Computers, Environment and Urban Systems 75 (2019) 244–253

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

A topological analysis of growth in the Zurich road network☆

Ylenia Casali*, Hans R. Heinimann

Singapore-ETH Centre, 1 Create Way, CREATE Tower #06-01, Singapore 138602, Singapore

ARTICLE INFO

Keywords:
Network
Betweenness centrality
Distribution
Criticality

ABSTRACT

The functioning of a city depends upon physical structures and services. Over time, a city evolves and the layout of its infrastructure is altered, providing a new arrangement of services for citizens. When designing road infrastructure, planners usually add or change existing roads, looking only at local consequences and ignoring the effects on the whole system. This results in an infrastructure that is developed without a clear understanding of how those alterations affect the global topology. We studied the development of road infrastructure in Zurich, Switzerland, modelled as a network. Our objectives were to 1) assess how non-spatial network topological metrics can identify changes in infrastructure over time and 2) determine any patterns of change in the spatial distribution of those metrics. We analyzed three types of betweenness centrality (BC) metrics with different weighting methods that depended upon structural properties, distance, and population distribution within the urban area. The study resulted in four main findings. First, the number of nodes and edges, together with the network diameter characterized the non-spatial aspect of development. Second, traditional connectivity metrics (alpha, beta, gamma indices) did not yield any changes in time due to their dependence on the average degree, which remained quite constant in infrastructure networks. Third, areas of high BC extended from 1955 to 2012 into a Y-shaped configuration, driven by the development of the main overarching national freeway system. Closeness centrality results show similar patterns. Fourth, the distribution of the 1955 normalized betweenness centrality values transformed into a more heavy-tail distribution in 2012, which implicated that the most critical nodes became even more critical. Future work will enhance this analysis by using more realistic assumptions and incorporating traffic survey data.

1. Introduction

Roads are the backbone of a city structure, enabling the flows of services, goods, and people. Over time, administrative projects progressively add new roads or modify existing ones in the urban landscape. The result is an evolving system that depends upon the decisions of planners. This process usually develops on the local scale and lacks a global view of the road system as a whole. Changes in road connections affect the topology of a network because they modify the internal relationships among network components. Hence, at some locations, topological characteristics vary with time. Alterations in topology lead to changes in the connectivity and vulnerability associated with a road system. Therefore, it is critical that planners investigate system behavior. Whenever the planning process modifies a road network, it creates a gap in the knowledge about new network properties that must be assessed if developers are interested to understand the physics behind the new system. Here, we focused on the problem of network growth, in particular, changes in topological properties over time.

Representations of road infrastructures as networks have been widely examined in fields of geographical and network theory (Barthélemy, 2011; Marshall, 2005). The foundation of topology theory was laid starting with studies by Euler in the 18th Century (Euler, 1736), and has continued to be developed since the 19th Century, when the term “topology” was first introduced (Listing, 1847). More recently, network topology concepts began to be applied in the 1960s for quantitative geography (Haggett & Chorley, 1969). Whereas the traditional stream of topological research looks at connectivity properties, a more recent stream consists of complex network studies. For the former type, Garrison and Marble (1962) and Kansky (1963) were the first to introduce connectivity measures (alpha, beta, and gamma indices) for assessing the structure of a transportation network. Medvedkov (1967) also used connectivity metrics to develop one of the initial approaches to analyzing the evolution of city networks in geographic areas. Since then, the relationship between network connectivity and land cover

☆ ETH Zurich, Future Resilient Systems at the Singapore-ETH Centre (SEC) which was established collaboratively between ETH Zurich and National Research Foundation (NRF) Singapore (FI 370074011), under the auspices of the NRF's Campus for Research Excellence and Techno-logical Enterprise (CREATE) programme.

* Corresponding author.
E-mail addresses: ylenia.casali@frs.ethz.ch (Y. Casali), hans.heinimann@env.ethz.ch (H.R. Heinimann).

https://doi.org/10.1016/j.compenvurbsys.2019.01.010
Received 6 August 2018; Received in revised form 5 December 2018; Accepted 29 January 2019
Available online 05 March 2019

development has been explored in places such as Thailand (Patarasuk, 2013). Levinson (2012) investigated how the network structure varies across some US cities by applying network as well as connectivity metrics. Centrality properties have been examined based on different street patterns (Crucitti, Latora, & Porta, 2006a; Crucitti, Latora, & Porta, 2006b). These network metrics have also been employed to describe the evolution of road networks in urban areas. For example, Barthélemy, Bordin, Berestycki, and Gribaudi (2013) evaluated how planned network transformations might affect topological measures, and they demonstrated that top-down planning influences the spatial distribution of centralities. The growth of London's road network has also been studied from the perspective of logistic functions (Masucci, Stanilov, & Batty, 2013). Monitoring of an urbanized area north of Milan revealed how the evolution of the road network can follow a pattern of densification and exploration, based on centrality values extracted for each road that was added during different developmental periods (Strano, Nicosia, Latora, Porta, & Barthélemy, 2012). On a larger scale, Erath, Löchl, and Axhausen (2009) analyzed the time evolution of the national road and railway systems in Switzerland by using topological metrics and transportation parameters.

Power laws are a common tool for characterizing the heavy-tail distributions of network properties. Lämmer, Gehlsen, and Helbing (2006) analyzed the road networks of the 20 largest German cities and they revealed very broad distributions of betweenness centrality (BC) depending upon the power-law exponent. More recently, Kirkley, Barbosa, Barthélemy, and Ghoshal (2018) investigated the BC distribution scales in 97 cities worldwide. Despite these advancements, network characterizations of urban road infrastructures are lacking in understanding how those topological properties evolve over time. Therefore, the main purpose of this current project was two-fold: 1) to determine how non-spatial topology metrics might explain the growth of a road network system, and 2) to assess how spatial topological metrics can identify patterns of change. Our case study focused on the city of Zurich between 1955 and 2012, utilizing only structural data and excluding any information about traffic flow.

2. Methodology

2.1. Study area

Our study area was the city of Zurich, located in the northeastern part of Switzerland (47.3769 N, 8.5417 E). As delineated by an administrative boundary that represents the political entity of the city, the current boundary was defined in 1934 and has not been changed over the years. It covers approximately 88 km^2 and comprises 12 districts at the northern end of Lake Zurich, where the Limmat River flows northward dividing the urban landscape into two parts. A second river, the Sihl River, passes through from the southwestern border of the city before flowing into the Limmat River at the city center. In 2016, the population was approximately 415,682 (Statistisches Jahrbuch, 2017).

2.2. Data

The Swiss Federal Office of Topography (Swisstopo) provided road vector data collected since 1988, plus a dataset of historical raster maps, with Zurich on sheet numbers 1071, 1091, and 1111. The vector data are line-geometrical features in a shapefile format while the raster maps are images in a TIFF format. Both sources are geo-referenced, projected to the CH1903 coordinate system. For the road vector data, we used swissTLM3D data for 2012 and VECTOR25 data for Years 1988, 1994, and 2000. We noted that these data showed some limitations. In particular, some connections at the motorway junctions changed their representation from 2000 to 2012, due to an alteration in the vector model. In addition, some roads within the 1988 data did not completely overlap with the same roads from the other road vector data. For the raster data, we selected the maps in 1955, 1962, 1970, 1976, and 1982 from the Old National Maps series (1:25,000 scale). Combining these resources produced a dataset spanning nearly 60 years of city evolution, from 1955 to 2012.

