TRANSPORT PROBLEMS 2015 Volume 10 Issue 2
PROBLEMY TRANSPORTU DOI: 10.21307/tp-2015-018

Keywords: metro network; graph theory; cluster analysis

Svetla STOILOVA*, Veselin STOEV
Technical University of Sofia
boulevard Kliment Ohridski 8, Sofia, Bulgaria
**Corresponding author*. E-mail: stoilova@tu-sofia.bg

AN APPLICATION OF THE GRAPH THEORY WHICH EXAMINES THE METRO NETWORKS

Summary. The graph theory gives a mathematical representation of transport networks and allows us to study their characteristics effectively. A research of the structure of metro system has been conducted in the study by using the graph theory. The study includes subway systems of 22 European capitals. New indicators have been defined in the research such as a degree of routing, a connectivity of the route, average length per link (which takes into account the number of routes), intensity of the route, density of the route. The new and the existing indicators have been used to analyze and classify the metro networks. The statistical method cluster analysis has been applied to classify the networks. Ten indicators have been used to carry out an analysis. The metro systems in European capitals have been classified in three clusters. The first cluster includes large metro systems, the second one includes small metro networks whereas the third cluster includes metro networks with only one line. The combination of both two methods has been used for the first time in this research. The methodology could be used to evaluate other existing metro networks as well as for preliminary analysis in the design of subway systems.

ПРИМЕНЕНИЕ ТЕОРИИ ГРАФОВ ДЛЯ ИССЛЕДОВАНИЯ МЕТРО СЕТЕЙ

Аннотация. Теория графов позволяет организовать математическое представление транспортных сетей и эффективно изучить их характеристики. В работе проводится исследование структуры системы метро с использованием теории графов. Исследование включает в себя 22 метро, которые находятся в европейских столицах. В статье определены новые показатели: степень маршрутизации; подключение маршрута; средняя длина на ссылку, которая учитывает количество маршрутов; интенсивность маршрута; плотность маршрута. Новые и существующие показатели используются для анализа и классификации городских сетей. Статистический анализ при помощи метода кластерного анализа применяется для классификации сетей. Десять показателей было использовано для анализа. Системы метро в европейских столицах подразделяются на три кластера. В первом кластере включены системы больших метро, во второй включены небольшие сети метро. Третий кластер включает сеть метро с одной линией. Сочетание двух методов используется в первый раз в этом исследовании. Данную методологию можно использовать для оценки других существующих городских сетей, а также для предварительного анализа при проектировании метро.

1. INTRODUCTION

Metro systems are the main type of public transport in many cities around the world. 76 European cities have metro networks which are fully constructed or planned for future expansion. In North and South America, there are 51 metro systems, in Asia – 65, in Africa - 3, Australia - 1. Worldwide Metropolitan is in three main varieties - classic subway (about 70%), light urban rail (light metro) and automatic metro. Classic metro has been called "Underground", "Subway", "U-Bahn" or‘’T-Bahn’’ in different countries around the world. The name "Metropolitan" (Metro) has been adopted by many countries.

The structure of different metro lines depends on the size of the city, the location of different regions, the density of the development and others. For cities with a population of 1.5 million citizens it is typical to have the following structure: linear (Warsaw, Helsinki), circular (Glasgow), diametrically (Sofia, Prague, Kiev), and X-shaped (Rome, Minsk). The type of metro network in larger cities with more than 1,5 -2 million citizens is diverse and could be defined as a diametrically-circular (Moscow), rectangular (Madrid), linear-rectangular (Oslo), mixed (Paris, Vienna), random (Copenhagen).

The graph theory has been used for many years by various researchers to describe the structure of the network of public transport, street networks, and others. In many journals, some of the metro systems in the world have been examined but not until now only European metro systems have been studied closely. This research includes only the classical metro networks located in the capitals due to the large number of metro networks in Europe

The object of the research is 22 metro systems of capitals cities in Europe. The aim of the study is:

- To apply the graph theory for studying metro networks;
- To examine the state and the structure of metro networks with indicators defined by the graph theory and categorize systems according to their network proprieties
- To classify the studied subway systems by using their network characteristics.

2. LITERATURE REVUE

The graph theory is inherently linked to transportation. A lot of researchers have used the graph theory to study the characteristics of transport networks. In [11] is explained the main graph theory concepts as well as various indicators have been introduced, such as traffic flow, network diameter, and other dimensionless ratios. The first introducing of three of the graph theory’s indicators directly linked to network design (circuits, degree of connectivity, and complexity) is made in [9]. In [12] is established a comprehensive series of new indicators, the line overlapping index, the circle availability, and network complexity. The main indices that represent the structural properties of a graph as such are beta index (a level of connectivity), alpha index (a measure of connectivity which evaluates the number of cycles in a graph), gamma index (connectivity), eta index (average length per link) and others, [12].

