Case Studies on Transport Policy 9 (2021) 1344–1357

Contents lists available at ScienceDirect

Case Studies on Transport Policy

journal homepage: www.elsevier.com/locate/cstp

Applying complex network theory to the analysis of Mexico city metro network (1969 – 2018)

Salvador Hernández González[a,*], Idalia Flores De La Mota[b]

[a] Dpto. Industrial, Campus II, Tecnológico Nacional de México en Celaya, Antonio García Cubas s/n., Celaya, Gto., Mexico
[b] Facultad de Ingeniería, Universidad Nacional Autónoma de México, Mexico

ARTICLE INFO

Keywords:
Metro
Complex networks
Mexico City

ABSTRACT

In the last fifty years, urban public passenger transport has undergone a huge transformation as cities grow, however, the transport system has not grown at the same speed. In the specific case of Mexico City, we analyzed the Metro public transport system, considering how it has changed over time; how it has become a network that can be analyzed from the perspective of complex networks, this approach is appropriate because it is a dynamic and stochastic network, which allows to study its evolution over time and how its evolution can be measured using the metrics for said networks. The results of the analysis show that the degree distribution in the Metro network has increasingly fitted in with the behavior of a scale-free network; the hub, betweenness and clustering indices also show that the stations with the highest values in each one of these measurements have been displaced from the central area and are now located in the eastern part of the city. In each period, the diameter of the network has increased at a variable rate. It must be pointed out that so far, the diameter has grown following a potential model without giving signs of an apparent contraction. Finally, an update of the current conditions of the subway as well as future work and recommendations are presented.

1. Introduction

Public transport is one of a city's vital infrastructures. Public transport networks are developed over long periods of time and imply huge investments on the part of governments; public transport systems also permit a coherent integration of different points in a particular area. Among the variants to be found in public transport systems is the network of Metropolitan Trains nowadays known by the shorter name of "Metro"; the "Metro" is an underground train that arises the response to the growing need to move large numbers of people at speed around crowded cities (Rosas-Gutiérrez, 2013).

There are relevant performance measurements that are very useful for administrators, builders and, in general, for those in charge of the management of these systems: stations with a larger number of connections, stations that act as intermediaries between stations are some examples of parameter issues; questions also arise such as: is it possible to have a metro network where all the stations are connected? What stations are the most important ones in the network? and are they always the same ones or do they evolve over time? How is the connectivity of the system affected when a new line is added? What are the network vulnerabilities?

In recent years, the Metro networks of a variety of cities have been analyzed using a complex networks approach (Derrible, 2012) to obtain information about the network's topology and answer some of the questions posed in the above paragraph. It is worth mentioning that the level of influx (passenger numbers arriving) at the stations is not in itself an indicator of particular station's relevance in the network's structure (Derrible, 2012).

Furthermore, Metro networks are not static entities, cities grow, so they require new routes, new stations to be opened and the creation of transfer points between routes.

Consequently, the network is growing in both extension and complexity as time goes by. Analyzing the evolution of Metro networks provides us with information for future investments as well as alternatives for the improvement of the network's structure. Unfortunately, at the present time, these studies are few and far between (Cats, 2017).

The purpose of this analysis is to show the evolution over time of the complexity of the metro network. Said analysis makes use of quantitative measurements and based on them, explains how decision-making, demographic, political and social circumstances all feed into a resulting structure.

* Corresponding author.
E-mail addresses: salvador.hernandez@itcelaya.edu.mx (S.H. González), idalia@unam.mx (I.F. De La Mota).

https://doi.org/10.1016/j.cstp.2021.07.008
Received 6 August 2020; Accepted 2 July 2021
Available online 14 July 2021

2. State of the art

In recent years, the complex networks approach to analyze transport networks has become a paradigm that has complemented the techniques that are regularly used in the design and analysis of these systems. Some of these techniques are routes optimization, heuristics approaches for the design of transport networks. In the specific case of metro networks, the papers of Gatusso and Miriello (2005), Derrible and Kennedy (2009), Derrible and Kennedy (2010) and Derrible (2012) on the world's metro networks have established that:

1. There were identified three phases in the evolution of metro networks; phase 1 networks were relatively simple in topology and no trends were apparent; phase 2 networks were relatively more complex and larger in size; phase 3 networks including some of the largest metros in the world. We notably found that a quasi-linear trend started to appear between phase 2 and 3 metro networks. Networks in the third phase have a 66% connectivity percentage (Derrible & Kennedy, 2010).

2. Networks can be categorized according to local coverage, regional coverage, and regional accessibility. This depends on the length of the network and the number of nodes (Derrible & Kennedy, 2010).

3. Metro networks can be categorized according to the different emphasis placed on factors in their design: the networks that favor connectivity; networks that favor the maximum number of transfers needed to go from one station to another, which is called directness, and ones that include both factors (Derrible & Kennedy, 2010).

4. There is a high level of correlation between the number of passengers and the number of transfers, network coverage and the number of stations to be gone through on a given journey (Derrible, 2012).

5. There is a positive correlation between the network's complexity and connectivity indicators. Complexity indicator is the ratio between the number of edges and the number of nodes. While connectivity indices are calculated as the ratio of links to nodes – the more links relative to nodes, the more connected (Gattuso and Miriello, 2005).

In general, it has been observed that metro networks have the properties of scale-free networks, as verified by Derrible and Kennedy (2010), who find two structures to be particularly relevant: scale-free networks and small-world networks. Having observed 33 metro systems throughout the world, most metros are scale-free networks (with scale factors ranging from 2.10 to 5.52) while small-world networks show atypical behavior, however they are growing.

The concept of scale-free network emerged in the late 1990′s also, from the work of Barabási and Albert (1999). In their case, they started to investigate the distribution of edges in real-life networks (i.e., number of connections per vertex) and noticed that instead of having bell-shapes, the distributions followed power laws; they notably looked at the structure of the World Wide Web and the US power grid. Derrible (2010).

Small worlds were introduced by Watts and Strogatz (1998); these networks have the particularity of being locally well connected while remaining close, with respect to degrees of separation, to all other parts of the network thanks to existence of a few supra-regional links. Derrible (2010).

In particular, the presence of transfer hubs (stations that have more than three lines) results in relatively large-scale factors. This analysis provides ideas/recommendations for improving the soundness of metro networks. The smallest networks must focus on creating transfer stations, generating cycles to provide alternative paths. For larger networks, few stations seem to have a certain monopoly on transfers, so it is important to create additional transfers, possibly on the periphery of the city centers; the Tokyo system noticeably seems to incorporate these properties.

Other studies have focused on analyzing the structure of the network (Stoilova & Stoev, 2015) and its vulnerability as the Nanjing metro (Deng et al., 2013), terrorist attacks (Sorin et al., 2015) the London metro infrastructure resilience (Chopra et al., 2016), the Mexico City metro as part of the complex transport public network (Flores-De La Mota & Huerta-Barrientos, 2017)

The study developed by Zhang et al. (2018), where the authors of analyze the network characteristics of three metro systems, and two malicious attacks are used to research the vulnerability of metro networks based on connectivity vulnerability and functionality vulnerability. While, at the same time, the characteristics and vulnerability of three metro networks are compared with each other. the results indicate that the proposed methodology is feasible and effective for researching vulnerability and exploring better structures for the metro networks are some examples of this type of analysis.

