EMERGENT HIERARCHICAL PROPERTIES OF A TRANSPORTATION NETWORK

ANDRÉS GÓMEZ LIÉVANO

Department of Industrial Engineering

APPROVED:

Roberto Zarama, Chair, Ph.D.

Jorge Acevedo, Ph.D.

Juan Alejandro Valdivia, Ph.D.

Alain Gauthier Sellier, Ph.D.
Decano de la Facultad de Ingeniería

To my

MOTHER, FATHER, and SISTER

with love.

EMERGENT HIERARCHICAL PROPERTIES OF A TRANSPORTATION NETWORK

by

ANDRÉS GÓMEZ LIÉVANO

THESIS

Presented to the Faculty of the Graduate School of

Universidad de los Andes

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Industrial Engineering

UNIVERSIDAD DE LOS ANDES

June 2008

Acknowledgments

I would like to express my deep-felt gratitude to my advisor, Dr. Roberto Zarama Urdaneta of the Industrial Engineering Department at La Universidad de los Andes in Bogotá, for his advice, encouragement, enduring patience, constant support, and friendship.

I am very grateful to my co-advisor, Dr. Juan Alejandro Valdivia, for receiving me into his research group, and for his guidance. He is such an admirable person and physicist, and I am very happy of having known him.

I also wish to thank Jorge Villalobos for the fruitful conversations, from which this work was enriched. Also for his open disposition, and for his guidance in some computational aspects this work needed.

Additionally, I want to thank the Sistemas Complejos group of La Universidad de Chile for receiving me and for providing me the means to work in the best of the academic environments.

Finally, I must thank my parents. All I am and have is thanks to them. I feel deeply and eternally grateful for having them, together with my sister, as my family. Their unconditional love and support have been my own strength all my life.

Abstract

Social transportation systems are remarkable examples of Complex Systems. Understanding their patterns and their dynamics in a holistic and global way is an obligated duty if we humans want to ever be capable of managing our social systems in an efficient way, friendly with our surrounding ecosystems and with ourselves.

The general purpose of this research project is to give a first and modest ground of understanding concerning transport phenomena in networks, from a complex system perspective. Topological properties of transportation networks have been the most studied ones, and the spatial ones have been often left aside. In this work it is presented and suggested a mathematical model that replicates certain characteristics of spatial transportation networks. We use the measure of betweenness centrality to approximate the traffic within the network, and we show that holes are crucial actors that affect the transport phenomena. Some important results are observed from the statistics of the betweenness centrality.

General guidelines for traffic assessment and road network management are given in light of the results shed by the model.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Organisms, artifacts, and organizations are all evolved structures. Even when human agents plan and construct with intention, there is more of the blind watchmaker at work than we usually recognize. What are the laws governing the emergence and coevolution of such structures?

— Stuart Kauffman, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity* [28]

Managing today's social systems is becoming extremely difficult; the number of elements they are made of is large and grows in time. Information, people, goods, and services are traveling around the globe in short time scales, and our human activities are now affecting the planet in a global and rapid manner. In the last fifty years we have come to see that social systems are emerging as truly organic, robust complex systems [36, 46, 18].

In order to steer our social systems towards a state of human and nature coexistence in a harmonic and efficient way, we have to understand them as holistic entities. The reductionist approach, i.e. studying the parts to infer the properties of the whole, is not enough.

Complexity has emerged in the last thirty years as a new kind of science [24, 28, 36], changing the traditional way of scientific thinking, from one where the properties of the whole system are considered to be the sum of the properties of its constituents, to a new point of view where one recognizes that in a system "the whole is more than the sum of its parts"[1] [48, 21, 14, 24, 51, 2]. Fractality, chaos, self-organization, hierarchical

[1]In [1], "the whole becomes not only more than but very different from the sum of its parts.". An

modularity, criticality, adaptation, and evolution, are among many of the concepts with wide applications in this new paradigm.

In the spatial networks considered in this research project edges do not define abstract relations (such as in friendship, collaboration, sexual, food, or acquaintances networks, where the edges represent some type of predefined interaction between individuals), but are real physical connections [10, 17]. Through these networks physical objects flow and hence it is essential to understand the underlying mechanisms governing the structure and function of these networks in junction with the properties of the systems in which they are embedded. Some of these spatial networks include neural networks, electric power grids, and all type of transportation systems such as biological circulatory systems, river, airport, and street networks (for more information see refs. in [10]). The study of transport phenomena in these spatial networks must take into account the properties of the objects in which they are embedded. The general problematic this research project seeks to investigate is the fact that many social and natural networks exist in space, and are embedded in physical systems which in turn have complex properties. We consider that this problem is not acknowledged in traditional planning and management activities [17].

In this work, we study the particular case of a simple planar network with "holes" inside, inspired by the fact that many transportation networks are embedded in objects that in turn display fractal properties, like road street networks that are embedded in cities that are considered to be fractal objects [7, 15, 41] that fill space in a non-trivial way. Although there is recent evidence that hierarchies emerge naturally in transportation networks [54], we argue that "holes" in a network are one additional plausible cause for the emergence of hierarchies. The question we address then is: how the size distribution of free spaces (e.g. green or un-built spaces in the case of a city road network) in a regular grid affects the transport properties in its nodes? The importance of an edge or a node can be characterized by the number of circulating objects passing through it within some time interval. This can be roughly approximated with the measure of betweenness centrality

interesting discussion about the meaning of *hierarchies* in science is presented in this reference.

[31]. In [31] it is stated that "the inherent structure of the road network topology itself has a tremendous effect on the emergence of road hierarchies". Studying specific mathematical models developed here, we go one step further and argue that free spaces in a regular planar network are one of the causes of hierarchy emergence in transport systems. Our final aim is to add a new layer of insight regarding decision-making, analysis, and management of transport systems.

This document is divided in six chapters, including this introduction. Chapter 2 contains a brief summary of the concepts and tools used in the complex systems approach that are applied in this work, and a brief state of the art in the area of transport and hierarchy in networks is included. A detailed description of the model used to validate the hypothesis is presented in Chapter 3, and in Chapter 4 we analyze it. In Chapter 5 we suggest some applications of the results shed by the model. Finally, we draw some conclusions in Chapter 6.

Chapter 2

Theoretical background

Roughly, by a complex system I mean one made up of a large number of parts that interact in a non-simple way. In such systems, the whole is more than the sum of the parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole.

— Herbert A. Simon, *The Architecture of Complexity* (1962) [48]

This chapter briefly presents and describes the mathematical tools and concepts used in the rest of the document. Complexity science is still a young area of research, and consequently no solid mathematical formalism exists yet that can handle many-component systems. The application of statistical physics to complexity phenomena is considered to be a promising attempt for understanding them in a theoretical and general manner [32, 46, 51]. Statistical mechanics has accomplished the development of a mathematical formalism that study, analyze, and understand many-particle systems at equilibrium (i.e. lots of particles interacting, in some way, between them and whose statistical properties, such as the number of particles, temperature and energy, do not change significantly in time), with the so-called *canonical ensembles* and the *partition function.* This discipline has been particularly successful in explaining the macroscopic behavior of some systems, such as phase transitions, from a "bottoms up" approach. Nevertheless, the challenge still remains: understanding open and far from equilibrium systems, such as economies, biology, and societies.

2.1 Power laws, hierarchies, fractality, and self-similarity

Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line.

— B.B. Mandelbrot, *The Fractal Geometry of Nature* [33]

Many social and natural systems exhibit emergent phenomena that are characterized by many layers of hierarchical organization, with complex structural, temporal and functional properties [33, 51, 46]. Very often, the complex dynamics that are displayed in a system consisting of many interacting elements are a result of self-organizing principles.

In the last 30 years the scientific community has witnessed an exploding new area of research called Complexity Science. It stands as a new paradigm for approaching physical, natural and social systems as holistic entities. Although there is no consensus about what a *complex system* is, it is often said that "Complex systems are systems with multiple interacting components whose behavior cannot be simply inferred from the behavior of the components.[...] Complex systems science is a new field of science studying how parts of a system give rise to its collective behaviors, as well as how the system interacts with its environment. Social systems formed by people, the brain formed by neurons, molecules formed by atoms, the weather formed by air flows— these are all examples of complex systems. By using mathematics to focus on pattern formation, and the question of parts, wholes and relationships, the field of complex systems cuts across all the disciplines of science, as well as engineering, management, and medicine."[1].

In the complex system approach, fractals, scaling, and power laws are ubiquitous. These concepts imply that activity in the system occurs through all scales, and this means that large events often play an important if not a leading role [51]. Events in complex systems

[1] citation of New England Complex Systems Institute, http://necsi.org/, by [46]

are seldom characterized by Gaussian distributions and are instead better described by *heavy tailed distributions* [50].

A quantity X is said to follow a power law distribution if the probability $p(x)\mathrm{d}x$ of measuring a particular value between x and $x + \mathrm{d}x$ varies as

$$\begin{aligned} p(X = x)\,\mathrm{d}x = p(x)\,\mathrm{d}x &= Cx^{-\alpha}\,\mathrm{d}x \\ &= \left(\frac{\alpha - 1}{x_{\min}}\right)\left(\frac{x}{x_{\min}}\right)^{-\alpha}\mathrm{d}x \end{aligned} \tag{2.1}$$

with C the normalization constant[2], $x_{\min}$ the minimum value of x for which the power law behavior holds, and the exponent α a positive number, often called the *critical* exponent [35]. In this work we will also be using the cumulative distribution $P_{\geq}(x)$, i.e. the probability of measuring a value greater than or equal to x:

$$\begin{aligned} P(X \geq x) = P_{\geq}(x) &= \int_x^{\infty} p(x')\,\mathrm{d}x' \\ &= \left(\frac{x}{x_{\min}}\right)^{1-\alpha}. \end{aligned} \tag{2.2}$$

There is a simple method for plotting the cumulative distribution called the *rank-ordering method* [51]. If the data contains N observations of the variable v, the method consists in sorting the N measured values in decreasing order $v_1 \geq v_2 \geq \ldots \geq v_N$. The integer part of $NP_{\geq}(v_n)$ is the expected number of values larger than or equal to v_n. But v_n is precisely the nth largest observed value and hence the following relation holds:

$$NP_{\geq}(v_n) = n. \tag{2.3}$$

The number n is called the *rank* of the value v_n. For plotting a cumulative distribution of the sample it is then sufficient to plot the ordered values v_n in the x-axis against the rank n divided by N in the y-axis. Cumulative histograms like this are sometimes called *rank/frequency plots*.

Estimating the exponent α from data is not a trivial task. If some measure is power law distributed, the histogram will appear like a straight line with a negative slope in a double

[2] $\int_{x_{\min}}^{\infty} p(x)\mathrm{d}x = 1$.

logarithmic scale plot (these are called *log-log plots* and the slope determines the exponent of the power law), and this suggested, for a long time, linear regressions methods for fitting the distribution. However, these methods of least-square fitting of the log-log transformed data grossly underestimates the standard errors of the slope (exponent) parameter [51], and the estimated parameter is not close to its true value (an illustration of this effect is shown in [35]); on the other hand, *maximum likelihood* estimation using *Bayesian* methods are a simple and reliable way for determining the exponent α [20, 35]. These methods use the following formula (see Appendix A.2), sometimes called the *Hill estimator* [51]:

$$\alpha = 1 + n\left(\sum_{i=1}^{n} \ln \frac{v_i}{v_{\min}}\right)^{-1}. \tag{2.4}$$

Again, $v_{\min}$ is usually not the minimum measured value but the minimum value for which the power law holds, and n is the corresponding number of values that are used in the estimation. The expected statistical deviation σ on Eq. (2.4) is (see Appendix A.2)

$$\sigma = \sqrt{n}\left(\sum_{i=1}^{n} \ln \frac{v_i}{v_{\min}}\right)^{-1} = \frac{\alpha - 1}{\sqrt{n}}. \tag{2.5}$$

Power law distributions are known to arise in a large number of physical, biological, economic and social systems [35, 24], and literature abounds with references to such distributions as "the signature of complexity" [39, 2, 51, 35, 28, 29]. Complexity pervades the natural world we see [49]; strong correlations, large fluctuations, self-similarity, extreme events, risks, and hierarchical organization are present in our world's phenomena [23, 51], and that is the reason why power laws are ubiquitous [35]. The following characteristics are some among several that determine this behavior:

- If X follows Eq. (2.1), the variance $\sigma^2 \equiv E(X^2) - (EX)^2$ does not exists for $\alpha \leq 3$. This is important since many mathematical models (e.g. Black-Scholes option-pricing formula) assume a finite variance [52, 25]. The absence of a finite variance, opens the possibility that a fluctuation or a perturbation affects all the system that is under observation [3, 4].

