# Urban mobility planning: a complex network approach.

Laura Lotero[1], Rafael Hurtado[2], Patricia Jaramillo[3]

[1] Doctoral Student, Departamento de Ciencias de la Computación y de la Decisión, Universidad Nacional de Colombia, Medellín, Colombia, llotero0@unal.edu.co

[2] Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia, rghurtadoh@unal.edu.co

[3] Departamento de Ciencias de la Computación y de la Decisión, Universidad Nacional de Colombia, Medellín, Colombia, gpjarami@unal.edu.co

## Abstract

Urban mobility is one of the most important aspects to be considered in urban planning, especially in developing countries, where population growth and car ownership rates are increasing significantly compared to those in developed countries. Mobility patterns in a city influence and are influenced by other city aspects in the short-term, such as congestion and pollution, and in the long-term by land use patterns. Here, we explore the complex network approach, in order to identify macroscopic properties of urban mobility from the analysis of the interaction of the microscopic components of the system. These macroscopic properties allow us for a better understanding of the systems and therefore a better urban planning process, which takes enormous relevance in rapidly growing or evolving cities such as most cities in developing countries. We aim to explain how to represent urban mobility network from available information, such as origin-destination surveys, how to calculate the main statistical or topological properties of the network and how to interpret them in terms of urban mobility and urban transportation planning. We apply these developments to the urban mobility network of the Medellin metropolitan area in Colombia, South America and give some insights for the improvement of the urban planning process.

**Keywords**: urban mobility, urban planning, complex networks

# Introduction

The relationship between transportation and other planning processes in a city is well known (Ortúzar & Willumsen, 2011). There is a complex and multidimensional relationship between the travel needs and the development of the city; transportation systems support the communication and interaction of people with other city sub-systems and therefore the comprehension of their structure and dynamics have important implications in urban planning and public policies.

A plausible way to represent transportation systems is by graph theory and its results applied to networks. This perspective, also known as complex network analysis or network science (Thadakamalla et al., 2006), has garnered the interest of researchers from several disciplines. In the particular case of urban mobility, networks provide the foundations for representing the interactions of the inhabitants of a city in terms of their travel patterns.

Complex network analysis has shown to be a promising approach to analyze complex systems, which are systems that are made of many interacting agents and that produce non-trivial emergent collective behavior that is not fully explained by the properties of the individual components (Amaral & Ottino, 2004; Thadakamalla, Kumara, & Albert, 2006). The fundamental insight from complex network analysis is that large scale networks are characterized by macroscopic properties rather than by the individual properties of the nodes and edges. These macroscopic or statistical properties have a huge influence on the dynamic processes taking place on the network (Thadakamalla et al., 2006), and it offers new methodological tools for the analysis and understanding of typical phenomena of urban transportation such as congestion. However, the study of transportation systems in the framework of complex networks seems relatively limited (Derrible & Kennedy, 2011).

# Methodology

The complex network methodology for analyzing large-scale networks is similar to the statistical physics approach to analyze complex phenomena. The basic idea is to obtain topological properties of the network related to macroscopic properties of the system in terms of statistical distributions of microscopic properties (Newman, 2010; Thadakamalla et al., 2006). Some of the main statistical properties are the nodal degree distribution, the average path length, the clustering coefficients and, in the case of weighted networks, the nodal strength and the link weight distributions (Table 1). These statistical properties of networks are meaningful in understanding the structure and dynamics of real systems and can be used for system planning or optimization. Here, we aim to explain to urban planners how to take the urban mobility complex network from information available, such as origin-destination surveys, and how to calculate the main statistical properties and how to interpret them in terms of urban mobility and urban transportation planning.

In the case of urban mobility, we can represent the fluxes of people from an origin to a destination zone by a weighted and directed network. The centroids of the origin-destination zones, predefined in the survey, are mapped into the node set $\boldsymbol{N}$, and we construct the adjacency matrix $\boldsymbol{A}$, where each element $a_{ij}$ provides information about the movement of people from the origin zone $i$ to destination zone $j$, and the weight $w_{ij}$ of each link represents the amount of travelers between zones and it provides additional information on the dynamics of the network. The resulting network can be represented as a graph $G=\{N,A\}$.