As regards the evolution and growth of the metro's structure, we can mention the following papers: The first one is Roth et al. (2012), that analyzes the evolution of systems with networks of 100 nodes in cities of over 10 million inhabitants. For most of these networks, the authors found that average degree of a node (station) within the core has a value of order 2.5 and the proportion of $k = 2$ nodes (where k is the average degree of the node) in the core is larger than 60 per cent. In most large urban areas, the network consists of a set of stations delimited by a 'ring' that constitute the 'core'. From this core, quasi-one-dimensional branches grow and reach out to areas of the city further and further from the core. The number of branches scales roughly as the square root of the number of stations, the current proportion of branches represents about half of the total number of stations, and the average diameter of branches is about twice the average radial extension of the core. In conclusion authors found several similarities between different subway systems for the worlds largest cities, despite their geographical and historical differences.

Cats (2017) analyzes the structure of the metro of the city of Stockholm, which alternates between periods of growth and contraction. It also found that, over time, the connectivity of the network has increasingly come to depend on a limited group of stations.

In these papers Roth et al. (2012) and Cats (2017), we can see that the networks do not grow at random, but rather their expansion is governed by public policy and the cities. These mechanisms are reflected in the distribution of the degree of nodes and the growth of the number of edges vs the number of nodes.

The evolution of others transport systems from the perspective of complex networks is reported in the following papers:

- One pioneering paper on traffic in a network viewed from the perspective of complex networks, is that of Daganzo (1994) who says that the applicability of kinematics theory helps predict traffic on the network and can be applied to evacuation plans to deal with disasters.
- Ding et al. (2015), examine the evolution of Public Urban Rail Transit Network of Kuala Lumpur and authors prioritized network protection and revealed that those nodes with the largest degree and the highest betweenness values are most important to the network's operation. These research findings have contributed to the design of the rail network. They also calculated the related network indices and topological network characteristics such as connection, clustering, and centrality.
- Strano et al. (2012) analyze the evolution of road networks in an area near Milan, finding that the layout of the paths has two phases: an early phase characterized by streets with a T structure (nodes with degree 3) and a second phase that indicates a change of paradigm in the planning streets with a cross structure (nodes with degree 4).
- Cats and Jenelius (2018) in their paper: Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability; assert that this study goes beyond the conventional topological analysis of complete link failures. Changes in network performance were examined in terms of total generalized passenger travel costs which were obtained from the agent-based model. and link importance to a dynamic-stochastic setting from the perspectives of both operators and passengers. which were obtained from the agent-based model. Vulnerability metrics were then

calculated for each line and critical segment in the Stockholm rapid PTN. The analysis performed in this study can support tactical planning of disruption management and planning mitigation strategies involving resource allocation and information provision.

There have been several studies made for the Mexico City metro, but none have used the complex networks approach; however, the authors of this article consider that it is important to apply this approach to this metro system given its worldwide importance, as shown by (Gallotti et al., 2016) who analyzed maps of the longest 15 rapid transit networks in the world, based on the total number of stations. They considered all the journeys that a passenger could make from point A to point B with two connections, and then they determined the fastest possible route for a particular journey. That framework is according to people behavior research that shows people can store up to four pieces of information in their working memory at a time, in this case, the place where the journey starts, its destination and two transfer stations.

The result was a "cumulative" complexity rating that placed the New York subway at the head of the group, earning it the title of "the most complex metropolitan system in the world" with Hong Kong metro rail network at the end of the list:

The following table gives the characteristics of each one of these networks

The number of routes N and connections K_{tot}, respectively, yield nodes and edges in the dual space. We list cities from most connections to fewest connections between different lines. *Diameter P* indicates the network diameter in dual space. It is equal to 2 for 10 of the 15 network and one additionally obtains a value of 2 in Paris if one cuts "3bis" (a four-stop line).

Next Fig. 1 shows the Metro network of Mexico City.

Fig. 1. Metro network, Mexico City.
Source: https://www.metro.cdmx.gob.mx/la-red/mapa-de-la-red

The advantages of using the complex network approach are several, as has already been described in the studies of other meters in the world, but fundamentally it allows to analyze the evolution of the network over time considering the metrics that allow to identify important nodes, the routes as well as vulnerabilities (Fig. 2).

By having complex network analysis, simulation can be used to consider possible scenarios as problems with some nodes disconnected, or bottlenecks, among others. This analysis is out of this article.

A disadvantage is that without reliable information about the network, the analysis cannot be carried out completely. Another disadvantage is that a detailed analysis of conditions that affect the system is not considered in a complex network analysis.

3. Basic concepts of complex networks.

This section describes some basic definitions of complex networks, their metrics and the usefulness of this information is presented for analyzing networks not only transport but a wide variety of areas. The words network and graph are used indistinctly.

3.1. Network

A network is an object made up by two sets, Nodes (N) and Edges (A). The set of nodes is a non-empty set, while a set of edges can be empty. A network is represented as follows (Trudeau, 1993):

If the edges have a direction, then the graph is called a directed graph or a digraph. The nodes are numbered $N = 1, 2, \ldots$, n; and the edges are numbered according to the pair of nodes they join (1, …, n-1).

3.2. Complex network

In the context of network theory, a complex network is a graph (network) with non-trivial topological features—features that do not occur in simple networks such as lattices or random graphs but often occur in networks representing real systems. (Barabasi, 2016).

3.3. Adjacency matrix

In mathematical terms, a network is represented by a matrix A with 0′s and 1′s, where (1) (Trudeau, 1993):

$$a_{ij} = \begin{cases} 1 & \textit{if there is an edge connecting nodes i and j} \\ 0 & \textit{Otherwise} \end{cases} \tag{1}$$

3.4. Degree of a node

Degree is the term used to refer to the number of edges that join a node i with others in the network. In an undirected network the total number of edges G, can be expressed as the sum of the node degrees k_i (2):

$$G = \frac{1}{2}\sum_{i=1}^{N} k_i \tag{2}$$

The factor ½ corrects the fact that in the sum (2) each edge is counted twice (Barabasi, 2016).

3.5. Average degree

An important property of a network is its average degree, which for an undirected network is in a network is obtained using the following expression (3):

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2G}{N} \tag{3}$$

3.6. Degree distribution

The degree distribution p_k, provides the probability that a randomly selected node in the network has degree k. Since p_k is a probability, it must be normalized (4):

$$\sum_{k=1}^{\infty} p_k = 1 \tag{4}$$

For a network with N nodes, the probability is calculated as follows (Barabasi, 2016):

$$p_k = \frac{N_k}{N} \tag{5}$$

In the expression (5), N_k is the number of nodes with degree k. In random networks (for example, the Erdös – Rény model) the degree follows a Poisson distribution, whereas in real networks it has been observed that the degree distribution fits a potential model (6):

$$p_k = k^{-a} \tag{6}$$

If we take the logarithm (7):

Fig. 2. Insurgentes metro station 1969. This station was a location in the futuristic movie Total Recall.
Source: https://www.mediotiempo.com/otros-mundos/inauguro-sistema-transporte-metro-ciudad-mexico-cdmx-fotos

$$\log(p_k) = -a\log k \quad (7)$$

In real networks it has been observed that the logarithm for the degree probability (k) has a linear behavior in respect of the degree logarithm (Newman, 2003).

