AN ANALYSIS OF SPATIAL STRUCTURE OF URBAN REGIONAL NETWORKS USING GIS

Vitalii Kostin*, Lena Halounová

Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

* corresponding author: `vitalii.kostin@fsv.cvut.cz`

Abstract. Road network is the foundation of the urbanization process. Initially, it represented an indispensable part not only for the very existence of cities but also for their further economic and social development. Gradually evolving over an extended period of time, networks acquire a certain pattern that can affect the functioning of the entire urban system. This article presents an analysis of the structural properties of the transportation networks across the thirteen large urbanized regions in the Czech Republic. Taking advantage of modern GIS technologies, we investigate the geometric and topological characteristics of road networks on detailed spatial data. The aim of this study is to analyse the qualities of transportation networks that arise from the interaction of their structural components. The results show that the properties of studied urban networks vary from region to region, however, we have determined some common patterns.

Keywords: spatial analysis; transportation network; structure; GIS.

1. Introduction

A city is a unique spatial and socio-economic system that provides the functioning of various aspects of society. Historically, cities served as spatial centres for economic growth and social transformation encouraging more and more people to move into urban areas. The success of urban life lies in the ability of cities to provide rapid movement and communication by bringing large number of people into close proximity [1]. However, economic benefits and developed social interactions would not be possible without the circulation routes ensuring the vitality of the urban space. The roads were initially the cause of the birth of cities and then turned into incentives for their development: they provide the exchange of material, cultural and intellectual values both between people and cities. Therefore, the study of the transportation networks can reveal an important properties of the urban system. Over the past few decades, many researchers from various fields, including urbanism, geography and economics, have focused their attention on this topic. Using different measures, scholars investigate the connectivity of streets in the transportation network [2, 3], the accessibility of certain locations in urban environment [4, 5] or the evolution of a street network over time [6]. From the methodological perspective, the graph theory remains the most popular approach in the study of complex networks, when the system is represented in the form of nodes (intersections) connected by edges (road segments). The main disadvantage, however, is the use of a simplified model of the urban space in calculations. The metric distance is often replaced by simple topological steps and the network itself is completely planarised (i.e. all lines break at their intersection points), which violates the real world connectivity rules such as overpasses.

Taking advantage of modern GIS technologies and the availability of detailed spatial data, we perform the analysis of structural properties on real networks while preserving their topological and geometric features. This paper is aimed at the quantitative description of the properties of large urban systems (at the regional scale) that emerge from their spatial form.

2. Study area and data

In this paper, we focus on the analysis of the structure of large urban networks. The study area is represented by thirteen regions or higher-level territorial self-governing units according to the current administrative division of the Czech Republic (Figure 1). Although Prague, as a capital city, belongs to the highest territorial unit or region, we decided to exclude it from the analysis in order to preserve the size similarity in our sample. In general, our study area covers a territory of $78371\,\text{km}^2$ for thirteen regions of the country with a total network length of $107616\,\text{km}$.