In table 1 we present the basic definitions of some common and broadly used properties of complex networks and their interpretation from the urban mobility context. From many interesting properties for urban planners we selected the nodal degree, the nodal strength and the link weight.

| **Network property** | **Definition** | **Urban mobility meaning** |
|---|---|---|
| Nodal degree | The nodal degree $k_i$ is the number of edges attached to node $i$. The degree distribution is the probability $p(k)$ of choosing a randomly selected node that has degree $k$.<br><br>In a directed network there are both in-degree and out-degree and their corresponding probability distributions. | The nodal degree $k_i$ of a zone $i$ is the number of other zones the people visit or the number of zones where they come from.<br><br>The degree distributions give information about the topology of the network related to the evolution of the system and how it is organizing. |
| Link weight | Link weight $w_{ijj}$ is the value or weight of the link between nodes $i$ and $j$. The link weight distribution $p(w)$ is the probability of choosing a randomly selected link with a weight $w$. | The link weight $w_{ij}$ indicates how many travelers are going specifically from origin $i$ to destination $j$.<br><br>The weight distribution gives information about the fluxes in the network and how they evolve. |
| Nodal strength distribution | The nodal strength $s_i$ is the sum of the weights of edges attached to node $i$. The strength distribution $p(k)$ is the probability of choosing a randomly selected node that has strength $s$.<br><br>In a directed network there are both in-strength and out-strength and their corresponding distributions | The nodal strength of node $i$ indicates how many travelers are coming to or leaving from the zone $i$.<br><br>The strength distributions give information about fluxes in the nodes and dynamics. |

Table 1. Definitions of the main network properties of a complex network

Other well-known and widely used network properties are the clustering coefficient of a node, that is the ratio of the number of edges that actually exist between the neighbors of that node and the total possible edges, and it characterizes the local transitivity in the neighborhood of a node; the clustering coefficient of the whole network that is the average over all the nodes clustering coefficients in the network, and the average path length that is the average of the shortest paths from each node to every other node in a connected network. These latter properties do not have a straight interpretation in terms of urban mobility and transportation planning but they can be analyzed in order to identify other interesting properties of the network that could be, for example, the *small-world* property (Watts & Strogatz, 1998).

The extraction of those distributions from an origin-destination survey or any other source of urban mobility data can be done by using one of the many specialized software such as *gephi* (Bastian & Heymann, 2009) or *pajek* (De Nooy, Mrvar, & Batagelj, 2005) which are open source software for network analysis, or by using network analysis packages for *R* or *phyton* such as *igraph* (Csardi & Nepusz, 2006), or even by the manipulation of a spreadsheet, which is available and known for many practitioners and urban planners. If the analyst chooses to use specialized software or packages the calculation of the distributions is straightforward but the analyst needs to have previous knowledge on the software and on the complex network centrality measures. On the other hand, if the analyst chooses to use a spreadsheet the steps are to calculate the single network properties for each node or link and then to calculate the probability density distributions and cumulative distributions of the nodal degree, nodal strength and link weights.

The analysis of the statistical properties of the network can be done for the entire mobility network or it can be done for separated analysis of different attributes of the trip, such as travel purpose, transportation mode, socioeconomic strata of the traveler, gender, etc. Segregated analysis allows for the comparison of different sub-classes of attributes of the trip, and therefore for specific planning purposes.

## Case study: Medellin metropolitan area.

We analyze the urban mobility of the Medellin metropolitan area, in Colombia, South America, which is a region that has about 3.5 million inhabitants distributed in 10 municipalities in 1,164 $km^2$. Medellin, that is the second city in importance in the country, is the main municipality of this region, containing about 67% of the population of the metropolitan area. Data about urban mobility was obtained in a recent research survey (AREA Metropolitana del Valle de Aburrá, 2006) that divided the area in 413 zones defined according to land use and mobility patterns. The figure 1 shows the location of the metropolitan area in the Americas.