3.7. Betweenness centrality

The importance of a node in respect of the frequency of its appearance on the shortest route between points i and j is measured by betweenness centrality (8). The more important a node is, the higher the proportion of routes that will use it (Barthélemy, 2011).

$$b_i = \sum_{j \neq k} \frac{n_{jk}(i)}{n_{jk}} \quad (8)$$

Where $n_{jk}(i)$ is the number of shorter routes connecting nodes j and k that use node i; while n_{jk} is the number of shorter routes between nodes j and k.

In transport networks, there are stations that more are often found within the route of a passenger's journey and this gives them a higher place in the hierarchy than others.

In information systems, nodes with high betweenness may have considerable influence within a network by virtue of their control over information passing between others. They are also the ones whose removal from the network will most disrupt communications between other nodes because they lie on the largest number of paths taken by messages.

This must not be confused with the concept of transfer node (hub).

3.8. Clustering coefficient

If, in a network, nodes A and C and nodes B and C are connected, there is a high probability of nodes A and C also being connected, forming a group or cluster. There are two types of cluster indices: global and local. The local measures the connectivity of nodes with their neighbors. The global index measures the total number of closed triangles in a network. The global coefficient is obtained by using the following expression (9):

$$C_\Delta = \frac{3 \times \#Triangles}{\#connected\ triples} \quad (9)$$

The clustering coefficient is the proportion of triangles in respect of the number of triplets of connected nodes. The calculation of this index is recommended when there are big variations in the degree of the nodes (Barthélemy, 2011).

3.9. Transfer nodes (Hubs)

A hub is a node with several edges that greatly exceeds the average.

A hub is a component of a network with a high-degree node. Hubs have a significantly larger number of edges in comparison with other nodes in the network.

3.10. Density

The density D of a network is defined as a ratio of the number of edges E to the number of possible edges in a network with N nodes (10).

$$Density = \frac{Edges}{(n(n-1)/2)} \quad (10)$$

Now then, over time the networks evolve, and the nodes form new connections with each other consequently become denser; the number of edges grows faster than the number of nodes following a law of potential growth (11):

$$A(t) \propto N(t)^{-a} \quad (11)$$

Where a is an exponent that determines the speed at which the network becomes denser.

Index α

Also known as the cyclomatic number; this provides the fraction of basic cycles or circuits in the network.

$$\alpha = \frac{E - N + 1}{1/2N(N-1) - (N-1)} \quad (12)$$

In equation (12) the numerator is the cyclomatic number μ, and the denominator can be associated with the maximum cyclomatic number possible as it is the maximum number of possible edges *1/2N (N-1)* minus the number of nodes in a tree graph *(N-1).*

For planar networks (i.e., two crossing edges necessarily create a new node), the maximum number of possible edges is *3(N-2),* and therefore the maximum cyclomatic number is greatly reduced to (*2 N-5);* hence (13):

$$\alpha = \frac{E - N + 1}{2N - 5} \quad (13)$$

As transportation networks are mostly planar, perhaps with the exception of airline networks, using the planar version of α is recommended. If we multiply α by 100, we get the percentage of possible cycles the network has Derrible (2010).

Index γ

With this index instead of looking at the ratio of cycles, it examines the ratio of actual to potential edges. This γ-index is sometimes referred to as connectivity and is defined as follows (14):

$$\gamma = \frac{E}{3N - 6} \quad (14)$$

This is the quotient for the number of edges of the network in respect of the maximum number of edges that said network can have. As the term connectivity has become generic in the literature, we prefer to refer γ as the degree of connectivity. If multiplied by 100, it is interpreted as the percentage of possible edges the network has. Derrible (2010).

Index β

Finally, the last indicator, the β-index, is derived differently. Instead of providing a measure of actual-to-potential property, it simply is the ratio of edges to nodes, in mathematical form (15):

$$\beta = \frac{E}{N} \quad (15)$$

The β-index is therefore the average number of connections per vertex; it is widely used in the scientific community. It has also been referred to as an indicator of complexity (i.e., the more connections per vertex the more complex). Derrible (2010)

Overall, these three indicators enable the understanding of different network characteristics and they have been used in various instances in the transport literature. Derrible (2010)

3.11. Average length of the route and diameter of the network.

The mean length of the route is the average number of nodes that must be visited in the network.

The diameter refers to the as the largest shortest-path to reach the two extremities of the network of all the calculated shortest paths in a network (i.e., take the two furthest nodes, the diameter is the shortest-path to go from one node to the other) (Table 1).

4. Evolution of the Mexico city metro

On September 4th, 1969 one of the most important transport works of Mexico City was opened: The Metro Collective Transport System.

According to the censuses, in 20 years the population of Mexico City doubled, going from 3.1 million inhabitants in 1950 to 6.9 million in 1970 (Plan Maestro del Metro). This meant that the means of transport

Table 1
Network characteristics of the largest connected component for the 15 largest metropolitan systems. Source Galloti et. al (2016).

City	Nodes	Edges	N	K_{tot}	P-diameter
New York	433	497	22	161	2
Paris	299	355	16	78	3
Tokyo	217	262	13	56	2
London	266	308	11	48	2
Madrid	209	240	12	38	2
Barcelona	139	165	11	37	2
Moscow	134	156	11	35	2
Seoul	420	466	12	35	3
Shanghai	239	264	11	35	3
Mexico City	147	164	11	31	2
Berlin	170	282	10	29	2
Chicago	167	222	8	25	2
Osaka	108	123	9	24	2
Beijing	163	176	13	21	4
Hong Kong	84	87	10	12	4

of the time were totally insufficient for moving the city's inhabitants; the need to have a cheap and accessible mass transport system was evident (STC-Metro, 2014).

Before this means of transport entered operations, 40 % of journeys were made within an area delimited Mexicó City mayors: Cuauhtémoc, Miguel Hidalgo, Venustiano Carranza and Benito Juárez (STC-Metro, 2014). This was a factor in designing the network so that lines 1 and 2 follow the routes of the main thoroughfares of the city and connect the residential areas with industrial and commercial sectors, as well as with Mexico City airport. The first expansion connected the north and east of the city with the center. In essence, this first expansion sped up the transfer of people to hospitals, bus terminals and universities (Rosas-Gutiérrez, 2013).

The results of the evolution of the network from 1969, the year when it opened, to the present day are presented below. The Metro network is treated as a closed system, so only the lines that form part of the so-called Collective Transport System, namely lines 1,2,3,4,5,6,7,8,9,12, A and B are considered. This does not include the Suburban Train that connects with the State of Mexico or the Metrobus (Soto-Patiño, 2010).

Table 2 shows the evolution of the length of the network. It is worth mentioning that the biggest expansion occurred between 1979 and 1989 when 69.7 km were added. After 1989, the number of additional kms decreased, from 1989 to 1999, 38.11 km were added, whereas from 1999 to 2009 saw the addition of only 7.6 kms. In 2012 the last line was opened, adding 21.11 km to the network.

To date, no stations in the network have been closed, and the expansions have increased the length of the network at an irregular rhythm (Figs. 3 and 4).

In respect of the number of new stations (nodes), there has been a significant decrease since 1989, with a recovery during the last period. It must be remembered that the opening of a new line does not necessarily imply that a considerable number of new stations are added as the new line uses several already established stations as transfer points. (Figs. 5 and 6).