Some authors have used the graph theory to study metro networks. In [5] and [6] are used three indicators such as coverage, directness, and connectivity to assess the overall properties of networks. There are introduced new indicators such as tau (directness) and rho (connectivity). Authors have analyzed 19 subway networks located around the world, [5, 8]. They are compared by using the annual numbers of boarding per capita as a performance indicator. In [6] has been adapted various concepts of the graph theory to describe characteristics of the State, Form and Structure of 33 metro systems. The complexity of metro systems and the impact of network size have been analyzed and the implications on robustness have been discussed, [7]. It uses three indicators relevant to ridership: coverage, directness, and connectivity. This study used the graph theory as a mathematical method to transform networks into graphs, from which relevant properties (e.g., links, nodes) were collected. The authors analyzed 19 subway–metro networks and developed three indicators to assess the overall properties of transit networks, linking them to ridership.

The graph theory and the Complex Network Theory are adopted to examine the connectivity, robustness and reliability of the Shanghai subway network of China, [15]. The subway network systems of four cities, i.e., Seoul, Tokyo, Boston and Beijing, are studied by using global and local efficiencies and the graph theory, [3]. The Complex Network Theory and the graph theory are adopted to analyze and calculate the vulnerability of metro network, [4].

All these studies indicate that the graph theory may be successfully used for examining the metro networks. In the papers, it has not been studied the effect of the number of routes in metro systems on their structure and the satisfaction of passengers. There is no comprehensive study which compares and classifies subway networks located in one continent (region).

3. METHODOLOGICAL APPROACH

3.1. A representation a metro network in a graph

A graph is a symbolic representation of a network and of its connectivity, [2, 14]. It implies an abstraction of the reality so it can be simplified as a set of linked nodes. The conversion of a real network into a planar graph is based on the following principles: every terminal and intersection point becomes a node; each connected node is then linked to a straight segment.

The metro network is presented in a graph $G = \{V, E\}$. It is a set of vertexes (nodes) (V) connected by edges (links) (E). The vertex is a terminal or an intersection point of a graph. It is the abstraction of a location. The edges are links between two stations. A link is the abstraction of a transport infrastructure which supports movements between nodes. Two types of vertices (nodes) have been defined: transfer and end-vertices. Transfer-vertices are transfer stations, where it is possible to switch lines without exiting the system regardless of the nature of the transfer which could be a simple cross platform interchange or a longer walk. End-vertices are the line terminals, where it is not possible to switch to another metro line. If a terminal actually hosts two lines, it is considered as a transfer-vertex. The ability to transfer is the determining factor to define the transfer-vertices.

An example of a representation of a metro network in a graph is shown in fig. 1. In Figure 1a it has been shown a real metro network, in Figure 1b this network is adapted into a graph structure whereas Figure 1c shows a presentation of a simple graph presented in the research.

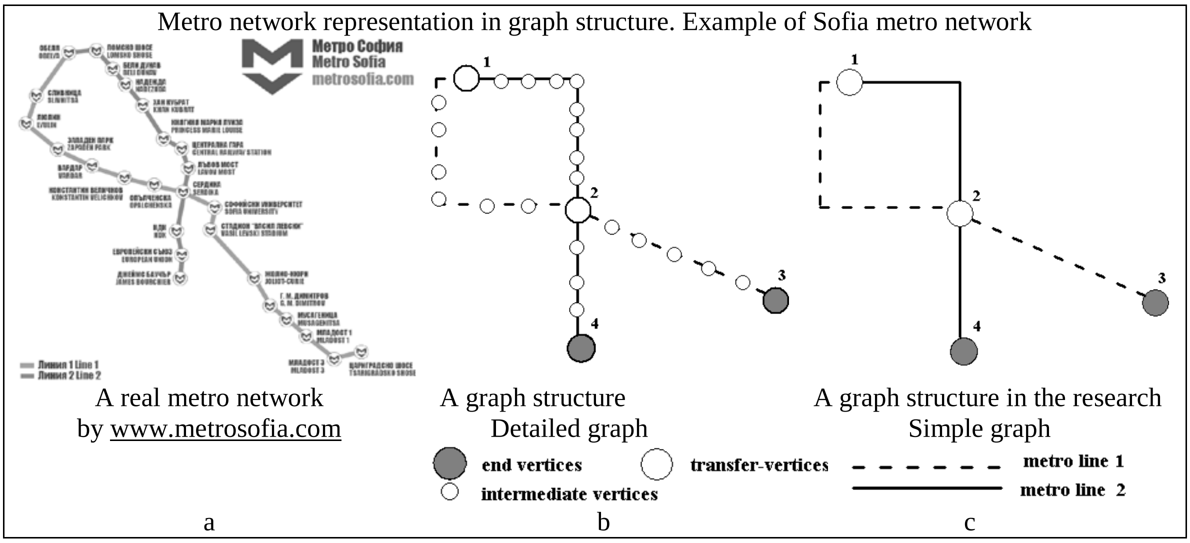

Fig. 1. Metro network representation in graph structure. Example of Sofia metro network
Рис. 1. Сеть метро представленнная в структуре графа. Пример сети метрополитена Софии

Edges are non-directional links. The edges are two types – single e_S and multiple e_M .

$$e = \frac{e_S + e_M}{2}, \qquad (1)$$

where: e is the total number of edges; e_S is the total number of single-use edges; e_M is the total number of multiple-use edges. In order to define the total number of edges e_S and e_M ,each edge must be reported twice- the first time from nod “i” to nod “j”, secondly from nod ” j “ to nod “i”. This is the reason why in formula (1) the sum of the total number is divided by two.