The data we use in this work were obtained from OpenStreetMap (OSM). The undoubted advantage of this service is the free access to extensive cartographic information from around the world. In recent years, the OSM is gaining an increasing popularity among researchers, government institutions and commercial structures [7]. In our case, the use of the OSM greatly facilitated the collection of data and allowed us to apply the methods of spatial analysis to a broad geographical unit. However, we should mention that this service also has several disadvantages. Since the OSM is created by users from all over the world, it is not possible to guarantee the quality of the final vector layer. The most common errors that can be found in the OSM data include disconnected roads (dangles), overlapped segments or pseudo connections, when one
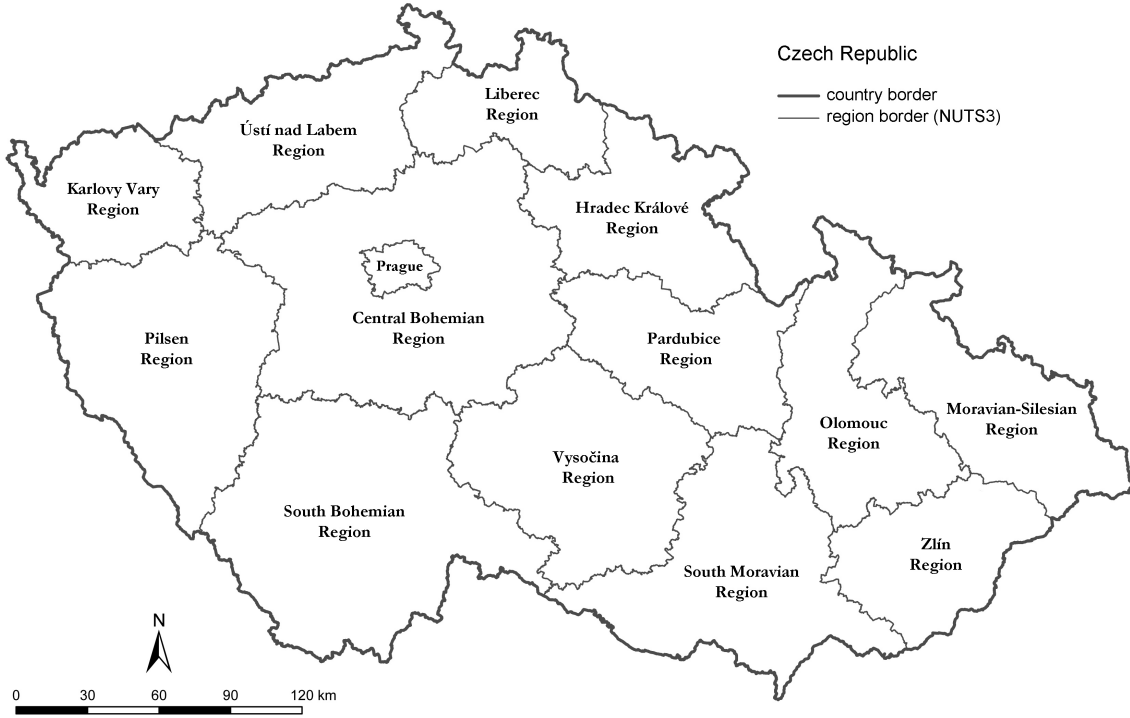
FIGURE 1. The location of study regions in the Czech Republic.

line is attached to another without a junction. Therefore, before using the OSM data in the analysis, it is necessary to check their quality for the presence of the mentioned topological errors. Another problem of the OSM data is a different degree of coverage within a certain area. Usually, highly urbanized areas are mapped in greater detail, than, for instance, some rural areas. In the case of the Czech Republic, most areas are covered very well thanks to the active community of the local OSM project. In order to access the OSM database, we used the service provided by Trimble [8], which allows users to download data for relatively large areas. The process of data collection included several steps. We first downloaded the vector layer for the entire road network in GeoJSON format. Then, we projected our data to S-JTSK Krovak East North coordinate system (EPSG: 5514) and stored them in a geodatabase (.gdb), keeping all the original attributes. We also checked the topology of networks for the mentioned problems such as overlapping segments and dangles. These steps were performed using ArcGIS 10.4 software.

3. METHODOLOGY

For the analysis of selected transportation networks, we used the principles and measures of a graph theory. This area of discrete mathematics has become very popular among researchers studying complex systems. The graph of our transportation networks consist of two main elements: edges that represent road segments and nodes that represent intersections as well as dead-ends. We use Model Builder in ArcGIS 10.4