2.3. Data pre-processing

Our dataset consisted of heterogeneous sources that had to be converted into a network representation. Before 1988, topographic maps were the only source of information about the Zurich road system. Our analysis required that all data be available in a vector format, which necessitated the conversion of the geo-referenced raster map. Road vector data were extracted for Years 1955, 1962, 1970, 1976, and 1982. To maintain data integrity and consistency across different points in time, we developed a backward branching procedure to convert the geo-referenced raster map into a vector format. We performed this process with ArcGIS 10.4 software (https://www.arcgis.com). Assuming that t was 1988 and t-1 was 1982, we made a copy of the road vector data in Year t and labelled it with the road vector data in Year t-1. The following steps were used:

1) overlay the t-1 map with the t road vector data;
2) select the roads that were not constructed in the t-1 map;
3) eliminate those selected roads from the t-1 vector data;
4) for roads present only on the t-1 map and missing in the t vector data because they had been re-arranged due to new urban planning, add them to the t-1 data; and
5) draw new road lines in the t-1 vector data according to the illustrations from the map for that year.

This method resulted in a consistent road vector dataset to model into a network representation. All of the road vector data extended within the Zurich administrative boundary, which was constant in our time frame. In 2012, bridges and tunnels accounted for no more than 2% of the total number of roads, which meant that the network was almost on the same plane. The delineation of roads as single entities can be achieved through various approaches (Ma, Omer, Osaragi, Samdberg, & Jiang, 2018; Marshall, 2005). Here, we defined a road as the contiguous space that geographically connects two crossroads. For processing the network data, we converted the line information about the road data into a mathematical node–edge structure (graph) by using the network topology toolbox of ArcGIS. This toolbox created the edge and node shapefile datasets from the road vector data that contained a geo-referenced set of lines and points. An ID number identified distinctly every point and line feature of the respective shapefile. Generally, nodes are points on a plane and edges are connections between a set of two node pairs. Two representations are commonly used to model road networks: primal and dual (Porta, Crucitti, & Latora, 2006). Here we used the primal representation, where junctions were nodes and road segments between junctions were edges. The dual representation stipulates the opposite. Although the latter has been well-studied (Batty, 2004), we selected the primal alternative so that the representation was more intuitive for the geographical space and real distances between roads. However, the dual representation has its own advantages. First, the distribution of node degree is more variable, making it easier for one to compare a road network with networks from other systems (e.g., biology, society, or technology) (Porta et al., 2006). Second, by modelling roads as nodes, the network topology represents the linkages of roads without requiring any information about changes in road directions at the junctions. Therefore, the dual network represents an informative point of view of the road–to–road structure (Rosvall, Trusina, Minnhagen, & Snepper, 2005). This property is useful when developing an analysis of community detection (Strano, Viana, Sorichetta, & Tatem, 2018).

For evaluating the topological metrics, we needed an edge list of the road networks, i.e., a list of node pairs belonging to the edges. This list was created using Python (https://www.python.org/) so that it could be further processed with Igraph (https://igraph.org/python/), a network analysis package. We developed a script that read the ID number of each point in the node shapefile and then assigned the respective node-pair IDs to each edge. The script ran within ArcGIS by applying the ArcPy package. We referred to the code from the Geographic Information Systems-Stack Exchange webpage (Tran, 2015), which we adapted to our data.

2.4. Non-spatial analysis

Our non-spatial analysis used topological metrics to characterize the structure of the network in years. We followed two descriptive perspectives – basic network metrics and basic connectivity metrics – and extracted those metrics via the Igraph package in Python. There, basic network metrics are the numbers of edges and nodes, the total road length, and the diameter length. Usually, many paths exist between two nodes. The shortest path is the one between a certain pair of nodes such that the total sum of the edge weights is minimal. The diameter, being the longest shortest path in a network, is a measure of network expansion in time. Using those principles, we then conducted a connectivity analysis by investigating the average node degree ($< k >$) of the whole graph and some common connectivity metrics, i.e., the alpha (α), beta (β), and gamma (γ) indices. Node degree, or the number of edges that touch a node, enables one to count how many roads are connected to a junction in a road network. The average node degree is a function of the total number of edges and nodes (see Eq. (1)). Connectivity indices were introduced in the 1960s during studies of transportation planning (Garrison & Marble, 1962; Kansky, 1963). These connectivity metrics, calculated by Eq. (2), are based on planar graph assumptions, which means that a network is represented in a two-dimensional space without any crossing edges. The α index characterizes the connectivity of a network as the ratio between the observed number of cycles and the maximum number of cycles, producing an interval of 0 to 1. The β index, describing the complexity of a planar graph, is the ratio of the number of edges to the number of nodes, and has an interval of 0 to 3. The γ index, with an interval of 0 to 1, represents the ratio between the observed number of edges and the maximum number of edges.

$$<k> = \frac{2E}{N} \tag{1}$$

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

$$\beta = \frac{E}{N}$$

$$\gamma = \frac{E}{3(N-2)}$$

where E is the number of edges and N is the number of nodes.

2.5. Spatial analysis

Spatial analysis determines which particular group of nodes–junctions is most critical to a road network. There, two spatial perspectives are taken: betweenness centrality (BC) and closeness centrality (CC) metrics. We examined changes in the node BC metrics from 1955 to 2012 and combined ArcGIS tools to visualize the results. Our definition of BC, as proposed by Freeman (1977), states that it depends upon network size, i.e., the total number of nodes. So that we could compare the BC values in networks with different dimensions in time, we calculated the normalized $BC(i)$ of a node i as in [3].

$$BC(i) = \frac{1}{(N-1)(N-2)} \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \tag{3}$$

where σ_{st} is the number of shortest paths going from a source node s to a target node t, $\sigma_{st}(i)$ is the number of shortest paths going from node s to node t that pass through node i,and N is the total number of nodes. The maximum number of total node pairs in an undirected network is $N(N-1)/2$ (Vragovíc, Louis, & Díaz-Guilera, 2005). Because node i cannot be an extreme of the shortest paths, the BC is normalized to the maximum number of possible pairs of nodes that becomes $(N-1)(N-2)/2$ (Crucitti et al., 2006a; Freeman, 1977).

We developed three types of centrality analyses. The first, "basic", evaluated topology based only on road-structure information and considered the graph described only by the adjacency list, without assigning any weight to the edges. However, because we were focused on an actual spatial infrastructure system, we wanted to integrate real measures from Zurich. Therefore, our other two types of analyses involved assigning weights at the edges to account for the geographical and demographic properties of the city. For the first weighted betweenness, we used as weights the lengths of roads to consider the geographical distances of paths within the network. After the actual lengths were determined by ArcGIS, we examined the normalized betweenness, including weights, in Python. For the second weighted betweenness, we computed as weights a proxy for the population distribution. The basic analysis assumed that the probability of exchange for a number of people in between two nodes is equal throughout the network. Because this assumption is not close to reality, we had to relax it by using another data source. Although relying upon traffic data would be the best way to describe the flow properties of a system, they were not readily available for our project. Therefore, we used a demographic distribution to assign weights to edges proportional to the number of people that might live close to each node. Within ArcGIS, we assessed the area of influence associated with the Voronoi cell (Okabe, Boots, Sugihara, & Chiu, 2000), which calculates a tessellation where every point inside a cell is closer to the corresponding node than to any other. The official website for the city of Zurich provided the population data distribution in each district from 1955 to 2012 (Statistisches Jahrbuch, 2015). From this, we identified the district for each cell and then applied Eq. (4) to compute the population at each node for every year examined here.

$$P_i = \frac{(P_d \cdot A_i)}{A_d} \tag{4}$$

where P_i is the population value at node i, P_d is the total population in district d where node i is located, A_i is the area within the Voronoi cell of node i, and A_d is the total area of district d where node i is located. In this way, we obtained a population value for each node. We then determined the population weight at each edge as the sum of the node populations at the node extremes of each edge. Based on this approach, we created weights that approximated the population traffic associated with the roads. Finally, we evaluated, via Python, the normalized BC with this new set of weights from 1955 to 2012.