Table 2
Number of stations (Nodes); number of edges; length of the network and its respective increases. Source: Compiled by the authors.

Period	Nodes	Edges	networklength (m)	Δ nodes	Δ edges	Δmeters
first opened – 1970	45	90	30,628			
Up to 1979	50	100	36,663	5	10,0	6035
Up to 1989	108	234	106,371	58	134,0	69,708
Up to 1999	140	312	144,486	32	78,0	38,115
Up to 2009	148	328	152,169	8	16,0	7683
To date	163	366	173,282	15	38,0	21,113

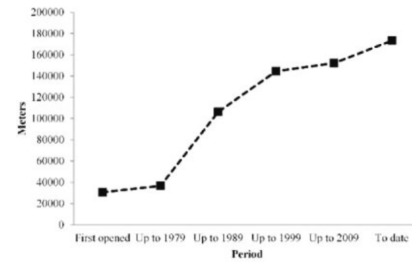

Fig. 3. Length of network per period.

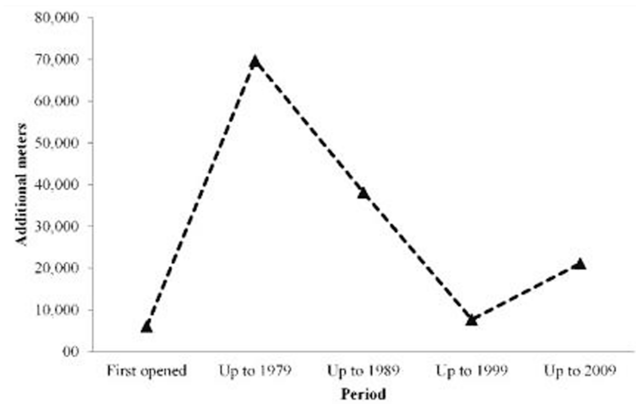

Fig. 4. Additional meters of network per period.

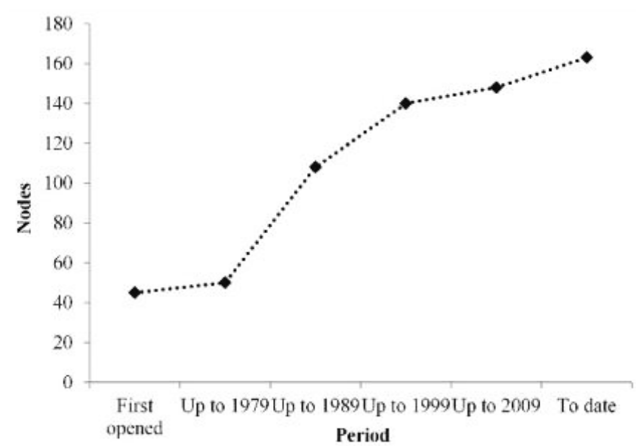

Fig. 5. Number of nodes per period.

The fact that stands out is that the network's expansions have favored the flow of people from the east of the city. This factor precisely coincides with the growth of the city towards in the east and the north. It must be pointed out that the same development has not happened in the south of the capital. The opening of the line 12 was a step towards better mobility in this area. We will return to this aspect later.

Index γ represents an improvement in the period from 1979 to 1999, in other words, the works carried out for 20 years always improved the entire network's connectivity. This behavior is also observable in the period from 2009 to now, which coincides with the opening of the last line. The works in these periods included the opening of connections between lines (Fig. 7).

In Fig. 7, connectivity parameter α shows that from 1970 to 1979 (extensions of lines 1 and 3) the network went from having 54.1% of the possible circuits to 53.7%.

After 1979 the index γ shows slight growth, going from 60.2%, in the

Fig. 6. Increase in respect of the previous period.

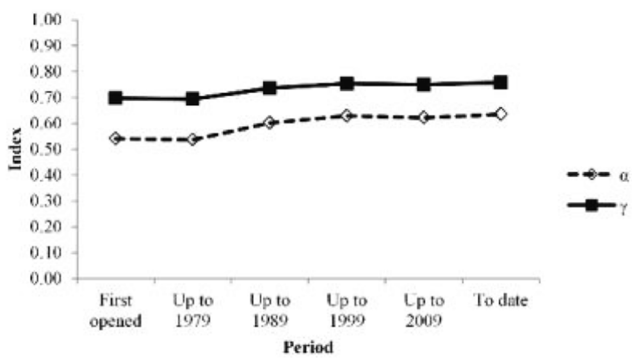

Fig. 7. Connectivity parameters α and γ.

period to 1989, and then 62.9 %, in the period up until 1999. In the period from 1999 to 2009, the value of the index drops slightly to 62.2% (coinciding with the enlargement of the final section of line B), with its growing to 63.6% (the opening of line 12 with several connections to other lines) during the last period.

Moreover, we observe in Fig. 8 that the number of possible circuits has followed a pattern of potential growth in respect of the number of nodes, at least until the opening of the last line.

Up to now, index γ presents a pattern of potential growth (Fig. 9). In this case it is particularly noteworthy that between 1970 and 1979, index γ went from 0.698 to 0.694 while between 1999 and 2009 it went from 0.754 to 0.749. In both cases the change coincides with the enlargement works solely aimed at completing the pre-existing lines, taking them to points where the connection with other lines is not considered. At the present time, the index has a value of 0.758, in other words, the metro network has something more than 75% of all the

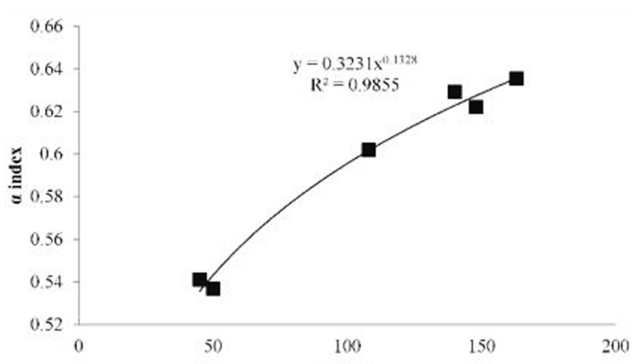

Fig. 8. Evolution of index α.

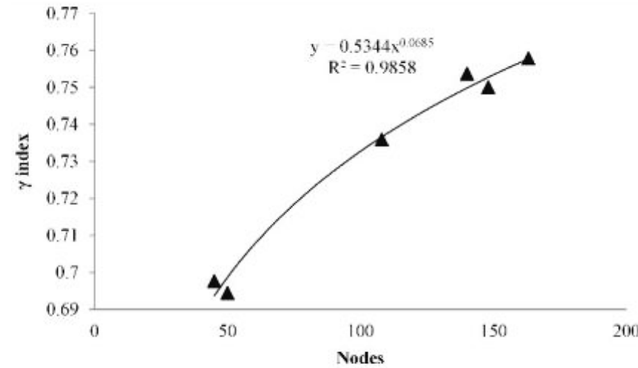

Fig. 9. Evolution of index γ.

possible edges (Fig. 9).

The growth of the network creates new connections that generate the process called densification. This means that the connectivity, in this case relates to ability to travel freely within the network, that is associated also with degree of mobility or density of transfer possibilities.