and Python script in order to construct the graph and calculate different measures on selected networks. The construction of a graph involves the transformation of the network so the nodes are formed at the intersection of roads and in dead-ends. However, we do not planarise the network entirely, but preserve the initial geometry of underpasses and overpasses. For instance, if two roads intersect and one of them is an overbridge, then the node will not be created at this intersection. After each graph of selected regions is formed, we analyse their structural properties and multiple geometric characteristics, such as road length, node density (ϱ_{nodes}), edge density (ϱ_{edges}) and intersection density ($\varrho_{\mathrm{intersections}}$). We apply the classical density concept, which measures the concentration of elements per unit area. In the case of edges, we calculate the ratio of the total edge length (L_{edges}) and the region area (Area). The degree of a node is the number of edges converging at it. We define the degree of each node in our graphs, and then find the average value ($\mathrm{Deg}_{\mathrm{avr}}$). In order to describe some properties of networks, we also determine the total number of nodes with a degree of 1 (N_{deg1}), nodes with a degree of 3 (N_{deg3}) and nodes with a degree of 4 (N_{deg4}).

The level of connectivity in a graph can be measured by the Beta index (β_{inx}). This measure defines the density of connections in the graph as the ratio [9]

$$\beta_{\mathrm{inx}} = \frac{N_{\mathrm{edges}}}{N_{\mathrm{nodes}}}, \qquad (1)$$

where N_{edges} is the number of edges in the network and N_{nodes} is the number of nodes in the network.

No.	Region	Population 1. 1. 2018	Area (km^2)	N_{nodes}	N_{edges}	$N_{\text{intersections}}$	Total L_{edges} (km)
1	Central Bohemian	1352795	10928.34	63945	83703	48304	18798.02
2	Moravian-Silesian	1205886	5430.52	35094	44836	25628	9755.99
3	South Moravian	1183207	7185.87	33173	43037	24892	9380.07
4	Ústí nad Labem	821080	5338.64	25803	33911	19710	7884.36
5	South Bohemian	640196	10058.16	27579	36055	21006	10912.84
6	Olomouc	633178	5271.53	21908	28133	16203	6762.49
7	Zlín	583056	3961.50	19078	23611	13385	5049.01
8	Pilsen	580816	7648.59	24075	31256	17910	8928.40
9	Hradec Králové	551089	4759.22	20157	26220	15156	6729.06
10	Pardubice	518337	4519.56	19236	24942	14403	6284.84
11	Vysočina	508916	6795.09	21164	27640	16125	8357.77
12	Liberec	441300	3163.61	17165	22647	13239	5185.26
13	Karlovy Vary	295686	3310.12	9929	12791	7340	3588.01

Table 1. Basic characteristics of the studied regions and their road networks.

Gamma index is another connectivity measure, which is determined by the ratio of real number of edges to the all possible edges in the network [10]:

$$\gamma_{\text{inx}} = \frac{N_{\text{edges}}}{3N_{\text{nodes}} - 2}, \qquad (2)$$

where N_{edges} represents the number of edges in the network, while N_{nodes} is the number of nodes in the network. The value of Gamma ranges from zero to one where a value of one indicates a complete network.

The organic ratio (OR) is a measure that allows analysts to determine whether a city or a region was planned or developed as self-organized. In most urban areas, the distribution of the node degree has a maximum value of 3 or 4. The organic ratio describes the distribution of the node degree as [11]:

$$\text{OR} = \frac{N_{\text{deg1}} + N_{\text{deg3}}}{N_{\text{nodes}}}, \qquad (3)$$

where N_{deg1} is the number of nodes with a degree of 1, N_{deg3} is the number of nodes with a degree of 3 and N_{nodes} is the number of nodes in the network.

On the one hand, planned cities have a fairly homogeneous network with a gridiron structure. Most nodes in such networks have a degree of 4, thus OR $\approx$ 0. In self-organized cities, on the other hand, the networks developed without imposing the standardized patterns, so their structures are characterized by the predominance of nodes with a degree of 3 and 1 (OR $\approx$ 1).

4. Results

The main characteristics of selected regions, sorted by the population, and their road networks are presented in Table 1.

