the neighbors of a given vertex and use a Min Heap to store the vertex/vertices not yet included in SPT. Min Heap is used to get the minimum distance vertex or vertices. Once the tree of paths is constructed, the shortest path will be selected as the path with the less distance between source and target. There may be more than one shortest path. Algorithm 4.1 displays the algorithm described previously.

Then, for generating the paths for our work we need to define the origin and destination points. We did this generation following two ideas [53], the first is the trivial one and generate the points randomly, the second is regarding some hotspots [54] in the city which represent

Algorithm 4.1 Dijkstra's algorithm.

Require: *Graph* and *source*

```
for vertex v in Graph.Vertices
    distance[v] ← INFINITY
    previous[v] ← UNDEFINED
    add v to Q
end for
distance[source] ← 0

while Q is not empty
    u ← vertex in Q with minimum distance[u]
    remove u from Q
    for neighbour v of u still in Q
        alt ← distance[u] + Graph.Edges(u, v)
        if alt < distance[v]
            distance[v] ← alt
            previous[v] ← u
        end if
    end for
end while
return distance[] and previous[]
```

the crowded spots such as residential or work zones. Once the origins and destinations are defined, we use Dijkstra's algorithm to get the paths. Now, we will explain in detail how both approaches were developed. For the first one, the idea is to select randomly N nodes as origins, and N nodes as destinations, where N represents the number of people using that transportation mode. If the origin and destination nodes are the same mean that the user does not commute, however we avoid this possibility since it would not contribute anything in the present work. The OD matrix then will have 0 in the diagonal. We use then the Dijkstra algorithm to generate the shortest paths. The procedure is summarized in Algorithm 4.2. The second approach is a bit more realistic since as we can read in the literature [55] and some studies using real data [56], the distribution of origins and destinations within a city may be fitted by an exponential distribution since there exists some hotspots, such as residential, and scholar or work places, which are more visited than other locations. In order to satisfied this idea, the origin and destination points are distributed as follow: we select randomly H_o number of origin hotspots and H_d number of destination hotspots. Then, using a Poisson distribution, the origins and destination points of the half of the users will be distributed around H_o and H_d, respectively, which will be different between them. The other half of users are distributed randomly as in the first approach because in the city there are also individuals that not commute only between their house and their work. The paths between these new points are generated using the Dijkstra algorithm as before. In Algorithm 4.3 we can see the algorithm explained.

Algorithm 4.2 Generate shortest paths with random origin and destination points.

Require: N *users* and *network*
  for $user = 1$ to $user = N$
    origin $\leftarrow$ select random node of the *network*
    destination $\leftarrow$ select random node of the *network*
    if $origin = destination$
      destination $\leftarrow$ select random node of the *network*
    end if
    path $\leftarrow$ connect origin and destination points with Dijkstra algorithm
    paths $\leftarrow$ append *path*
  end for
  return *paths*

Algorithm 4.3 Generate shortest paths with some hotspots as origins and destinations.

Require: *N users*, *network*, *hotspots*
  for *user* $= 1$ to *user* $= N/2$
    origin $\leftarrow$ select a node of *network* following a Poisson distribution around a *hotspots*
    destination $\leftarrow$ select a node of *network* following a Poisson distribution around a *hotspots*

    path $\leftarrow$ connect origin and destination points with Dijkstra algorithm
    paths $\leftarrow$ append *path*
  end for
  for *user* $= N/2$ to *user* $= N$
    origin $\leftarrow$ select random node of the *network*
    destination $\leftarrow$ select random node of the *network*
    Be sure *origin* and *destination* nodes are different
    path $\leftarrow$ connect origin and destination points with Dijkstra algorithm/method
    paths $\leftarrow$ append *path*
  end for
  return *paths*

4.2.2 Congestion

Up to now, we have explained how the network (2 graphs) and the paths (2 distributions) have been generated, however we still haven't pointed how relate them with the population game model. Basically, the idea is to extract some information from the network using the paths and the traffic flow theory explained in Section 3.4. The specific information we talk about is the baseline commuting time, as well as the cost related to the congestion: direct and indirect commuting time that are computed for every transport mode. These values will be the input in the replicator equation, Equation 4.4, and they will be updated every iteration since the users switch between transport modes.

In the present work we assume only two transportation modes for simplicity, bicycles and cars. We choose them because they are very common methods used for commuting in the cities, and by consequence they compete for the space. Although they have different characteristics, for example, cars are the most used for the speed to reach the destination, however also cause problems such as the pollution generated both short and long term. In contrast, bicycles are a bit more friendly with the environment and can cause a good humor in the individual. Coming back to the goal of this work, we must compute the congestion generated by bicycles and cars. In order to do it, we have to define some quantities based on the literature [57], and shown in

Parameter	Value
Bike free-flow speed	3.34 m/s
Car free-flow speed	8.34 m/s
Edge length	100 m
Bike length	2 m
Car length	5 m
Safe bike distance	2 m
Safe car distance	3 m

Table 4.1: Values of physical parameters that characterize the bicycles and cars in a city.

table Table 4.1. These quantities allow us to calculate, for example, the time that a transport mode (car or bike) needs to cross an edge by

$$t_B = \frac{block\ length}{free - flow\ speed}.$$

For bicycles, $t_B^B = 30\ s$, and for cars, $t_B^C = 12\ s$. Even more important, the defined quantities allow to compute the capacity of each road segment using Equation 3.21. In this case, the number of lanes is $N_l = 1$, and the time interval is $\Delta t = 1\ hour = 60\ mins$. The latter assumption is crucial and will be explained in detail in Section 4.3. Then, the capacity for bicycles is

$$C_B = \frac{100m}{2m + 2m} \times 60 = 1500,$$

and the capacity for cars is

$$C_C = \frac{100m}{5m + 3m} \times 60 = 750.$$

As the links in the network are considered similar, all of them will be assigned the same capacities.

We have defined all the ingredients to cook the model. It is a static one for simplicity where the paths are defined in a period of one hour. Before to proceed with the computation of the costs, it is important to remark that these values are not global parameters but that they will be computed every time iteration since the users switch the commuting methods and the congestion changes. To compute the **baseline commuting time** (BCT) we have two approaches, the paths-dependent and the paths-independent. The paths-independent approach is easier and can be still considered as a global parameter since it depends only on the *average shortest path* and the size of the network, and is given by

$$B = \textbf{average shortest path length} \times t_B, \tag{4.9}$$

where t_B is the time required by the method to cross a street. On the other hand, the paths-dependent approach will changes every time iteration since the number of paths (or users) changes along the time. However, it is still similar with respect to the path-independent. For computing, we do an average of the multiplication between the time that a transport mode needs to cross an edge and the length of each path, as follow,

$$B = \frac{1}{N}\sum_{i=1}^{N} len(i) \times t_B, \tag{4.10}$$

where N is the number of user. Now, we follow with the congestion costs using the traffic flow theory. Firstly we compute the volume-over-capacity (VoC) ratio for every edge in the network using Equation 3.20 and summarized in Algorithm 4.4. In general this quantity measure the level of congestion in every edge. For the **direct commuting time** case, a small penalty (between the 1% and 5% of the time required to cross an edge) is defined. The penalty is expressed in function of the time, t_B, that each transportation mode require to cross an edge, and by consequence they will be different between bicycles and cars. Specifically we define small penalties in both cases, 0.1 seconds for bicycles and 0.5 seconds for cars, because in general cars will cause more impact in the congestion. In order to preserve the idea of MFD where the velocity decreases linearly with respect to the users density in each edge, this penalty is multiplied by the congestion ratio, $penalty \times VoC$, so that, the trajectory of each user will experience an extra time to every edge, which will increase the journey time with respect to the congestion. Finally the direct cost is the average of the extra time of all the users, as shown in Algorithm 4.5.

In the case of **indirect commuting time**, the analysis must also preserve the linearly decreasing velocity with respect to the density. As we know, the indirect cost is smaller than the direct one, for this we are going to generalize the penalty approach using Table 3.1. In this case the penalty is the same for bicycles and cars, and it is very small with a value of 0.1. It will be governed by the LoS class and will increase linearly depending on the LoS value in each edge. Then, in the most congested edges, $VoC > 1.0$, the extra time will be of 0.5. In the edges with few

Algorithm 4.4 Congestion ratio algorithm

Require: $N\,paths$ and $capacity$
  $counts \leftarrow$ Count the number of users in each edge
  for $edge = 1$ to $edge = counts$
    $CongestionRatio \leftarrow list(\frac{edge}{capacity})$
  end for
  return $CongestionRatio$

Algorithm 4.5 Direct commuting time algorithm

Require: $N\,paths$, $CongestionRatio$ and $penalty$
  $DirectCost \leftarrow [\,]$
  for $path = 1$ to $path = N\,paths$
    $ExtraTime \leftarrow 0$
    for $edge = 1$ to $edge = length(path)$
      $ExtraTime \leftarrow ExtraTime + (penalty \times CongestionRatio[edge])$
    end for
    $DirectCost \leftarrow DirectCost + ExtraTime$
  end for
  $Dcost \leftarrow \frac{1}{N}\sum_{1}^{N} DirectCost$
  return $Dcost$

congestion, $VoC < 0.6$, the extra time will be of 0.0. As usual, we do the previous analysis for each user and at the end the indirect cost is the average of the extra time of all the users. The idea is visualized better in Algorithm 4.6.