The normalized closeness centrality was calculated per [5]:

$$CC(i) = \frac{N-1}{\sum_{i \neq j} l_{ij}} \tag{5}$$

where l_{ij} is the shortest-path length between node i and node j. This metric measures how close a particular node is to all other nodes in a network (Crucitti et al., 2006a).

2.6. Distribution analysis

In the distribution analysis, we analyzed the time evolution of the BC distributions with the R package (https://www.r-project.org/). After calculating the empirical distribution of those BC values, we tested the hypothesis that the empirical data are a sample from a parametric distribution function. Previous studies had investigated power-law distributions of BC data in urban road networks (Kirkley et al., 2018; Lämmer et al., 2006). Mathematically, a variable x follows a power law if its probability distribution is $p(x) \sim x^{-\alpha}$ [6], when $> x_{min}$. Clauset, Shalizi, and Newman (2009) have provided a more detailed review of this power-law analysis. In contrast, other researchers concluded that exponential and Gaussian distributions represent BC distributions (Crucitti et al., 2006b). Goodness-of-fit statistics for continuum distributions usually rely upon three types: 1) Cramér–von Mises, 2) Kolmogorov Smirnov, or 3) Anderson-Darling. The Anderson-Darling statistics are of particular interest when the tails are as important as the entire body of the distribution.

3. Results and discussion

3.1. Characterization of the road system

Our analysis covered several perspectives. First, we looked at five scalar metrics, in particular basic network and connectivity, to characterize the network for the city of Zurich. We then examined betweenness centrality to describe the extension of that network between 1955 and 2012. Finally, we studied the statistical distributions of the topological metrics, aiming to understand their variability over time. Although those perspectives focused only on endogenous properties of the road network, we assumed that the embeddedness of the Zurich network into the larger road system, which evolved concurrently, affected the development of our system under investigation. This embeddedness of the Zurich road networks within the regional context is illustrated in Fig. 1, which features the large-scale topology of the local motorway system that was developed between 1967 and 2009. For example, the A1 motorway is the primary West–East connection in Switzerland. It was originally intended to pass through Zurich by forming the "head of a Y", connecting to the A3 motorway, which built the "tie of the Y" (Keresztes, 1985). Although the arms of this Y were opened between 1967 (A3) and 1974 (A1 East), the heart of this Y was never realized due to public resistance.

Consequently, the northern bypass of the A1 motorway was opened in 1985, while the western bypass–Uetliberg Tunnel (A3) and link to

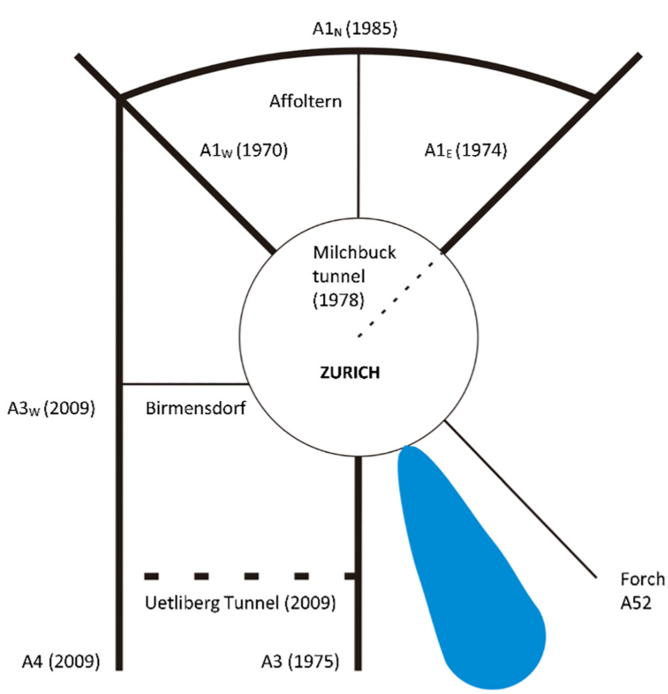

Fig. 1. Main motorway connections for Zurich (marked with circle). Lines represent roads and motorways. Years of construction are displayed in brackets for any infrastructure section. A1, A3, and A4 are Swiss national motorways. For Y-shaped structure of Zurich network, $A1_W$ is western branch and $A1_E$ is eastern branch of A1 motorway. $A1_N$ and $A3_W$ are northern and western bypass, respectively, that connect overall system. Forch road links city with A52 motorway.

the A1 West motorway (A3/A4) core became operational in 2009. The "right arm" of the Y, which connects the city center to the A1 East motorway, was completed in 1978 by opening the "Milchbucktunnel". In addition to the national motorway system, the city of Zurich has three other links to the region. The first, A52, connects with the southeastern area of the region while a southern link connects with the inner part of Switzerland (i.e., Zug and Lucerne). A western link connects the city with the Reuss valley to the southwest. Before this motorway system was built, those links were of primary importance because they carried a heavy traffic load. However, major changes occurred between 1970 and 1978 with the opening of the trunk and arms of the Y motorway system, which we hypothesized had an effect on the development of the Zurich road network.

3.2. Basic network metrics

In the first part of our analysis we looked at the total length of roads, numbers of nodes and edges, and diameter to characterize the urban network development. From 1955 to 2012, the total length grew by 224 Km (1041 Km to 1265 Km); the number of nodes, from 5112 to 6629; and the number of edges, from 7749 to 9883. The diameter, which was the 1% of the total length, also increased from 15.7 Km to 16.9 Km during that time span. Although those growth rates were not linear, we calculated increases of 2.5% for total road length, 3.4% for number of nodes, 3.2% for number of edges, and slightly less than 1% for diameter. We detected no correlation between changes in those parameters and infrastructure development over time. Diameter showed the highest increase between 1976 and 1982, when new infrastructure was being installed on the $A1_E$. Those changes could be estimated by the average shortest paths, which scale as $N^{1/d}$ in a regular lattice where d is the dimension of the lattice, or as $\log(N)$ in a random small-world network (Barthélemy, 2011). Here, we learned that the diameter grew more closely to $N^{1/3}$ than to $\log(N)$. The mean absolute error was 2.2 for $N^{1/3}$ and 7.4 for $\log(N)$. Overall, when combined with changes in diameter, the increase in road length and, consequently, nodes and edges, meant that the network structure developed internally. Because this study area was delineated by political boundaries that did not change during our period of analysis, such development was not caused by expansion, but rather by increased densification of the urban area. Barthélemy et al. (2013) found that, between 1836 and 1888, the road network for Paris (342 Km^2) grew by approximately 300 Km in length and added 3000 nodes. During that same 52-year period, the surface density of increases in Zurich was approximately 3 Km per Km^2 in total length and 17 nodes per Km^2 versus 9 Km and 88 nodes per Km^2 in Paris. This meant that, within that time span, the size of Paris was enlarged by a factor of approximately 3 in length density and by approximately five-fold in node density while Zurich experienced a relative smaller scale of development. We might explain this by noting that, in the core of some major European cities, more urban transformation and, thus, road construction, occurred in the 19th Century than in the late 20th Century.