The Central Bohemian Region represents the largest urban area in terms of all indicators presented in Table 1, which is not surprising due to its geographical location. A significant impact on the network

structure here has Prague as a main centroid for the interstate highways. In terms of size, the Liberec Region has the smallest area, however, the Karlovy Vary Region is distinguished by low values of other parameters. This western region of the country has a rich environmental resources (forests, cliffs, mineral water springs), including landscape protected areas, which, together, have an impact on the features of the road network. Another important factor influencing the characteristics here is the presence of military training area Hradiste Doupov, which extends over an area of $332\,\text{km}^2$. Such areas occupy vast territories, however, most of the roads located there represent tracks (usually unpaved) that we do not take into account in the analysis. When comparing the two extremes from Table 1, we can see that the largest urban area has six times as many nodes and intersections than the Karlovy Vary Region, six times as many road segments and its road network is five times longer. In general, the average size of an urban area in the Czech Republic is $6000\,\text{km}^2$, its road network is $8300\,\text{km}$ long and has 26000 nodes, of which 19500 are crossroads.

The length of roads is an important attribute that displays spatial configuration of the network. Long roads usually have the highest category and represent the backbone of the network, while short roads allow to trim off the distance between locations. The length may also reflect the degree of urbanization as the edge length typically decreases in areas with a high built-up density. Figure 2 shows the chart of the road lengths distribution for four regions of the Czech Republic. We decided not to overload the chart with data for all regions, since the form of the length distribution is similar for the whole area.

Common to all studied networks is the trend towards shorter edges, which is considered more efficient and sustainable since networks with short segments offer greater variety of routes. The distribution of

Region	ϱ_{nodes} $(1/\text{km}^2)$	$\varrho_{\text{intersections}}$ $(1/\text{km}^2)$	ϱ_{edges} (km/km^2)	Deg_{avr}	β_{inx}	γ_{inx}	N_{deg1}	N_{deg3}	N_{deg4}	Organic OR
Central Bohemian	5.851	4.420	1.720	2.608	1.309	0.436	15641	42211	5968	0.905
Moravian-Silesian	6.462	4.719	1.797	2.546	1.278	0.426	9466	22686	2888	0.916
South Moravian	4.616	3.464	1.305	2.587	1.297	0.432	8281	22131	2676	0.917
Ústí nad Labem	4.833	3.692	1.477	2.618	1.314	0.438	6093	17440	2217	0.912
South Bohemian	2.742	2.088	1.085	2.606	1.307	0.436	6573	18755	2214	0.918
Olomouc	4.156	3.074	1.283	2.553	1.284	0.428	5705	14615	1562	0.928
Zlín	4.816	3.379	1.275	2.467	1.238	0.413	5693	12198	1165	0.938
Pilsen	3.148	2.342	1.167	2.588	1.298	0.433	6165	15558	2295	0.902
Hradec Králové	4.235	3.185	1.414	2.592	1.301	0.434	5001	13415	1709	0.914
Pardubice	4.256	3.187	1.391	2.582	1.297	0.432	4833	12799	1582	0.917
Vysočina	3.115	2.373	1.229	2.601	1.306	0.435	5039	14530	1557	0.925
Liberec	5.426	4.185	1.639	2.628	1.319	0.439	3926	11809	1394	0.917
Karlovy Vary	2.999	2.217	1.084	2.559	1.288	0.429	2589	6568	747	0.922