The average degree (β) of the metro network, up to now, shows a tendency to "settle". In 1999 it reached an average value of 2.229, while in the following two periods its value changed to 2.216 and 2.245, respectively (Fig. 10). The average degree of the network is found in the interval $2 \leq degree \leq 2.245$. By way of comparison, it has been observed that the behavior of the mean degree, in the case of road networks, is linear (Strano et al., 2012).

One conclusion that can be drawn from this result is that the mean degree is not constant. On the contrary, it has changed over time although a frequent assumption is that the mean degree in networks does not experience any significant changes.

Table 3 shows the equations that correlate the number of edges with the number of nodes in the network. In this case, the result indicates that growth has followed an exponential trend, which shows the speed of growth that has not been constant throughout the evolution of this network.

The Fig. 11 shows how the network's density evolved for the exponential case; each point represents a period; the equation's exponent is 1.0941. This is known as the densification exponent (Leskovec et al., 2007).

Once new stations are incorporated, additional kilometers are added to the metro network. At the present time, the average distance between stations is 1213.8 m (Fig. 12); this value is 1.46 times the average length when the first three lines were opened, namely 836 m. This result shows that the network is being designed with an increasingly longer distance between stations, reflecting the fact that this transport system is a means

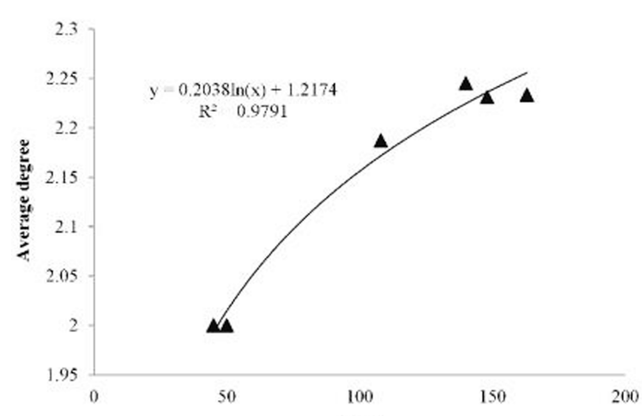

Fig. 10. Evolution of the average β degree of the network vs number of nodes.

Table 3
Regression models for the number of edges vs number of nodes.

Growth	Equation:
Linear	$Edges = 2.3377x - 16.476, R^2 = 0.9988$
Exponential	$Edges = 1.3918x^{1.0941}, R^2 = 0.9999$

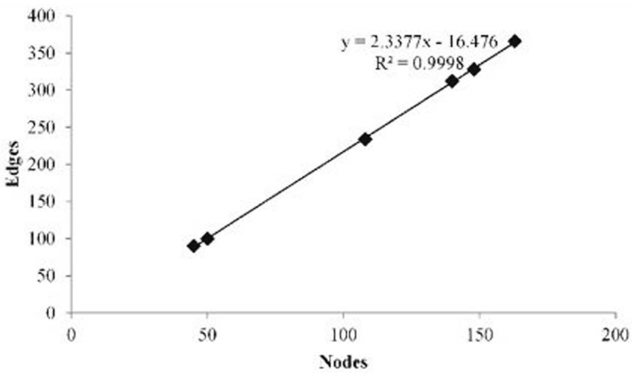

Fig. 11. Evolution of the number of edges vs number of nodes.

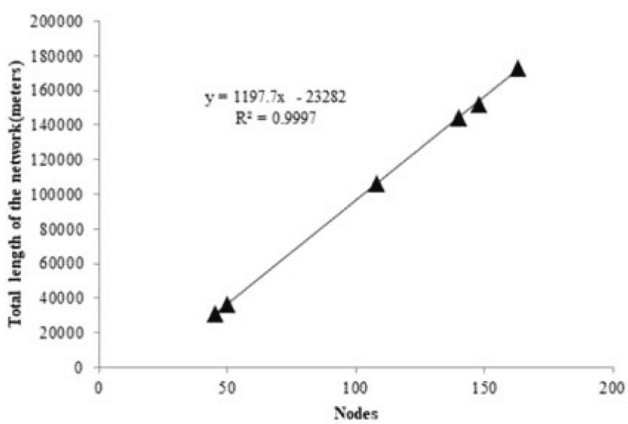

Fig. 12. Total length of the network vs number of nodes.

of bringing people from farther and farther away.

By way of comparison, the average length of European systems is 1000 m, making them very dense systems. Next table 4 shows the equations for the linear and exponential cases that correlate the length of the network with the number of nodes. So far, the kilometers of the network have grown in a linear manner in respect of the number of stations of the network, with an average increase of 1197.7 m f or each new node (station).

4.1. *Degree distribution in the network*

Figure 13 illustrates the evolution of the histogram of frequencies corresponding to the degree of the nodes in the network for each period.

The shape of the bar chart shows a heavy tail to the right, corresponding to the higher-grade nodes. That means, the distributions of

Table 4
Regression models of the total length of the network vs number of nodes.

Growth	Equation:
Linear	$Km = 1197.7x - 23282, R^2 = 0.9997$
Exponential	$Km = 190.85x^{1.3408}, R^2 = 0.9988$

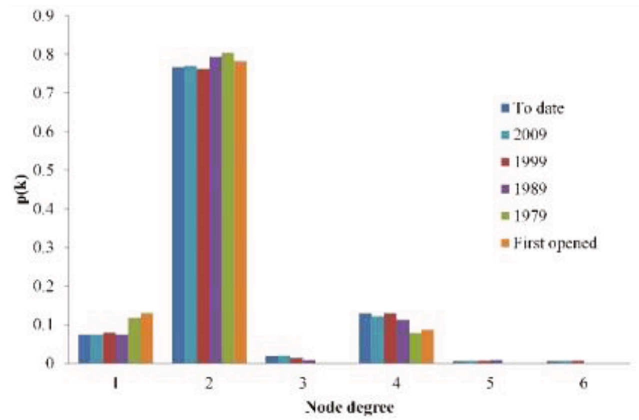

Fig. 13. Bar chart of the network that shows the probability of a node having degree *k*.

number of edges per node (i.e., degree distribution) follow a power law; in other words, few nodes have many edges, and many nodes have few edges.

Figures 14 and 15 shows $p(k)$ in respect of k in logarithmic scale for the periods corresponding to the opening and the present day. We can see the progress over 49 years of the structure of the network to a scale-free network structure; this behavior indicates that the network has not evolved randomly; on the contrary, the decisions taken so far show it is a real network. This probability is related to the power low mentioned in the last paragraph.

Now then, from the perspective of the number of edges, in the present configuration the Chabacano station has 6 edges (center area, connected with 50% of the lines) and Tacubaya has 5 edges (in the west, connections with 41.67 % of the lines); the Pantitlán station (located to the east) has 4 edges (connected with 30% of the lines) and the Pantitlán station has the exclusive functions of a terminal station.

In Fig. 16 we observe the behavior of the maximum degree of the nodes in respect of the number of nodes in the network for each period. Until now, this has corresponded to a linear growth. In scale-free networks the maximum degree grows polynomially in respect of the number of nodes in the network (Barabasi, 2016).

The Hub value has changed throughout the evolution of the network; as we have already said, the urbanization of the city mainly happened in the east and north of the city.