Algorithm 4.6 Indirect commuting time algorithm

Require: $Npaths$, $CongestionRatio$ and $penalty$
  $IndirectCost \leftarrow [\,]$
  for $path = 1$ to $path = Npaths$
    $ExtraTime \leftarrow 0$
    for $edge = 1$ to $edge = length(path)$
      if $CongestionRatio[edge] < 0.6$
        $ExtraTime \leftarrow ExtraTime + (0 \times penalty)$
      else if $CongestionRatio[edge] < 0.7$
        $ExtraTime \leftarrow ExtraTime + (1 \times penalty)$
      else if $CongestionRatio[edge] < 0.8$
        $ExtraTime \leftarrow ExtraTime + (2 \times penalty)$
      else if $CongestionRatio[edge] < 0.9$
        $ExtraTime \leftarrow ExtraTime + (3 \times penalty)$
      else if $CongestionRatio[edge] < 1.0$
        $ExtraTime \leftarrow ExtraTime + (4 \times penalty)$
      else if $CongestionRatio[edge] > 1.0$
        $ExtraTime \leftarrow ExtraTime + (5 \times penalty)$
      end if
    end for
    $IndirectCost \leftarrow IndirectCost + ExtraTime$
  end for
  $Icost \leftarrow \frac{1}{N} \sum_{1}^{N} IndirectCost$
  return $Icost$

The algorithms displayed previously, 4.2, 4.3, 4.4, 4.5, 4.6, are generic. It means that they can be adapted for any number of users, any kind of network, and any kind of commuting method. For the latter we only need to define a proper *penalty* considering the impact of the transportation mode in the traffic of the city.

The last idea we explored in this work is the role of introducing parking spots in the cities, summarized in Algorithm 4.7. This is not strictly related to the congestion so we will consider it as extra time in the BCT. Specifically in the BCT for cars since the user needs to park them in specific places, while bicycles can be parked in any place. The idea is simple: some edges are selected randomly that will be considered as parking spot edges. If the destination of the car users is one of the nodes corresponding to any edge marked as parking, nothing happens. Otherwise, we will add a penalty of 3 minutes for every edge between the destination and the closest parking spot. These 3 minutes represent the average time that a user needs to walk across an edge of 100 meters. The analysis is done for every user and the extra time due to the parking is the average of all the users. In that sense, if there exist many parkings the extra time is small. Finally, we add this time into the BCT for cars.

Algorithm 4.7 Extra time due to parkings algorithm

Require: G, *destinations* and *parkingNodes*
  $DistToPark \leftarrow [\,]$
  $parkCost \leftarrow 0$
  for $destNode = 1$ to $destNode = destinations$
    for $parkNode = 1$ to $parkNode = parkingNodes$
      if *destNode* not in *parkingNodes*
        $distance \leftarrow$ compute number of edges between *parkNode* and *destNode*
        $DistToPark \leftarrow$ append *distance*
      end if
    end for
  end for
  for $d = 1$ to $d = DistToPark$
    $parkingCost \leftarrow$ get the smaller distance of every user d
    $parkCost \leftarrow parkingCost \times 3$
  end for
  $Pcost \leftarrow \frac{1}{N}\sum_{1}^{N} parkCost$
  return *Pcost*

Once the baseline, direct and indirect commuting time have been computed we go back to the replicator equation using these results as inputs. At this point we face against a tricky aspect related with the time scale because for the latter results (commuting time extracted from the network), the time scale is smaller than for the former results (time interval for the replicator equation). In the next section, Section 4.3, we describe what we did for combining both models.

4.3 Modeling modal choice in a network framework

At this point, we have a model based in a population dynamics where the minimum time step is one day. On the other hand, the network model that computes the congestion costs in minutes.

The natural question is, how can we combine both models? The way we found to do it was considering that the congestion and, by consequence, the direct and indirect cost explained in Subsection 4.2.2 were computed for an specific hour. This is reflected in Equation 3.21 when the capacity is defined, in a sense that we use the capacity of a street during *one* hour. For the next day, we understood that the same hour is studied.

Once the capacity is well defined, we can generate the paths and compute the congestion, direct and indirect costs. The costs should be converted into day units in order to use in the replicator equation. The replicator equation indicates the number of users who change their transportation mode from one day to the next. When we know how many users switch from one transportation mode to the other, the last step is to switch them manually in the spatial network. Basically, we identify the users/paths - the number that the replicator equation suggests - with the largest commuting time and switch their transportation mode. For the next iteration, namely the next day, we compute again the baseline, direct and indirect commuting time for solving the replicator equation once more. The complete model is easier visualized with the flow chart showed in Figure 4.5.

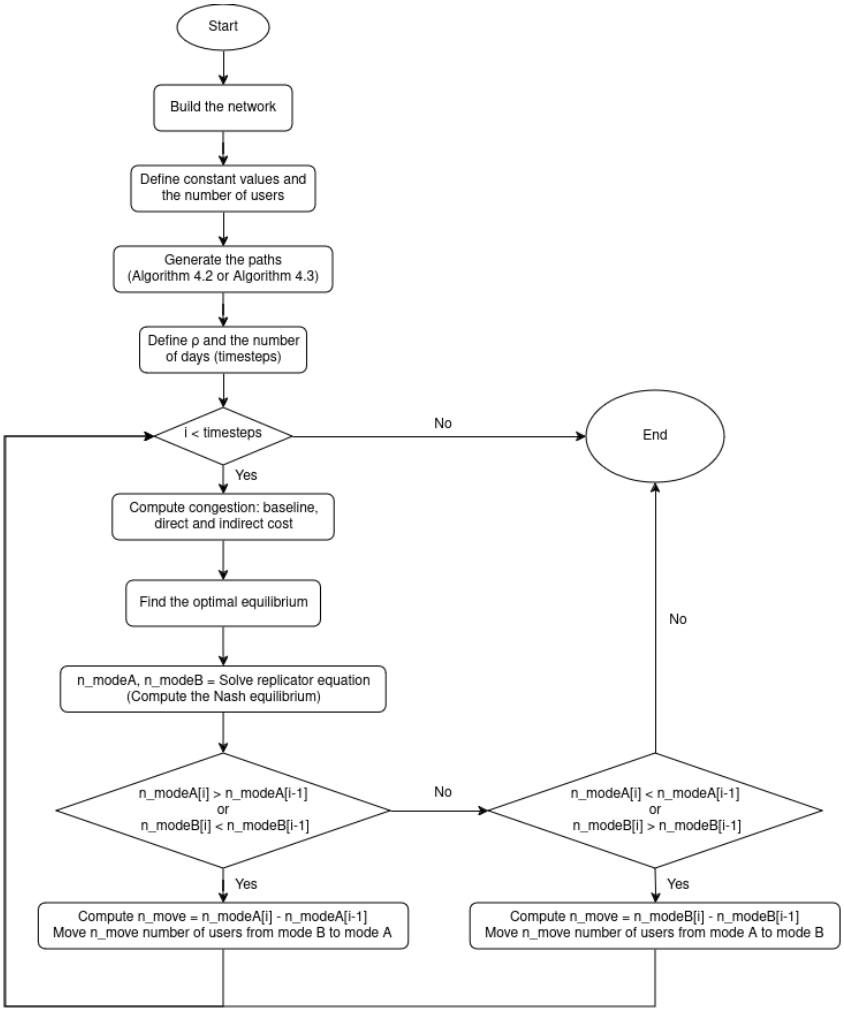


Figure 4.5: Flow chart of the combination of population game model and network model.