3.3. Connectivity metrics

The second part of our analysis investigated the development of traditional connectivity indices that were introduced in the 1960s (Garrison & Marble, 1962; Kansky, 1963). These indices rely only upon the numbers of nodes (N) and edges (E). They characterize the degree of connectedness and completeness of a planar network by using ratios: the α index, by the ratio of the observed over the maximum number of circuits; the β index, by the ratio of number of edges over the number of nodes; and the γ index, by the ratio of the observed to the maximum number of edges.

Considering that

$$\alpha = \frac{(E - N) + 1}{(2N + 5)} \tag{7}$$

and that the average degree $< k >$ of a network is:

$$E = \frac{N < k>}{2} \tag{8}$$

substituting E in [7] with [8], and letting N approach infinity, yields the limit of [7]:

$$\alpha_{lim} \propto \frac{<k>}{4} - \frac{1}{2} \tag{9}$$

The limit of the gamma index:

$$\gamma = \frac{E}{3(N - 2)} \tag{10}$$

Running similar calculations, but substituting E in [10] with [8], and then letting N approach infinity, this yields:

$$\gamma_{lim} \propto \frac{<k>}{6} \tag{11}$$

Assuming that the numbers of edges and nodes of a network are large enough, [9] and [11] indicate that the alpha and gamma indices are a function of the average degree only. Thus, $\beta = E/N$ can always be calculated by $\beta = \ < k > /2$. Table 1 illustrates how these indices for the Zurich road network developed between 1955 and 2012. The assumption that a planar network is a good approximation of the road network is reasonable because the number of crossing roads (e.g., bridges and tunnels) is only approximately 2% of the total. Over time, these indices were almost constant, i.e., 0.25 for α, 1.50 for β, and 0.50 for γ. This raised the question of whether this result was a consequence of connectivity or the index definitions themselves. Patarasuk (2013) made similar observations from a study of the road network in Lop Buri Province, Thailand. If we took the limits ([9] and [11]) of the alpha and gamma indices and the average degree of the Zurich network, which was approximately 3.0, then α_{lim}became 0.25 and γ_{lim}0.50. Those results, as expected from the theoretical relationships ([9], [11]) were consistent with the indices calculated with [7] and [10]. The value for the average degree of our road network was approximately 3.0, which fell within the same range of 3 to 4 observed earlier (Courtat, Gloaguen, & Douady, 2011). Considering that man-made road networks seem to have an average degree distribution that does not vary significantly in time, the three indices found here would not be suitable for characterizing topological changes over time, even if the network grew (see Table 1). In contrast, by using a dual representation of road networks, the connectivity results provided different outcomes because the node degree distribution was more variable. In addition, the relationships among indices were constant, which meant that they contained exactly the same information content for large networks. Therefore, knowing one index allowed us to derive the other two easily. Considering an average degree of 3.0, and dividing [11] by [9], the relationship between γ and α took an expected value of exactly 2.0 (Table 1). The same was true for the ratio of β to α, which took a value of exactly 6.0.

3.4. Centrality results

In the third part of our study, we investigated spatial changes in patterns of network growth by looking at more complex topological metrics. The concept of betweenness centrality was first introduced by Freeman (1977) to describe complex networks. The node BC evaluates the shortest paths between every pair of nodes and assigns a ratio to each node after counting how many shortest paths pass through it. This approach allows BC to become a proxy of flows. We used two steps to investigate this centrality: basic and weighted. Basic BC analysis first examined only the simple topological structure of a network by considering just the adjacency of roads. Fig. 2 illustrates the spatial distribution of the basic BC in 1955 and 2012. During that time span, the number of nodes (N) rose, meaning that the network was enlarged. Because increasing the number of nodes meant that there were more shortest paths, then the value of BC at each node would also depend upon network size. To compare more easily the different results from each year, excluding the effect of size changes, we needed to normalize the BC values. To do so, we divided them by the largest number of shortest paths, set at $(N - 1)(N - 2)$, for an entirely connected network with $(N - 1)(N - 2)/2$ numbers of node pairs. This generated a maximum normalized BC value of 0.11 in 1955 and 0.12 in 2012 (Table 2). Afterward, we scaled all of the values within a range of 0 to 1 and divided by the maximum normalized number for each year. This produced five classes of equal interval. Heat areas were indicated by classes with values higher than 0.2. The center of Zurich was consistently characterized by heat areas above 0.8. In neighboring regions, those heat areas extended mainly in four directions connecting the city center with the Southwest, Northwest, Southeast, and North-Northeast sections, all of which were apparent on the 2012 map. In contrast to 1955, when the heat areas were scattered around the city center, this distribution was more narrowed in 2012. Fig. 2 illustrates the nodes with high BC values, which we designated as hotspots. They occurred primarily on the accesses to bridges and tunnels that crossed the rivers and railway lines. Their development progressed from the center toward the Northwest section along the Limmat River, and southwestwardly along the Sihl River. We then looked at changes in spatial distributions and noted a discontinuity between 1955 and 2012. For example, distinct patterns for distribution of heat areas were found from 1955 to 1976 and from 1982 to 2012. These hotspots and heat areas moved from a northerly direction in 1955 to a northeasterly direction in 2012 because infrastructure on the $A1_E$ was constructed in the 1970s. Those building projects included the Milchbucktunnel in 1978 as well as a northeasterly shift in 1982 for some hotspots located on its accesses. The infrastructures connecting Zurich with the northeastern periphery were already identifiable as an area of urban development in the 1990s (Widmer, 1995).

In their study of Paris, Barthélemy et al. (2013) showed that hotspots appeared at some junctions after the Haussmann planning works, which modified road configurations. Likewise, our results indicated that the distribution in Zurich was more in the direction of the motorway, forming a Y-shape inside the city after the 1970s. The BC values

Table 1
Basic topological metrics calculated for city of Zurich for 9 years selected between 1955 and 2012. The annual average growth rates (AAGR) are displayed in the right column.

		1955	1962	1970	1976	1982	1988	1994	2000	2012	AAGR
Total length	(km)	1041	1070	1158	1196	1220	1241	1270	1257	1265	2.5%
# Nodes		5112	5337	6068	6280	6297	6484	6642	6517	6629	3.4%
# Edges		7749	8054	9115	9392	9459	9697	9951	9731	9883	3.2%
Alpha		0.26	0.26	0.25	0.25	0.25	0.25	0.25	0.25	0.25	−0.6%
Beta		1.52	1.51	1.50	1.50	1.50	1.50	1.50	1.49	1.49	−0.2%
Gamma		0.51	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	−0.2%
Diameter length	(km)	15.7	15.6	15.5	15.6	16.3	16.3	16.6	16.6	16.9	1.0%
Average degree		3.03	3.02	3.00	2.90	3.00	2.90	2.90	2.90	2.90	−0.20%