The stations with the most relevance because of the quality of their connections with other nodes were initially the stations of Balderas (connection with lines 1 and 3), Hidalgo (connection with lines 2 and 3), Pino Suárez (connections with lines 1 and 2). Salto del Agua (connections with lines 1 and 8) and Juárez is not a station with connections; however, it appears on the list because of its vicinity to Balderas,

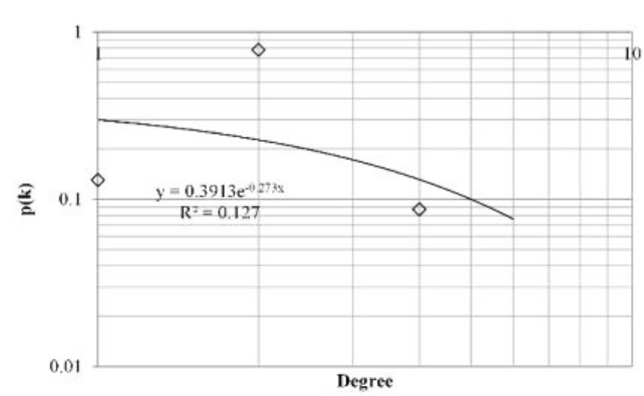

Fig. 14. Probability p(k) vs Degree for the network at its opening.

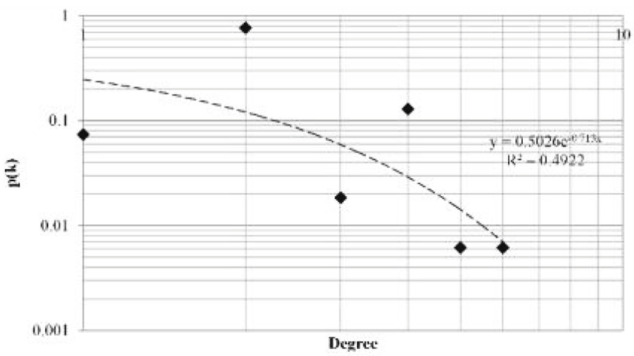

Fig. 15. Probability p(k) vs degree for the present metro network.

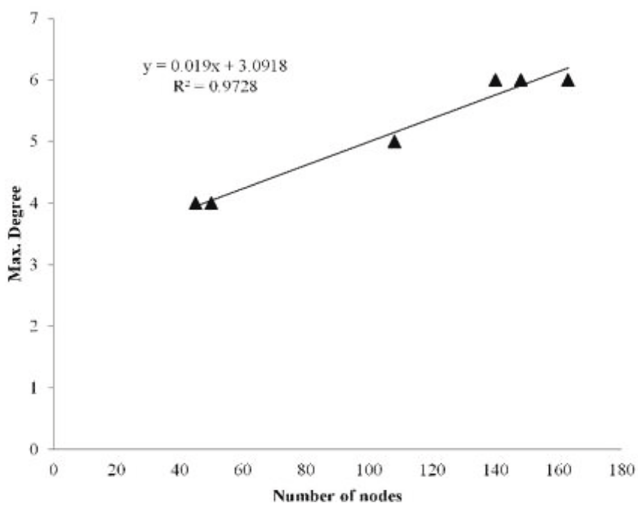

Fig. 16. Maximum degree vs number of nodes.

Hidalgo, and Pino Suárez.

With the opening of new lines, the hubs were geographically displaced in the eastern part of the city and, since 1999, are now the stations of Chabacano (lines 2, 9 and 8), Jamaica (lines 4 and 9), Candelaria (lines 1 and 4), Morelos (lines 4 and B) and San Lázaro (lines 1 and B) (Table 5).

4.2. Betweenness

The average betweenness value has undergone an evolution going from 0.56 when lines 1, 2 and 3 opened, to 0.35 in 2012 when line 12 was put into operation. The value of this index shows that betweenness is distributed in the metro network and is not concentrated on one group of stations (as is the case in the Stockholm metro network, Cats, 2017).

This behavior has been interpreted as a "democratization" of the importance of stations in the network (Table 6).

From1969 to 1979, the stations of Pino Suárez, Hidalgo, Balderas, San Antonio Abad, and Merced had the highest betweenness values (Fig. 17).

The common denominator of these stations is that is found within or very close to what is known as the Historical Center. During the following period (1989), we observe a displacement to the south and west of the capital, with the stations of Centro Médico, Chabacano, Hidalgo, Jamaica and Tacubaya being the ones with the highest relevance.

After 1999, the central stations were displaced by stations in the eastern part of the city until we come to the present configuration, where Jamaica, Chabacano, Candelaria, Centro Médico and Fray Servando stations predominate (Fig. 18).

To have a better idea please check the maps at the Appendix.

What particularly stands out is the fact that:

1. From the perspective of influx, the stations Jamaica, Candelaria, Chabacano, Centro Médico and Fray Servando represent 2.78% of the

Table 5
Most important hubs according to period.

Period	Station	Value of Hubs
First opened	Balderas	0,4043
	Hidalgo	0,3816
	Juárez	0,3199
	Pino Suárez	0,297
	Salto del Agua	0,257
Up to 1979	Balderas	0,399
	Hidalgo	0,389
	Juárez	0,3221
	Pino Suárez	0,288
	Salto del Agua	0,2556
Up to 1989	Chabacano	0,3788
	Jamaica	0,342
	Pino Suárez	0,2877
	Candelaria	0,2639
	San Antonio Abad	0,2539
Up to 1999	Chabacano	0,41
	Jamaica	0,314
	Candelaria	0,286
	Morelos	0,248
	San Lázaro	0,243
Up to 2009	Chabacano	0,41
	Jamaica	0,314
	Candelaria	0,286
	Morelos	0,248
	San Lázaro	0,243
To date	Chabacano	0,41
	Jamaica	0,314
	Candelaria	0,286
	Morelos	0,248
	San Lázaro	0,243

Table 6
Stations with more betweenness per period.

Period	Node	Betweenness
first opened – 1970	Pino Suárez	0,6125
	Hidalgo	0,371
	Balderas	0,362
	San Antonio Abad	0,3329
	Merced	0,3044
Up to 1979	Pino Suárez	0,57398
	Hidalgo	0,454
	Balderas	0,361
	San Antonio Abad	0,306
	Merced	0,2789
Up to 1989	Centro Médico	0,3416
	Chabacano	0,33
	Hidalgo	0,242
	Jamaica	0,233
	Tacubaya	0,223
Up to 1999	Chabacano	0,364
	Jamaica	0,322
	Centro Médico	0,288
	Lázaro Cárdenas;	0,214
	Candelaria	0,189
Up to 2009	Chabacano	0,3515
	Jamaica	0,3238
	Centro Médico	0,2749
	Lázaro Cárdenas;	0,217
	Candelaria	0,209
To date	Jamaica	0,3317
2018	Chabacano	0,287
	Centro Médico	0,229
	Candelaria	0,2168
	Fray Servando	0,197

Fig. 17. Lattice of the Metro in 1970. The size of the nodes indicates their importance with respect to the shortest routes between nodes.