5

Results

In this chapter, we will present how the network model and the population game work together, the effects on the evolution of the individual decision-making process about transport methods in a city whit different constraints such as some changes in the topology of the network, in the initial modal shares, and in the methodology for generating paths. We are going to present the results in a chronological way how they were obtained. But firstly, in Section 5.1 a brief summary of the model is displayed as well as a brief discussion about the parameters used. Section 5.2 show the first result we found considering a network with triangular lattice. Then, we asked what happen if we change the initial modal share, considering the same network studied before. In Section 5.3 we display the effect related with changing the modal share, both the initial modal share and adding more users when the system reached the equilibrium. Next idea was to change the network and use a quadratic lattice instead of a triangular one varying also the initial modal shares. The results are displayed in Section 5.4. Until this point all the results showed in the previous section were generated with random paths. However, in real cities there exist specific areas that are busier than others. In Section 5.5 we analyze the dynamics of switching transport methods considering some hotspots in the cities. In Section 5.6 we introduce the idea of parking spots and how they affect to the dynamics.

5.1 Setup and Parameters

Our model aims to study the effects of changing the topology of a generic city on the individual's decision-making regarding the method of transportation. The dynamics of the individual's decision making is described by the replicator equation (Equation 4.4). It is at this point when we introduce the network (Section 4.2), which is a theoretical network that describes a city, and the traffic flow theory (Section 3.4) in order to compute the baseline commuting time, the congestion (Algorithm 4.4), the direct cost (Algorithm 4.5) and the indirect cost (Algorithm 4.6). Depending on the constraints used with respect to the network (Figure 4.4), paths generation (Algorithm 4.2, Algorithm 4.3) and initial modal share, the previous values will give different results, as it expected. These results will be the input in the replicator equation which indicates the number of users that switch from one transport method to the other. We solve the equation only for the exactly next time step, and, manually, we switch the number of pointed paths, regarding the ones with the larger congestion. After the manual changes, once more we compute the costs to solve the dynamical equation. The process is better visualized in Figure 4.5. Finally, we want to test the impact of changing the topology of the network in the dynamics.

However, we have an issue with respect to the timescales because the population game is in terms of days while the congestion obtained in the network model is hour dependent. It means that the timescale of population game model is bigger than the timescale of the network model. To solve this issue we focus on the network model and in the capacity definition, Equation 3.21. Thus, we make a crucial assumption saying that the studied timescale per day is only about one hour, and then, the street capacity is defined in that sense. The hour of the day is not considered for the present work. Finally, we understand that at every time step in the population game, we are studying the same hour.

After defining the time scale of the model, we define other parameters such as the radius of the city and the initial modal share. The radius is a dimensionless value that helps only to generate the network, and it is understood as the number of edges existing from the center to the border in horizontal or vertical directions. Basically, it determines the size of the network. In order to set this parameter we use a physical value as the area of a city, and for the present work we considered a middle-size city about 16 km^2. In Table 5.1, we display the radius of each network for getting the assumed area. We also compute some parameters related to the size of the network such as the number of nodes and edges, as well as the diameter and the average shortest path length (ASPL). When we multiply them by the length of the edge length, defined

Lattice	TRIANGULAR	QUADRATIC
Radius	20	25
Nodes	1951	1681
Edges	5700	3280
Diameter (km)	5.00	8.00
Area (km^2)	**16.24**	**16.00**
ASPL (km)	2.32	2.73
BCT bikes (min)	11.62	13.67
BCT cars (min)	4.65	5.47

Table 5.1: Physical values related to an area close to $16\ km^2$. The radius required to get this area and some other values extracted from the network are displayed. ASPL means average shortest path length, and BCT is the baseline commuting time that is given in minutes.

in Table 4.1, we provide a physical meaning and in Table 5.1 is expressed in kilometers. On the other hand, if we take the topological average shortest path length and multiply, using equation Equation 4.9, by the time that a transport method needs to cross an edge which was defined in Subsection 4.2.2, we obtain the path-independent baseline commuting time (BCT) which can be considered as a global parameter and it is used only to understand the relation across the networks. For the model we use the path-dependent baseline commuting time that will change every iterations. In any case, we see that these commuting times for bicycles and cars are similar in both networks. Finally, we want to remark that although the network with triangular lattice is slightly bigger than the one with the quadratic lattice, both the ASPL and by consequence the baseline commuting times for bicycles and cars show a smaller value. This is because the geometry of the networks since the the triangular-lattice one has a higher degree.

Continuing with the setting of the initial modal shares, we found in the literature some studies in cities using mobile phone data [58, 59, 60, 56] where suggest that the number of commuters in a hour varies between 4000 and 10000. However, we should consider that these values also includes other types of transport modes such as public transport or pedestrians. Thus, we use values around those to evaluate the model.

Once the network and the paths (given by the number of users) were generated, we have to compute the direct and indirect cost. For computing we introduce the idea of congestion (Equation 3.20) which is related with the volume of the transportation in every edge. Besides, we also introduce the idea of a penalty for linearize the velocity and accomplish the MFD. The penalty is given in seconds and is a small portion, less than 5%, of the time required for a method to cross an edge. Since cars take up more space, and interact more in the city because they are wider used, the penalty for them is defined slightly higher than the one for bicycles. This is evidenced in the direct commuting time computation where we simply multiply the penalty times the congestion ratio. For the case of indirect commuting time the penalty is equal for both methods and small because the indirect cost has less impact than the direct cost. The cost is multiplied for a factor following the Level of Services (Table 3.1) idea.

5.2 Triangular-lattice network

Now we have all the ingredients to evaluate the dynamics based on the spatial network. In this brief section we show the first result using an initial modal share, and we also explain in detail how to read the plots since the most of them showed in this chapter will have the same structure.

We started the work using the triangular-lattice network, basically, for simplicity and because it makes the system anisotropic. The majority of cities has a round shape, however there are also some exceptions that we are going to study later using a quadratic-lattice network. Anyhow, the first result, Figure 5.1, we used an initial modal share of 1700 bicycles and 2300 cars, for a total of 4000 users. For generating the paths we are using the random approach described in Algorithm 4.2. We assigned $\rho = 0.6$ which represent how fast the users take the decision to change the transportation mode and it will be the same for all the results along the chapter. Thus, the figure is made of two panels. The left panel shows the frequencies of populations along the time. If the individuals population using cars is larger than the bicycle population, they are expected to decrease over time since the cost of using cars is larger. Otherwise, if the individuals population of bicycles is larger, the dynamics remains unchanged. This will be explored in detail in the next section. The left panel also shows in a box the values for sub-optimal (Equilibrium) and optimal (Optimal) equilibria, where the first value is the number of bicycles and the second for cars: $[n\_bicycles, n\_cars]$. Given that specific modal share ($x_0 = [1700, 2300]$) we notice that the dynamics converge almost to the same number of users, while the optimal equilibrium suggests more number of individuals using bicycles.

On the other hand, the right panel displays the commuting time in minutes for each transport method and the mean commuting time both sub-optimal and optimal. We notice that the commuting cost of cars are decreasing faster than the commuting cost of bicycles increases. We also notice that the sub-optimal mean commuting time has a similar behavior with the more

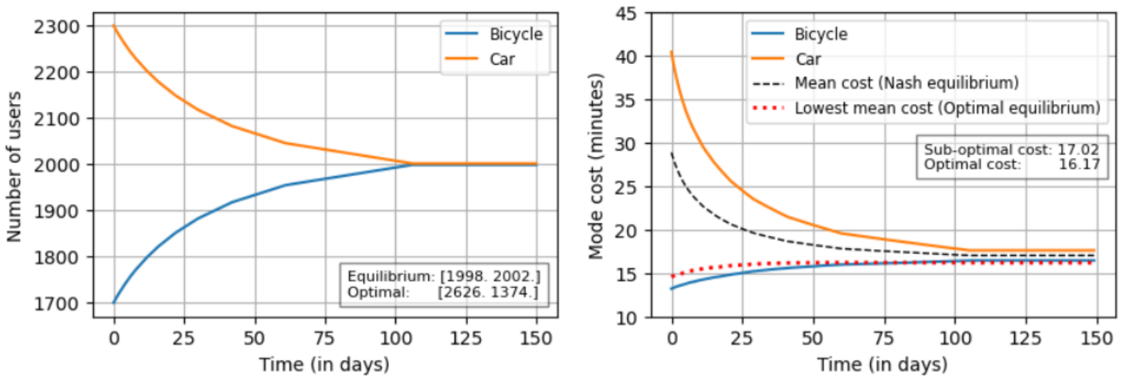


Figure 5.1: Modal share distribution versus time measured in 'days' scale for N = 4000 individuals, considering random paths, a triangular-lattice network, and two transport modes: bicycles and cars. The left panel indicates the number of users and their frequency, pointing the optimal and sub-optimal modal share. The right panel shows the commuting time in minutes for each transport mode, as well as the optimal and sub-optimal mean commuting time experienced by all users.

expensive method, and the optimal mean commuting time has a similar behavior with the less expensive method. At the end everything converge in a similar point. In the box we see the values at the last iteration and it is expressed in minutes. As expected, the optimal cost is lesser, although the difference is few. Finally, the complete plot shows that, given a triangular-lattice network with paths generated randomly and 4000 individuals commuting in one hour, the best users' distribution is 2626 for bicycles and 1374 for cars, which leads to an average commuting time for all the individuals of 16.17 minutes. Any other modal share will increase the latter time.