Figure 1 - The Medellin metropolitan area.
Source: AREA Metropolitana del Valle de Aburra - http://www.metropol.gov.co/contenidos.php?seccion=14

We analyze the whole urban mobility that includes work, study, shopping, and other travel purposes, and then we segmented our analysis by socioeconomic strata of travelers, which is a driver for household income, in order to make comparisons between the travel patterns in the different income levels of the traveling population.

We found that the probability distributions for the nodal degree and the strength of nodes fit to exponential functions. The figure 2 and figure 3 show the complementary cumulative distribution for the in-degree and out-degree and the in-strength and out- strength, respectively, in a log-normal scale and the fit parameters for the exponential decay function.

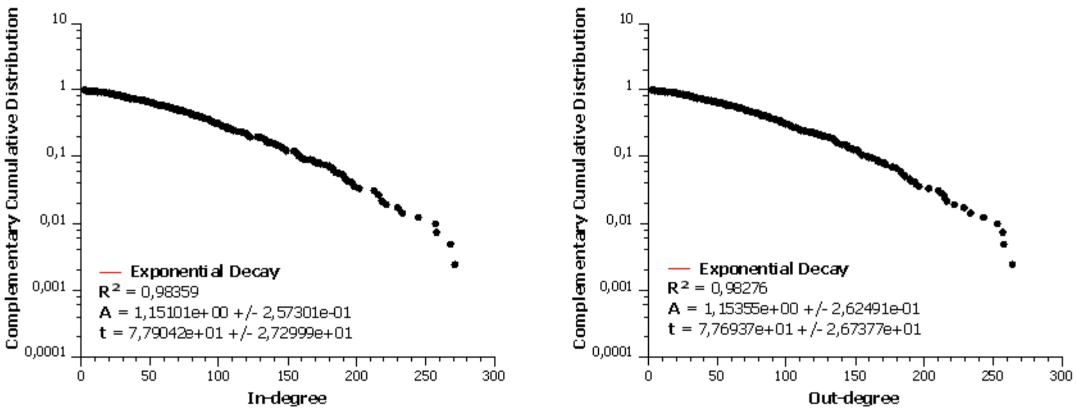


Figure 2 - In and out degree distributions for the urban mobility network of the Medellin metropolitan area.

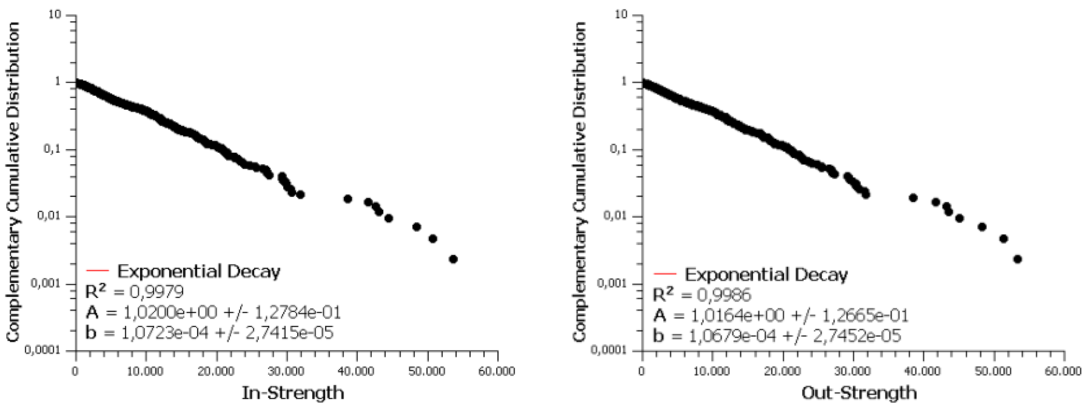


Figure 3- In and out strength distributions for the urban mobility network of the Medellin metropolitan area.