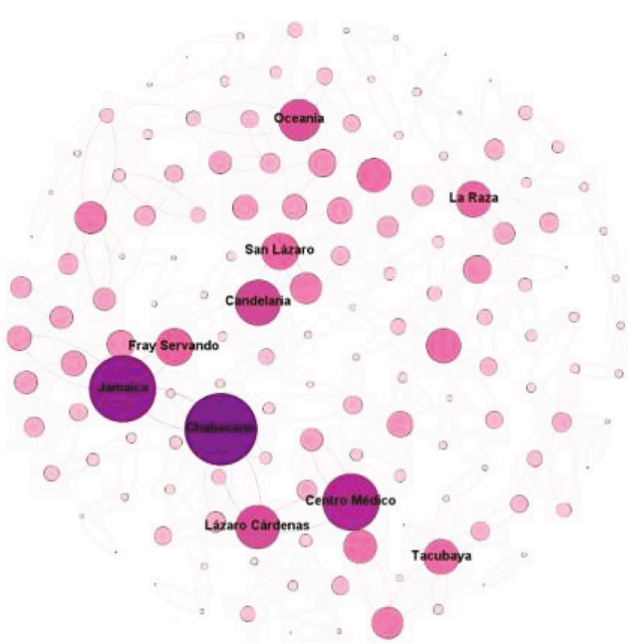

Fig. 18. Lattice of the Metro network in 2018, the size of the node indicates its importance in respect of the shortest routes between nodes.

total passengers (tickets bought), in other words the entry of passengers is low in comparison to terminal stations, such as Pantitlán; however, there is a very high probability of the route the passengers take crossing through some of these stations,

2. These stations (Point 1) belong to lines 4, 9, 2 and 1; these lines coincide with the routes of three of the city's main arterial thoroughfares.

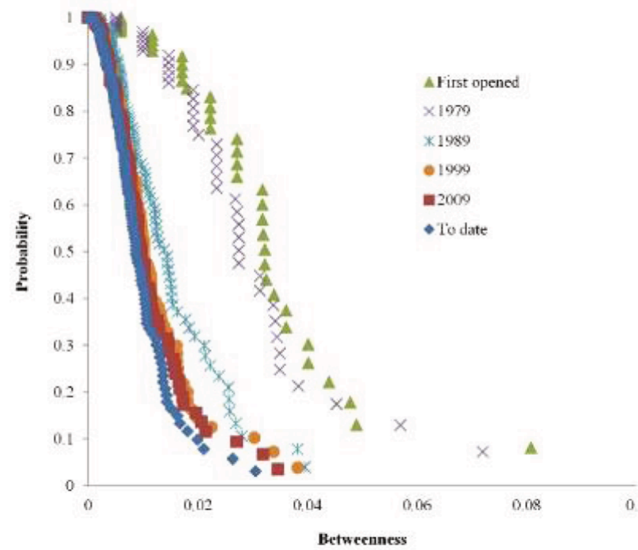

Fig. 19. Betweenness probability.

3. The change of stations with greater betweenness since the metro was inaugurated until today, coincides with the city's growth towards the east, where there is a huge residential zone with high population density, together with Mexico City's International Airport and the city's main wholesale market.

4. This change is also an indicator of socio-economic level because most of the inhabitants of these areas require public transport.

This betweenness result is different from the one reported in Derrible (2012) that situates the stations with the highest betweenness values to the north and west of the city; this is because our paper includes the entire network and not just the transfer stations.

This leads us to conclude that, in terms of connectivity, the most important stations are no longer the ones in the historic downtown area but are now to be found in the eastern part of the city.

Figure 19 shows the distribution, according to period 1969–2018, of the betweenness value. Unlike, for example, the Stockolm network where the change towards a greater dependence on some stations (Cats, 2017), in the Mexico City network the relative importance of the stations tends to have been spread out, which decreases the saturation of the trains, as passengers have several options for their journeys.

The clustering value (measured as the tendency to form triangles) always has low values in this network, which is evidence of the absence of groups, this behavior encompasses the period from the opening of the first three lines (1969) until 1999 with the appearance of a (still unique) triangle formed by the stations Candelaria – San Lázaro – Morelos, which is in the eastern part of the city (Table 7 and Fig. 20).

4.3. Diameter of the network

The purpose of the transport network is to connect distant points in an orderly fashion. One way of measuring the expansion of a network is by looking at the network's diameter. In the case of the evolution of the metro, the diameter has followed a potential growth. When the network opened, the longest journey requires 21 nodes, whereas the diameter has reached a value of 37; in other words, the longest route implies a run of 37 nodes (Table 8).

Leskovec et al. (2007) showed that the diameter of a network tends

Table 7
Clustering coefficient.

Stations	Cluster
Candelaria	0,166
Morelos	0,166
San Lázaro	0,166

to contract as the number of nodes increases; however, in this case the Metro network has not followed this pattern of behavior (Fig. 21) (Leskovec et al., 2007).

Moreover, we can see in Fig. 22 that the mean route length has been following a linear behavior in its evolution. When the network opened, the mean length was 8.24 nodes, whereas it now has a value of 11.989 nodes.

For 2018 the metro presented these data:

Total passengers transported: 1,647 million 475 thousand 013 users

Total courtesy access granted: 208 million 383 thousand 433

Total energy consumed (estimated): 786 million 772 thousand 431 kW

Station with less influx: Deportivo March 18, Line 6 with 681,350 users

Busiest station: Pantitlán Line “A” with 40 million 850 thousand 325 users

Kilometers of the Network in service: 226,488

Kilometers traveled: 44 million 075 thousand 690.54

Service: 365 days.

5. Conclusions

In this paper we analyzed the evolution of the network of the Mexico City Metro using different network metrics.

The number of kilometers in the metro network has increased;

Table 8
Diameter of the Metro network.

Period	Nodes	Diameter
Opening – 1970	45	21
Up to 1979	50	21
Up to 1989	108	23
Up to 1999	140	27
Up to 2009	148	29
To date	163	37

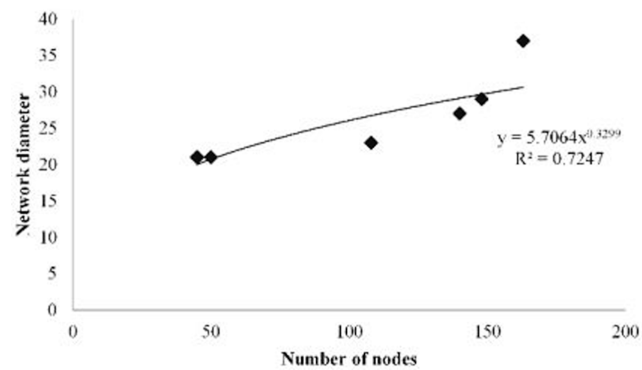

Fig. 21. Diameter of the network vs number of nodes.

however, we have detected a period when the growth slowed down. No stations have been closed so far; they are all still open. Owing to the growing urbanization of the city to the east and north, the new lines have followed a pattern of growth that is mainly located in the east and most of the network’s parameters reflect this.

The connectivity indicators α, γ and β as well as the number of edges

Fig. 20. Triangle formed by three stations.