5.3 Effect of different modal shares

In this section we show the effect of changing the modal shares. We start playing with the initial modal share and then we play with the modal share once the system has converged. In the previous section we use 4000 individuals distributed in a sense that there exists a larger number of cars users than bicycles users, because the cost on cars is higher. A higher cost implies that the method is not the preferred choice and it will decrease along the time. In Figure 5.2, the dynamics with different population size are displayed, but keeping the idea that the cars distribution should be higher, at the moment. The motivation of this choice is due to the penalty because for cars is higher. Later we will show what happen if we use equal number of users of more bicycle users than cars. Continuing with the image, the first thing we noticed is that the higher individuals number the more favorable to the transport method switch. It is explained because more users generate more congestion being the system uncomfortable to reach a fast equilibrium and pushing more individuals to switch into methods. We also notice that the difference between the optimal and sub-optimal equilibria is smaller as much as the individuals number increase. It is understandable because as the individuals number increases, the space and the options (paths) for commuting decrease. In all the cases the optimal costs are still smaller but at some point, when the system is overwhelm, these costs will be the same as we can see in Figure 5.3. Another effect that we noticed is the difference on the number of users in the same method between the sub-optimal and the optimal, which decreases as the total users number increases. For example, for 3000 users the difference between the optimal and sub-optimal for bicycles is 707, for 4000 is 628, for 5000 is 440, for 6000 is 365 and for 7000 is 195. In the opposite sense, this difference increases for cars.

Now we break the idea of distributing the cars users as the larger population and in Figure 5.4 we show what happen if in the initial modal share the distribution is the same or more bicycles users. As we expected the dynamic is poor and the system reach easily the equilibrium. In the case of using more bicycles than cars, the former still have less cost and increase the number of users. In any case, when we compare them with the inverse initial modal share (Figure 5.1) the average commuting time of the system does not experience a bit difference, it is reduced in less than 1 minute only. On the other hand, taking attention in the case of 5000 users with a different initial modal share: top plot in Figure 5.4 and second plot in Figure 5.2, we can notice two things: firstly, the bigger initial number of bicycle the larger bicycle users at the end because following the replicator dynamics ideas, there will be more users that replicate that strategy. The second idea is that more bicycles benefit to reduce the average commuting time.

As we can see, the previous examples where the number of bicycles users are higher or equal is tasteless because the higher cost and congestion is caused by the cars. Now we come back to the first treatment where the number of cars is larger and we explore what happen if we change the modal share once the system already reached the equilibrium. In order to develop this idea, the initial modal share employed is 2600 users for bicycles and 3100 for cars, giving a total of 5700 individuals. Then, specifically 20 time steps after the system reached the equilibrium, new 300 users were added in order to get a population of 6000 users. In the first case we added 225 bicycles and 75 cars, and the system reached the new equilibrium very fast, exactly the next time step. In the second case we added only bicycle users and the system took a couple of days to reach the new equilibrium, (Figure 5.5). In the case where we added cars the average commuting time is higher than in the case where we added only bikes which is in agree with what we have been explaining along the project because cars generate more congestion and more cost.

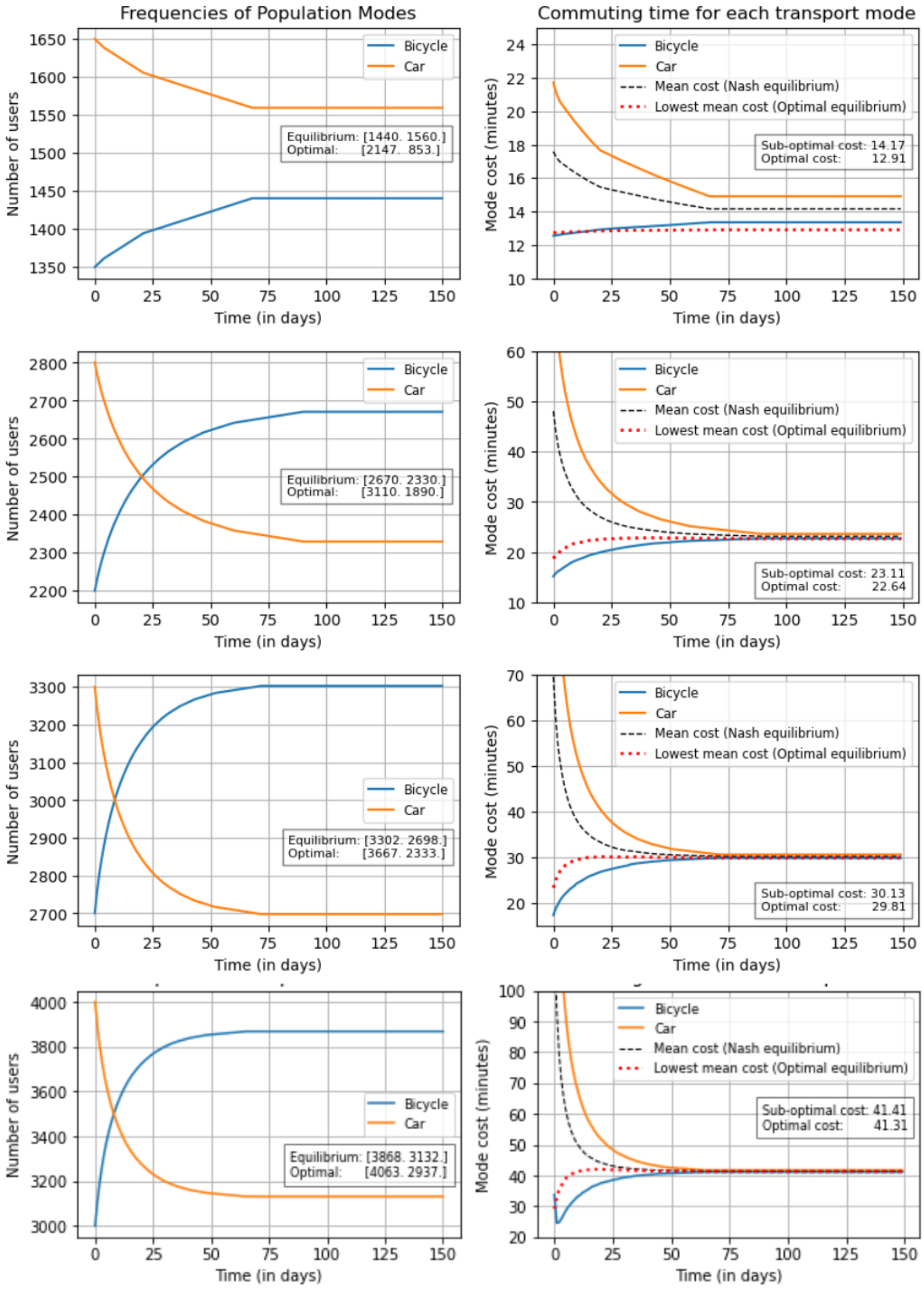


Figure 5.2: Dynamics of the individual decision-making process considering a triangular-lattice network and random origins and destinations, but changing the number of the total population. First plot is with 3000 users, second with 5000, third with 6000 and fourth with 7000. The plot with 4000 users is showed in Figure 5.1.

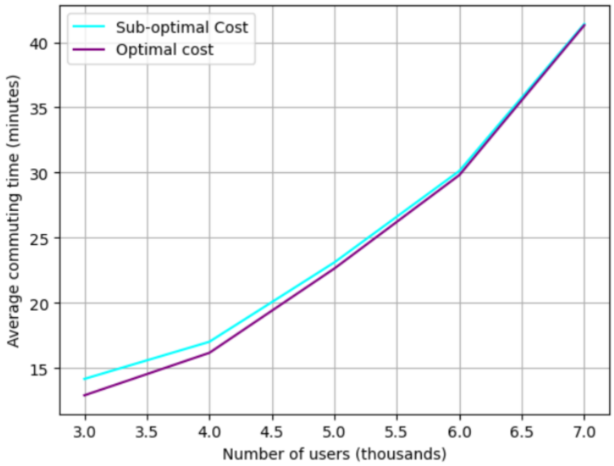


Figure 5.3: Relation between the total number of users and the average commuting time of the system. The values was taken from the previous results (Figure 5.1 and Figure 5.2) which means that it is for a triangular-lattice network with random paths. The modal share has more number of car users than bicycle users.