Once we have identified the distribution of the nodal degree and nodal strength, we can model the system represented by an adjacency network in analogy to a canonical ensemble of exponential random graphs with a conserved quantity (equations [1-3]). An ensemble means that a model is defined to be not a single network, but a probability distribution over many possible networks (Park & Newman, 2004).

$$Max\, S = -\sum_{G \in G_N} P(G) \ln P(G) \qquad [1]$$

$$subject to \qquad \sum_{G} P(G) x_i(G) = \langle x_i \rangle \qquad [2]$$

$$\sum_{G} P(G) = 1 \qquad [3]$$

The exponential form of the probability distributions for degree and strength defines an exponential random graph model. The exponential random graph is the distribution over a specified set of graphs $G_N$ that maximizes the entropy $S$ (equation 1) subject to a condition that the expected values of the observable $\langle x_i \rangle$ (be it the node degree or the node strength) matches the measured values in the real network, where $x_i(G)$ is the value of $x_i$ in graph $G$ (equation 2) and a condition for the normalization of probabilities (equation 3).

As we analyze the sub-graphs of socioeconomic strata of the travelers, we see that the probability distributions for the nodal degree and nodal strength are similar to those of the complete network. Figure 4 presents the in-degree and out-degree complementary cumulative distribution functions in a log-normal scale, where one can see that exponential fits are suitable, and more interestingly one can see that the highest- and lowest-income travelers have a similar distribution with a steeper slope (est 1 and est6). Something alike happens to the medium-low and medium-income travelers, which are the majority of the population in the metropolitan area (est 2 and est 3), that have a similar distribution with a less pronounced slope. The same analysis applies for the nodal strength, that is total amount of trips incoming or leaving a zone, presented in figure 5.

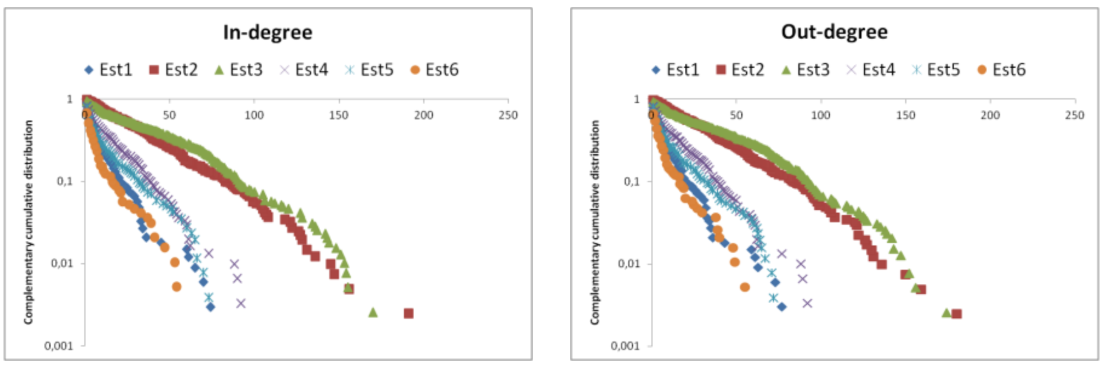


Figure 4 - nodal degree distribution for the different socioeconomic strata of travelers