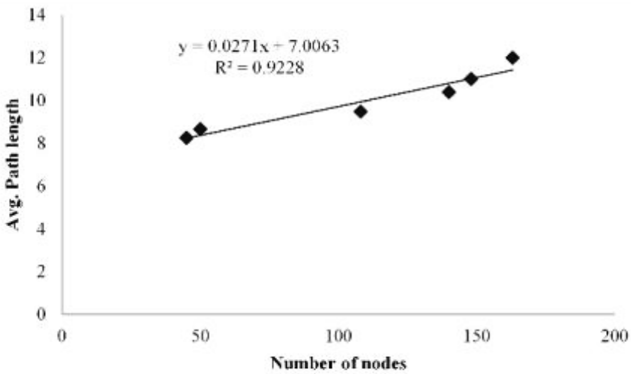

Fig. 22. Mean length of the route vs number of nodes.

in the network show a potential evolution in respect of the number of nodes that are typical of free-scale networks. They show that the expansion works were executed at different speeds in each period. In fact, we observe periods where the expansion works were only aimed at finishing the final sections of the network, only adding missing stations, without any transfers. Moreover, the total length of the network has followed a linear behavior in respect of the number of nodes.

In the case of maximum degree of the network, the growth has followed a linear behavior.

As time goes by, the degree distribution in the Metro network has increasingly fitted in with the behavior a scale-free network; the hub, betweenness and clustering indices also show that the stations with the highest values in each one of these measurements have been displaced from the central area and are now located in the eastern part of the city. In the case of betweenness (betweenness), the results show that the metro network in Mexico does not depend on a group of stations.

In each period, the diameter of the network has increased at a variable rate. It must be pointed out that so far, the diameter has grown following a potential model without giving signs of an apparent contraction.

The research presented in this article only shows how the metro network has grown, as well as its topology since its opening in 1969, so more research is required to consider these problems.

Finally, in accordance with the 2018–2030 Metro Master Plan, the network is proposed to expand based on a proposal presented in the Master Plan, that was supported by a research based on simulation using the EMME4 software, unfortunately this proposal does not consider a complex network approach, what could give more elements to the future expansion, since based on it, an agent-based simulation could give more information or at least to validate the simulation that has been done.

Appendix

References

Barabási, A., Albert, R., 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512. https://doi.org/10.1126/science.286.5439.509.

Barabasi, A.L., 2016. Network Science. Cambridge University Press, Cambridge.

Barthélemy, Marc, 2011. Spatial networks. Phys. Rep. 499 (1–3), 1–101. https://doi.org/10.1016/j.physrep.2010.11.002.

Cats, O., 2017. Topological evolution of a metropolitan rail transport network: the case of Stockholm. J. Transp. Geogr. 62, 172–183. https://doi.org/10.1016/j.jtrangeo.2017.06.002.

Cats, O., Jenelius, E., 2018. Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability. Transportmetrica B: Transport Dynamics 6 (2), 77–96. https://doi.org/10.1080/21680566.2016.1267596.

Chopra, S., Dillon, T., Billec, M., Khanna, V., 2016. A network-based framework for assessing infrastructure resilience: a case study of the London metro system. Retrieved from J. Royal Soc. Interface 1–11. http://rsif.royalsocietypublishing.org/content/13/118/20160113.

Daganzo, C., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. B 28 (4), 269–287. https://doi.org/10.1016/0191-2615(94)90002-7.

Deng, J., Li, Q., Lu, Y., Yuan, Y., 2013. Topology vulnerability analysis and measure of urban metro network: the case of Nanjing. J. Netw. 8 (6), 1350–1356.

Derrible, S. (2010). The Properties and Effects of Metro Network Designs.Ph. D. tesis, University of Toronto.

Derrible, S., 2012. Network centrality of metro system. PLoS One 7 (7), 1–10. https://doi.org/10.1371/journal.pone.0040575.

Derrible, S., Kennedy, C., 2009. Network analysis of world subway systems using updated graph theory. Transp. Res. Record: J. Transp. Res. Board 3 (2112), 17–25. https://doi.org/10.3141/2112-03.

Derrible, S., Kennedy, C., 2010. Characterizing metro networks: state, form, and structure. Transportation 37 (2), 275–297. https://doi.org/10.1007/s11116-009-9227-7.

Ding, R., Ujang, N., Hamid, H., Wu, J., 2015. Complex network theory applied to the growth of Kuala Lumpur's public urban rail transit network. PLoS One 10 (10), e0139961. https://doi.org/10.1371/journal.pone.0139961.

Flores-De La Mota, I., & Huerta-Barrientos, A. (2017). Simulation-Optimization of the Mexico City Public Transportation Network: A Complex Network Analysis Framework. In M. Mujica-Mota, & I. Flores-de-la-Mota, Applied Simulation and Optimization (Vol. 2, pp. 43-79). Mexico City: Springer.

Gallotti, R.P., Porter, M.A., Barthelemy, M, 2016. Lost in transportation: Information measures and cognitive limits in multilayer navigation. Sci. Adv. 2 (2), e1500445. https://doi.org/10.1126/sciadv.1500445.

Gattuso, D., Miriello, E., 2005. Compared analysis of metro networks supported by graph theory. Netw. Spatial Econ. 5 (4), 395–414. https://doi.org/10.1007/s11067-005-6210-5.

Leskovec, J., Kleinberg, J., Faloutsos, C., 2007. Graph evolution: densification and shrinking diameters. ACM Trans. Knowl. Discovery Data 1 (1), 1–41. https://doi.org/10.1145/1217299.1217301.

Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Rev. 45 (2), 167–256. https://doi.org/10.1137/S003614450342480.

Plan Maestro del Metro 2018-2030 https://metro.cdmx.gob.mx/storage/app/media/Metro%20Acerca%20de/Mas%20informacion/planmaestro18_30.pdf.

Rosas-Gutiérrez, J. (2013). La movilidad de los pasajeros en el Sistema de Transporte Colectivo Metro 1997-2012, y su repercusión en el área Metropolitana de la Ciudad de México. Mexico City: Master's Thesis, UNAM.

Roth, C., Kang, S.M., Batty, M., Barthelemy, M., 2012. A long-time limit for world subway networks. J. Royal Soc. Interface 9 (75), 2540–2550. https://doi.org/10.1098/rsif.2012.0259.

Sorin, M., Dehmer, M., Pickl, S., 2015. Network exploratory analysis on subway transportation systems against complex threats including a human factors perspective. Procedia Manuf 3, 6593–6598. https://doi.org/10.1016/j.promfg.2015.07.711.

Soto-Patiño, M.E. (2010). La red del Metro y su eficiencia en distancias físicas. Mexico City: Bachelońs Thesis, UNAM.

STC-Metro. (2014). El Metro es de todos. Ciudad de México: Gobierno del Distrito Federal.

Stoilova, S., Stoev, V., 2015. An application of the graph theory which examines the metro networks. Transp. Probl. 10 (2), 35–48. https://doi.org/10.21307/tp-2015-018.

Strano, E., Nicosia, V., Latora, V., Porta, S., Barthélemy, M., 2012. Elementary processes governing the evolution of road networks. Sci. Rep. 2 (296), 1–8. https://doi.org/10.1038/srep00296.

Trudeau, R., 1993. Graph Theory. Dover, New York.
Watts, D., Strogatz, S., 1998. Collective dynamics of "small-world" networks. Nature 393 (6684), 440–442. https://doi.org/10.1038/30918.
Zhang, J., Wang, S., Wang, X., 2018. Comparison analysis on vulnerability of metro networks based on complex network. Physica A 496, 72–78. https://doi.org/10.1016/j.physa.2017.12.094.