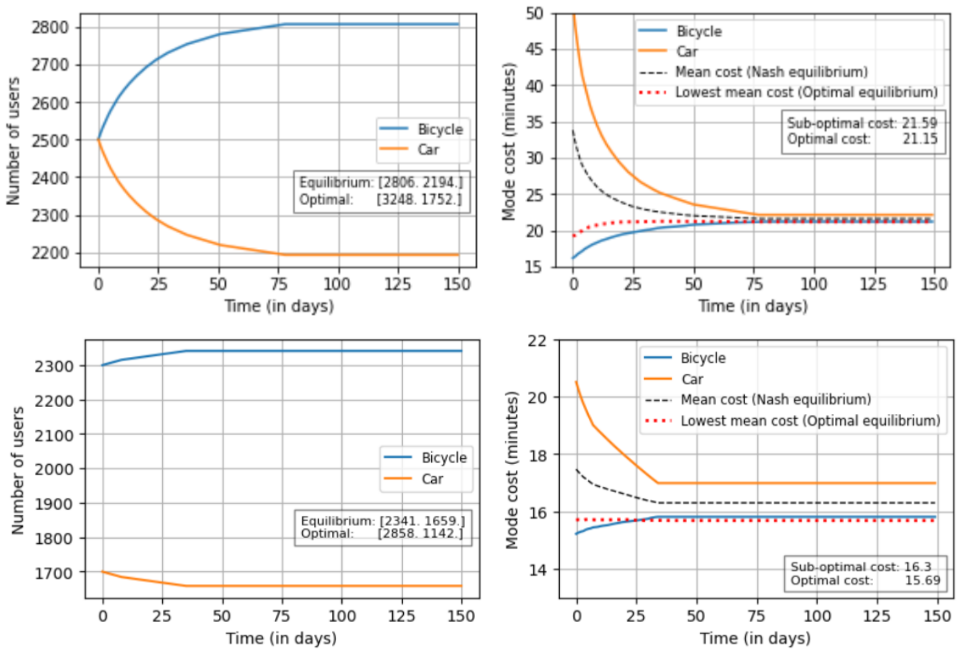


Figure 5.4: Dynamics of the individual decision-making process considering a triangular-lattice network and random origins and destinations. Unlike the previous results, in the top plot the initial modal share is the same for both methods. In the bottom there are more bicycles users.

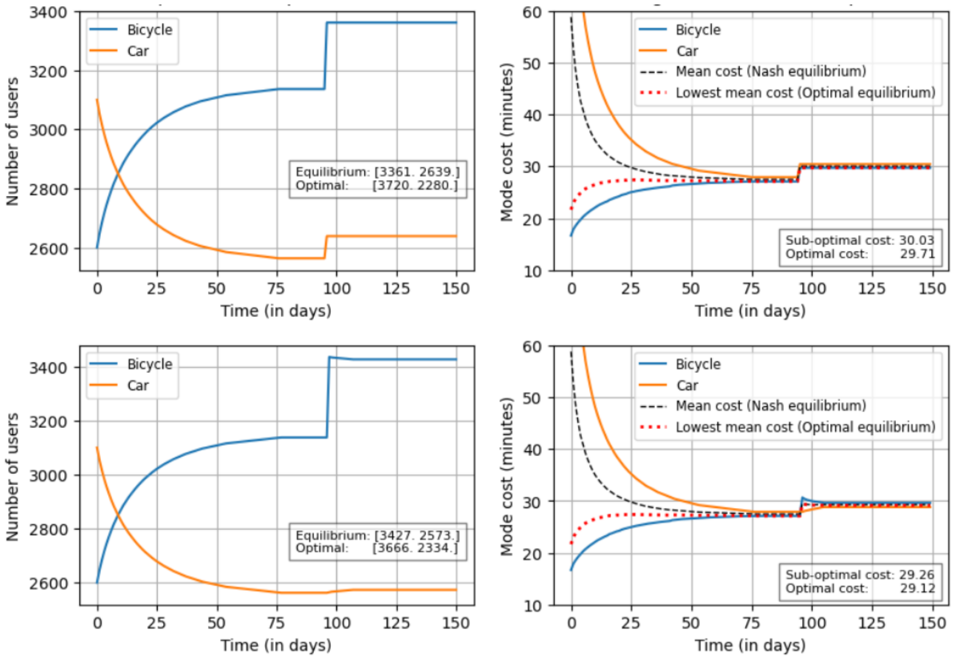


Figure 5.5: Dynamics of the individual decision-making process considering a triangular-lattice network, random origins and destinations, and initial modal share of $[2600, 3100]$. Once the system reached the equilibrium, new users were introduced. In the first plot 225 bicycles users and 75 cars users were added. In the second the addition was only 300 new bicycle users.

If we compare with the third plot of Figure 5.2 where there is the same total number of individuals, 6000, we see that the dynamics, the sub-optimal and optimal average commuting times are very similar. It suggests that the addition of new users has a very small or even no impact in the evolution of the decision-making into the transportation methods.

5.4 Quadratic-lattice network

The next idea we explored was to change the lattice of the network based on contemporary cities. We chose a quadratic arrangement (Figure 4.4.b.) that could represent street networks in US cities. In this chapter we are going to present the result using this lattice. The origins and destinations for path generation are randomly located. We are going to use firstly the same total number of users as in Figure 5.1 to compare the model. Latter we also shows what is the effect of changing the initial modal share.

Figure 5.6 displays the result of using the same initial modal share. We notice that the model using the quadratic-lattice network converge faster, which suggests that this network is more

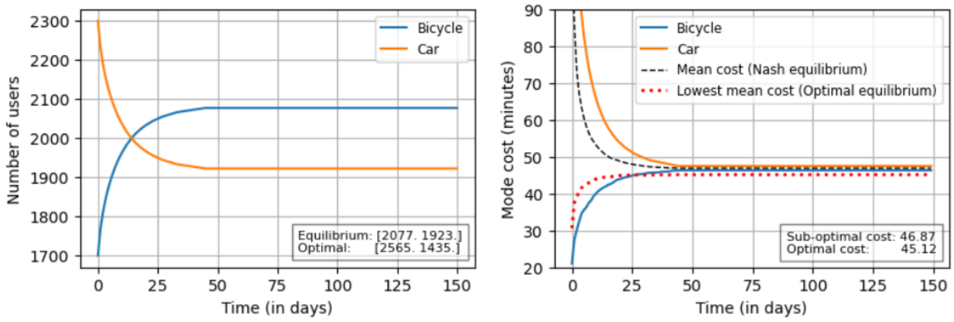


Figure 5.6: Modal share distribution versus time measured in 'days' considering the individual decision-making process for $N = 4000$ individuals, quadratic-lattice network and random origins and destinations.

limited and offers less options. We can confirm it when compare the average commuting time because it is almost 3 times higher. Making this lattice very problematic for the purpose of reducing the commuting time. In this sense the number of users must be reduced.

Now we explore the effect of treating with different initial modal shares, Figure 5.7. The result is similar as before (Section 5.3): the larger number of users the better dynamics (more users switch) and the larger cost. In this case (quadratic-lattice network) we obtain huge average commuting time values with less number of users. Confirming the limitation of the network, and once more it is due to the spatial degree of the nodes in the network.

5.5 Paths generation with hotspots

Beyond the topology of the network, we also asked ourselves what happen if the origins and destinations are generated in a different way. Until now we have generated randomly, but according the literature the cities have some hotspots which are more crowded than other areas. Thus, using Algorithm 4.3 we performed the model with these constraint for both triangular-lattice and quadratic-lattice networks with initial modal share of 1700 bicycles and 2300 cars, in order to compare with the previous results in Figure 5.1 and Figure 5.6.