- the average $\langle x\rangle \equiv EX$ does not exist for $\alpha \leq 2$. This is also important, given that the traditional statistical analysis frequently uses averages. It is a known fact ([35] and references therein) that wealth in a country follows a power law, but by the above observation, wealth in a country *may* not have an average, and indicators such as GDP per capita (*Gross Domestic Product per capita income*) may be misleading.

- Power laws are also called *scale-free* laws, due to the property that a scale increase in the measure of a variable x by a factor k in the distribution is, in relative terms, independent of the value x. This means that if $f(x) = Cx^{-\alpha}$, then $f(kx)/f(x) = k^{-\alpha}$ which is clearly independent of x. This property implies that if, for example, the spatial correlation function between two measures follows a power law (such as in first order phase transitions in which the *correlation length* or *characteristic scale* diverges at the critical point), then the strength of the correlation is independent of scale, or is scale-free, in relative terms. In general, the power law is the only distribution that is the same at whatever scale we look at [35].

- Lévy laws are *stable distributions* whose asymptotic behavior approaches a power law. The sum of independent identical distributed (i.i.d.) power law random variables tend to a Lévy law [51].

Power laws and complexity are strongly related with the concept of *hierarchy* [39]. Hierarchies arise in nature in different forms: spatial, temporal, ontological, structural, or functional hierarchies are some among many. In the context of this research project, by hierarchy we mean the large range of scales in which a same phenomenon affects some system. The recognition of this emergent property, brings forward the possibility of observing complex behavior in a system through many scales.

Associated with the concepts of hierarchy and power laws are the concepts of *fractality* and *self-similarity*. Since Euclid until the 20^{th} century, the scientific community regarded the shapes we see in nature, such as mountains, coastlines, rivers or clouds, as simply irregular, amorphous, random and disordered. However, in the nineteen-seventies and nineteen-eighties the *fractal* term was coined, and a deep hidden order was discovered: *scale-invariance* and *fractionary dimensions*.

According to the inventor of the word, Benoît Mandelbrot [33], the term fractal refers to objects that display self-similar features (exact or statistical) across a large range of scales that are described by a *fractal dimension*. It is not intended here to give an exhaustive revision of fractal concepts, but only a brief exposition of their main characteristics.

Rigorously, there are several fractal dimensions that can be defined (some of them equivalent). The dimension we are going to identify is given by the following relation between the length ϵ and the number $N(\epsilon)$ of times we use ϵ to measure the object size M. It is called the *box-counting dimension* D:

$$N(\epsilon) \propto \frac{M}{\epsilon^D} \sim \epsilon^{-D}. \tag{2.6}$$

This relation holds for regular objects: $N(\epsilon) \propto L/\epsilon^1$ for a line, $N(\epsilon) \propto A/\epsilon^2$ for a smooth surface, $N(\epsilon) \propto V/\epsilon^3$ for a 3D object, and this definition is thus coherent with euclidean geometry. The dimension is estimated then as

$$D = \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln(1/\epsilon)}. \tag{2.7}$$

We note that the fractal dimension is nothing but the exponent in a power law, and clearly determines scale-free properties. Fractal dimensions give us information about how an object is occupying space. A dimension $D \in (1, 2)$ means that the object is a line that somewhat "fills" the plane, and a dimension $D \in (2, 3)$ means a surface that in some way is "filling" space; typically what we see is a fragmented line and a rough surface, respectively. In addition to this "irregular" nature of fractal objects, is the property of self-similarity mentioned above. This is a consequence of the scale-free property of the power law in Eq. (2.6). An illustration of a fractal with dimension $D = \ln 4/\ln 3 \approx 1,262$ is shown in Fig. 2.1.

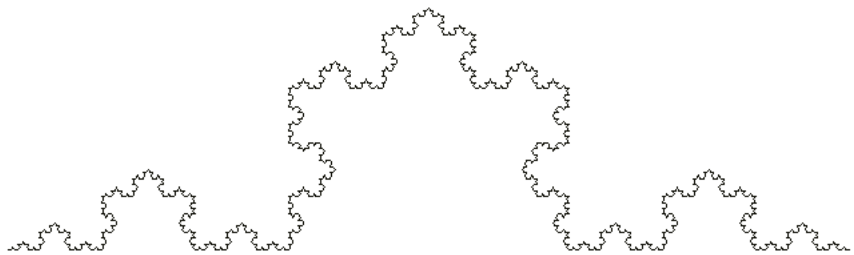


Figure 2.1: Illustration of the Koch curve.

2.2 Networks

The principal aim of this section is to define the notation that we will use through this document, along with some historical background on graph theory.

Graph theory was developed by Euler in 1736 while solving the Königsberg's Bridge Problem, and it grew rapidly through the years as a solid and developed branch of pure Mathematics. However, it has received plenty of attention from the complexity science community, since there is accumulated evidence that systems across many disciplines (when represented using nodes as actors and links as relations between them) display the same statistical archetypal properties. Two of the most known phenomena are the *small-world effect* and the *scale-free structure*. The study of graphs has shifted since the second half of 20$^{\text{th}}$ century from a formal theory that studies exact mathematical properties to a field of research that investigates the statistical properties of *complex networks* that represent natural and social systems [36].

In formal terms, a network can be represented as a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$. A graph is defined as a collection of a nonempty set $\mathcal{V} = \{1, \ldots, i, \ldots, n\}$ of vertices and a set $\mathcal{E} = \{e_1, \ldots, e_k, \ldots, e_m\}$ of edges that connect pairs of nodes in $\mathcal{V}$. An edge $e_k = (i, j)$ exists if there is a link between node $i \in \mathcal{V}$ and node $j \in \mathcal{V}$.

A network can be represented in mathematical terms using what is called the *adjacency matrix*. This is a square $n \times n$ matrix, where the elements a_{ij} are equal to 1 if there is a link coming from the vertex i to the vertex j, and 0 if i and j are disconnected[3]. The adjacency matrix of an undirected graph is symmetrical: $a_{ij} = a_{ji}$. If we assume that self-loops are absent, then the diagonal elements are equal to zero: $a_{ii} = 0$.

In the following list we mention some of the important concepts in the study of networks.

- *Degree of node i*: The number of edges connected directly to the node i. It is usually denoted by k_i, and is calculated as

$$k_i = \sum_{j=1}^{n} a_{ij}. \tag{2.8}$$

  It is one of the measures of how central a node is.

- *Path*: Sequence of vertices $v_1 v_2 \dots v_{s-1} v_s$, such that there is an edge connecting vertex v_i with v_{i+1} for all $i \in \{1, \dots, s-1\}$.

- *Geodesic path*: The shortest path through the network between two nodes. There may be multiple geodesic paths between two nodes. The *geodesic distance*, or simply the *distance*, between two nodes, is defined as the length of the geodesic path connecting them.

- *Clustering coefficient*: Measure of the probability in a network that nodes i and k are connected, given that node i is connected to j, and j to k. It is denoted by C and can be written as [34]

$$C = \frac{6 \times \text{number of triangles in the network}}{\text{number of paths of length two}}. \tag{2.9}$$

- *Betweenness centrality of node i*: Centrality measure of a node. It counts how many geodesic paths pass through node i. It is given by

$$C^B(i) = \sum_{j,k \in \mathcal{V}, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}}, \tag{2.10}$$

[3]In principle, the value of a_{ij} could be any non-negative value, representing the weight of the edge. Since generally one is interested in the topological properties of a network, the values are usually taken as 0 or 1.

where n_{jk} is the total number of geodesics between nodes j and k, and $n_{jk}(i)$ the number of those geodesics passing through i.

We encourage the reader to review Ref. [36] for a wide and extensive compilation of the most important results in the study of networks.

2.3 Transportation networks

Our aim in this section is to introduce the concept of multiple centrality assessment, which sets a ground methodology for the study of transportation networks.

The methodology presented here apply to transportation networks in general, but given that our interest lies in the study of *social* systems, and since urban studies have received a great amount of attention from the network scientific community, we focus thus in city road networks.

In 1984, the notion of 'space syntax' was born with the seminal work of Bill Hillier and Julienne Hanson in the social science study of the use of space and architecture [22]. This methodology changed the main stream of research in urban studies and design from a theoretical one to a more practical one, with a coherent and consistent application of network theory to urban spaces [38]. The conventional space syntax analysis of road patterns relies on what is called the *dual* network representation, in which nodes represent streets and edges represent intersections that link different street pairs. This form of representation determines topological relations rather than spatial relations, and distance is measured in consequence in 'steps' rather than in meters. This methodology has been useful to identify some structural properties of social spaces that were hidden, revealing, for example, patterns that differentiate old cities from new ones [22]. One of the shortcomings of the dual representation is that it is not clear how streets should be turned into nodes, i.e. it is not clear whether a long street with the same name or a single street segment should be turned into a node.

By contrast to the dual representation, the *primal* representation stands on a more intuitive ground and a direct abstraction of the system develops from it: it consists of nodes that represent intersections or point locations and edges that represent physical links between this locations.

In Ref. [38] the concept of *multiple centrality assessment* (MCA) is defined as a methodology for the primal analysis of *centralities* on urban street systems. By this they mean that there is no single centrality measure that can characterize a spatial system: “Centrality is a multifaceted concept that, in order to measure the ‘importance’ of single actors, organizations, or places in complex networks, has led to a number of different indices” [38]. Porta *et. al* [38] classify such indices into four categories for being central: being near, being between, being straight to, and being critical for the others. This indices translate directly into the measures of closeness centrality, betweenness centrality, straightness centrality, and information centrality (the reader interested in the corresponding equations should revisit refs. [38, 13, 37]). In the same spirit, they continue and argue that

> A new approach to the network analysis of centralities in geographic systems is therefore appearing. Its three pillars are (1) primal graphs; (2) metric distance; (3) many different indices of centrality. As such, we may well name it multiple centrality assessment. Offering a set of multifaceted pictures of reality, rather than just one, MCA leads to more argumentative, thus less assertive, indications for action.

Results presented in [13, 37, 38] suggest that “the primal approach is a more comprehensive, objective, realistic, and feasible methodology for the network analysis of geographic systems such as those of streets and intersections” [38]. This is an assertion relevant to our work since we are going to show that spatial structure underlying transportation networks affect the traffic inside it.

In [31, 54, 42, 47] it is shown in several different ways that street hierarchies emerge naturally in a road network, from a primal approach perspective. Our aim in this work is to show a plausible mechanism for such emergence, and to analyze it.

For the purposes of this research project, the measure of betweenness centrality will be enough to validate our hypothesis which states that the underlying physical properties of an object affect the distribution of traffic in the transportation network embedded in it.

2.4 Complex systems' standard methodology

The principal tool we can use to detect regularities in a system is the construction of probability distributions. When analyzing complex systems one must study the system as a whole and probability distributions are the first quantitative approximation in this quest [50].