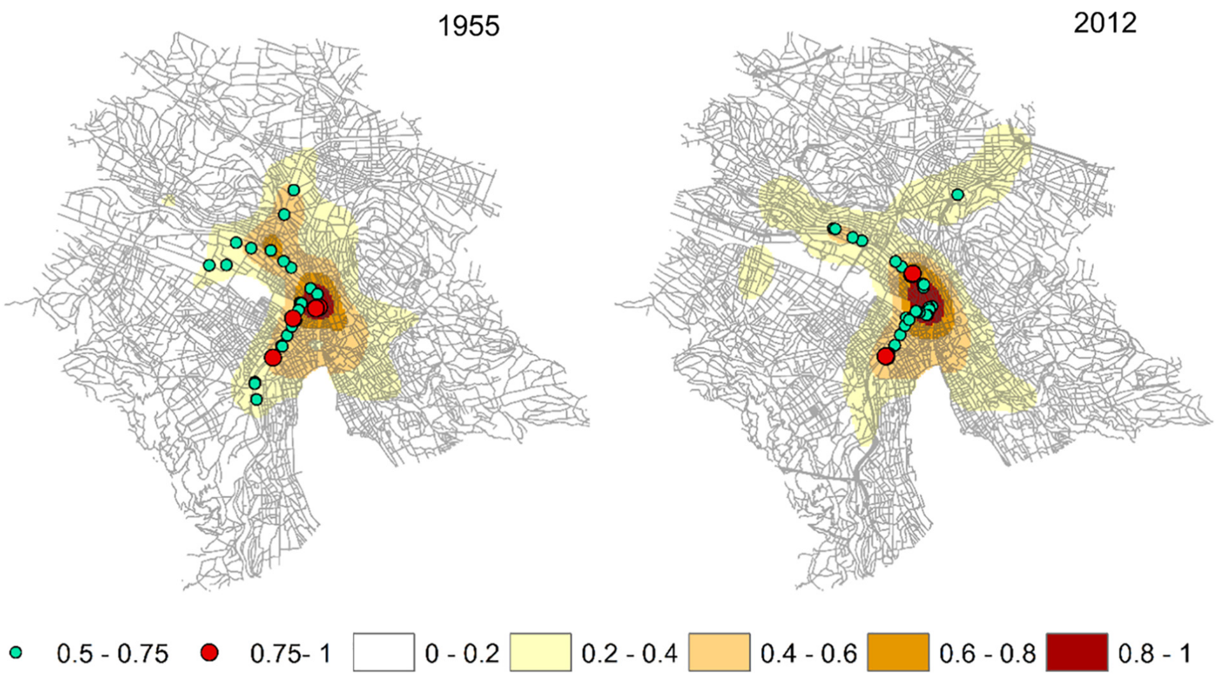

Fig. 2. Spatial distribution of node betweenness centralities (BCs) in Zurich road network in 1955 and 2012. For each year, centrality was evaluated for each network node, dividing number of shortest paths through node by total number of shortest paths within network. Values of betweenness were normalized with (N-1)(N-2), where N is number of nodes for given year. Normalized values were scaled on range [0–1] for each year by dividing them by maximum value of normalized BC. Dots on maps indicate BC hotspots, with different colors representing spatial distribution of betweenness, as evaluated with Kernel density function.

were consistently high in the center. This finding confirmed the results of Barrat, Barthélemy, and Vespignani (2005), who reported that, when space is a critical factor, central nodes tend to get closer to the barycenter, which in our case was approximated by the Zurich city center. Although they did not contradict our theory, the hotspots on bridges and tunnels instead meant, in fact, that those infrastructures acted as shortcuts that would easily connect two otherwise-separated areas. Therefore, rivers and railway lines could serve as geographical constraints that formed junctions, which, if closed, would affect all of the flow in a city.

Urban road networks are designed to support the flow of individuals and goods. Because the basic BC considers only the topological structure of roads and does not include information about the characteristics of flows, we developed our new BC analysis to obtain a 'closer to reality' proxy for flows. To do so, we added geographical layers that were then evaluated by the weights of the edges, their geographical distance from junctions (nodes), and the number of persons who could be connected to the network structure. Fig. 3 displays the results in 1955 and 2012 when the actual lengths of roads were assigned to the weights of the edges. Heat areas were scattered around the center, where values were the highest. Those hotspots then spread and outlined routes inside the city. Although similar, the patterns of distribution expanded over time. When compared with the basic BC, no hotspots were detected to the southwest along the Sihl River. Instead, hotspots increased over time in the southeastern direction toward the route to Forch, as well as along the northwestern route between the north and south sides of the

Table 2
Characteristics of BC hotspots from 1955 to 2012. Two ranges of normalized values were considered: 0.50 to 0.75 and 0.75 to 1. For each selected year, number of hotspots for each range, their sum, and maximum values of normalized BC are displayed, along with annual average growth rate (AAGR). Change in growth rate (%) is expressed as difference in number of hotspots from Year t to Year t-1.

Standard betweenness	1955	1962	1970	1976	1982	1988	1994	2000	2012	AAGR
No. 0.50–0.75 values	23	22	23	16	19	16	14	23	24	4%
No. 0.75–1 values	4	6	8	3	3	3	1	5	4	42%
Sum	27	28	31	19	22	19	15	28	28	5%
Norm. max. value	0.11	0.10	0.10	0.12	0.16	0.16	0.17	0.13	0.12	2%

Betweenness, distance-weighted	1955	1962	1970	1976	1982	1988	1994	2000	2012	AAGR
No. 0.50–0.75 values	91	121	135	112	112	111	106	136	133	6%
No. 0.75–1 values	19	29	30	18	19	18	18	29	28	9%
Sum	110	150	165	130	131	129	124	165	161	6%
Norm. max. value	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	− 1%

Betweenness, population-weighted	1955	1962	1970	1976	1982	1988	1994	2000	2012	AAGR
No. 0.50–0.75 values	36	26	32	27	16	23	18	20	24	− 1%
No. 0.75–1 values	21	20	11	9	4	3	5	4	4	− 13%
Sum	57	46	43	36	20	26	23	24	28	− 6%
Norm. max. value	0.12	0.14	0.13	0.14	0.18	0.18	0.19	0.19	0.19	6%

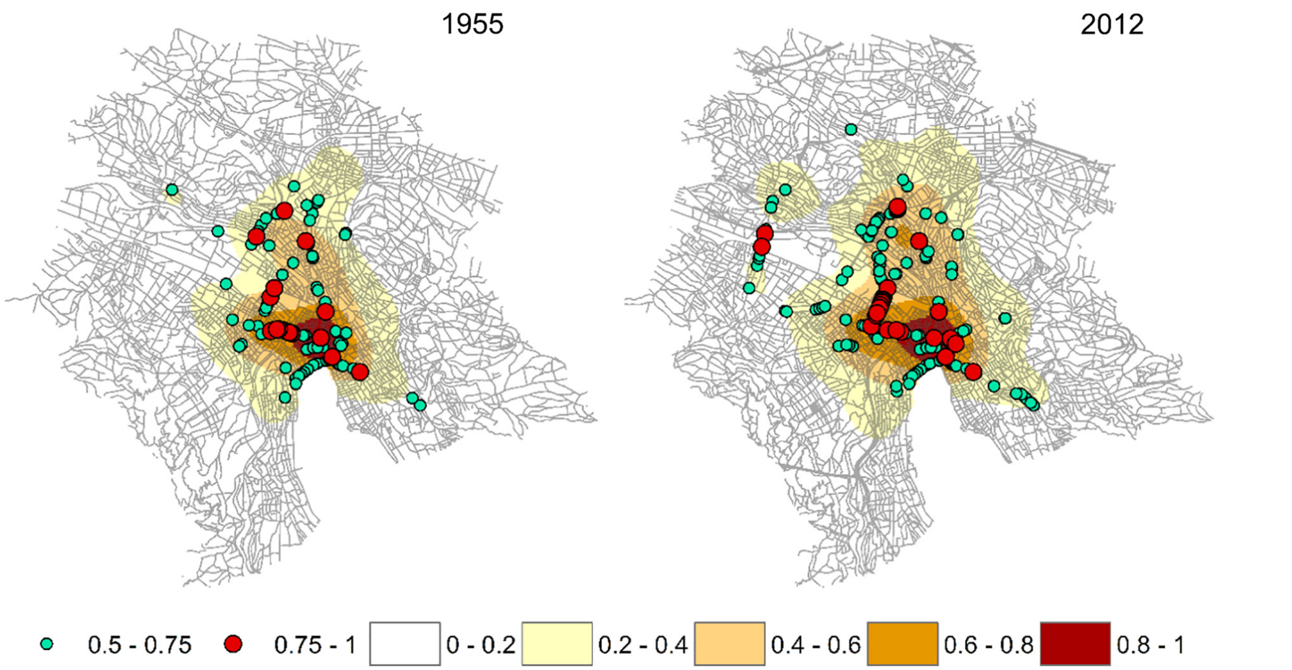

Fig. 3. Spatial distribution of BC nodes weighted with lengths of Zurich network roads in 1955 and 2012. Weight between any 2 nodes was equal to length of road between them. Betweenness values were normalized with (N-1)(N-2), where N is number of nodes for each year. Values were scaled in range of 0 to 1.