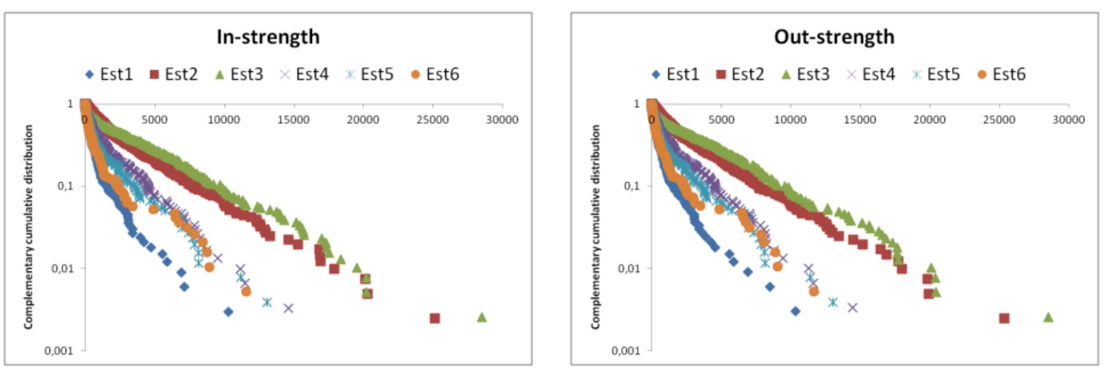


Figure 5 - nodal strength distribution for the different socioeconomic strata of travelers

We also analyze the distribution for the weights of links. We find that this probability distribution follows a power law. This result is similar to the one obtained for inter-urban mobility (De Montis, Barth, Chessa, Vespignani, & Costruzioni, 2008). We interpret the power law distributions as related to a preference rule that governs the self-organization of urban commuting flows.

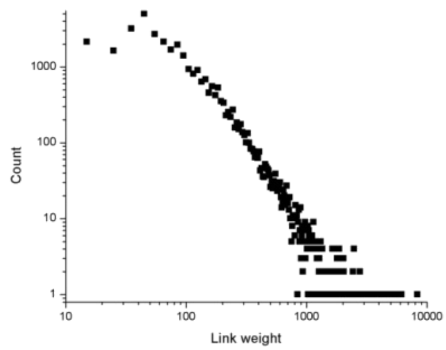


Figure 6 - Probability distribution for the link weights

## Discussion on development issues

The main objective of the complex network analysis is to understand the processes taking place on the network, and this also applies for urban mobility. Once the analyst, the urban planner or the person interested, has gained a broader understanding of the system, he can optimize the processes in the network and make better plans and policies.

We found that the nodal degree of the mobility network in the Medellin metropolitan area, i.e., the number of zones connected to other zones of the region by traveling people follows an exponential distribution which is a process of maximizing the entropy of the system, and this behavior is the same for the whole network as it is for socioeconomic strata sub-graphs. Similarly, the nodal strength also follows an exponential distribution. From these empirical results, we found that the system undertakes an organization process in order to maximize its entropy. With this in mind, and after the validation from the complex network approach, one can use the entropy maximization models such as the spatial interaction models (Wilson, 1970, 2010) and others that have been widely used in urban and transportation planning and that are linked to operational research to improve urban and transportation plans.

To know the structure and the properties of the network allow the urban planners for identifying more complex dynamics in the network. For example, a power law distribution, such as the one found in the link weights in the mobility network of the case study, implies that urban mobility flows are scale-free and self-organizing, and according to other results found for scale-free networks, we could state that urban mobility is robust for the random failure but highly vulnerable to targeted attacks on the important elements of the network, in this case the most weighted links (Albert, Jeong, & Barabási, 2000; Barabási & Albert, 1999)

Other possible applications of this complex networks results to urban planning are those related to public health issues such as the analysis of epidemics spread in the city, since urban mobility works as a vector for many diseases such as the flu. To know the structure of the network allows for identifying the most important travel patterns that could lead to disease spreading (Brockmann, 2012; González, Hidalgo, & Barabási, 2008).

Also, urban planners are interested in how the structure of the network affects the accessibility of the zones in the region (Caschili & De Montis, 2013). These analysis for the Medellin metropolitan area are our future work in order to help in the urban planning decision making process and the development of the city.