We refer the reader to [50] for a complete discussion regarding the use of probability distributions in the complex system approach. However, we cite the following text on which the standard methodology is described:

> A central property of a complex system is the possible occurrence of coherent large-scale collective behaviors with a very rich structure, resulting from the repeated non-linear interactions among its constituents: the whole turns out to be much more than the sum of its parts.
>
> [...] [A] first standard attempt to quantify and classify the characteristics and the possible different regimes consists in
>
> 1. identifying discrete events,
> 2. measuring their sizes,
> 3. constructing their probability distribution.
>
> The interest in probability distributions in complex systems has the following roots.
>
> - They offer a natural metric of the relative rate of occurrence of small versus large events, and thus of the associated risks.
> - As such, they constitute essential components of risk assessment and prerequisites of risk management.

- Their mathematical form can provide constraints and guidelines to identify the underlying mechanisms at their origin and thus at the origin of the behavior the complex system under study.
- This improved understanding may lead to better forecasting skills, and even to the option (or illusion(?)) of (a certain degree of) control [45, 44].[4]

In Chapter 4 this methodology becomes evident; there we characterize betweenness centrality measurements made from our model through its distribution.

[4] Taken from [50, page 3]

Chapter 3

Model

Most if not all complex systems are fractal, displaying organization or order on all scales. [...]
The central issue is how fractal systems emerge and evolve [...]
— Michael Batty, *Cities and Complexity* [6]

In this chapter we present a simple model of a transportation network from which we measure the distribution of betweenness centrality as an indicator of how often a node is visited. Corridors of high betweenness emerge when irregularities are included. Before we explain the model, some conceptual aspects are discussed.

3.1 Conceptual framework

We assume that two types of processes shape the natural and social systems we observe: *centralized* processes and *self-organized* processes. By centralized processes we mean those (internal or external) mechanisms that affect some part of a system in a direct manner, independent of the response of the system itself. One such example are the orders given by the High Command of a military group: all the system (the collection of soldiers under his command) is affected by his decisions; in principle the soldiers' reactions do not affect those orders. By self-organized processes we mean those that come from the interaction of the elements of the systems itself. For example, fractality in a city is a property that emerges as a consequence from the self-organizing processes that underlie urban growth [16, 15, 7]. Properties of a system whose underlying mechanism are self-organizing processes are, usually, hard to control, and are intimately related to non-linear dynamics. Complexity science helps us to identify which processes are product of self-organizing principles, and which are not.

The distinction made above is necessary since we are going to model transport systems that are subjected to both types of processes; the properties of the objects in which they are embedded come from self-organized and decentralized processes, but the planning and management we give to them are centralized. It is in this context that we suggest, for example, that pavement, traffic light times, and the geometry of streets in a city, must take into account the properties that emerge from the road network (Chapter 5).

The following models are meant to give intuition and insight about the phenomenon of transport in spatial networks. They do not intend to give exact numerical results with direct application to real systems, and we do not attempt to translate the results presented in this research project into action as such. The models do imply that some mechanisms are at play in transport networks that should be accounted for in management and design, and we do insist that in understanding transportation networks "we should seek a baseline of intelligence that has its origins in complexity theory" [6].

3.2 Description of the model

In the models presented in this section the *topological* and *metrical* properties of the transportation network are strongly related.

We begin with a regular grid of $n \times n$ nodes, as shown in Fig. 3.1[1]. This is the simplest representation of a planar transportation network we will work with. This representation, in which intersections are represented as nodes and "streets" as edges is called the "primal" representation [38] (see Section 2.3). To keep the model simple, we assign a unit length to every edge.

[1]The stretched appearance of the networks in the figures in this document is a visual effect only, due to the graphical functions used to represent them.

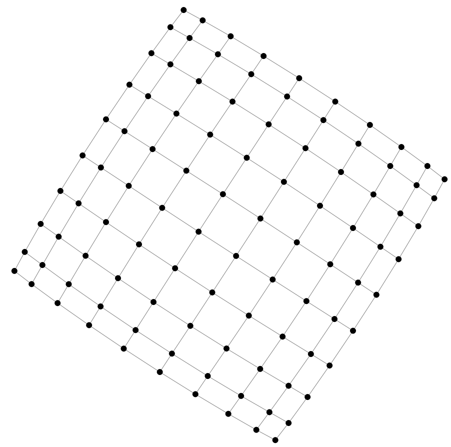


Figure 3.1: Example of a 10×10 regular grid.

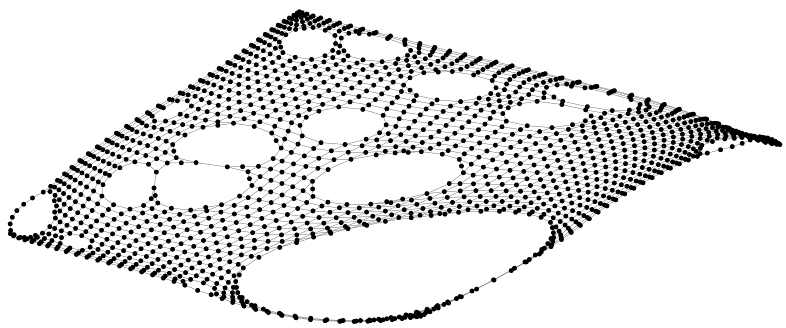
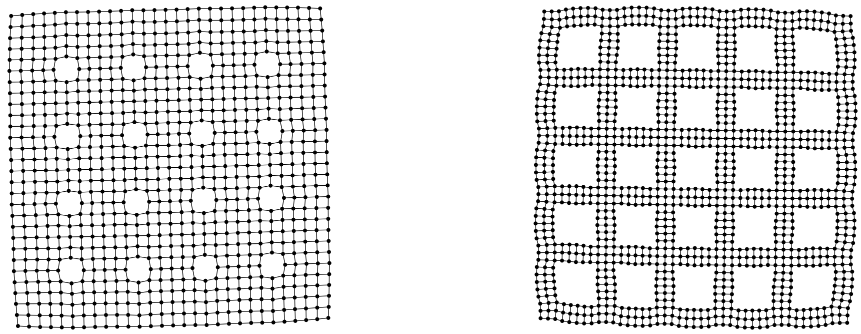


Figure 3.2: Example of a 40×40 network with holes whose size follow a power law.

The next step is to add holes to the grid (by a hole we mean the absence of nodes in some spatial location of the network). The addition of holes (or removal of nodes in practical terms) serves not only to imitate the fact that cells sizes vary in natural networks, but also to take into account the spatial heterogeneity distribution of nodes. Thus, the resultant network aims to model in a simple and straightforward way a planar transport network. In Fig. 3.2 an example of a grid with holes is shown.

There are many ways in which holes can be introduced into a regular grid, but we will study only three of them: equal sized and regularly spaced holes, equal sized and randomly spaced holes, and power law sized and randomly spaced holes. The full explanation and



(a) (b)

Figure 3.3: Regular model: (a) Sixteen holes ($n_c = 4$) of size one ($h_s = 1$) spaced by five nodes ($r_s = 5$). (b) Twenty five holes ($n_c = 5$) of size 25 ($h_s = 5$) spaced by three nodes ($r_s = 3$).

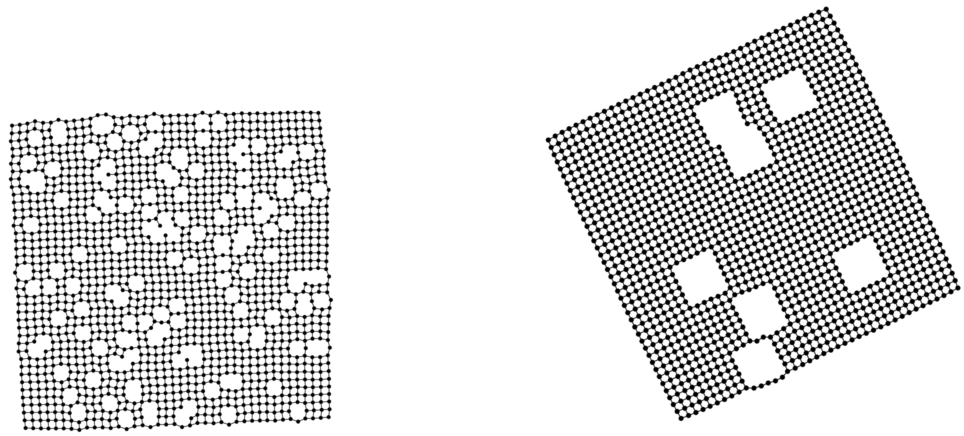
description of these procedures will be presented in the following sections.

3.2.1 Equal sized and regularly spaced holes

The first procedure we propose to introduce holes within the grid, is in a regular manner; we introduce holes of equal sizes, in a periodic way. This could be compared with a highly planned transport network. The holes could represent green spaces, in the case of a road street network.

In the model, we can vary the size of the holes h_s^2 (which, in practice, is the number of nodes removed from the graph), the spacing between them r_s (i.e. the number of nodes separating free spaces), and their total number n_c^2 in the grid. Two examples are shown in Fig. 3.3.

This way of generating a network, is going to be useful as a reference from which we can compare the following two procedures of node removal. From this point forward, we will refer to this model as the *regular model*.

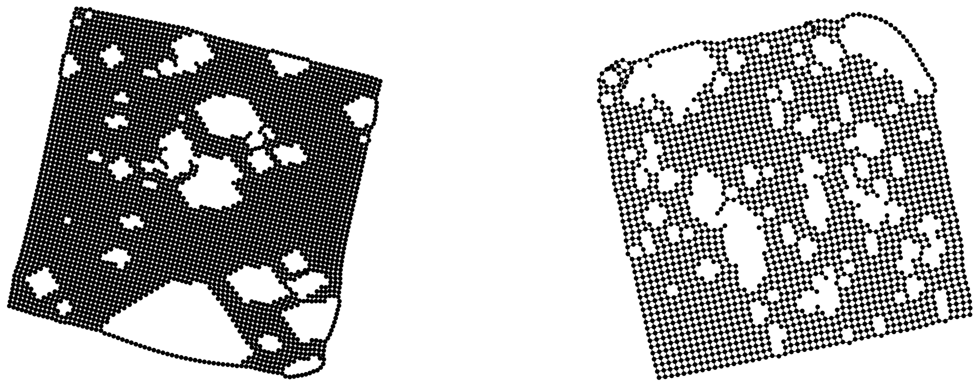


(a) (b)

Figure 3.4: Random model: (a) Network with parameters $n = 40$, $h_s = 1$, and $f = 0.9$. (b) Network with parameters $n = 40$, $h_s = 5$, and $f = 0.9$. Note that both networks have the same $n^2 f \approx 1440$ nodes.

3.2.2 Equal sized and randomly spaced holes

An additional feature we can introduce in the model, is to randomize the placement of holes. In this manner, we no longer have control on the spacing between them. The parameters that the model receives are the size of the grid n (the total number of nodes before the insertion of holes is n^2, given that we always work with grids with a square form), the size of the holes h_s (the total number of nodes removed in each hole is h_s^2), and the fraction f of nodes that remain after the placement of the holes (in practice, then, the number of nodes that are removed from the grid is $n^2(1 - f)$). In Fig. 3.4 are shown two examples of networks produced by this model. Also notice in the figure that the model allows for different holes to be joined; this juxtaposition creates holes with greater size. We will refer to this model as the *random model*.



(a)

(b)

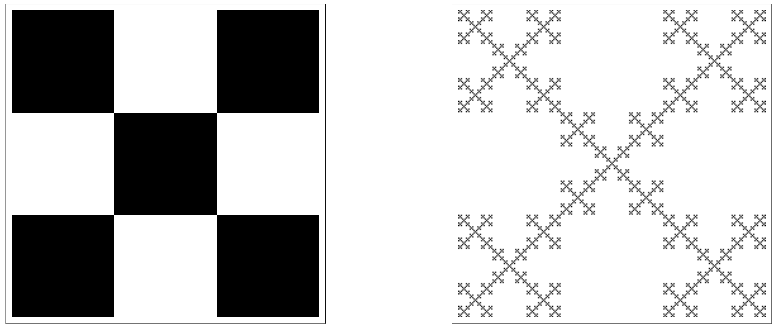
Figure 3.5: Fractal model: (a) Network with parameters $n = 70$, $\alpha = 1.4$, and $f = 0.8$. (b) Network with parameters $n = 40$, $\alpha = 2.3$, and $f = 0.8$.