TABLE 2. Structural analysis of regional transportation networks.
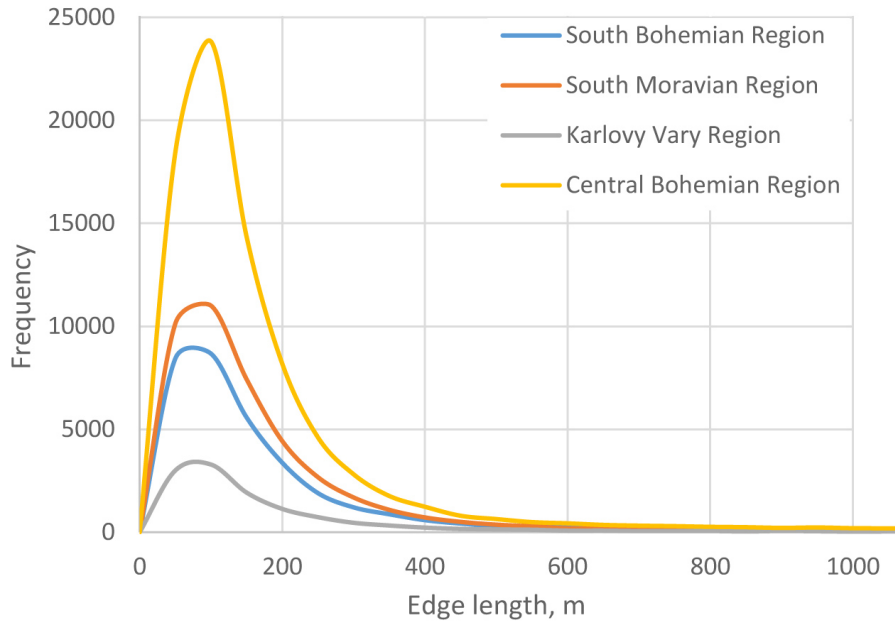

FIGURE 2. Edge length distribution for selected regions.

road lengths has a similar shape for all thirteen regions with a single peak in the range of 50 to 150 m. If we examine the lengths in more detail, however, we will notice greater differences between networks. The highest number of short segments in the network can be found in the South Moravian Region, where 93 % of roads lie in the length range of 1–500 m, reflecting the typical size of a dense urban environment. In the same region, 4.5 % of roads have a length between 500–1500 m and 2.5 % are longer than 1500 m. The longest roads, in contrast, can be found in the South Bohemian Region, where 86 % of edges are in the range of 1–500 m, 9 % have a length between 500-1500 m and 5 % are longer than 1500 m. The Karlovy Vary Region also belongs to those with elongated roads. Its network has 88 % of roads that belong in the range of 1-500 m, while 8 % are in the range of 500-1500 m

and 4 % are longer than 1500 m. In the Central Bohemian Region 91 % of edges lie in the length range of 1-500 m, 6 % have a length between 500-1500 m and 3 % are longer than 1500 m.

In Table 2, we present results of calculations on measures described in the previous section.

The Moravian-Silesian Region has the densest road network, leading in all three density indicators (nodes, intersections and edges per km² of the territory). At the other end of the range is the South Bohemian Region, which, however, has a specific landscape with a predominance of wooded areas (in the south-west) and agricultural land. Being the second largest region of the country, it has the lowest values for all density indicators. In terms of the overall efficiency, this means that its network has an enlarged cell structure, which increases the travel distance between locations.
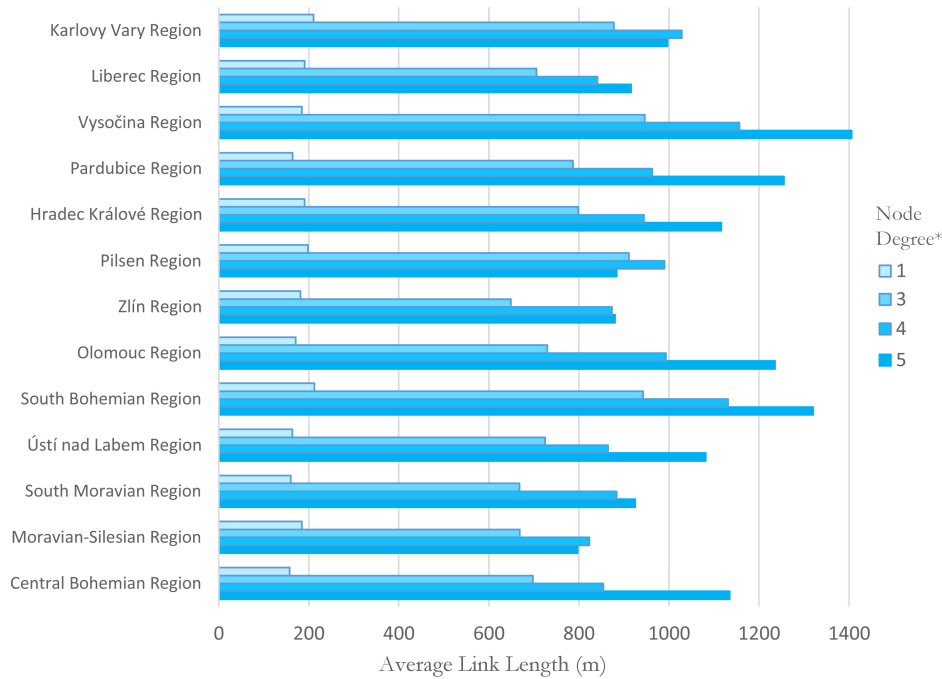
FIGURE 3. The distribution of the average length of edges that terminate at nodes with the corresponding degree.