## Conclusions

Complex network analysis is a promising methodology for the study of large-scale networks that exhibit emergent behavior that is not completely explained by the analysis of the single parts of the network but that can be characterized by the macroscopic or statistical properties of the network. To understand the structure means a gain in the comprehension of the dynamics taking place in the network. For urban planners, it is important to know the structure and the topologic properties of the network in order to optimize any process on the city.

We analyzed the urban mobility network of the Medellin metropolitan area in order to discover the patterns and the statistical properties of the network, in the search of improvement for urban transportation planning, which influences the urban planning and therefore the development of a city. We applied a complex network analysis in order to characterize the network as a whole and to compare the mobility dynamics of the population by socioeconomic strata, obtaining that both aspects of the network's topology, degree and strength, are relevant in urban mobility and merit further studies in order to introduce this perspective in urban planning. Future work includes analyzing other aspects of the transportation planning such as robust optimization of the network given the structure of the urban mobility in the region.

## References


Albert, R., Jeong, H., & Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–82. doi:10.1038/35019019

Amaral, L. A. N., & Ottino, J. M. (2004). Complex networks: Augmenting the framework for the study of complex systems. *European Physical Journal B*, *38*(2), 147–162. doi:10.1140/epjb/e2004-00110-5

AREA Metropolitana del Valle de Aburrá. (2006). *Capítulo 2: Diagnóstico. Formulación del Plan Maestro de Movilidad para la Región Metropolitana del Valle de Aburrá. Informe Final* (pp. 102–188).

Barabási, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(October), 509–512. Retrieved from http://www.sciencemag.org/content/286/5439/509.short

Bastian, M., & Heymann, S. (2009). Gephi : An Open Source Software for Exploring and Manipulating Networks, 361–362.

Brockmann, D. (2012). Spotligth on mobility. *Nature News & Views*, *48*, 40–41.

Caschili, S., & De Montis, A. (2013). Accessibility and Complex Network Analysis of the U.S. commuting system. *Cities*, *30*, 4–17. doi:10.1016/j.cities.2012.04.007

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*. Retrieved from http://www.necsi.edu/events/iccs6/papers/c1602a3c126ba822d0bc4293371c.pdf

De Montis, A., Barth, M., Chessa, A., Vespignani, A., & Costruzioni, S. (2008). The structure of Inter-Urban traffic: A weighted network analysis. *Geographical*, 1–12.

De Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. *Network* (Vol. 40, p. 362). Cambridge University Press. Retrieved from http://www.mendeley.com/research/exploratory-social-network-analysis-with-pajek-1/

Derrible, S., & Kennedy, C. (2011). Applications of Graph Theory and Network Science to Transit Network Design. *Transport Reviews*, *31*(4), 495–519. doi:10.1080/01441647.2010.543709

González, M., Hidalgo, C., & Barabási, A. (2008). Understanding individual human mobility patterns. *Nature*, *453*(7196), 779–82. doi:10.1038/nature06958

Newman, M. E. J. (2010). *Networks: an introduction* (p. 720). New York, New York, USA: Oxford University Press.

Ortúzar, J. de D., & Willumsen, L. (2011). *Modelling transport* (4th ed., p. 606). Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9781119993308

Park, J., & Newman, M. E. J. (2004). The statistical mechanics of networks. *Phys. Rev. E*, *70*(6), 066117. doi:10.1103/PhysRevE.70.066117

Thadakamalla, H. P., Kumara, S., & Albert, R. (2006). Complexity and Large-Scale Networks. In R. Ravindran (Ed.), *OPERATIONS RESEARCH AND MANAGEMENT SCIENCE HANDBOOK*. Taylor & Francis Group.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–2. doi:10.1038/30918

Wilson, A. (1970). The use of the concept of entropy in system modelling. *Operational research quarterly*, *21*(2), 247–265. Retrieved from http://www.jstor.org/stable/10.2307/3008157

Wilson, A. (2010). Entropy in Urban and Regional Modelling: Retrospect and Prospect. *Geographical Analysis*, *42*(4), 364–394. doi:10.1111/j.1538-4632.2010.00799.x