3.2.3 Power law sized and randomly spaced holes

This is the last procedure we employ to remove nodes in the grid. This time, the sizes of holes are power law distributed, and they are randomly positioned within the lattice. The parameters that the model receives are the size of the grid n, the exponent α of the power law, and the fraction f of nodes that are not removed. Two different networks created in this way are shown in Fig. 3.5[2]. We will refer to this last model as the *fractal model*.

The motivation for distributing the holes in this way is the fact that many networks are embedded in fractal objects, and in fractals a hierarchy of "white" or "free" spaces emerge. We can see clearly this phenomenon for the case of a Sierpinski carpet in Fig. 3.6.

[2]Because we are going to work with exponents smaller than three, $\alpha < 3$, the sizes of the holes do not have a finite variance. And in some cases, $\alpha < 2$, and thus there is not a size average. For these reasons, there are cases where it is impossible to control the input parameter f. Also, it has to be noted that the networks are not too large, hence there are not to many holes to add, and in consequence, the sample of sizes do not converge strictly to a power law distribution.



(a) (b)

Figure 3.6: (a) First iteration in the creation of the Sierpinski carpet. We see four major white spaces. (b) Sierpinski carpet in its sixth iteration. We see already a hierarchy of white spaces: few large, and many small ones.

The number of white spaces $N^{fr}(\lambda_n)$ of size λ_n in a fractal scales like a power law

$$N^{fr}(\lambda_n) \sim \lambda_n^{-\delta} \tag{3.1}$$

where δ is an exponent which may differ from the box-counting dimension D of the fractal itself [16]. For the case of the Sierpinski carpet shown in Fig. 3.6, $\delta = D = \ln 5 / \ln 3 \approx 1.46$. The fact that fractals display this hierarchy in their free (non-occupied) spaces, motivates us to implement a simplification of this phenomenon into the model of the transportation grid.

This phenomenon is clearly seen in cities, where green or un-built spaces exist in a large range of scales: there are a lot of small un-built spaces, such as backyards of the order of $1\,m^2$, and few large (many times larger than the small ones) green spaces of the order of $10^4\,m^2$, such as parks.

Generating numbers that are power law distributed is not a trivial task, but a function can be built in such a way that a random number Y uniformly distributed between one and zero, transforms into a number with a power law distribution.

Let X be the number we want to follow a power law distribution $p(x) = Cx^{-\alpha}$. If Y is a number that follows a uniform distribution between zero and one, we look for a function

g such that $X = g(Y)$. First, we notice that

$$P_{\leq}(y) = y, \tag{3.2}$$

and

$$P_{\leq}(x) = \int_{x_{\min}}^{x} p(x')\mathrm{d}x' = \frac{C}{1-\alpha}\left(x^{-\alpha+1} - x_{\min}^{-\alpha+1}\right) \tag{3.3}$$

$$= P(X \leq x) = P(g(Y) \leq x) = P(Y \leq g^{-1}(x)). \tag{3.4}$$

Using Eq. (3.2) and replacing the normalization constant $C = (\alpha - 1)x_{\min}^{\alpha-1}$, we conclude:

$$\frac{C}{1-\alpha}\left(x^{-\alpha+1} - x_{\min}^{-\alpha+1}\right) = g^{-1}(x) = y$$
$$x = x_{\min}\left(1-y\right)^{\frac{1}{1-\alpha}}. \tag{3.5}$$

In the model, we choose $x_{\min} = 1$ to be the minimum size of a hole.

Chapter 4

Analysis and discussion

Here we analyze the models of transportation networks proposed in Chapter 3. We show that an exponential and a power law tail, in the distributions of betweenness centrality, emerges for the models presented in Sections 3.2.2 and 3.2.3, respectively. We believe that the results presented in this section affect the management and planning activities concerning transportation networks.

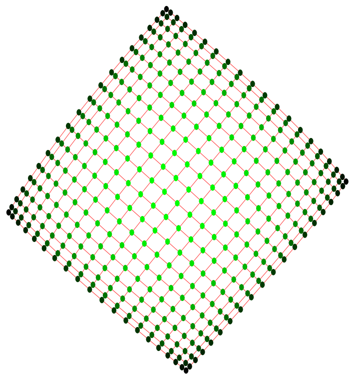
4.1 Betweenness centrality

Among the measures that indicate central and important elements in a network (see Section 2.2), such as the *degree* or *closeness*, *betweenness centrality* has the advantage that it gives information of the global network as well. It is based on the idea that a node is central if it is a participant of many of the shortest paths (*geodesics*) that connect pairs of nodes. The betweenness centrality of node i is defined as [13, 53, 34]

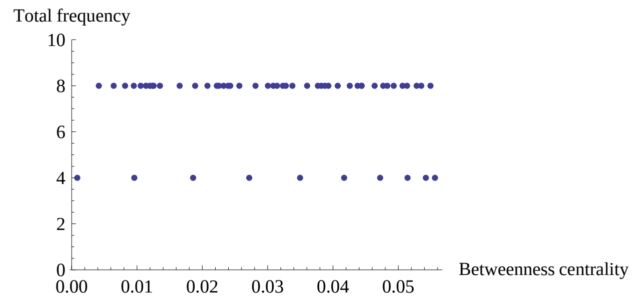
$$C^B(i) = \frac{2}{(N-1)(N-2)} \sum_{j,k \in \mathcal{V}, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}}, \tag{4.1}$$

where n_{jk} is the number of geodesic paths between nodes j and k, and $n_{jk}(i)$ is the number of geodesic paths connecting j and k and passing through node i. The factor $\frac{2}{(N-1)(N-2)}$ is a normalization term: very often [53] quantities are normalized by their highest possible value; thus, in a network with N nodes, $(N-1)(N-2)/2$ is the highest value of betweenness centrality a node can have.[1]

[1]This happens in a *star* graph, where all vertices are individually connected to one unique central node. It is a trivial combinatorial problem to find all the geodesics that pass through this central node; that



(a)



(b)

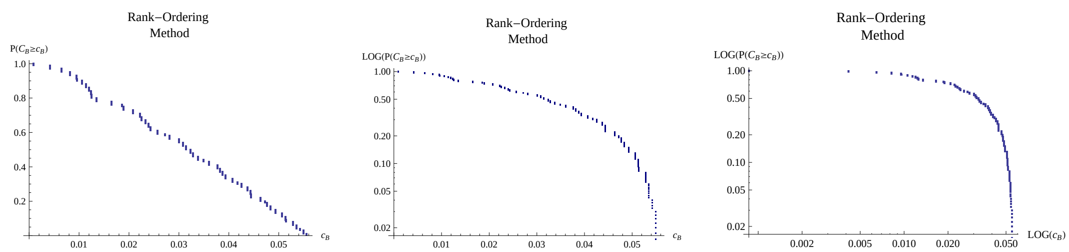
Figure 4.1: (a) Visual representation of a 20×20 regular grid. Nodes with highest betweenness centrality are pictured with light green and nodes with the lowest with black. (b) Total frequency $N(C_B = c_B)$ of the betweenness centrality of the grid in (a).

We follow the approach in [31], where it is argued that the number of elements (e.g. cars) passing through a link or node within some time interval can be roughly approximated with the measure of betweenness centrality. Here we assume also, in the same way they do, that origin-destination pairs are equally distributed. These assumptions "allow for estimating the implications of the network topology on the spatial distribution of traffic flows" [31]. Reference [19] affirms that they "have evidence that betweenness approximations can help to construct better highway-node hierarchies for road networks".

There are computational limitations concerning the calculation of betweenness centrality for large networks [8, 40, 30, 9]. For this reason, the statistics taken from the models of Chapter 3 are only for relatively small networks, with a total number of nodes of the order of hundreds.

4.2 Statistics of betweenness centrality

If we calculate the betweenness centrality of every node in the graph of Fig. 3.1, we can see that the spatial distribution within the network is trivial (Fig. 4.1(a)). Due to the symmetry of the graph, the more interior nodes are the ones with the highest betweenness, and it can be seen that groups of four and eight nodes with the same betweenness (Fig. 4.1(b)) appear with decreasing value from the inside to the outside.



(a) (c) (b)

Figure 4.2: (a) Cumulative distribution $P(C^B \geq c^B)$ of a normal grid. Same cumulative distribution in (b) a log plot and (c) a log-log plot.

Since we are going to work from now on with the cumulative distribution $P(C^B \geq c^B)$ (frequency of nodes that have a betweenness centrality C^B greater than or equal to the value c^B), we show in Fig. 4.2 the corresponding distribution for the case of the grid illustrated in Fig. 4.1(a); we also show the same cumulative distribution in Fig. 4.1(b) using a *log* plot (this will reveal an exponential distribution if a straight line is observed), and on a *log-log* plot in Fig. 4.1(c) (this will be useful to rapidly identify tails in the distribution that fall like power laws).

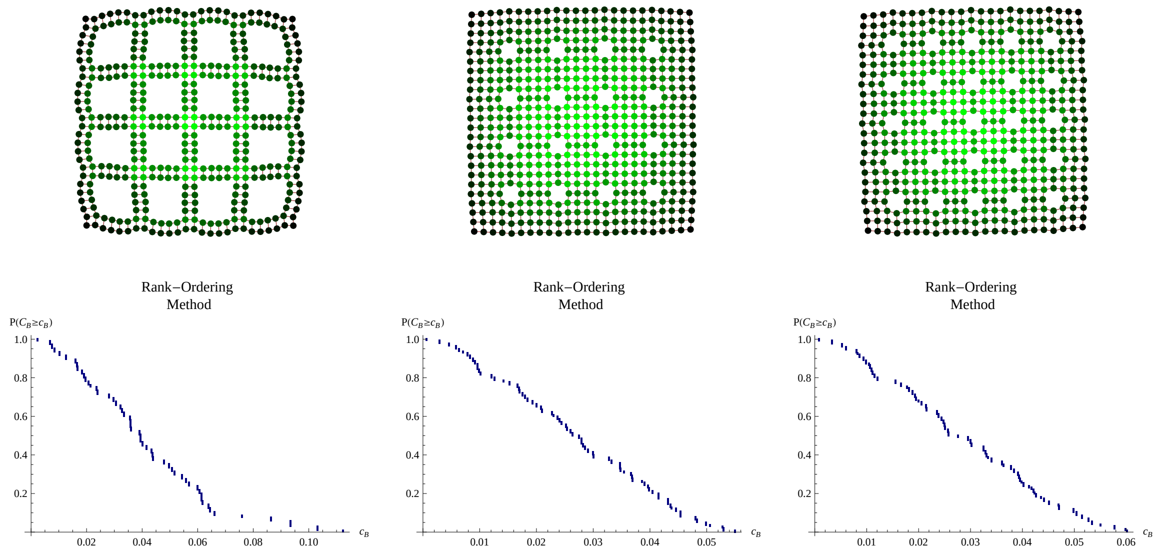
From Figs. 4.1 and 4.2 we can make the following observation: the betweenness centrality distribution in regular lattices like the one shown in Fig. 3.1 is, for practical purposes, uniform. The way in which this distribution changes with the inclusion of holes in the grid is what we will be exploring in the next sections.

number is precisely $\binom{N-1}{2} = \frac{(N-1)(N-2)}{2}$.

4.2.1 *Regular model*

The betweenness centrality distribution for the network with regular holes of equal sizes is also uniform. This is not surprising given that the mechanisms of transport in the network does not change essentially from the original lattice.