The single edge shows that vertices are connected. The multiple edges show that there is more than one specific line between two vertices. If two consecutive vertices are linked by two or more edges, this is considered as a single edge (e_S =1) and a multiple one (e_M =1).The total number of edges equals 2 in this case. If two consecutive vertices are linked by one edge, this is considered as a single edge (e_S =1) и the total number of edges is e=1.

For the purpose of the study, metro networks of European capitals are represented by stations where two or more lines have been crossing each other (transfer nodes) as well as start and end stations of each line (end nodes).

3.2. Indicators for study metro networks

In order to study the indicators *it must be taken into account the following* При изследване на показателите трябва да се имат впредвид следните уточнения:

- A metro line is an infrastructural track which connects a starting point with a finishing point and it has a definite number of stations.
- A metro route is an organization of trains’ movement between a starting station and a finishing station and it consists of one or more than one metro lines. In most European metro networks metro routes coincide with metro lines. In the case of Sofia metropolitan a metro route consist of two metro lines.

3.2.1. Complexity (beta index)

The main network indicators which have been developed are complexity β and a degree of connectivity γ , [5, 6, 12]. A state refers to the current development phase of a metro network. The complexity is expressed by the relationship between the number of links (e) divided by the number of nodes (v). The complexity β is determined by the formula, [6]:

$$\beta = e/v, \qquad (2)$$

where: v is the sum of the transfer-vertices v_T and the end-vertices v_E ; e is the number of edges.

$$v = v_T + v_E \quad . \qquad (3)$$

A connected network with one cycle has a value of 1.

3.2.2. A degree of connectivity (gamma index)

The degree of connectivity γ calculates the ratio between the actual numbers of edges to the potential number of edges; that is if the network is 100% connected. The value of γ is between 0 and 1, where a value of 1 indicates a completely connected network. This indicator is a measure of the evaluation of a network in time. For planar graphs with $v \geq 3$, the degree of connectivity is calculated by the formula, [12]:

$$\gamma = \frac{e}{3(v-2)} \quad . \tag{4}$$

The structure of metro networks is a planar graph in most cases because there is not any connectivity between all peaks within the network. In this case, the following formula has been use (4).

In non planar graphs the degree of connectivity γ is calculated as follows, [12]:

$$\gamma = 2e / v(v-1) \quad . \tag{5}$$

Metro networks which are consisted of nods connected completely have A non planar graph. For example Warsaw metro network is linear and it consist of two nods and one edge. In this case. γ =1. It could be easily checked by using the formula (5) that a non planar graph with 3,4, and 5 nods γ =1 and in the case of number increase up to 6 then $\gamma < 1$.

The two indicators a complexity β and a degree of connectivity γ show the structural differences between two networks of an equal size.

3.2.3. An average length per edge (Eta index)

Adding new nodes will cause the eta index to decrease as the average length per edge declines.

$$\eta = \frac{L}{e} \quad , \tag{6}$$

where: L - is the total metro network route length, km; e – the total number of edges in metro network.

This ratio indicates the intensity (density) of the stations in the network.

3.2.4. Connectivity (rho)

The network structure is presented by connectivity ρ. This indicator measures the intensity and the importance of connections (i.e. transfers) in a metro system. This indicator is the relationship between the net numbers of transfer possibilities divided by the number of the transfer stations. It is calculated by the formula, [5, 6]:

$$\rho = (v_T^c - e_M) / v_T \quad , \tag{7}$$

where: v_T^c is the total number of transfers in the transfer vertices; e_M is the total number of multiple edges; v_T is the total number of transfer vertices in a metro network.

The total number of transfers in the transfer vertices is the sum of the number of metro lines going through a transfer station minus one. A transfer station sharing two transit lines offers one transfer possibility, another sharing three lines offers two possibilities, and so on.

This indicator calculates the total number of net transfer possibilities. The ratio indicates the average connectivity of each transfer node in the network. The advantage of this indicator is that it provides information about the stations where more transfers from one line to another could be done, i.e. it crosses more than two metro lines.

3.2.5. Directness (tau) τ

This indicator is proportional to the maximum number of transfers and it is related to the number of lines n_L, [5, 6]

$$\tau = \frac{n_L}{\delta} \quad , \tag{8}$$

where: n_L is the number of metro lines; δ is the maximum number of transfers in a diameter (i.e. the longest route).

The above coefficients, which are defined by the Theory of Graphs and are introduced into [5, 6], characterize the structure of the network and its specific characteristics.

3.2.6. Indicators of routing

New indicators such as a degree of routing, connectivity of the route, average length per edge (which takes into account the number of routes), intensity of the route, density of the route have been introduced in this research. These new factors have great impact on the categorisation and the evaluation of a metro network based on their routes. The coefficients describing the routes in a metro network show the total transport satisfaction.

The degree of routing a subway network gives a greater degree of satisfaction of transport to passengers.

$$g = \frac{i}{3(v-2)} \quad , \tag{9}$$

where: g - is the degree of routing; i - is the total number of route arcs in the metro network; v - is the total number of vertices in the metro network.