For Figure 5.8 we have used 10 hotspots selected randomly, where the top figure is for triangular-lattice network and bottom figure is for quadratic-lattice network. Once more we notice that the quadratic-lattice network converges faster and shows a higher commuting time. Comparing the final average commuting time (sub-optimal) using the same initial modal share distribution, in the case of triangular lattice we have 17.02 minutes (Figure 5.1) when the origins and destinations are random, while when the hotspots are introduced the time increases to 22.62

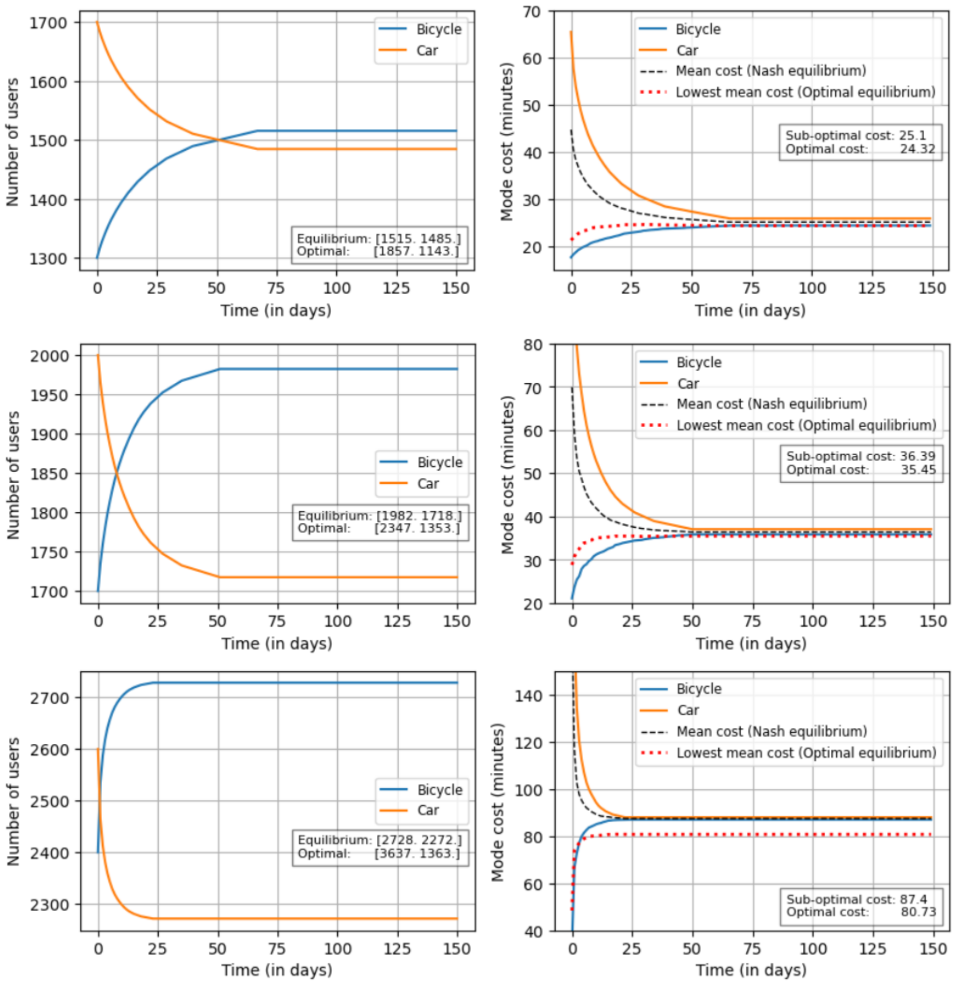


Figure 5.7: Dynamics of the individual decision-making process considering a quadratic-lattice network and random origins and destinations, but changing the number of the total population. First plot is with 3000 users, second with 3700 and third with 5000. The plot with 4000 users is showed in Figure 5.6.

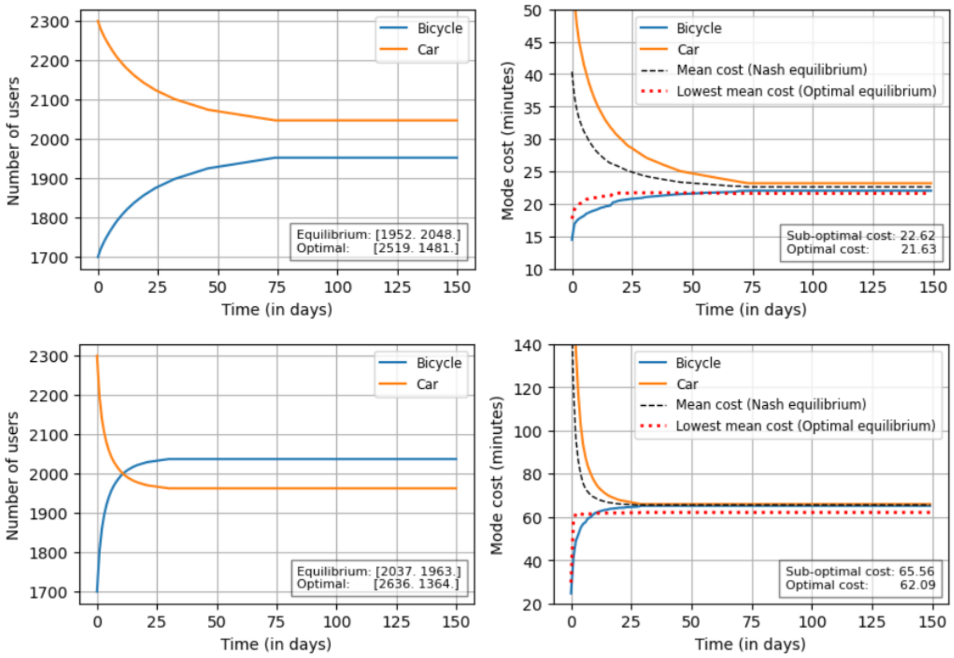


Figure 5.8: Dynamics of the individual decision-making process considering $N = 4000$ individuals distributed as follow: 1700 bicycles and 2300 cars. The origins and destinations points were generated regarding some hotspots (Algorithm 4.3) along the city. In the top plot we use the triangular-lattice network and in the bottom plot we consider the quadratic-lattice network.

minutes. Something similar happens with the quadratic-lattice network: increases from 46.87 minutes (Figure 5.6) with random OD to 65.56 when some hotspots are introduced. Then, we can say that when some hotspots were added the average commuting time is larger. And could be explained because the options of paths are reduced and use the central paths, which creates more congestion in the edges near to the hotspots.

5.6 Effect of parking spots

The final idea implemented in this work is about the parking spots using Algorithm 4.7. Understanding what is the impact of adding or removing some parking spots in the city could help to take public policies around the urban infrastructure. We have to remark that this is a first and very basic approach. Figure 5.9 displays the results which can be compared with Figure 5.1 because the same constraints were established. Same network with triangular lattice, same paths generated randomly and same initial modal share of $x^0 = [1700, 2300]$. Clearly, we notice a high impact when we introduced parking spots. The extra time added in the BCT causes the

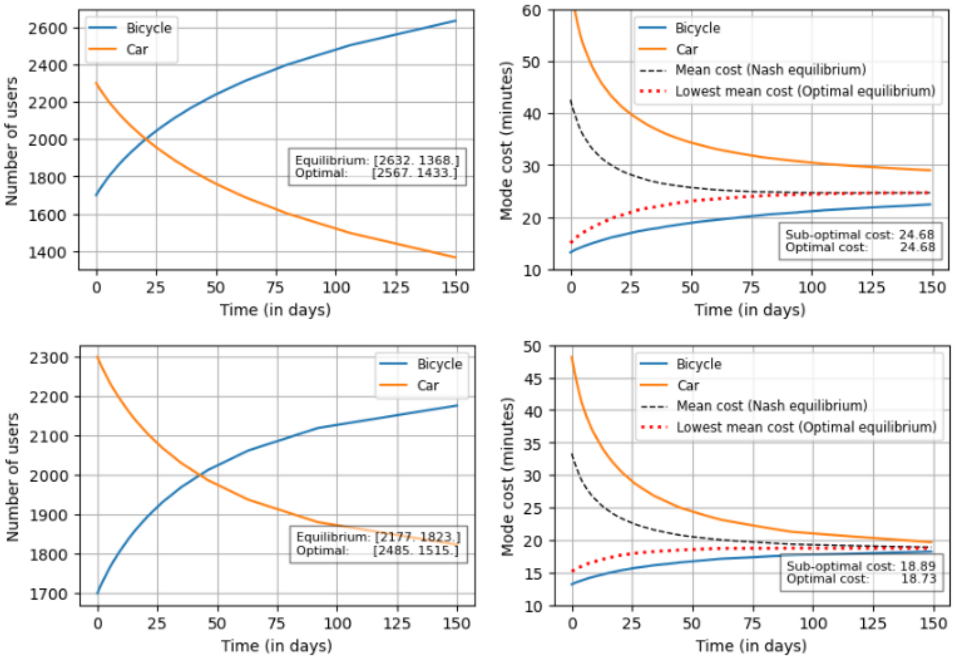


Figure 5.9: Dynamics of the individual decision-making process considering parking spots for $N = 4000$ individuals distributed as follow: 1700 bicycles and 2300 cars. The network has a triangular lattice, and the origins and destinations were generated randomly. In the first plot we consider 25 parkings, while in the second 150.

system to take too long to converge as well as increases the average commuting time. For the first plot we introduce only 25 parking edges which is a very small number considering the size of the network and the number of cars. In that case, we can say that if there are not enough spots the optimal and sub-optimal costs are equal. For the second plot, the number of parking spots increase in 125 in comparison with the first plots. Thus, given the same arrangement, and increasing 125 parking spots, which correspond approximately only to the 12 $\%$ of the nodes, the average commuting time decreases considerably and it is almost similar to the base scenario without parking spots.