Three different examples of networks are shown in Fig. 4.3 with their cumulative distributions of node betweenness centrality. The plots are in linear scale, and a uniform distribution can be seen.



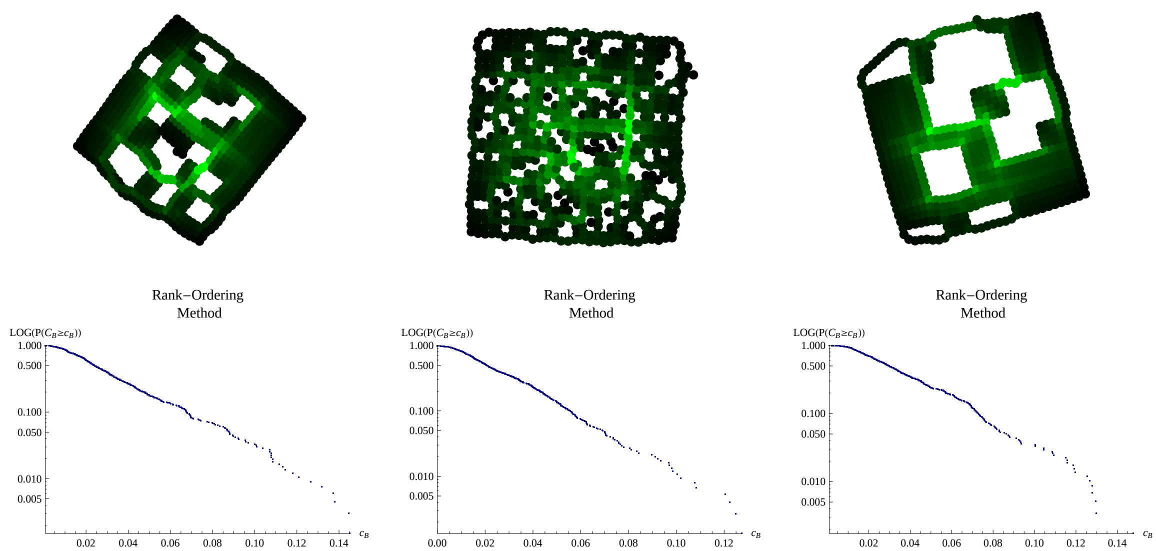
(a) (b) (c)

Figure 4.3: Examples of different networks generated with the *regular model* (top) with their cumulative distributions in linear plots (bottom). (a) Network with $n_c = 4$, $r_s = 2$, and $h_s = 4$. (b) Network with $n_c = 4$, $r_s = 4$, and $h_s = 1$. (c) Network with $n_c = 5$, $r_s = 3$, and $h_s = 1$.

These distributions are very similar to the one in Fig. 4.2(a), and thus we can conclude that the inclusion of holes in this way does not change essentially the transport phenomena from a regular normal grid.

4.2.2 *Random model*

This model introduce random features in the networks, and therefore the distribution of betweenness centrality must be taken from an ensemble of many networks. This was done with samples of 10 networks of $n = 30$, for the values $h_s = 1, \ldots, 7$ and $f = 0.6,\ 0.7,\ 0.8,\ 0.9$. In Fig. 4.4, we show three examples of networks produced by this model with their corresponding betweenness distribution, and in Fig. 4.5 we show the cumulative distributions of the ensembles for some values of h_s and f. All the cumulative distributions in this section are plotted in log-linear scales.



(a) (b) (c)

Figure 4.4: Examples of different networks generated with the *random model* (top) with their cumulative distributions in log-linear plots (bottom). (a) Network with $n = 30$, $h_s = 4$, and $f = 0.7$. (b) Network with $n = 30$, $h_s = 1$, and $f = 0.8$. (c) Network with $n = 30$, $h_s = 7$, and $f = 0.6$. Node sizes have been enlarged to facilitate the visualization.

We observe for this model an exponential distribution of the betweenness centrality; one recognizes it because it appears like a straight line with negative slope in a log-linear plot.

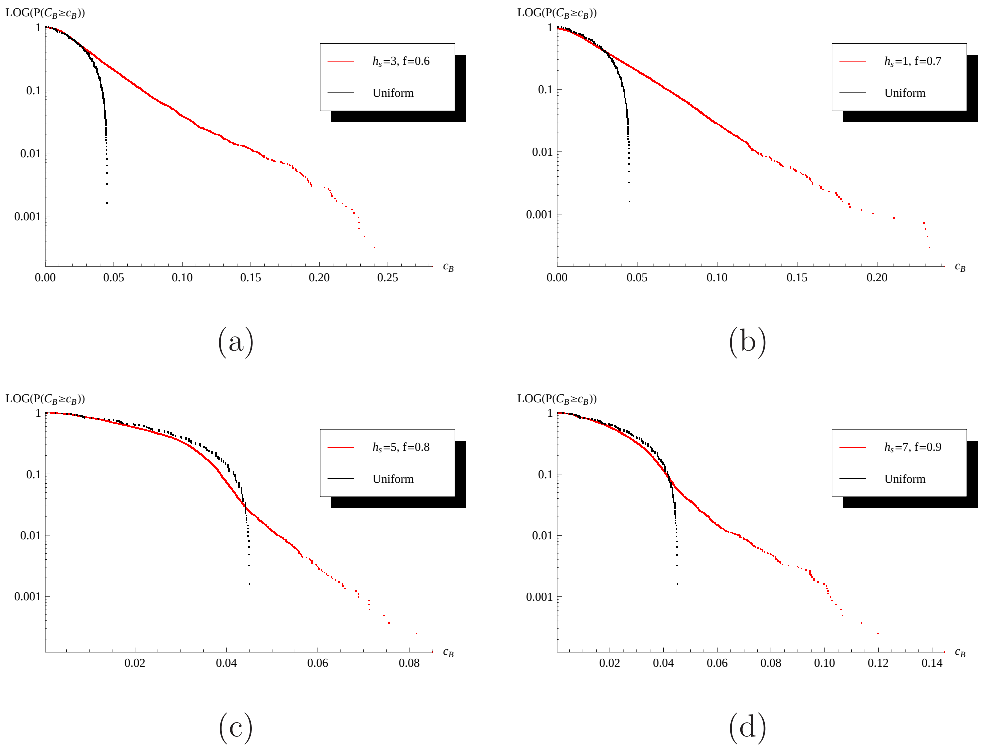


Figure 4.5: Betweenness centrality distributions of different instances of the random model (red); each distribution is obtained from an ensemble of 10 random realizations. The graph in black is the uniform distribution of the betweenness centrality for a 25×25 regular lattice.

The cumulative distributions of betweenness centrality in the random model appears to follow an exponential distribution of the form

$$P_{\geq}(c^B) = A\,\mathrm{e}^{-c^B/\mu} \tag{4.2}$$

where $A = \mathrm{e}^{c^B_{\mathrm{min}}/\mu}$ is the normalization constant (with c^B_{min} the minimum value for which the distribution holds).

An exponential distribution for some variable in a system implies that a characteristic scale exists. This characteristic scale is associated with the parameter μ. The estimator of μ from the data can be derived using maximum likelihood methods (see Appendix A), and is given by

$$\mu = \frac{1}{N}\sum_{i=1}^{N}(c_i^B - c_{\min}^B) \tag{4.3}$$

with an error in the estimation

$$\sigma_\mu = \frac{\mu}{\sqrt{N}}, \tag{4.4}$$

where N stands for the number of values among the data set for which the exponential distribution holds.

Although it is not fundamental to us in this work to calculate exactly the values of μ (and thus we are not going to give much attention to goodness-of-fit tests), we are going to calculate some values to learn, at least qualitatively, about the behavior of the distribution, given parameters f and h_s of the model (see Table 4.1). What we observe is that, in general, μ diminish its value when h_s is fixed and f increases.

$\hat{\mu}$ ($c_{\min}^B$)	$f = 0.6$	$f = 0.7$	$f = 0.8$	$f = 0.9$
$h_s = 1$	0.0450 (0.00)	0.0283 (0.01)	0.0171 (0.04)	0.0075 (0.04)
$h_s = 4$	0.0349 (0.00)	0.0205 (0.02)	0.0154 (0.06)	0.0087 (0.03)
$h_s = 7$	0.0508 (0.07)	0.0273 (0.06)	0.0357 (0.06)	0.0167 (0.06)

Table 4.1: Estimated values of μ for several values of the parameters f and h_s. The number in parenthesis is $c_{\min}^B$, the minimum value for which the exponential distribution holds; the values below $c_{\min}^B$ appear to follow the same uniform distribution observed on regular lattices.

Interestingly, we observe also a heavy tail that emerge from the exponential distribution for high values of h_s. This suggests a transition in the tails of the distribution from an exponential to a power law. This will become clear when we compare the different models (Section 4.3.1).

4.2.3 *Fractal model*

Again, the statistics for this model were taken from samples of 10 realizations of the same parameters $\alpha = 1.3, 1.5, 1.7, 1.9, 2.1, 2.3$, and $f = 0.6, 0.85$. Two single realizations are shown in Fig. 4.6. All the cumulative distributions in this section are plotted in log-log scales.

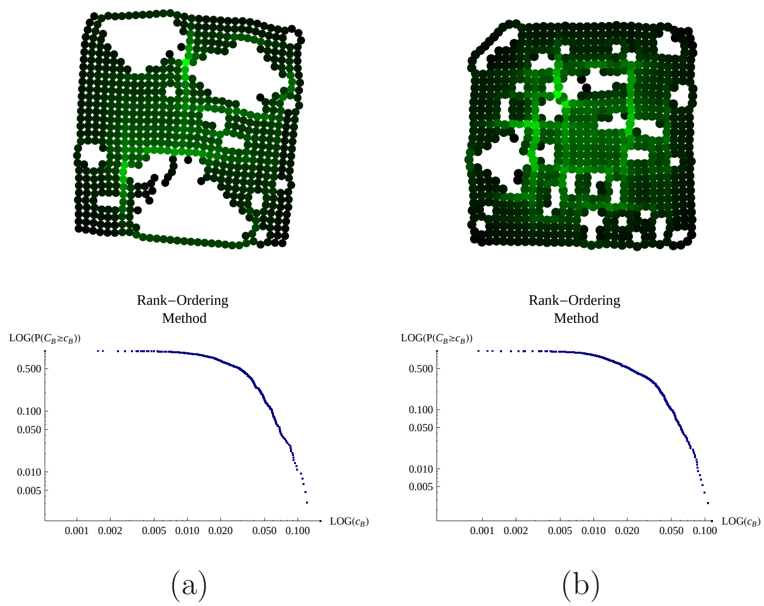


Figure 4.6: Examples of different networks generated with the *fractal model* (top) with their cumulative distributions in log-log plots (bottom). (a) Network with $n = 30$, $\alpha = 1.7$, and $f = 0.6$. (b) Network with $n = 30$, $\alpha = 2.3$, and $f = 0.85$.

There are some realizations of the model that display a distribution of their node betweenness centrality more close to an exponential than to a power law. This supports the fact that there is indeed a transition between an exponential distribution and a power law, from the random model to the fractal model. Some of the ensembles of realizations that display a power law tails are shown in Fig. 4.7. The distributions follow the function

$$P_{\geq}(c^B) = \left(\frac{c^B}{c^B_{\min}}\right)^{1-\beta}, \tag{4.5}$$

and in Table 4.2 some values of the exponent β are estimated. Even thought we do not have enough data in the table to make a strong assertion, there is an apparent tendency of the exponents β to decrease as the values for α increase.