The total number of metro routes is a sum of the route edges i_S and the multiple routes edges i_M.

The coefficient g considers the degree of connectivity of the routes in the transferring vertices of a metro network. The increase in the number of arcs of routes will cause an increase in the number of the vertices.

The coefficient of connectivity of the routes b is the ratio of the total number of arcs route to the total number of vertices in a metro network. It takes into account the connectivity of the routes in the structure of the network.

$$b = \frac{i}{v} \quad . \tag{10}$$

For the coefficients g and b and is valid the following:

$$g \geq \gamma \ ; b \geq \beta \quad . \tag{11}$$

A network representation of the Oslo's metro network is a clear example of the difference between the coefficients introduced in the research and the existing factors.

In figure 2, it is shown a presentation of the Oslo's metro system as a graph. Table 1 presents its matrix of edges. Table 2 presents the matrix of edges describing the routes in a metro network.

The total number of edges (routes) is equal to the sum of the single edges (routes) and the multiples edges (routes). In the matrices, one arc is passed twice for each of the two directions. When determining the total number of arcs, the sum is divided by two. This applies to e_S, e_M, i_S, i_M, e and i. If the edges between two nodes pass through different infrastructures, multiple arc is not counted. They are accounted as single arcs.

The edge 6-7 in fig.2 could be looked up-close in order to make the difference between e and I. In this case three metro lines go between 6 and 7. This shows two edges- a single one (e_s=1) and a multiple one (e_m = 1). This means that the total number of edges is e=2 and it is written down in cell

6-7 in table 1. The indicator 1 reports the real number of going-through metro lines between 6 and 7. In that case “ i ” is written as i=3.

Table 1

A matrix of edges for the Oslo’s metro network

From/To	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	e_S	e_M	e
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
3	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	3	1	4
4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1
5	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
6	0	0	0	0	1	0	2	0	0	0	2	0	0	0	0	0	0	0	0	3	2	5
7	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	2	2	4
8	0	0	2	1	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	4	3	7
9	0	0	0	0	0	0	0	2	0	2	0	2	0	0	0	0	0	0	0	3	3	6
10	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	2	2	4
11	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	2	1	3
12	0	0	0	0	0	0	0	0	2	0	1	0	1	0	0	0	0	0	0	3	1	4
13	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1
14	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	1	0	0	0	3	2	5
15	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	2	0	3	2	5
16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	2	2	4
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	1	2
Total	1	1	4	1	1	5	4	7	6	4	3	4	2	5	5	1	1	4	2	38	22	60

Only one metro line goes through edge 6-5 in fig. 2. A value of 1 is written in table 1 in cell 6-5. The same value is written in table 2 as well. Two metro lines go through edge 6-11 this shows that one single and one multiple edges are reported. So, in table 1 in cell 6-11 a value of 2 is written down. In table 3 in cell 6-11 the real number of going-through metro lines is written down which is 2. In this case, when two metro lines go through two neighbor peaks of the graph then the number of edges (singles and multiples) and the number of route edges is the same.

As a conclusion it could be pointed out that in the first case (tab. 1) when one line goes through two neighbor peaks of a graph, only one edge is defined. But if two or more than two lines go through, a single or a multiple edge is defined. In the second case scenario (tab. 2), when one line goes through two neighbor peaks of a graph, it is reported as their real number as a multiple route edges. The coefficient a for average length per edge, which takes into account the number of routes, determines the intensity of the routes in a metro network.

$$a = \frac{L}{i} \quad , \tag{12}$$

where: L is the total length of the metro network, km.

As the value of the coefficient is smaller, the more intense is a metro network of routes. A low coefficient indicates saturated with routes metro network.

For the coefficients η and a is valid the following:

$$a \le \eta \quad . \tag{13}$$

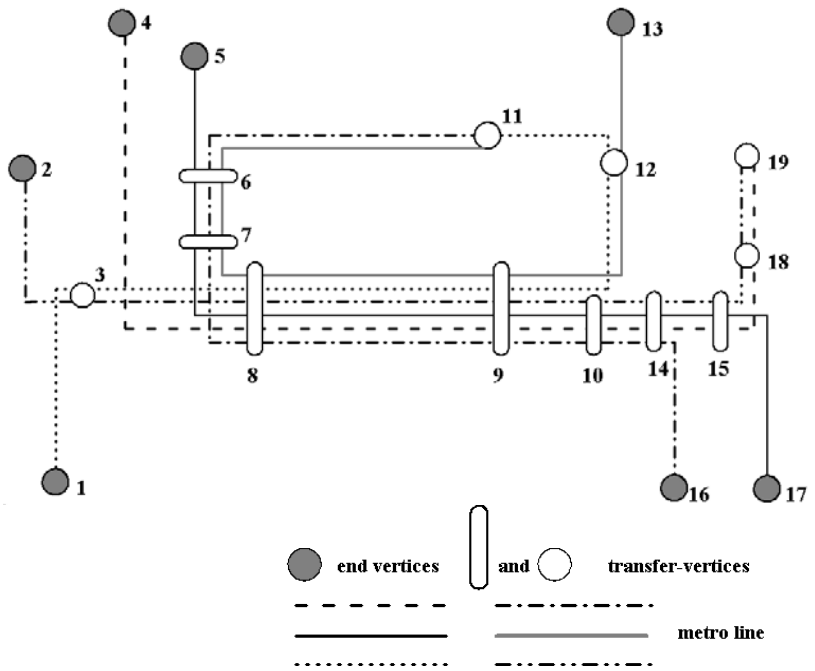

Fig. 2. A representation of a metro network in a graph structure. An example of the Oslo's metro network
Рис. 2. Метро сеть представление в структуре графа. Пример сети метрополитена Осло