This is confirmed by the average edge length in the network, which, for the South Bohemian Region, has the highest value of 303 m. The opposite extremum for the average edge length can be found in the Zlín Region with a value of 214 m and Moravian-Silesian Region with the average length of 218 m.

Node degree is a straightforward parameter to characterize vertices and connectivity in the network. The values of an average node degree for the Czech regions lie in a narrow span between 2.467 and 2.628, with the average value of 2.579 for the whole Czech Republic. It is interesting to note that the value of 2.579 for the entire country is lower than the average node degree of 2.86 for the interstate road network in the US presented in [12]. The US interstate network does not include a dense grid of local urban roads, however, its value of the average degree is still somehow bigger than that of the Czech network with local roads. Usually, in highly urbanized areas, the transportation network tends to be a parallel rectangular shape with the predominance of 3-way and 4-way intersections. This is primarily due to the human understanding of the form, the economic feasibility and ease of construction. The overall value of 2.579 indicates irregular web-like structure of the network with a large number of 3-way intersections, some dead-ends and a low count of rectangular grids. To confirm that last assumption, we can also investigate the pattern of the network development using the organic ratio (OR) and node degree values. Urban networks with a predominance of degree 3 and 1 can be considered as organic, i.e. they grew gradually following the natural development of urban areas without a strict typology. The opposite case would be structures with a typical

degree of 4 or higher, they are coherent and consist of planned patterns (grid-like). The organic ratio for all our networks has a value greater than 0.9 with an average of 0.92 for the entire network. This value shows that 90 % of the territory of the Czech Republic is self-organized. Most of the Czech cities have a strong historical core, so the transportation network evolved independently from settlement-to-settlement without a strict intervention from authority.