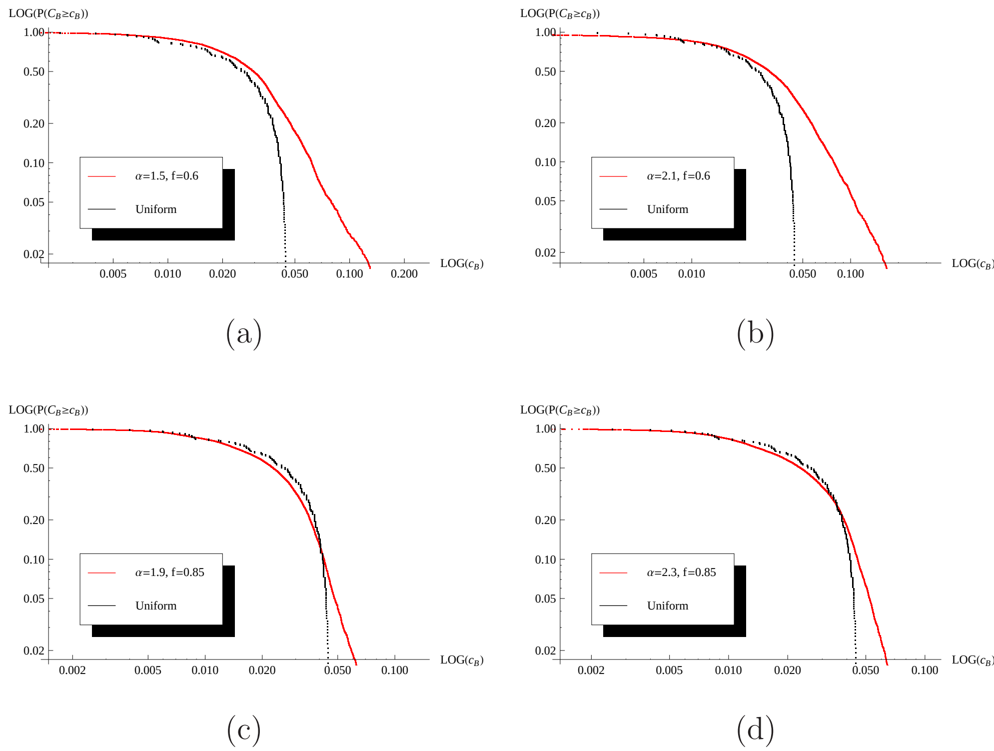


Figure 4.7: Log-log plots of the fractal model, that display power law tails in the distribution of the betweenness centrality (red); each distribution is taken from an ensemble of 10 random realizations. The graph in black is the uniform distribution of betweenness centrality of a 25×25 regular lattice.

4.3 Discussion of results

Transportation networks exist in real space, and are, by definition, embedded in physical objects. In the simple models studied in this research project, the principal abstract entity that serves to model this fact are the holes in these networks; they represent the absence of nodes and edges of the network in the object in which they live. The results presented in this chapter show evidence that leads to believe that holes are responsible for non trivial distributions of traffic (approximated by the measure of betweenness centrality) within the network.

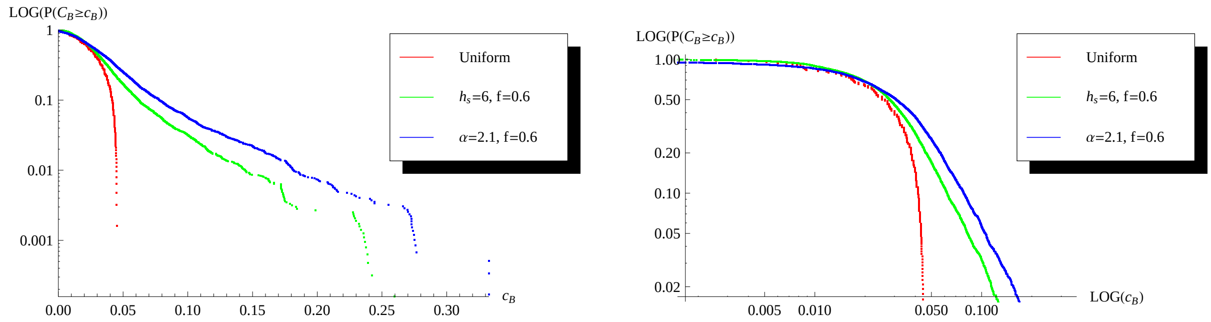
On one hand, the holes induce an exponential distribution when they have the same

$\hat{\beta}$ (c^B_{min})	$f = 0.6$	$f = 0.85$
$\alpha = 1.3$	4.22 (0.04)	6.83 (0.01)
$\alpha = 1.5$	3.46 (0.04)	5.64 (0.02)
$\alpha = 1.7$	3.77 (0.04)	5.56 (0.01)
$\alpha = 1.9$	3.48 (0.05)	5.82 (0.02)
$\alpha = 2.1$	3.25 (0.05)	$\mu =$*0.0147 (0.06)*
$\alpha = 2.3$	$\mu =$*0.0407 (0.00)*	$\mu =$*0.0137 (0.06)*

Table 4.2: Estimated exponent β of the power law distributions tails for several values of the parameters f and α. The value in parenthesis is c^B_{min}, the minimum value for which the power law distribution holds; the values below this appear to follow the same uniform distribution as for the regular lattice. The emphatized values are three cases for which the distribution was closer to an exponential than to a power law tail.

size. This can be understood from another perspective: when the sizes of the holes have a characteristic scale, the betweenness centrality of nodes also have a characteristic scale, defined by the parameter μ in the distribution. On the other hand, when holes have sizes distributed according to a power law, betweenness centrality distribution display a heavy tail that follows a power law. In other words, when holes lack a characteristic scale, betweenness centrality displays scale-free properties.

Of course, the behavior of traffic in real transportation networks is dynamical in its nature, and changes in time. Here, the model is static, and hence the measure of betweenness centrality is a static property of the model. Nevertheless, the measure of betweenness centrality can be seen as an approximation of the aggregated traffic that flows within the network over long periods of time. Thus, a large value in the betweenness of a node should be interpreted as an indication of how central and important is that location regarding transport phenomena for that particular network.



(a) (b)

Figure 4.8: Comparison of the betweenness centrality distribution for a regular lattice (red), a random model (green), and a fractal model (blue). (a) On a log-linear scale and (b) on log-log scale. All distributions were taken from networks with aproximately the same number of nodes (≈ 540).

4.3.1 Comparisons between models

From Sections 4.2.1, 4.2.2, and 4.2.3 we learned that when holes are randomly introduced within a planar regular lattice, a heavy tail in the distribution of betweenness centrality appear, meaning that some nodes are becoming more central than others. In particular: when holes have the same size, the distribution approximates to an exponential distribution; and when the sizes are power law distributed, the tails become more rightly skewed and a power law tail emerges.

To illustrate the process by which the distribution of betweenness centrality in our models changes from a uniform, then to an exponential, and finally to a power law, we fix the total number of nodes and f and plot the distributions for the different models in Fig. 4.8.

4.3.2 Emergence of hierarchies

An additional important result that can be observed from the different models is the emergence of corridors with high betweenness centrality. Visual examples of such phenomenon are shown in Fig. 4.9.

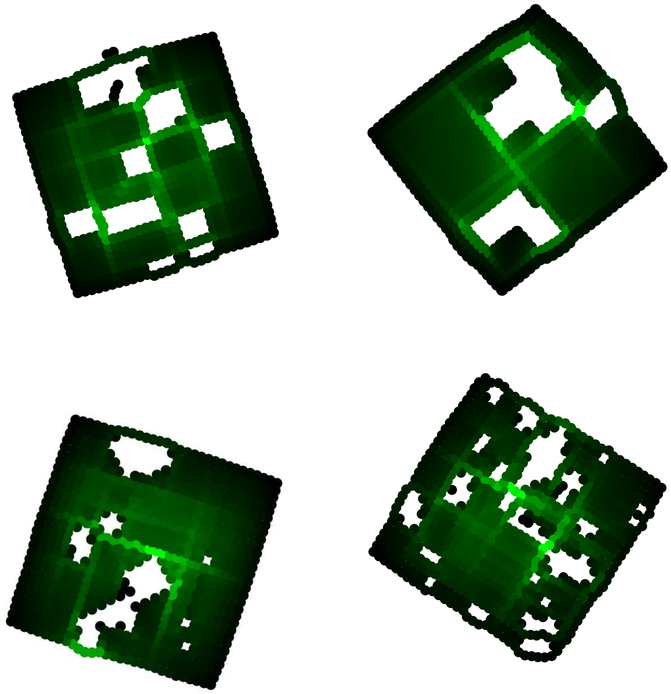


Figure 4.9: Examples of networks generated by the different models on which corridors of high betweenness centrality emerge. Since betweenness centrality is correlated to high traffic, these corridors should be interpreted as important lines of transport (e.g. arterial and main roads in a city).

As can be observed, central lines of high betweenness emerge naturally from the models. The emergence of hierarchies of roads in the transportation networks studied in this work, suggest that holes in the objects in which they are embedded may play a crucial role in the transport phenomena inside it. Identifying such (natural) principal routes could be very important, if one wants to manage traffic in an efficient manner.

Chapter 5

Analogy: the case of a road street network

In the Introduction we mentioned several examples of spatial transportation networks. Traffic problems and transportation networks are *mutually dependent* phenomena, and we face them every day in cities. This chapter is intended to be a brief conceptual guideline for management and design, and for this reason, we contrast and discuss the case of city road networks with the results found in the preceding chapters. However, we are conscious that any ideas we put forward in this chapter, based on results presented previously on this document, must be validated by the criteria of an expert in urban transportation problems. Only under such critical eye can the consequences, usefulness, meaning, and relevance of this chapter be of practical interest.

5.1 Urban problematic

In [43] the architect Germán Samper expresses the problematic regarding urban road networks in the following manner:

> [...] la característica de la modernidad en la ciudad actual está simbolizada en su estructura vial, y especialmente aquellas vías que deben permitir un gran volumen de tráfico a una gran velocidad. Toda la tecnología de la ingeniería vial se debe aplicar a estas vías que deben conformar un red o malla en todas las direcciones y que permitan un transporte rápido y seguro. Hoy estas vías no cumplen estos dos requisitos. Son relativamente *angostas*, con *muchos cruces* y pretenden servir a los edificios aledaños.

> Por tanto la circulación es *lenta e insegura*. En cambio estas vías penetran los sectores residenciales, introduciendo el *caos*, el *desorden* y la *contaminación*. La jerarquización de vías de la ciudad debe ser revisada. Hoy en día, se clasifican por su *ancho*, por su *capacidad de transporte*, pero toda la red vial parte de un mismo principio: calzadas centrales para vehículos y angostos andenes para peatones, y así hasta la más angosta vía se diseña con un sentido de prioridad al vehículo, que ha dado lugar a la aparición del término "deshumanización" de la ciudad. En un alto porcentaje la humanización tiene que ver con el tratamiento de las vías.

More specifically for the case city of Bogotá, [11] affirms that:

> En los últimos años la demanda vial ha crecido por el aumento del número de vehículos automotores, se puede decir que la oferta es bastante inferior a la demanda de transporte y de tránsito vial, esto ha traído como consecuencia, particularmente en la Ciudad de Bogotá, incrementos en la congestión, demoras, accidentes y problemas ambientales, bastante mayores que los considerados aceptables.
>
> [...] Esta situación plantea la necesidad de diseñar una infraestructura vial que optimice las exigencias presentadas por la circulación vehicular, teniendo como objetivo principal proporcionar un sistema que brinde eficiencia, y sea a su vez seguro, económico y que esté acorde a los recursos disponibles.

These authors point out that a good assessment and management of traffic and roads are essential for a sustainable city, in environmental and social terms. In the practice, they are revealing, among other things, the need for answers to the questions: which are the important roads, and what makes them important?

Regarding the issues expressed above, we believe that the methodology presented in this work along with its results could improve transport assessment and road management.