Limmat River. This prominent distribution along routes confirmed the results of Crucitti et al. (2006a), who showed that the spatial distribution of BC for several cities captured the continuity of urban routes across numerous intersections.

We also used weights as a proxy to estimate the flow of people passing through each edge (Fig. 4).

Our approach utilized the population distribution in space when assigning weights that accounted for the passage of people inside the network structure. The population at each node was assumed to be proportionate to the total population of the district in which that node was located. We first identified the influence area of a node via the Voronoi tessellation method. Afterward, we determined the number of persons at each node by comparing the population density of the node area with the density of the entire district. Finally, we computed the edge population as the sum of the numbers of persons at the edge's node extremes. The heat areas were higher in the center of Zurich. As the

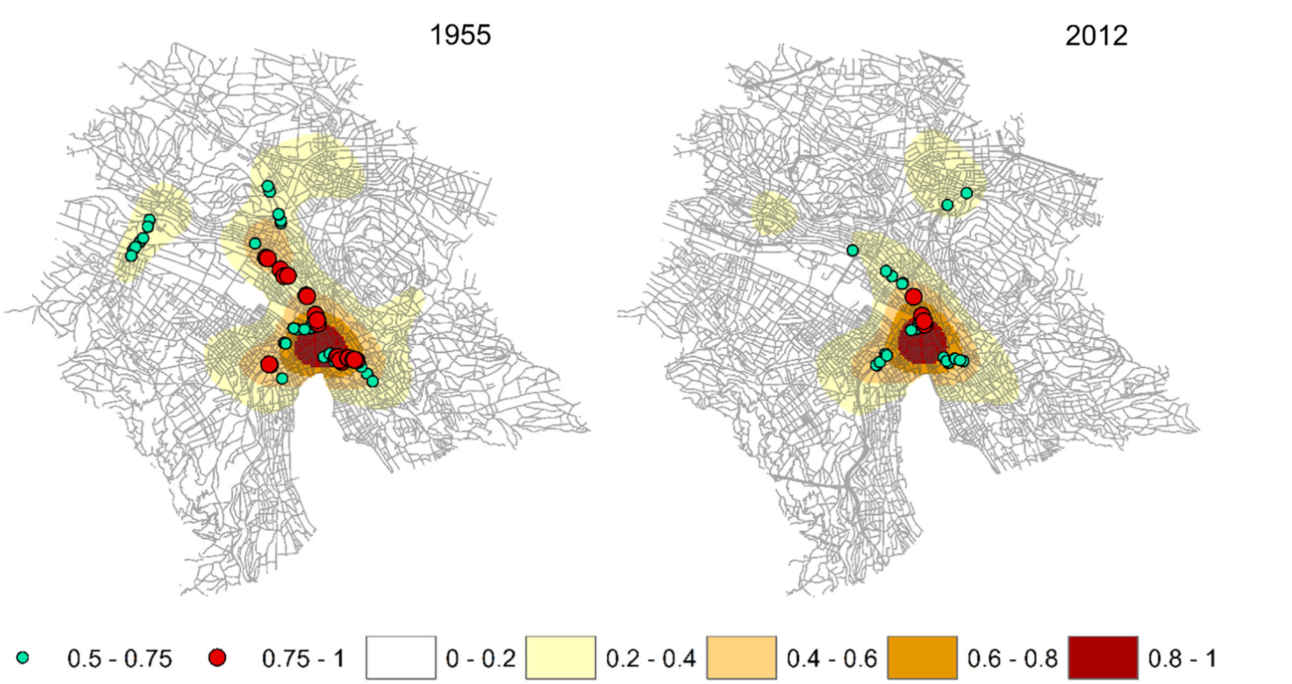

Fig. 4. Spatial distribution of node BCs weighted using population data for Zurich road network in 1955 and 2012. Population weights were derived by first determining Voronoi cells for each node. Population in each node was calculated as $P_i/A_i = P_d/A_i$, where P_i and A_i are population and area, respectively, of cell i, and P_d and A_d are population and area, respectively, for district during given year. Weights between 2 specific nodes were evaluated as sum of their populations. Betweenness values for each year were normalized separately with (N-1)(N-2), where N is number of particular node. Normalized values were scaled on range of 0 to 1.

population expanded to the southwest, northwest, and southeast, hotspots developed toward the Sihl River, the Limmat River, and the route to Forch, respectively. We looked at the differences and noticed that, in 1955, the BC distributions were northward but had shifted by 2012 more toward the northeast. Based on this new pattern post-1982, and in line with the results from our BC analysis, we determined that this movement was a consequence of the development of the A1 east motorway system and construction of the Milchbucktunnel.

Finally, we compared the different BC results according to the number of hotspots. Table 2 shows that the normalized maximum values were almost constant (−1% rate of growth) in the length-weighted BC, but were increased by 2% in the basic BC and by 6% in the population-weighted BC. The number of hotspots increased by 5% and 6% for the basic BC and length-weighted BC, respectively, but decreased by −6% for the population-weighted BC. We detected no specific trend in time. While looking at the values between 1976 and 1982, we noticed a change in the BC spatial distributions but could not identify any particular variation in the number of hotspots. Nevertheless, even slight changes indicated that the internal structure and, thus, network connectivity, was altered over time. For the population-weighted results, an increase in the maximum values for BC, combined with a decrease in the number of hotspots, meant that some nodes became more central with time. Overall, a rise in the number of hotspots or an increase in more-centralized nodes indicated a loss of network robustness because more vulnerable junctions existed that, if closed, would affect system performance.

We extracted the closeness centrality (CC) to complete our spatial analysis based on topological metrics. Closeness centrality can contribute information regarding changes in the distribution of accessibility. The metric identifies which nodes are the closest to the others along the shortest-path distances. As in our BC analysis, we calculated the normalized values, using formula (5), and ranged the results along a 0 to 1 interval. Fig. 5 displays the CC results for 1955 and 2012. Values were highest in the central areas and lowest as one moved closer to the administrative boundary. This result meant that the core of the city is structurally closer than the areas at the boundaries, possibly due to the particular shape of Zurich, which developed around the city center. The 2012 map showed an increase in CC values within the northeastern portion of Zurich, results that were not present in the 1955 map.

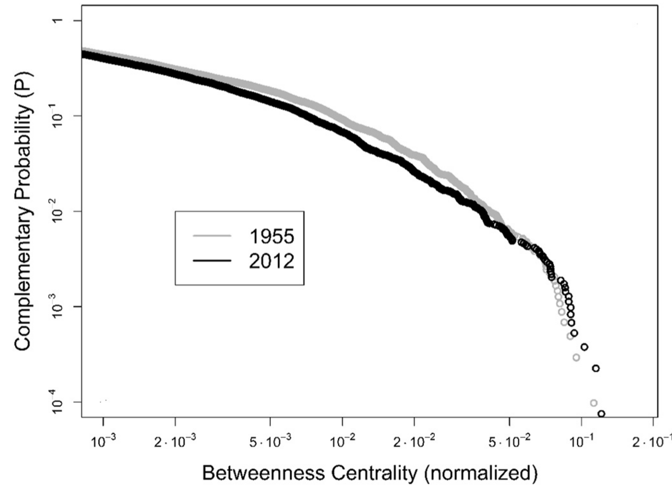

Fig. 6. Distributions of basic normalized BCs in 1955 and 2012. Exceedance probability was lower in 2012 than in 1955 (difference of up to 0.06), where the 2 curves switched positions so that probability was higher for 2012 than for 1955.