The range of node degree values (N_{deg1}, N_{deg3}, N_{deg4}) represent a very homogeneous structure. The difference between the maximum and minimum value is only 6 % on average. For example, in the Liberec Region, 23 % of nodes have a degree of 1, while in the Zlín Region, this value is 30 %. The number of nodes with a degree of 4 is very low for all regions. These vertexes represent intersections in which 4 roads are connected and their high percentage in the network indicates its gridiron pattern. In the Pilsen Region network, 10 % of nodes have a degree of 4, however, this value is still very low for the network to be considered as a regular grid. In the Zlín Region, only 6 % of nodes connect 4 roads, which is the smallest value among the studied regions. The average region of the Czech Republic has 25 % of dead-ends, 66 % of nodes with a degree of 3 and only 8 % of nodes with a degree of 4. The average value of Beta is 1.3, which is relatively low and corresponds with a high number of 3-way intersections. In the context of sustainability and safety rate, the cross intersections are more dangerous. T- and Y-intersections often have fewer accidents and injuries than 4-way intersections. The main reason for that is having far less points of possible collision when compared to 4-way
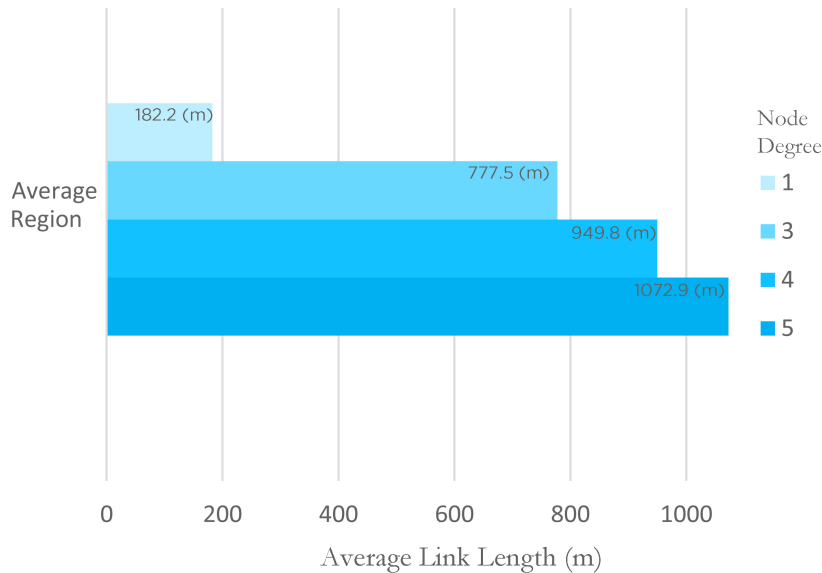
FIGURE 4. The distribution of the average length of edges that terminate at nodes with the corresponding degree in the average region.

junctions [13]. We also assume that the priority of movement is clearer at 3-way junctions, so there is a less chance of a driver making a mistake. They also accelerate the traffic flow in the network due to a lower waiting time. Similarly to Beta index, the values of Gamma are also associated with the degree of nodes. For example, a large number of dead-ends (i.e. junctions in which the network is interrupted) is also reflected in low values of Gamma. As we can see from Table 2, the lowest value of Gamma can be found in the Zlín Region, where there is a high proportion of dead-ends. However, the aforementioned Liberec Region has the highest Gamma value and the lowest number of dead-ends.

Another important observation can be made by analysing the relationship between the degree of nodes and the length of the attached road segments. The degree of a node, in some way, represents the variety of possible routes, while the edge length determines the amount of space that must be overcome in order to reach the desired location. In this regard, for each degree value present in our networks, we will consider the distribution of the average length of roads that terminate at them. Figure 3 represents a chart in which the horizontal axis shows the average edge length for the four classes of the node degree.

The form of the relationship looks similar for the studied regions. Roads that lead to dead-ends are characterized by a short length; after that, however, the edge length increases sharply for the degree of 3 and continues to grow with the increasing node degree (Figure 3). This trend is well-visible on the chart of the average values of the same indicators (Figure 4).

Such behaviour in the distribution of quantities can mainly be explained by the scale of our networks. Earlier, we identified that urban regional networks in the Czech Republic are self-organized, with a great predominance of 3-way intersections and, in general, have a web-like pattern with irregular cells. These characteristics are also combined with a polycentric structure, where cities and small settlements represent nodules in the web with a relatively high density of roads. The connection between these centres provided by the long edges that have an influence on the overall picture. The most frequent 3-way intersections mainly connect residential or secondary roads, while many of 4-way and 5-way intersections connect local networks with highways or appear as junctions on the 1st and 2nd class roads between cities. Usually, at least one elongated segment terminates at such a junction, which causes an increase in the average length of edges that are attached to the nodes with higher degree. It is not common, however, for dead-ends to appear on arterial roads. The links that lead to dead-ends are very short in length, since most of them belong to residential networks and are often truncated by private land or represent a driveway.