5.2 Cities

In [38] we find the following statement:

> [. . .] today old neighborhoods are often underestimated in their most fundamental values: they might be considered picturesque, even attractive, but their *structure* is not so valuable: it is *disordered*. Against this modernist stigmatization, a whole stream of counterarguments have been raised since the early 1960s in the name of the 'magic' old cities (Jacobs, 1993). The claim was not just about aesthetics: it was about livability. The modern city is hard to live in. The social success of an urban settlement *emerges* from the complex, uncoordinated interaction of countless different routes and experiences in a suitable environment. Is this a nostalgic claim to a prescientific era? Jane Jacob argued, following Weaver (Jacobs, 1961; Weaver, 1948), that cities are complex-organized problems and, as such, in order to be understood, they require to be approached with a new science: only by means of the new science of complexity can the 'marvelous' complex order of the old city be revealed that, unlike the Euclidean geometry, is not visible at a first glance, is not imposed by any central agency, but, rather, sprouts out from the uncoordinated contribution of countless agents in time. That order, Jacobs concluded, is *the order of life*: that is why it fosters human life in cities; it is that order which builds the sustainable city of the future (Newman and Kenworthy, 1999).

This quote reveals the feeling that a new way for understanding cities is needed. One of the first approaches in this spirit sees cities as fractal objects, and was formally developed in 1994 within two books: *La fractalité des structures urbaines* by Pierre Frankhauser [15] and *Fractal Cities: A Geometry of Form and Function* by Michael Batty and Paul Longley [7]. In these books, several fractal dimensions of different cities around the world were calculated, and it was argued that "[. . .] l'analyse fractale donne la possibilité d'étudier, dans une même structure, un phénomène d'organisation interne *à travers les échelles*. Cette méthode, proposée par B. Mandelbrot, a permis de découvrir des lois d'ordre interne dans des structures souvent désignées comme amorphes, complexes ou irrégulières, telles que des textures."[1] [15]. The results presented in Chapter 4 suggest that fractality is one of the aspects that generates many other emergent properties in cities; in particular, a hierarchical organization of urban roads.

5.3 Comparison of the model with real cities

We compare, mainly, our different models with their corresponding results with those of [13], [38], [12], and [31]. All four references study different centrality measures for several world city road networks, including betweenness centrality.

Our model is able to reproduce the exponential distribution of betweenness centrality seen in *self-organized* cities such as El Cairo and Venice [13, 12]. It also displays scale free properties, reported for cases such as the one for the city of Dresden [31].

There is also evidence that the distribution of cell areas (blocks) in a city follows a power law $P(A) \sim A^{-\alpha}$, particularly with $\alpha \approx 1.9$ in the case of Dresden [31, 5]. When introducing holes in our network, we are directly changing the size of the cells. This leads directly to the fact that in the model of Section 3.2.3, cell areas are distributed according to a power law.

In Fig. 5.1 a real street pattern is shown. The ressemblance with the model exposed in this document is evident.

These facts allow us to establish a validation of our insights about the role of holes in transportation networks.

[1] "[...] fractal analysis gives the possibility of studying, in one single structure, the phenomenon of internal organization *through scales*. This method, proposed by B. Mandelbrot, has been useful in the discovery of laws of internal order in structures often designated as amorphous, complex or irregular, like textures." (translated by the author)

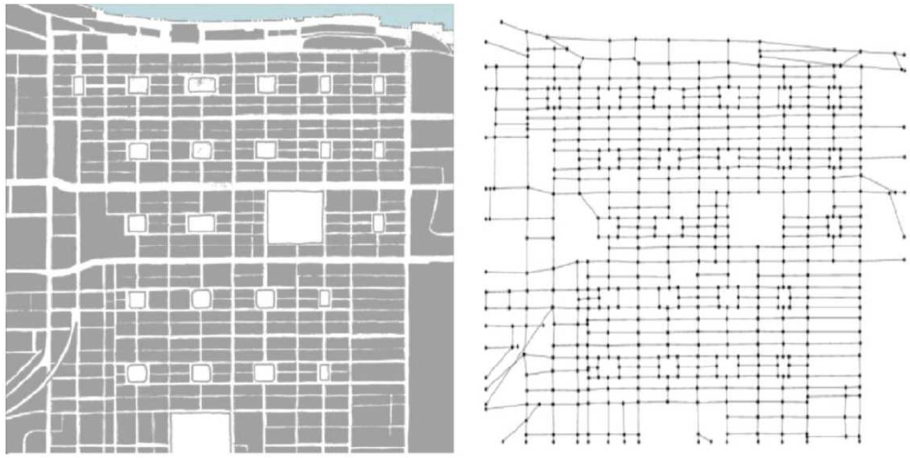


Figure 5.1: The urban pattern of Savannah as it appears in the original map (left), and reduced into a spatial graph (right). Taken from [10].

5.4 Proposed methodology for real cities

Assuming one possess the graph that represents the road network in a city, the following steps should stand as a general methodology (supplementary to the existing ones) for the assessment of transport phenomena in urban studies:

1. Construct the distribution of the sizes of holes in a city, such as parks or un-built spaces. This should provide information about the general scaling properties of transport in the road network, and should help answer the question if whether there exist or not a characteristic traffic volume in the system.

2. Construct the exact distribution of betweenness centrality of the road network. This should reveal the spatial localization of the most important and central nodes (intersections), and the principal routes inside the city.

Additional analysis should follow the MCA methodology (see Section 2.3 and references therein).

5.5 Quality of the pavement

Betweenness centrality measures could be useful for the structural design of roads [11]; knowing which road segments and road intersections are *central* could improve the efficient allocation of resources in their construction. And knowing how parks, green spaces and free spaces in a city affect, according to the research work presented in this document, the transportation phenomena within the road network, could give even more information in the design activity; different types of asphalt could be used in construction, depending on the indices given by the betweenness of the street or the intersection, and of the closeness to parks or large un-built spaces.

5.6 Traffic lights and the geometry of the streets

In general, the total period of a traffic light system in an intersection is related positively to the amount of cars that cross that intersection (transit volume)[2]: more cars, longer the period. Identifying which intersections have high betweenness centrality could be useful to calibrate and assess these times.

The emergent routes, such as the ones shown in Fig. 4.9, are expected to mobilize large volumes of traffic. Thus, the geometry (e.g. the width) of the streets that compose these principal routes, together with a coherent programming of traffic lights along it, can be critical for the functioning of the transport within the road network.

[2]By 'total period of a traffic light system in an intersection' we mean the time by which the sequence of 'greens' and 'reds' of the traffic lights returns to a previous state.

Chapter 6

Conclusions

The model presented in this research project may be regarded as a modest step in the pursue of understanding traffic and transport phenomena in spatial networks, using a complex system approach. Our results show that the distribution of the sizes of holes inside a planar network strongly affects the distribution of betweenness centrality of the nodes that compose it. The underlying motivation for this inclusion of holes is, as mentioned several times along this work, to simulate the structural properties of real objects.

Evidence that general and universal laws exist in many apparently different complex systems [51, 35, 28, 27, 29, 14, 24, 2, 23], allows us to study simple models in search of new general collective emergent behavior that can suggest non-traditional solutions to problems faced by today's society such as terrorism [24, 26], stock market crashes [52, 25], economic growth [46], climate change [51], cancer, HIV [28], and the increasing expansion and hard to control growth of cities with its overpopulation, traffic jams, and pollution problems [6, 24, 15, 46, 32].

To understand how holes affect traffic within a planar network, three models were proposed. A first one in which holes were introduced in a regular and periodic way, a second one in which holes were randomly placed but had the same size, and a third one where the holes were randomly placed but their sizes were power law distributed. The main outcome of these models is that an exponential tail in the distribution of betweenness centrality emerges for the second model and a power law tail emerges for the third one. Additionally, principal and central routes emerge naturally inside the network for these models.

We can infer the distribution of traffic within a transportation network through the characterization of the holes inside it. Regarding cities, we can conclude from this research that a global understanding of holes (distribution of sizes) in the road network gives global information about the traffic (betweenness centrality distribution). The distribution of hole sizes is related to the distribution of traffic within the city.

It would also be desirable to further explore how the weight of nodes relates to the transport phenomena.

We cite [13]: "The spatial distribution of C^B nicely captures the continuity of prominent urban routes across a number of intersections, changes in direction, and focal urban spots". Our findings suggest that holes may be one of the causes responsible for the emergence of important roads in street networks. Additionally that the size distribution of the holes affect the shape of the distribution of betweenness centrality of these important locations of a city.

Finally, this work is not exhaustive and further research should be done: relaxing some of the assumptions, digging out analytical results, and expanding the number of analyzed data sets.

References

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

[2] Per Bak. *How nature works: the science of self-organized criticality*. Springer-Verlag Telos, 175 Fifth Avenue, New York, 1996.

[3] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of 1/f noise. *Phys. Rev. Lett.*, 59(4):381, July 1987.

[4] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Phys. Rev. A*, 38(1):364, July 1988.

[5] Marc Barthélemy and Alessandro Flammini. Modeling urban street patterns. *Phys. Rev. Lett.*, 100:138702, April 2008.

[6] Michael Batty. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. The MIT Press, London, 2005.

[7] Michael Batty and Paul A. Longley. *Fractal Cities: A Geometry of Form and Function*. Academic Press, San Diego, 1994.

[8] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[9] Ulrik Brandes. On variants of shortest-path betweenness centrality and their generic computation. Preprint submitted to Social Networks, November 2007.

[10] Alessio Cardillo, Salvatore Scellato, Vito Latora, and Sergio Porta. Structural properties of planar graphs of urban street patterns. *Phys. Rev. E*, 73:066107–1, 2006.

[11] S. L. Uribe Celis. Manual de diseño geométrico para vías e intersecciones urbanas. *Memos de investigación, Universidad de los Andes, http://hdl.handle.net/1992/942*, 2006.

[12] Paolo Crucitti, Vito Latora, and Sergio Porta. Centrality in networks of urban streets. *Chaos*, 16:015113, 2006.

[13] Paolo Crucitti, Vito Latora, and Sergio Porta. Centrality measures in spatial networks of urban streets. *Phys. Rev. E*, 73:036125–1, 2006.

[14] Gary William Flake. *The Computational Beauty of Nature: Computer Explorations Fractals, Chaos, Complex Systems, and Adaptation*. The MIT Press, 1998.

[15] Pierre Frankhauser. *La Fractalité des Structures Urbaines*. Anthropos, Economica, Paris, 1994.

[16] Pierre Frankhauser. The fractal approach. a new tool for the spatial analysis of urban agglomerations. *Population: An English Selection*, 10(1):205–240, 1998.

[17] M. T. Gastner and M. E. J. Newman. The spatial structure of networks. *Eur. Phys. J. B*, 49(2):247–252, 2006.

[18] Richard J. Gaylord and Louis J. D'Andria. *Simulating Society*. Springer-Verlag, New York, 1998.

[19] Robert Geisberger, Peter Sanders, and Dominik Schultes. Better approximation of betweenness centrality. In SIAM, editor, *Proceedings of the 10th Workshop on Algorithm Engineering and Experiments (ALENEX'08)*, 2008.

[20] M. L. Goldstein, S. A. Morris, and G. G. Yen. Problems with fitting to the power-law distribution. *Eur. Phys. J. B*, 41:255–258, 2004.

[21] Hermann Haken. *Synergetics: Introduction and Advanced Topics*. Springer-Verlag, Berlin, 1983.

[22] Bill Hillier and Julienne Hanson. *The Social Logic of Space*. Cambridge University Press, Cambridge, 1984.

[23] Henrik J. Jensen. *Self-Organized Criticality*. Cambridge University Press, Cambridge, 1998.

[24] Neil Johnson. *Two's Company, Three is Complexity*. Oneworld Publications, 2007.

[25] Neil F. Johnson, Paul Jefferies, and Pak Ming Hui. *Financial Market Complexity: What Physics Can Tell Us About Market Behaviour*. Oxford University Press, London, 2003.