Table 2

A matrix of routes for the Oslo's metro network

From/To	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	i_S	i_M	i	z
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
3	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	2	4	1
4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
5	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
6	0	0	0	0	1	0	3	0	0	0	2	0	0	0	0	0	0	0	0	1	5	6	2
7	0	0	0	0	0	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	6	6	2
8	0	0	2	1	0	0	3	0	6	0	0	0	0	0	0	0	0	0	0	1	11	12	3
9	0	0	0	0	0	0	0	6	0	4	0	2	0	0	0	0	0	0	0	0	12	12	3
10	0	0	0	0	0	0	0	0	4	0	0	0	0	4	0	0	0	0	0	0	8	8	2
11	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2	3	1
12	0	0	0	0	0	0	0	0	2	0	1	0	1	0	0	0	0	0	0	2	2	4	1
13	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0
14	0	0	0	0	0	0	0	0	0	4	0	0	0	0	3	1	0	0	0	1	7	8	2
15	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	2	0	1	5	6	2
16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	0	4	4	2
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2	1
Total	1	1	4	1	1	6	6	12	12	8	3	4	1	8	6	1	1	4	2	16	66	82	22

The following relationship is valid for metro networks for which the graph arcs do not pass more than two lines:

$$g = \gamma\ ; b = \beta\ ;\ a = \eta \quad . \tag{14}$$

The coefficient of density of the routes r shows what the density of multiple route arcs in a metro network is.

$$r = \frac{i_M}{v} \quad , \tag{15}$$

where: i_M is the total number of multiple route arcs in a metro network.

The coefficient of intensity of the routes u shows the difference between arcs where routes have bigger than two.

$$u = \frac{z}{i_M} \quad , \tag{16}$$

where: z is the number of sections with multiples arcs. A section is a trace between two neighbor peaks of a graph

As an example for defining the values of "i_m" and "z" a peak 6 in fig. 2 could be looked up-close.

Nod 6 connects with nod 5 by one route edge, with nod 11 by two multiple edges and with nod 7 by three multiple route edges. This could be found in tab.2. The number of participants with multiple edges is 2(6-7 and 6-11), therefore z=2. The total number of multiple edges is i_S=1. The total number of route edges is $i = i_s + i_m = 6$.

For metro networks where in the graph structure for all sections pass not more a line is valid: $u = 0$. For metro networks which have arcs with more than two routes between two neighbouring nodes the value of u is $0 < u < 1$.

4. APPLICATION. A STUDY OF THE INDICATORS FOR 22 EUROPEAN METRO NETWORKS

A study of 22 capitals' metro networks of European countries has been conducted by using the above indicators of the graph theory as well as the new ones that takes into account the routes. Only subway lines have been studied for these metro networks due to the lack of sufficient information necessary for examining the whole subway routes.

Table 3 shows the value of all coefficients of the examined European metro systems.

Complexity β depends on the number of subway lines, transfer and end nodes. By increasing the complexity of metro network and crossings of metro lines, its value increases. For example: London subway's complexity is β = 1,77; Paris' metro is β =1,59; the Moscow's is β =1,56. This index has a low value in a metropolitan linear structure such as Warsaw's - β = 0,5.

The indicator of a degree of connectivity γ shows to what extent metro lines have contact to each other. The metro network in Warsaw has the highest degree of connectivity (γ =1). This value clearly represents the linear structure of the metro network with only one line.

Average length per link η has a value greater than 1. Networks with a small number of arcs and small length of lines have a higher value of the coefficient. These networks are in Warsaw (η =22), Minsk (η =8,85), Kiev (η =7,36) and Sofia (η =7,78).

Connectivity ρ depends on the metro network. This ratio shows the level of the average connectivity of each transfer node in a network. Connectivity is 0 in linear networks such as the metro network in Warsaw. In this case, there are no transfer nodes. With an increasing number of metro lines and complexity of the network, the value of the connection is increased (for Paris ρ =1,47). The results from the study have shown values of this coefficient from 0 to 1,47. The maximum value of this coefficient is for the Paris' metro network because it has larger number of transfer units compared to the other capitals' networks in the study.