6
Conclusion

In this work, we developed a theoretical model to investigate the impact of topological variations in spatial networks on the average fitness parameter in the replicator equation, which governs the evolution of decision-making processes in transportation systems. By employing a spatial network as a simplified representation of an urban environment and integrating traffic flow theory, we characterized essential parameters such as congestion levels and commuting times. These metrics served as the foundation for linking network structure to individual transportation choices. Specifically, the commuting time was used as the fitness parameter in the replicator equation, driving the adoption of commuting methods, where a method i became increasingly favored if its commuting time was lower than the system-wide average. Consequently, modifications to network topology altered the commuting times, reshaping the dynamics of decision-making processes across the population.

We compared two network types with distinct topologies: a triangular lattice network with an average degree of $k_{triangular} = 6$ and a quadratic lattice network with an average degree of $k_{quadratic} = 4$. By simulating random origin-destination pairs and incorporating hotspots to represent areas of concentrated urban activity, we examined the influence of these structures on congestion, commuting times, and the dynamics of modal switching.

The analysis of the triangular lattice network revealed that larger populations facilitated smoother transitions between transport methods, driven by the emergent dynamics of modal switching. As the population size increased, the difference between optimal and suboptimal equilibria diminished, suggesting that larger populations enable more efficient outcomes. Moreover, when the initial modal share was balanced or slightly favored bicycles over cars, the system converged to equilibrium more rapidly. This result highlighted the efficiency of bicycles, attributed to their lower commuting costs.

In contrast, the quadratic lattice network produced higher average commuting times and greater congestion levels, indicating that this topology was less robust to increases in commuter numbers. The triangular lattice exhibited greater resilience, maintaining lower congestion levels and supporting larger populations more effectively. These results emphasize the critical role of network topology in shaping transportation efficiency, with denser connectivity offering significant advantages in urban mobility.

The introduction of hotspots significantly increased congestion in both network types, leading to slower modal switching and longer commuting times. This highlighted the strain imposed on transportation systems by an uneven distribution of urban activity. To mitigate these effects, a sectorized urban design, in which services are distributed into sub-clusters, could reduce the need for long-distance commutes and alleviate congestion in busy areas.

Parking constraints were shown to play a pivotal role in transportation dynamics. A lack of sufficient parking spaces caused a sharp decline in car usage, encouraging alternative modes such as bicycles. Notably, even a small allocation of parking spaces proportional to the network size resulted in a significant reduction in average commuting times, underscoring the potential of strategic parking management to optimize urban mobility.

A key overarching conclusion of this study is that incorporating bicycles as a primary commuting method markedly reduces average commuting times in mid-sized cities when competition is limited to cars and bicycles. This finding reinforces the importance of promoting sustainable and efficient transportation modes in urban planning.

This study provides a foundational framework for analyzing the interplay between urban topology and individual decision-making in transportation systems. The results demonstrate how network structure, congestion dynamics, and resource allocation collectively shape commuting patterns and equilibria in modal choices. Moving forward, future research could expand the model by incorporating real-world data, such as actual road networks and commuter flows, to enhance its practical relevance. Incorporating more refined traffic flow dynamics, such as temporal variations during peak hours or weather conditions, would allow for more realistic simulations of urban mobility systems. Additionally, expanding the framework to include other transport modes, such as pedestrian traffic, public transit, or emerging technologies like e-scooters, would provide a richer understanding of multimodal systems. Introducing urban policy considerations, such as congestion pricing, dedicated bike lanes, or public transit investments, would further deepen insights into strategies for designing sustainable and efficient cities.

By bridging theoretical population dynamics with practical urban scenarios, this work establishes a versatile platform for exploring the relationship between city topology, individual decision-making, and transportation efficiency. The insights derived here provide valuable guidance for developing resilient, sustainable, and efficient urban mobility systems.

References

[1] W. W. Zachary, "An information flow model for conflict and fission in small groups," *Journal of Anthropological Research*, vol. 33, no. 4, pp. 452–473, 1977. [Online]. Available: https://doi.org/10.1086/jar.33.4.3629752

[2] D. Rhodes, S. Tomlins, S. Varambally, V. Mahavisno, T. Barrette, S. Kalyana-Sundaram, D. Ghosh, A. Pandey, and A. Chinnaiyan, "Probabilistic model of the human protein-protein interaction network," *Nature Biotechnology*, vol. 23, pp. 951–959, 2005.

[3] T. Afrin and N. Yodo, "A survey of road traffic congestion measures towards a sustainable and resilient transportation system," *Sustainability*, vol. 12, p. 4660, 06 2020.

[4] A. Rabl and A. de Nazelle, "Benefits of shift from car to active transport," *Transport Policy*, vol. 19, no. 1, pp. 121–131, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967070X11001119

[5] M. Keall, R. Chapman, P. Howden-Chapman, K. Witten, W. Abrahamse, and A. Woodward, "Increasing active travel: results of a quasi-experimental study of an intervention to encourage walking and cycling," *Journal of Epidemiology & Community Health*, vol. 69, no. 12, pp. 1184–1190, 2015. [Online]. Available: https://jech.bmj.com/content/69/12/1184

[6] M. D. Keall, C. Shaw, R. Chapman, and P. Howden-Chapman, "Reductions in carbon dioxide emissions from an intervention to promote cycling and walking: A case study from new zealand," *Transportation Research Part D: Transport and Environment*, vol. 65, pp. 687–696, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136192091830419X

[7] R. Prieto Curiel, H. González Ramírez, M. Quiñones Domínguez, and J. P. Orjuela Mendoza, "A paradox of traffic and extra cars in a city as a collective behaviour," *Royal Society Open Science*, vol. 8, no. 6, p. 201808, 2021. [Online]. Available: https://royalsocietypublishing.org/doi/abs/10.1098/rsos.201808

[8] S. Çolak, A. Lima, and M. C. González, "Understanding congested travel in urban areas," *Nature Communications*, vol. 7, p. 10793, 2016. [Online]. Available: https://doi.org/10.1038/ncomms10793

[9] J. De Vos, P. Mokhtarian, T. Schwanen, V. Van Acker, and F. Witlox, "Travel mode choice and travel satisfaction: bridging the gap between decision utility and experienced utility," *Transportation*, vol. 43, pp. 771–796, 09 2016.

[10] Y. Liao, J. Gil, R. H. M. Pereira, S. Yeh, and V. Verendel, "Disparities in travel times between car and transit: Spatiotemporal patterns in cities," *Scientific Reports*, vol. 10, p. 4056, 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-61077-0

[11] J. Correa, A. Schulz, and N. Stier-Moses, "On the inefficiency of equilibria in congestion games," vol. 3509, 06 2005, pp. 167–181.

[12] T. Roughgarden, *Selfish Routing and The Price of Anarchy*, 01 2005, vol. 74.

[13] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.162.3859.1243

[14] N. Levy, I. Klein, and E. Ben-Elia, "Emergence of cooperation and a fair system optimum in road networks: A game-theoretic and agent-based modelling approach," *Research in Transportation Economics*, vol. 68, pp. 46–55, 2018, sI: Frontiers in Transportation. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S073988591630018X

[15] R. Prieto Curiel, H. González Ramírez, and S. Bishop, "A ubiquitous collective tragedy in transport," *Frontiers in Physics*, vol. 10, 2022. [Online]. Available: https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2022.882371

[16] L. Lambertini, *Game Theory in the Social Sciences: A reader-friendly guide*. Routledge, 2011.

[17] W. H. Sandholm, "Chapter 13 - population games and deterministic evolutionary dynamics," ser. Handbook of Game Theory with Economic Applications, H. P. Young and S. Zamir, Eds. Elsevier, 2015, vol. 4, pp. 703–778. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444537669000136

[18] ——, *Population Games and Evolutionary Dynamics*. The MIT Press, 2010. [Online]. Available: http://www.jstor.org/stable/j.ctt5hhbq5

[19] M. A. Nowak, *Evolutionary Dynamics*. Belknap Press of Harvard University Press, 2006.

[20] J. M. Smith and M. Smith, "Evolution and the theory of games : In situations characterized by conflict of interest , the best strategy to adopt depends on what others are doing," 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:260882253

[21] E. Parzen, *Stochastic Processes*. Dover Publications, 2015.