3.5. Distribution of betweenness centrality

The fourth part of our analysis examined the variability in topological properties (i.e., node BCs) for the Zurich road network at the beginning (1955) and end of our time series (2012). Because BC is not an absolute metric but depends upon the size of a network, any comparability requires the normalization of BC values. This we performed with the maximum number of shortest paths, which equaled $(N-2)$ $(N-1)$, where N was the number of nodes in a network. From 1955 to 2012, the mean normalized BC value decreased by approximately 19% (from 0.0036 to 0.0029). Whereas the 50% quantile decreased by approximately 25% (0.0008 to 0.0006), the 95% quantile declined by

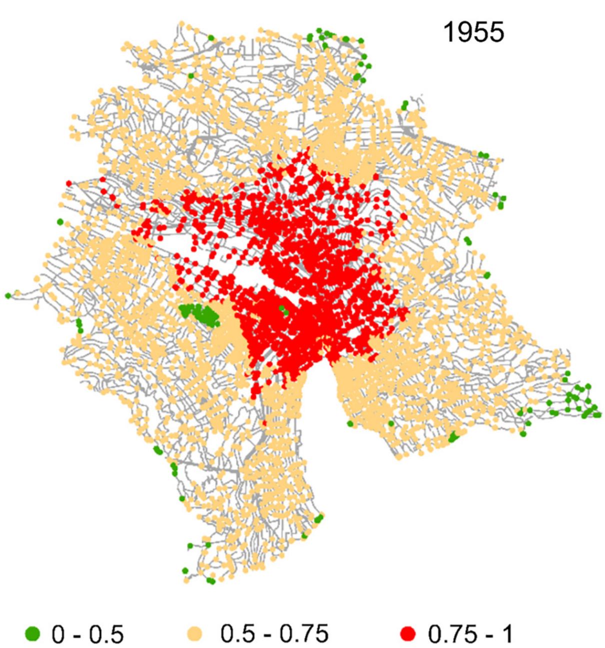

Fig. 5. Spatial distribution of normalized closeness centrality (CC) in Zurich road network in 1955 and 2012. For each year, centrality was evaluated for each network node, dividing by total sum of shortest-path lengths within network. Values of closeness were normalized with (N-1), where N is number of nodes for given year. Normalized values were scaled on range [0–1] for each year by dividing them by maximum value of normalized CC. Different colors on maps indicate different ranges of CC values.

approximately 50% (0.02 to 0.01). Concurrently, skewness rose approximately 30%, from 5.2 to 6.8. Those results indicated that the shape of the distribution changed between 1955 and 2012, and that the densities were higher for smaller BC values. Fig. 6 illustrates these complementary distribution functions, with the exceedance probability $1 - P(BC)$ being a function of normalized BC values. The distribution function for 2012 showed lower exceedance probabilities for values up to 0.06, the point at which we observed a switch. This meant that, at the tail of the curve, the 2012 distribution had larger exceedance probabilities when compared with 1955. Furthermore, this demonstrated that the relative importance of the majority of nodes decreased from 1955 to 2012, while a small number of nodes with high BC values (i.e., the tail) increased in significance. Those values in the tail matter when trying to understand the relative importance of nodes closer to the upper extremes that have a higher incidence of shortest paths. Our findings indicated, therefore, that the relative importance of these nodes increased from 1955 to 2012. Previous researchers have reported that BC distributions follow a parametric function, such as exponential, Gaussian, and power-law distributions. Here, we tested whether our BC values exhibited an exponential, log-normal, or power-law distribution. Performing a goodness-of-fit test, we found that the hypothesis that the empirical distribution is a sample of a theoretical distribution had to be rejected in favor of the exponential and power-law distributions. Although the goodness-of-fit for the log-normal distribution performed better than the other distributions, it still rejected the hypothesis that the empirical observations are an example of a log-normal distribution. These findings were in contrast with those from other studies with a huge dataset of cities that revealed distributions that were either scale-free power-law (Lämmer et al., 2006) or the tails of the BC distributions well approximated by a truncated power-law distribution (Kirkley et al., 2018). For example, Lämmer et al. (2006) indicated that the node BC for some German cities follows a scale-free power law over the entire range of data. Low values for the exponent could be interpreted as a high concentration of shortest paths occurring at the most important intersections. Our results partially confirmed those described by Crucitti et al. (2006b), who found that BC did not show a power-law distribution but instead followed an exponential or Gaussian distribution. In contrast, our BC results were closer to a log-normal distribution.

4. Conclusion

The purpose of this study was to: 1) examine how non-spatial network metrics could explain the historical development of the road infrastructure system of Zurich, and 2) investigate how spatial BC metrics, defining the spatial relations of nodes and edges, might identify patterns of change. We obtained the following major results. First, joint assessment of four basic network metrics – numbers of edges and nodes, total length of roads, and network diameter – adequately characterized the development of the road network. Second, connectivity metrics (i.e., alpha, beta, and gamma indices; Kansky, 1963), which are usually used to monitor transportation networks, detected almost no changes over our test period, from 1955 to 2012. This small variability is explained by the fact that connectivity indices are a function of the average degree in large networks, which remained constant here (at approximately 3) over the years. The absence of spatial referencing by such metrics affected our ability to evaluate network connectivity. Therefore, our results implied that other metrics, embedding the spatial relationships between nodes and edges, should be studied.

Third, our BC results indicated that most of the central nodes were spatially distributed along three main connections forming a Y-shape that linked the city with the national freeway outside the urban boundary. Closeness centrality results showed similar spatial distribution patterns. Beginning with the 1982 maps, hotspots were noted at the access points to the A1 east motorway in the northeastern portion of Zurich. This change reflected some major alterations in infrastructure introduced after the 1970s to improve accessibility. When we added the road-length and population weights to the BC implementations to conduct more ‘close to reality’ evaluations of flows, we acquired different results depending upon the kind of weights selected. Hotspots often occurred along narrow passages, making access to tunnels and bridges a key component when planning network expansions. Values calculated for our length-weighted BC yielded the highest hotspot values along routes that crossed bridges for rivers or railway lines, all locations that are often critical because they link parts of the city that otherwise would be disconnected due to associated natural and artificial gaps. These outcomes extend the discussion presented by Crucitti et al. (2006a), who identified hotspots along the most popular routes in four cities. However, we could not determine from that earlier research why those centralities produced such results. Here, we introduced the topic of how the particular geographical morphology of a city and the development of its internal road structure can lead to a specific distribution of centralities. By defining the most important nodes in an infrastructure system, one might control most of the shortest paths. Consequently, one can determine in which year the configuration of the city was less critical with regard to centrality values.

Finally, the distribution patterns of normalized betweenness centrality changed over time, with a decreasing exceedance probability for lower BC and an increasing exceedance probability for high BC. This potential switch indicated that the tail of the distribution became more fat and critical for a small number of nodes. We could not find any published reports of similar behavior to make a comparison. Furthermore, our data trends did not follow any of the distribution patterns tested here, i.e. power-law, exponential, or log-normal. In contrast, previous researchers have found that BC distributions in an urban road system follow a power-law distribution (Lämmer et al., 2006) or exponential or Gaussian distributions (Crucitti et al., 2006b).