Table 3

Values of indicators for European metro systems

City	L, [km]	n_L	v	e	i	γ	g	β	b	ρ	r	η	a	τ	u
Amsterdam	32,7	4	9	13	14	0,62	0,66	1,44	1,55	1,25	1	2,52	2,34	4	0,89
Athens	53,2	3	10	11	11	0,42	0.42	1	1	1	0	5.32	5.32	3	0
Berlin	147,4	10	31	41,5	41.5	0,48	0,48	1,34	1,34	1,18	0	3,55	3,55	3	0
Brussels	32,2	3	9	16	18	0,76	0,86	1,78	2	1	1,56	2,01	1,79	4	0,86
Bucharest	69,3	4	11	16	18	0,59	0,67	1,45	1,64	0,8	0,36	4,33	3,85	2	1
Budapest	33	3	10	10	10	0,42	0,42	1	1	1,33	0	3,3	3,3	4	0
Copenhagen	21	2	5	4	4	0,67	0,67	1	1	0,5	0,5	5,25	5,25	0	1
Helsinki	21	1	4	4,5	4,5	0,75	0,75	1,13	1,13	0,5	0,5	4,67	4,67	0	1
Kiev	66,2	3	9	9	9	0,43	0,43	1	1	1	0	7,36	7,36	3	0
Lisbon	44,3	4	13	14	14	0,42	0,42	1,07	1,07	1	0	3,16	3,16	2	0
London	402	11	56	99	104,5	0,61	0,62	1,77	1,87	1,13	1,24	4,06	3,85	6	0,92
Madrid	286,3	13	46	79	79	0,6	0,6	1,72	1,72	1,26	0	3,62	3,62	4,33	0
Minsk	35,4	2	5	4	4	0,44	0,44	0,8	0,8	1	0	8,85	8,85	2	0
Moscow	309,4	12	40	62,5	62,5	0,55	0,55	1,56	1,56	1,12	0,1	4,95	4,95	6	1
Oslo	62	5	19	30	49	0,59	0,96	1,58	2,58	1,42	1,74	2,07	1,27	6	0,67
Paris	218	16	66	105	105	0,55	0,55	1,59	1,59	1,47	0,12	2,16	2,16	4,67	1
Prague	59,1	3	9	9	9	0,43	0,43	1	1	1	0	6,57	6,57	1.5	0
Rome	41,6	2	5	7	7	0,46	0,46	1	1	1	0	5,94	5,94	2	0
Sofia	31,2	2	4	4	4	0,67	0,67	1	1	0,5	0,5	7,8	7,8	2	0
Stockholm	105,7	3	18	18	18	0,35	0,35	1	1	1.14	0,11	6,22	6,22	3	1
Vienna	74,6	5	19	24,5	24,5	0,48	0,48	1,29	1,29	1,1	0	3,04	3,04	2,5	0
Warsaw	22,6	1	2	1	1	1	1	0,5	0,5	0	0	22,6	22,6	0	0

In the study directness τ has a value between 0 and 6. Developed metro networks with a large number of lines and transfer units have a high value of the coefficient. For example: Moscow, London and Oslo ($\tau = 6$), Paris ($\tau = 4{,}67$), Madrid ($\tau = 4{,}33$). Small metro networks with two intersecting lines have $\tau = 2$. Those are in Minsk and Rome. For Warsaw's metro network directness τ is 0 because it consists of only one line.

The number of arcs between two nodes in a metro system for the introduced new coefficients (g, b, r, a, u) is the number of routes. For example, if three metro lines pass two nodes with coefficients $\beta, \gamma, \eta, \rho$, the value for the number of multiple links is "2", while for the new coefficients is taken "3". This specificity gives an idea of the intensity of routes between two nodes (stations) in a metro network. These new coefficients are always different from those from the graph theory where the number of multiple arcs between two nodes is more than two.

The coefficient "g" considers the degree of connectivity of the routes between nodes in a metro network. The value of this ratio for studying metro networks is from 0 to 1.

The values of the coefficient‘’ b "in the research are in the range between 0,5 and 2,58. The largest value has the Oslo's metro network. It has a large number of metro lines (routes) in an arc.

When the values of ‘’g’’ and ‘’b’’ are higher, then the network has more than two routes. Such are the networks of Brussels, Budapest, London, Oslo, Prague. The most intense of routes is the metro network of Oslo.

The coefficient "a" accounts for intensity of routes in a metro network. It has a value from 1,27 to 22,6.

The coefficient “r” shows the density of multiple route arcs in a metro network. The more lines pass through an area, the greater is the coefficient value. The highest value has the metro network of Oslo (1,74), Brussels (1,56), London (1,24). For networks in which any section does not have more than one route, r =0. Networks with a coefficient $r \geq 1$ are saturated of routes and have a high density of multiple arcs.

The coefficient “u” shows the intensity of the routes. Large metro networks with a large number of sections, where many lines pass have an average value of the coefficient such as the one of Oslo’s metro network ($u = 0{,}67$)

5. CLUSTER ANALYSIS

The Cluster analysis is a suitable method for a classification of the examined metro networks into groups by using different factors. It is multi-measurable statistical analysis for a classification of units into groups, preliminarily unknown, based on numerous characteristics in relation to these units [13]. The number of examined factors is greater than 2. The statistics theory suggests different methods of clusterisation.