[26] Neil F. Johnson, Mike Spagat, Jorge A. Restrepo, Oscar Becerra, Juan Camilo Bohórquez, Nicolas Suárez, Elvira Maria Restrepo, and Roberto Zarama. Universal underlying ongoing wars and terrorism. *http://xxx.lanl.gov/abs/physics/0605035*, 2006.

[27] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. New York. Oxford University Press, 1993.

[28] Stuart A. Kauffman. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford University Press, USA, New York, 1995.

[29] Stuart A. Kauffman. *Investigations*. Oxford University Press, New York, 2000.

[30] Maksim Kitsak, Shlomo Havlin, Gerald Paul, Massimo Riccaboni, Fabio Pammolli, and H. Eugene Stanley. Betweenness centrality of fractal and nonfractal scale-free model networks and tests on real networks. *Phys. Rev. E*, 75:056115–1, 2007.

[31] Stefan Lämmer, Björn Gehlsen, and Dirk Helbing. Scaling laws in the spatial structure of urban road networks. *Physica A*, 363:89–95, 2006.

[32] Hernán A. Makse, José S. Andrade Jr., Michael Batty, Shlomo Havlin, and H. Eugene Stanley. Modeling urban growth patterns with correlated percolation. *Phys. Rev. E*, 58(6):7054, December 1998.

[33] Benoit B. Mandelbrot. *The Geometry of Nature*. W. H. Freeman, San Francisco, 1983.

[34] M. E. J. Newman. The structure and function of complex networks. *arXiv:cond-mat/0303516 v1*, Mar 2003.

[35] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *arXiv:cond-mat/0412004 v3*, May 2006.

[36] Mark E. J. Newman, Albert-László Barabási, and Duncan J. Watts. *The Structure and Dynamics of Networks*. Princeton Studies in Complexity. Princeton University Press, New Jersey, 2006.

[37] Sergio Porta, Paolo Crucitti, and Vito Latora. The network analysis of urban streets: A dual approach. *Physica A*, 369:853–866, 2006.

[38] Sergio Porta, Paolo Crucitti, and Vito Latora. The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design*, 33:705–725, 2006.

[39] Denise Pumain. *Hierarchy in Natural and Social Sciences*, volume 3. Springer, Dordrecht, 2006.

[40] Matthew J. Rattigan, Marc Maier, and David Jensen. Using structure indices for efficient approximation of network properties. In ACM, editor, *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge Discovery and Data mining*, 2006.

[41] Nikos Salingaros. *Principles of Urban Structure*. Tecne Press, 2005.

[42] Nikos A. Salingaros and Bruce J. West. A universal rule for the distribution of sizes. *Environment and Planning B: Planning and Design*, 26:909–923, 1999.

[43] Germán Samper. *Recinto Urbano: La Humanización de la Ciudad*. Fondo Editorial Escala, Bogotá, 1997.

[44] J. B. Satinover and D. Sornette. Illusion of control in a brownian game. *Physica A (http://arxiv.org/abs/physics/0703048)*, 2007.

[45] J. B. Satinover and D. Sornette. "Illusion of control" in minority and parrondo games. *European Journal of Physics B (http://arxiv.org/abs/0704.1120)*, 2007.

[46] Frank Schweitzer. *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*. Springer Series in Synergetics. Springer, Berlin, 2003.

[47] Frank Schweitzer, Werner Ebeling, Helge Rosé, and Olaf Weiss. Optimization of road networks using evolutionary strategies. *Evolutionary Computation*, 5/4:419–438, 1998.

[48] Herbert A. Simon. The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), December 1962.

[49] Ricard Solé and Brian Goodwin. *Signs of Life: How Complexity Pervades Biology*. HarperCollins Publishers, New York, 2000.

[50] D. Sornette. Probability distributions in complex systems. Department of Management, Technology and Economics, ETH Zurich (http://arxiv.org/abs/0707.2194).

[51] Didier Sornette. *Critical Phenomena in Natural Sciences*. Springer Series in Synergetics. Springer, Heidelberg, 2nd edition, 2003.

[52] Didier Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, New Jersey, 2003.

[53] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge University Press, Cambridge, 1994.

[54] Bhanu M. Yerra and David M. Levinson. The emergence of hierarchy in transportation networks. *Annals of Regional Science*, 39(3):541–553, 2005.

Appendix A

Maximum likelihood parameter estimation

In this appendix we derive the formulas used in the estimation of statistical parameters, using maximum likelihood methods based on Bayesian interpretation. We take inspiration from [35] for this brief presentation.

A.1 Exponential distributions

Consider the normalized exponential distribution

$$p(x) = \left(\frac{1}{\mu}\right) \mathrm{e}^{-\frac{1}{\mu}(x - x_{\min})}. \tag{A.1}$$

Given a set of n values x_i, the probability that those values are distributed according to Eq. A.1 is proportional to the *likelihood* $P(x|\mu)$ of the set, where

$$P(x|\mu) = \prod_{i=1}^{n} p(x_i) = \left(\frac{1}{\mu}\right)^n \mathrm{e}^{-\frac{1}{\mu}\sum_{i=1}^{n}(x_i - x_{\min})} \tag{A.2}$$

$$P(x|\mu) = \left(\frac{1}{\mu}\right)^n \mathrm{e}^{-\frac{1}{\mu}b}, \tag{A.3}$$

where $b = \sum_{i=1}^{n}(x_i - x_{\min})$.

To find the most plausible value of μ given the observed set of values, we need to calculate the probability $P(\mu|x)$ of a particular value μ given the data $\{x_i\}$. By Bayes' theorem

$$P(\mu|x) = \frac{P(x|\mu)\,P(\mu)}{P(x)}. \tag{A.4}$$

The prior probability of the data $P(x)$ is fixed since it refers to the fixed observed values, and does not vary in our calculations. Also, in the absence of additional information, we have to assume that the prior probability $P(\mu)$ of the parameter μ is uniform (given that it could have, with equal probabilities, any value $0 < \mu < \infty$), and in consequence, independent of μ. Hence, $P(\mu|x) \propto P(x|\mu)$. Typically, the maximum likelihood method for estimating parameters from data uses the *log-likelihood* function, denoted by $\mathcal{L}$, which is the logarithm of $P(x|\mu)$. The log-likelihood function is then equal, to within an additive constant, to $\ln P(\mu|x)$, and is given by:

$$\begin{aligned}\mathcal{L} &= \ln\left[\left(\frac{1}{\mu}\right)^n \mathrm{e}^{-\frac{b}{\mu}}\right] \\ &= -n\ln\mu - \frac{b}{\mu}.\end{aligned} \tag{A.5}$$

The method then consist in calculating the value μ that maximizes $\mathcal{L}$ (and in consequence it also maximizes the likelihood $P(\mu|x)$ since the logarithm is a monotonic increasing function). That is, to set $\partial\mathcal{L}/\partial\mu = 0$:

$$\frac{\partial\mathcal{L}}{\partial\mu} = -\frac{1}{\mu}\left(n - \frac{b}{\mu}\right), \tag{A.6}$$

which means that (knowing that μ has a finite value)

$$\begin{aligned}\mu &= \frac{b}{n} \\ \mu &= \frac{1}{n}\sum_{i=1}^{n}(x_i - x_{\min})\end{aligned} \tag{A.7}$$

is the value that maximizes the likelihood. In the case that $x_{\min} = 0$, $\mu = \langle x\rangle$.

We also want to know the expected error $\sigma_\mu^2 = \langle\mu^2\rangle - \langle\mu\rangle^2$ in our estimation. From Eq. A.5 it can be deduced that

$$P(\mu|x) \propto \mathrm{e}^{-n\ln\mu - \frac{b}{\mu}}, \tag{A.8}$$

so that

$$\langle\mu\rangle = \frac{\int_0^\infty \mu\, P(\mu|x)\mathrm{d}\mu}{\int_0^\infty P(\mu|x)\mathrm{d}\mu} = \frac{\int_0^\infty \left(\frac{1}{\mu}\right)^{n-1}\mathrm{e}^{-b/\mu}\mathrm{d}\mu}{\int_0^\infty \left(\frac{1}{\mu}\right)^{n}\mathrm{e}^{-b/\mu}\mathrm{d}\mu}, \tag{A.9}$$

and

$$\langle\mu^2\rangle = \frac{\int_0^\infty \mu^2\, P(\mu|x)\mathrm{d}\mu}{\int_0^\infty P(\mu|x)\mathrm{d}\mu} = \frac{\int_0^\infty \left(\frac{1}{\mu}\right)^{n-2} \mathrm{e}^{-b/\mu}\mathrm{d}\mu}{\int_0^\infty \left(\frac{1}{\mu}\right)^{n} \mathrm{e}^{-b/\mu}\mathrm{d}\mu}. \tag{A.10}$$

Making the variable substitution $\beta = b/\mu$, and knowing that $\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$,

$$\langle\mu\rangle = \frac{\left(\frac{1}{b}\right)^{n-2}\Gamma(n-2)}{\left(\frac{1}{b}\right)^{n-1}\Gamma(n-1)} \tag{A.11}$$

$$\langle\mu^2\rangle = \frac{\left(\frac{1}{b}\right)^{n-3}\Gamma(n-3)}{\left(\frac{1}{b}\right)^{n-1}\Gamma(n-1)}, \tag{A.12}$$

and by the property $\Gamma(x+1) = x\,\Gamma(x)$, we finally have

$$\langle\mu\rangle = \frac{\sum_{i=1}^n (x_i - x_{\min})}{n-2} \tag{A.13}$$

$$\langle\mu^2\rangle = \frac{\left[\sum_{i=1}^n (x_i - x_{\min})\right]^2}{(n-2)(n-3)}. \tag{A.14}$$

The variance of μ is then

$$\begin{aligned}\sigma_\mu^2 &= \frac{\left[\sum_{i=1}^n (x_i - x_{\min})\right]^2}{n-2}\left(\frac{1}{n-3} - \frac{1}{n-2}\right) \\ &= \frac{\left[\sum_{i=1}^n (x_i - x_{\min})\right]^2}{(n-2)^2(n-3)}.\end{aligned} \tag{A.15}$$

In most cases, n is large, so

$$\sigma_\mu \approx \frac{\sum_{i=1}^n (x_i - x_{\min})}{n^{3/2}} = \frac{\mu}{n^{1/2}}. \tag{A.16}$$

A.2 Power law distributions

The power law distribution is given by

$$p(x) = \frac{\alpha - 1}{x_{\min}}\left(\frac{x}{x_{\min}}\right)^{-\alpha}, \tag{A.17}$$

already in its normalized form.

Following the same steps as before, we calculate the likelihood of a data set of n values x_i,

$$P(x|\alpha) = \prod_{i=1}^{n} p(x_i) = \left(\frac{\alpha - 1}{x_{\min}}\right)^n \prod_{i=1}^{n} \left(\frac{x_i}{x_{\min}}\right)^{-\alpha}. \tag{A.18}$$

By Bayes' rule,

$$P(\alpha|x) = \frac{P(x|\alpha)\, P(\alpha)}{P(x)} \tag{A.19}$$

Again, $P(x)$ is fixed, and the lack of additional information about α force us to assume a uniform probability distribution $P(\alpha)$ (for the values $1 < \alpha < \infty$ so that $p(x)$ is normalizable), so $P(\alpha|x) \propto P(x|\alpha)$. Working with the log-likelihood

$$\begin{aligned} \mathcal{L} &= \ln P(x|\alpha) \\ &= n \ln(\alpha - 1) - n \ln x_{\min} - \alpha \sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}}, \end{aligned} \tag{A.20}$$

we set $\partial \mathcal{L}/\partial \alpha = 0$ to find the value of α that maximizes the likelihood of the set:

$$\alpha = 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}}\right]^{-1}. \tag{A.21}$$

By a procedure very similar to the one presented for the exponential distribution, one finds that the expected error in our calculation of α is

$$\sigma_\alpha = \frac{\alpha - 1}{n^{1/2}}. \tag{A.22}$$