[22] H. Ohtsuki and M. A. Nowak, "The replicator equation on graphs," *Journal of Theoretical Biology*, vol. 243, no. 1, pp. 86–97, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022519306002426

[23] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1, pp. 145–156, 1978. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0025556478900779

[24] J. Hofbauer and K. Sigmund, "Evolutionary games and population dynamics," 1998. [Online]. Available: https://api.semanticscholar.org/CorpusID:85023742

[25] R. Cressman and Y. Tao, "The replicator equation and other game dynamics," *Proceedings of the National Academy of Sciences*, vol. 111, no. supplement_3, pp. 10 810–10 817, 2014. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1400823111

[26] P. L. Erdos and A. Rényi, "On the evolution of random graphs," *Transactions of the American Mathematical Society*, vol. 286, pp. 257–257, 1984. [Online]. Available: https://api.semanticscholar.org/CorpusID:6829589

[27] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.286.5439.509

[28] W. R. Black, *Transportation: A Geographical Analysis*. Guilford Press, 2003.

[29] J. Añez, T. De La Barra, and B. Pérez, "Dual graph representation of transport networks," *Transportation Research Part B: Methodological*, vol. 30, no. 3, pp. 209–216, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0191261595000240

[30] M. Barthélemy and A. Flammini, "Modeling urban street patterns," *Physical Review Letters*, vol. 100, no. 13, Apr. 2008. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.100.138702

[31] T. H. Grubesic and A. T. Murray, "Spatial–historical landscapes of telecommunication network survivability," *Telecommunications Policy*, vol. 29, no. 11, pp. 801–820, 2005, network Security, Survivability and Surveillance. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0308596105000881

[32] W. Gou and Z. Jin, "Understanding the epidemiological patterns in spatial networks," *Nonlinear Dynamics*, vol. 106, pp. 1059–1082, 2021.

[33] G. A. Pagani and M. Aiello, "The power grid as a complex network: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 11, pp. 2688–2700, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378437113000575

[34] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. Bebber, M. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, "Rules for biologically inspired adaptive network design," *Science (New York, N.Y.)*, vol. 327, pp. 439–442, 01 2010.

[35] W. L. Garrison, "Connectivity of the interstate highway system," *Papers in Regional Science*, vol. 6, no. 1, pp. 121–137, 1960. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1056819023000349

[36] F. Hillebrand, M. Luković, and H. J. Herrmann, "Perturbing the shortest path on a critical directed square lattice," *Phys. Rev. E*, vol. 98, p. 052143, Nov 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.98.052143

[37] A. Schrijver, "On the history of the shortest path problem," 2012. [Online]. Available: https://api.semanticscholar.org/CorpusID:15086733

[38] S. Bera and K. V. K. Rao, "Estimation of origin-destination matrix from traffic counts: the state of the art," *European Transport Trasporti Europei*, no. 49, pp. 2–23, 2011. [Online]. Available: https://ideas.repec.org/a/sot/journl/y2011i49p2-23.html

[39] D. Rhoads, C. Rames, A. Solé-Ribalta, M. C. González, M. Szell, and J. Borge-Holthoefer, "Sidewalk networks: Review and outlook," *Computers, Environment and Urban Systems*, vol. 106, p. 102031, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0198971523000947

[40] A. Boumahdaf, M. Broniatowski, Émilie Miranda, and A. Le Squeren, "A behavioral probabilistic model of carrier spatial repositioning decision-making," *Transportation Research Part C: Emerging Technologies*, vol. 153, p. 104194, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X23001833

[41] H. K. Evans, B. D. Greenshields, D. Schapiro, and E. L. Ericksen, "Traffic performance at urban street intersections," *Journal of the American Statistical Association*, vol. 44, p. 142, 1949. [Online]. Available: https://api.semanticscholar.org/CorpusID:121640662

[42] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261508000180

[43] K. Zaidi, M. Milojevic, V. Rakocevic, and M. Rajarajan, "Data-centric rogue node detection in vanets," 09 2014, pp. 398–405.

[44] D. Helbing, "Traffic and related self-driven many-particle systems," *Rev. Mod. Phys.*, vol. 73, pp. 1067–1141, Dec 2001. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.73.1067

[45] T. R. Board, E. National Academies of Sciences, and Medicine, *Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis*. Washington, DC: The National Academies Press, 2022. [Online]. Available: https://nap.nationalacademies.org/catalog/26432/highway-capacity-manual-7th-edition-a-guide-for-multimodal-mobility

[46] D. Othayoth and K. K. Rao, "Investigating the relation between level of service and volume-to-capacity ratio at signalized intersections under heterogeneous traffic

condition," *Transportation Research Procedia*, vol. 48, pp. 2929–2944, 2020, recent Advances and Emerging Issues in Transport Research – An Editorial Note for the Selected Proceedings of WCTR 2019 Mumbai. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146520306074

[47] Q. D. of Transport and M. Roads, "Cost-benefit analysis manual and cba6," Feb 2011, last version March 2021. [Online]. Available: https://www.tmr.qld.gov.au/business-industry/Technical-standards-publications/Cost-Benefit-Analysis-Manual

[48] F. He, X. Yan, Y. Liu, and L. Ma, "A traffic congestion assessment method for urban road networks based on speed performance index," *Procedia Engineering*, vol. 137, pp. 425–433, 2016, green Intelligent Transportation System and Safety. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877705816003040

[49] U. S. B. of Public Roads, *Traffic Assignment Manual for Application with a Large, High Speed Computer*, ser. Traffic Assignment Manual for Application with a Large, High Speed Computer. U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, 1964, no. v. 37. [Online]. Available: https://books.google.it/books?id=AvNUR_O_JEcC

[50] H. Spiess, "Conical volume-delay functions," *Transportation Science*, vol. 24, 1990. [Online]. Available: https://api.semanticscholar.org/CorpusID:55843734

[51] G. A. Davis and H. Xiong, "Access to destinations: Travel time estimation on arterials," Aug 2007. [Online]. Available: https://conservancy.umn.edu/items/b704d677-4f72-46e6-89fc-c5511ae89aaf

[52] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, Dec 1959. [Online]. Available: https://doi.org/10.1007/BF01386390

[53] M. Lee, H. Barbosa, H. Youn, P. Holme, and G. Ghoshal, "Morphology of travel routes and the organization of cities," *Nature Communications*, vol. 8, no. 1, Dec. 2017. [Online]. Available: http://dx.doi.org/10.1038/s41467-017-02374-7

[54] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, p. 779–782, Jun. 2008. [Online]. Available: http://dx.doi.org/10.1038/nature06958

[55] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, and M. Tomasini, "Human mobility: Models and applications," *Physics Reports*, vol. 734, pp. 1–74, 2018, human mobility: Models and applications. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037015731830022X

[56] T. Louail, M. Lenormand, O. G. Cantu-Ros, M. Picornell, R. Herranz, E. Frias-Martinez, J. J. Ramasco, and M. Barthelemy, "From mobile phone data to the spatial structure of cities," *Scientific Reports*, vol. 4, p. 5276, 2014. [Online]. Available: https://doi.org/10.1038/srep05276

[57] J. Tang and H. R. Heinimann, "A resilience-oriented approach for quantitatively assessing recurrent spatial-temporal congestion on urban roads," *PLOS ONE*, vol. 13, no. 1, pp. 1–22, 01 2018. [Online]. Available: https://doi.org/10.1371/journal.pone.0190616

[58] L. Alexander, S. Jiang, M. Murga, and M. C. González, "Origin–destination trips by purpose and time of day inferred from mobile phone data," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 240–250, 2015, big Data in Transportation and Traffic Engineering. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X1500073X

[59] C. M. Schneider, V. Belik, T. Couronné, Z. Smoreda, and M. C. González, "Unravelling daily human mobility motifs," *Journal of The Royal Society Interface*, vol. 10, no. 84, p. 20130246, 2013. [Online]. Available: https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2013.0246

[60] M. Lenormand, M. Picornell, O. G. Cantú-Ros, A. Tugores, T. Louail, R. Herranz, M. Barthelemy, E. Frías-Martínez, and J. J. Ramasco, "Cross-checking different sources of mobility information," *PLOS ONE*, vol. 9, no. 8, pp. 1–10, 08 2014. [Online]. Available: https://doi.org/10.1371/journal.pone.0105184

Acknowledgments

I would like to thanks to my advisors, Andrea and Riccardo, for the opportunity and for encouraging me to perform this journey. As well as the whole CHuB group that more than a group is a family. Also, I would like to thank Pitufibanda for their advices and the time. To Rafael Correa for believing in young people. And finally, I have to thank to all the people that has been supporting me these years.