The dispersion analysis could be used for an approximate evaluation of the clusterisation’s results as well as for determining the roles of each variables used for clusters’ establishment. The determination of the statistical importance of different factors is done by using the F criterion (Fisher’s criterion).

$$F \geq F_T \qquad , \tag{17}$$

where: F is the empirical value of the criterion resulted from the dispersion analysis, F_T is the theoretical value when the level of risk $\alpha = 0{,}05$ and the number of degrees of freedom, $k_1 = n- 1$; $k_2 = m - n$; m is the number of observations (22 metro networks), n is the number of examined factors (number 10). The theoretical criteria F_T is determined by using the tables for F distribution. F is defined as a relationship between two independent values of dispersion.

$$F = \frac{S_1^2}{S_2^2}\ ;\ \ S_1^2 = \frac{\sum_{i=1}^{n}(\overline{x_i} - \overline{x})^2 . m_i}{n-1}\ ;\ S_2^2 = \frac{\sum_{i=1}^{n}\ \sum_{j=1}^{m}(x_{ij} - \overline{x_i})^2}{m-n}, \tag{18}$$

where: S_1^2 is the between-group value of total dispersion; S_2^2 is the inner-group value of total dispersion; x_{ij} is the value of j-factor for i-metro network.

On one hand, the Fisher’ criterion’s evaluation determines which factors are significant for the study, on the other it does not dismiss those other factors which are used for clusterisation but does not satisfy the condition (17). The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences between cases in different clusters.

A method for hierarchical clustering has been used in the study. The main advantage of this method is that the determination of a unit into a specific cluster is definitive. Hierarchical clustering is performed by the agglomerative method of average linkage between groups. For the distance-type measures it is chosen the Squared Euclidean distance, [13].

Table 4 shows the value of F criterion for the examined indicators which are defined by the graph theory. The theoretical value of F criterion is defined by using standard tables and is F_T =2,8 при степени на свобода k_1=*22-10=12 и* k_2 *= 10-1=9*.

Table 4

Values of F criterion for the examined indicators

Indicator	γ	g	β	b	ρ	r	η	a	τ	u
F criterion	16,36	9,69	6,05	4,27	17,76	0,38	94,75	87,93	3,06	0,81

From the results, shown in Tab. 4, it could be defined by the importance of each one of the factors when a classification has been conducted. The indices intensity of the route and density of the route have value of F criterion smaller than theoretical.

SPSS (Statistical Package for Social Science) software has been used for carrying out the study with a cluster analysis. A dendrogram of the formed clusters and their respective elements are shown in figure 3.

The results indicate that metro systems of the considered European capitals can be classified into three groups:

- A cluster of complex subway networks. This cluster contents 12 subways. In this group are: Amsterdam, Berlin, Brussels, Bucharest, Budapest, Lisbon, London, Madrid, Moscow, Oslo, Paris and Vienna.
- A cluster of simple subway networks. This cluster contents 9 subways. In this group are: Athens, Copenhagen, Helsinki, Kiev, Minsk, Prague, Rome, Sofia and Stockholm.
- A cluster of networks that have only one metro line. This cluster contents 1 metro network. This is the subway of Warsaw.