5. CONCLUSION

We presented the results of a spatial analysis of sufficiently large urbanized regions that cover the territory of the Czech Republic. Macro-scale analysis allows us to analyse the main characteristics of the network, to reveal similarities and differences of urban systems, to describe the form of relationship of their elements and to track some historical imprints imposed on the urban form at a large geographical unit. These findings may provide urbanists and transportation engineers with new information on the development and structural properties of Czech regional networks. The results show that the properties of studied urban networks vary from region to region, however, at the overall we have determined some common patterns. The development of urban areas in the Czech Republic occurred

through self-organization. This is confirmed by the values of the organic ratio and the significant predominance of nodes with the degree of 3 in all regions of the country. Such urban systems grow independently without a strict planning or imposing a certain pattern. Gradually spatial evolution and interaction led to the unification of settlements, thereby giving a form to the entire urban system. On the macro-scale, self-organization leads to another common trend for our regions, when the average length of roads increases with the degree of nodes. We explain this by the fact that most nodes with a high degree connect local networks to arterial roads (highways, 1st and 2nd class roads) through at least one long segment, which affects the average value. Self-organized transportation networks are considered to be safer due to the simplicity of road connections, where T- and Y-intersections predominate. This also leads to a lesser waiting time at such crossroads, which can speed up the traffic. However, in the planned networks (e.g. gridiron), the length of the road segments is shorter due to the high connectivity of the roads. This is reflected primarily in a greater variety of routes and may reduce the total fuel consumption. We would like to note that large urbanized systems were formed through the integration of small areas and it is in the full competence of designers to shape the local environment. However, any decisions taken now may, in the future, affect the structure and efficiency of the entire system.

References

[1] Vuchic, V. R. *Transportation for livable cities.* Center for Urban Policy Research, New Jersey, 1999.

[2] Jiang, B., and Claramunt, C. (2002). Integration of Space Syntax into GIS: New Perspectives for Urban Morphology. *Transactions in GIS*, 6(3), pp. 295-309. DOI:10.1111/1467-9671.00112

[3] Crucitti, P., Latora, V., and Porta, S. (2006). Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(1), 015113. DOI:10.1063/1.2150162

[4] Martínez Sánchez-Mateos, H. S., Sanz, I. M., Francés, J. M., and Trapero, E. S. (2014). Road accessibility and articulation of metropolitan spatial structures: the case of Madrid (Spain). *Journal of Transport Geography*, 37, 61-73. DOI:10.1016/j.jtrangeo.2014.04.003

[5] Sevtsuk, A., and Mekonnen, M. (2012). Urban network analysis. A new toolbox for ArcGIS. *Revue internationale de géomatique*, 22(2), 287-305. DOI:10.3166/rig.22.287-305

[6] Strano, E., Nicosia, V., Latora, V., Porta, S., and Barthélemy, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports*, 2(1). DOI:10.1038/srep00296

[7] Ma, D., Sandberg, M., and Jiang, B. (2015). Characterizing the Heterogeneity of the OpenStreetMap Data and Community. *ISPRS International Journal of Geo-Information*, 4(2), 535-550. DOI:10.3390/ijgi4020535

[8] Trimble Data Marketplace. [2018-05-22], `https://www.trimble.com/`

[9] Rodrigue, J., Comtois, C., and Slack, B. (2013). *The Geography of Transport Systems.* 3rd Edition, Routledge, New York.

[10] Scott, D., Novak, D., Aultman-Hall, L., and Guo, F. (2006). Network Robustness Index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14(3), pp.215-227.

[11] Courtat, T., Gloaguen, C., and Douady, S. (2011). Mathematics and morphogenesis of cities: A geometrical approach. *Physical Review E*, 83(3). DOI:10.1103/physreve.83.036106

[12] Gastner, M. T., and Newman, M. E. J. (2006). The spatial structure of networks. *The European Physical Journal B*, 49(2), 247–252. DOI:10.1140/epjb/e2006-00046-8

[13] Wolhuter, K. (2017). *Geometric design of roads handbook.* Boca Raton: CRC Press, Taylor & Francis.