Our findings have implications for planners, policy-makers, and scientists. Historically, road networks were extended and improved in a “piecemeal” fashion, without looking at the overall topology. Changes in an urban road system can be detected through BC metrics. Therefore, from a robustness point of view, developers must analyze the impact of network expansions on the vigor of an entire network if they are to guide the whole project in the direction of greater robustness. Centrality metrics can identify which areas of the city are the most critical. Policy-makers can then use that knowledge to formulate their strategic plans. Employing topological metrics constitutes a quantitative tool for evaluating the physical structure of a city, and planners can take that approach to complete their analysis. New plans can then integrate that information to assess the effects of possible changes in road structure. For scientists, our study contributes to determining how topological metrics can be applied to understand geographical processes. These findings support the use of topological metrics to characterize road systems and to detect properties not easily measurable by spatial metrics alone.

Our analysis was based on the strong assumption that a network adequately represents the flow of services between any pair of edges. This meant that, with respect to location and distance, each pair would have the same probability of exchanging goods and services. For our purposes, we slightly relaxed this assumption by using population-weighted BC metrics, which assumed that pairs of population clusters had a similar probability of exchange. However, real-world behavior demonstrates that such exchanges are established according to locational preferences for work, housing, leisure, etc. Consequently, our present study is a proxy for examining the exchange of goods and services within a city, but this approach must be improved for future investigations. One step could be applying the “first law of geography”, which states that the probability of exchange between two nodes depends on distance, with that probability diminishing as the distance increases (Tobler, 1970). A second step would be to use traffic survey data so we can obtain a more realistic picture of exchange rates between nodes.

This study begins a preliminary discussion about changes in the distribution patterns of normalized betweenness centrality over time. Some questions are still open. A first step would be to examine whether patterns of exceedance probability appear in the development of other cities. The second step would be to expand the investigation about the relationship between changes in those probabilities and actual development within a city.

Acknowledgements

This work is supported by the Future Resilient Systems research programme at the Singapore-ETH Centre (SEC), which was established collaboratively between ETH Zurich and the National Research Foundation (NRF) Singapore (FI 370074011), under the auspices of the NRF's Campus for Research Excellence and Technological Enterprise (CREATE) programme. A special thank goes to Monika Niederhuber who supported us in the acquisition and preprocessing of the spatial data.

Declarations of interest

None.

References

Barrat, A., Barthélemy, M., & Vespignani, A. (2005). The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment, 2005*(05).

Barthélemy, M. (2011). Spatial networks. *Physics Reports, 499*(1–3), 1–101.

Barthélemy, M., Bordin, P., Berestycki, H., & Gribaudi, M. (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific Reports, 3*, 2153.

Batty, M. (2004). *A new theory of space syntax. Center for advanced spatial analysis.* London: University College. http://www.casa.ucl.ac.uk/publications.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review, 51*(4), 661–703.

Courtat, T., Gloaguen, C., & Douady, S. (2011). Mathematics and morphogenesis of cities: A geometrical approach. *Physical Review E, 83*, 036106.

Crucitti, P., Latora, V., & Porta, S. (2006a). Centrality in networks of urban streets. *Chaos, 16*, 015113.

Crucitti, P., Latora, V., & Porta, S. (2006b). Centrality measures in spatial networks of urban streets. *Physical Review E, 73*, 036125.

Erath, A., Löchl, M., & Axhausen, K. W. (2009). Graph-theoretical analysis of the Swiss road and railway networks over time. *Networks and Spatial Economics, 9*, 379–400.

Euler, L. (1736). Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Petropolitanae, 8*, 128–140.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35–41.

Garrison, W. L., & Marble, D. F. (1962). The structure of transportation networks. *Technical report.* Washington DC: U.S Army Transportation Command.

Haggett, P., & Chorley, R. J. (1969). *Network analysis in geography.* London, UK: Edward Arnold.

Kansky, M. (1963). *Structure of transportation networks: Relationships between network geometry and regional characteristics.* Chicago, IL, USA: University of Chicago Press.

Keresztes, J. (1985). Zur Eröffnung der Nordumfahrung Zurich N20 und des Milchbucktunnels SN1. *Schweizer Ingenieur und Architekt, 103*, 539–540.

Kirkley, A., Barbosa, H., Barthélemy, M., & Ghoshal, G. (2018). From the betweenness centrality in street networks to structural invariants in random planar graphs. *Nature Communications, 9*, 2501.

Lämmer, S., Gehlsen, B., & Helbing, D. (2006). Scaling laws in the spatial structure of urban road networks. *Physica A, 363*, 89–95.

Levinson, D. (2012). Network structure and city size. *PLoS One, 7*(7), e29721.

Listing, J. B. (1847). Vorstudien zur Topologie. *Vandenhoeck und Ruprecht.* Germany: Göttingen.

Ma, D., Omer, I., Osaragi, T., Samdberg, M., & Jiang, B. (2018). Why topology matters in predicting human activities. *Environment and Planning B: Urban Analytics and City Science, 0*(0), 1–17.

Marshall, S. (2005). *Streets and patterns* (1st ed.). New York, NY, USA: Spon Press.

Masucci, A. P., Stanilov, K., & Batty, M. (2013). Limited urban growth: London's street network dynamics since the 18th Century. *PLoS One, 8*(8), e69469.

Medvedkov, Y. V. (1967). An application of topology in central place analysis. *Regional Science Association: Papers XX.* Hague Congress.

Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. (2000). *Spatial tessellations: concepts and applications of Voronoi diagrams* (2nd ed.). New York, NY, USA: Wiley.

Patarasuk, R. (2013). Road network connectivity and land-cover dynamics in Lop Buri Province, Thailand. *Journal of Transport Geography, 28*, 111–123.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A primal approach. *Environment and Planning. B, Planning & Design, 33*, 705–725.

Rosvall, M., Trusina, A., Minnhagen, P., & Snepper, K. (2005). Networks and cities: An information perspective. *Physical Review Letters, 94*(2), 028701.

Statistisches Jahrbuch der Stadt Zürich 2015. Stadt Zürich. (2015). https://www.stadt-zuerich.ch/prd/de/index/statistik/publikationen-angebote/publikationen/Jahrbuch/statistisches-jahrbuch-der-stadt-zuerich_2015.html (accessed 19 July 2018).

Statistisches Jahrbuch der Stadt Zürich 2017. Stadt Zürich. (2017). https://www.stadt-zuerich.ch/prd/de/index/statistik/publikationen-angebote/publikationen/Jahrbuch.html (accessed 19 July 2018).

Strano, E., Nicosia, V., Latora, V., Porta, S., & Barthélemy, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports, 2*, 296.

Strano, E., Viana, M. P., Sorichetta, A., & Tatem, A. J. (2018). Mapping road network communities for guiding disease surveillance and control strategies. *Scientific Reports, 8*, 4744.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*, 234–240.

Tran, J. K.. Standalone python script to split a polyline with a point layer. Geographic information systems-stack exchange. (2015). https://gis.stackexchange.com/questions/154708/standalonepython-script-to-split-a-polyline-with-a-point-layer (accessed 15 December 2017).

Vragovíc, I., Louis, E., & Díaz-Guilera, A. (2005). Efficiency of informational transfer in regular and complex networks. *Physical Review E, 71*, 036122.

Widmer, B. (1995). Development of the region-the case of Zürich Nord. *Cities, 12*(2), 97–105.