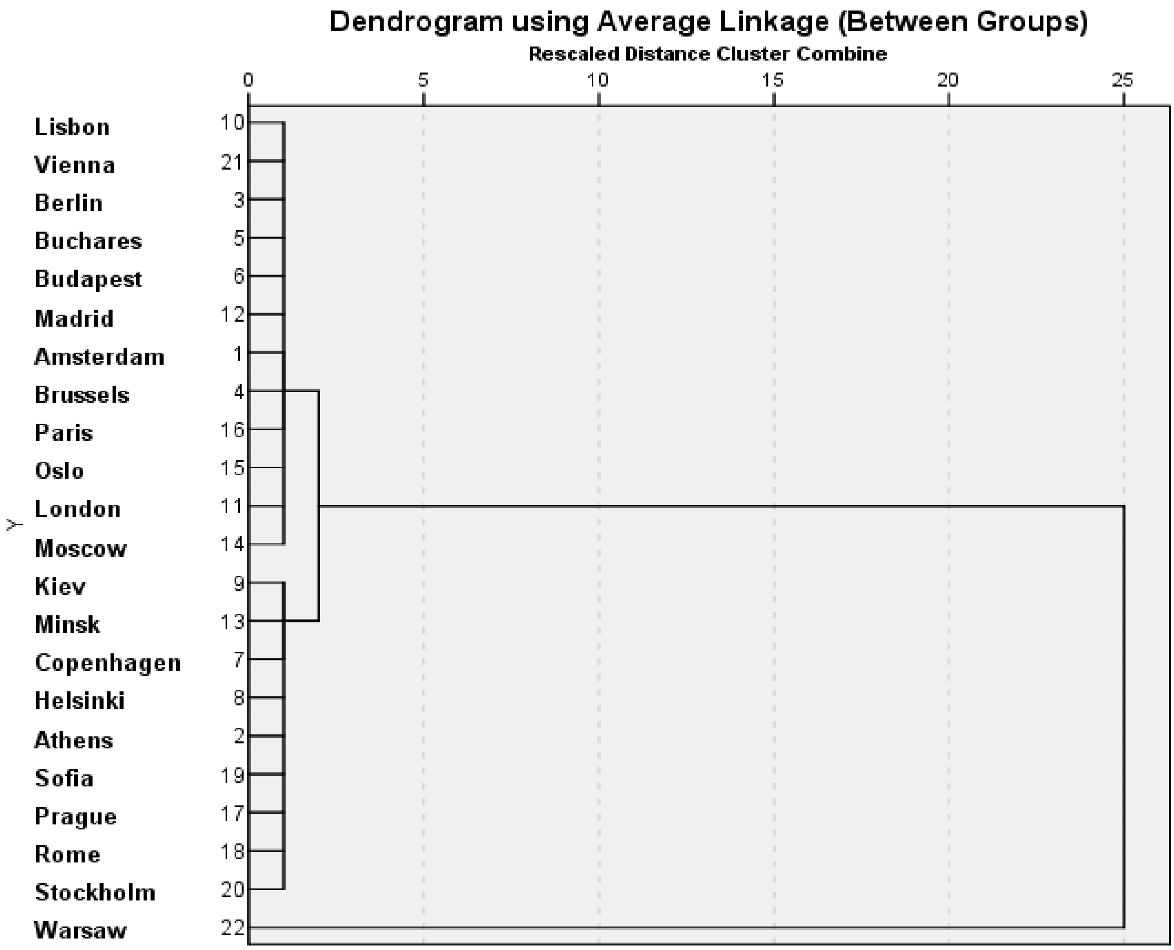

Fig. 3. Dendrogram of cluster tree of subway in European capitals
Рис. 3. Дендрограмма кластерного дерева для метрополитенов в Европейских столицах

6. CONCLUSIONS

The study has shown the following results:

- The factors for classifications of metro networks have been defined-these factors allow us to evaluate the stage of development of the examined systems.
- The graph theory has been applied to characterise the metro networks and to define the indicators of the state and the structure of a metro network.
- A cluster analysis has been used for classification of the metro networks. The classification has been conducted by using 10 different factors.
- European capitals' metro networks are divided into three groups.
- The application of the Cluster analysis allows us to evaluate the stage of development of metro systems.
- The factors which define the state and the structure of a network are important for grouping metro networks.
- In comparison with the method of grouping metro networks into linear, diametrical, mixed infrastructures, the classification shown in the research, based on the Theory of Graphs and Cluster Analysis, allows a complex evaluation of form, structure, connectivity and routes of metro systems.

Acknowledgement

This research is conducted in relation to the execution of a contract № 142ПД0019-04 "A simulation modelling of technological processes in main metro stations of the Sofia's subway". The research has been funded by Technical University of Sofia, Bulgaria.

References

1. Barberillo, J. & Saldana, J. Navigation in large subway networks: An informational approach. *Physica A: Statistical Mechanics and its Applications*. 2011. Vol. 390. No. 2. P. 374-386.
2. Bondy, J. & Murty, U. *Graph theory with applications*. New York. Amsterdam. Oxford: North-Holland. 1976.
3. Chang, K. & Kim, K. & et al. Subway Networks in Cities. *Journal of the Korean Physical Society*. 2006. Vol. 48. P. S143-S145.
4. Deng, Y. & Qiming Li, Q. & Ying Lu, Y. & et al. Topology vulnerability analysis and measure of urban metro network: The case of Nanjing. *Journal of networks*. 2013. Vol. 8. No. 6. P. 1350-1356.
5. Derrible, S. & Kennedy, A. Network analysis of world subway systems using updated graph theory. *Transportation Research Record: Journal of the Transportation Research Board*. 2009. No. 2112. P. 17-25.
6. Derrible, S. & Kennedy, A. Characterizing metro networks: state, form, and structure. *Transportation*. 2010. No. 37. P. 275-297.
7. Derrible, S. & Kennedy, A. The complexity and robustness of metro networks. *Physica A*. 2010. Vol. 389. P. 3678-3691.
8. Derrible, S. & Kennedy, A. Evaluating, Comparing, and Improving Metro Networks. *Transportation Research Record: Journal of the Transportation Research Board*. 2010. No. 2146. P. 43-51.
9. Garrison, W. & Marble, D. Factor-analytic study of the connectivity of a transportation network. *Papers of the Regional Science Association*. 1964. Vol. 12. No. 1. P. 231-238.
10. Musso, A. & Vuchic, V. Characteristics of metro networks and methodology for their evaluation. *Transportation Research Record: Journal of the Transportation Research Board*. 1988. No. 1162. P. 22-33.

11. Kansky, K. *The structure of transportation networks: Relationships between network geography and regional characteristics*. Chicago: University of Chicago. Research Paper No. 84. 1963.
12. Rodrige, J. *The geography of transport systems*. New York: Routelage. 2013.
13. Tan, P. & Steinbach, M. & Kumar, U. *Introduction to Data mining*. Minnesota: Iniversity of Minnesota. 2006.
14. Tutte, W. *Graph theory*. Cambridge: Cambridge University Press. 2001.
15. Zhang, J. & Xu, X. Networked analysis of the Shanghai subway network in China. *Physica A: Statistical Mechanics and its Applications*. 2011. Vol. 390. Nos. 23-24. P. 4562-4570.

Received 13.02.2014; accepted in revised form 26.05.2015