URBAN ROAD TRANSPORT NETWORK ANALYSIS: MACHINE LEARNING AND SOCIAL NETWORK APPROACHES

Emre Kuşkapan [1], M. Yasin Çodur [2], Ahmet Tortum [3], Giovanni Tesoriere [4], Tiziana Campisi [4,*]

[1]Department of Civil Engineering, Engineering and Architecture Faculty, Erzurum Technical University, Erzurum, Turkey
[2]College of Engineering and Technology, American University of the Middle East, Kuwait
[3]Department of Civil Engineering, Engineering Faculty, Ataturk University, Erzurum, Turkey
[4]Faculty of Engineering and Architecture, Kore University of Enna, Enna, Italy

*E-mail of corresponding author: tiziana.campisi@unikore.it

Resume

Traffic congestion is one of the most significant problems in urban transportation. It has been increasing, especially in regions close to intersections. Several methods have been developed to reduce the traffic congestion. One of the analysis methods is social network analysis (SNA). This method, which has increased use in transportation, can quickly identify the most central intersections in transportation networks. Improvements to central intersections, identified in a road network structure, speed up the traffic flow across the entire network structure. In this study, the Istanbul highway transportation network has been examined and values for a series of network centrality measures have been calculated using the SNA. The accuracy and error scales of the centrality values were compared using a machine learning algorithm. The Bonacich power centrality has been the best performance. Based on the study results the most central intersections in Istanbul have been determined.

Available online: https://doi.org/10.26552/com.C.2022.4.A232-A245

Article info

Received 7 May 2022
Accepted 20 September 2022
Online 17 October 2022

Keywords:

central intersections
machine learning
transportation planning
urban transportation

ISSN 1335-4205 (print version)
ISSN 2585-7878 (online version)

1 Introduction

The traffic congestion is one of the important problems in daily life, in cities. Traffic jams can prevent people from getting where they want to go on time. In order to prevent this situation, it is necessary to carry out studies that would minimize the congestion. A wide range of detailed information is required to estimate congestion and to evaluate solutions for transport planning. The traditional method is to calculate the traffic manually and to report the traffic situations. This method's application has significantly decreased in recent years. Traditional models can require a lot of time and effort; consequently, they are not economical [1-2]. In addition, traffic congestion negatively impacts drivers through increased travel time and delayed destination arrival. To reduce these negative effects, many transport studies were conducted. For example, improvements at intersections where the traffic congestion is intense can provide more permanent solutions for the road network [3]. At the same time, to understand the network performance and design the best transportation planning, it is imperative for transport planners to consider many factors (for instance traffic volume, location of intersections and the connection of roads). Travelers' travel times are variables affected by traffic incidents and bad weather conditions. Therefore, it is difficult to analyze and model different traffic conditions [4].

Various numerical traffic data must be provided for these analyses and models. Thanks to development of technology and intelligent transportation systems, vehicle counting hoses, motion detector cameras, computers and counting units used, it is very easy to obtain traffic volume data [5]. With development of these counting methods, traffic counts and analyses have become quite practical. In the past several years, the analysis of network traffic has become a subject of continuous research in various sub-fields of computer networks. Many researchers have implemented effective network traffic algorithms for analysis and the prediction of network traffic [6-7]. Similarly, technical information used in these analyses is useful to formulate the relationship between the traffic composition and the ring road safety and to formulate recommendations for designing the vehicle safety improvement strategies

[8]. As an example of this situation, analyses to investigate the effects of hazardous substances on road transportation and the effects of risk factors leading to accidents can provide theoretical support for adoption of effective measures to reduce the risks associated with road transportation [9]. Likewise, traffic problems can be seriously prevented through the correct usage and analysis of traffic data. Thanks to this, delays can be minimized by reducing the traffic congestion on the roads during the traffic peak hours. Additionally, the optimum order can be achieved by examining how drivers behave and react to bottlenecks with varying traffic capacities from time to time, user balance and optimal traffic order of the system. For this reason, it is very important to provide the traffic data and conduct analyses using this data [10].

2 Literature review

Traffic analysis methods have gained prominence in recent years. Social network analysis (SNA) technique has recently been utilized to analyze traffic data for traffic analysis. The SNA is frequently used in many fields, such as sociology, anthropology, social psychology, communication, economics and mathematics. The SNA, which is an interdisciplinary field of study, examines the structure of communities, tries to describe the network structure and visualizes the relationships that cannot be easily observed between the communities by modeling the existing connections [11].

The SNA-based approach focuses on links and relationships within the community. It can be observed that studies on SNA have proliferated in recent years and researchers' interest in the subject remains strong. Some of the studies in which the SNA has been applied in the field of transportation are included in this study. In one of the existing studies, the dynamic development of taxi sharing behaviors and the social network structure focused on taxi sharing were investigated. The advantages of taxi sharing based on a social network, an increasing match ratio and a satisfaction level comparable to the satisfaction level associated with trip-based methods were demonstrated [12]. In another study, Pritchard et al. [13], underlined the importance of considering the social aspects of accessibility and mobility when planning for urban transport systems using the SNA. Cheng et al. [14], investigated the use of centralities to identify central nodes in a transport network to improve the design of the transport network and address network failures. They implemented these centralities in the Singapore subway network and compared them to traditional topology-based centralization measures. In another study, a road network was constructed using the SNA. Three types of centralities were calculated: degree centrality, closeness centrality and betweenness centrality. These node centralities explain how important a node is in a network [15]. Watts and Witham, [16] used the SNA method to analyze the communication networks of organizations that promote sustainable transport policy in the UK. At the end of the study, they found that different types of centrality, such as degree and betweenness centrality, have stronger relations to influence than closeness centrality. Networks offer many different roles for organizations that need to be effective. For this reason, organizations should carefully consider which role they want to play and how they can properly position themselves in the context of their network. El-adaway [17], considered the most important step in the implementation of the SNA in the field of transportation. He used the SNA to analyze the intersections in Louisiana in the United States, modeling the intersections as nodes and modeling the streets as edges. Using the volume maps for the connection forces between the node points, he obtained four different centrality measures for all the intersections in each case study. By interpreting those centrality values, he identified the most central intersections in the road network structure. The results indicated that the SNA intersection modeling aligns with current congestion studies and Regional Transportation Planning decisions in Louisiana. At the end of the study, the SNA is cited as an important method for analyzing and improving transport networks. El-adaway et al. [18], also mentions that the use of the SNA, which is practical and economic, will create a new perspective that will allow for making more effective infrastructure network decisions when adopted alongside traditional transport network analysis techniques.

When similar studies on this line of research are examined, few studies on SNA in transportation engineering have been conducted. This paper presents one of the few studies to have focused on this issue. In addition, unlike in other studies, the effect of a new road, added to a road network, on the whole network structure was examined. Using the traffic volume data, two different case studies were investigated from before and after the construction of the 3rd Bosphorus Bridge and the Northern Ring Motorway. The five different centrality values, obtained for both cases, were compared by using support vector machine, which is a machine learning algorithm and the most central intersections in the Istanbul road network were detected. Using the SNA as the first analysis step to identify central intersections will assist decision-makers in better focusing their more detailed analyses using traditional methods in such areas only compared to the overall network. This will significantly decrease the invested resources in transportation planning and will create a more integrated and comprehensive perspective for evaluating the transportation networks.

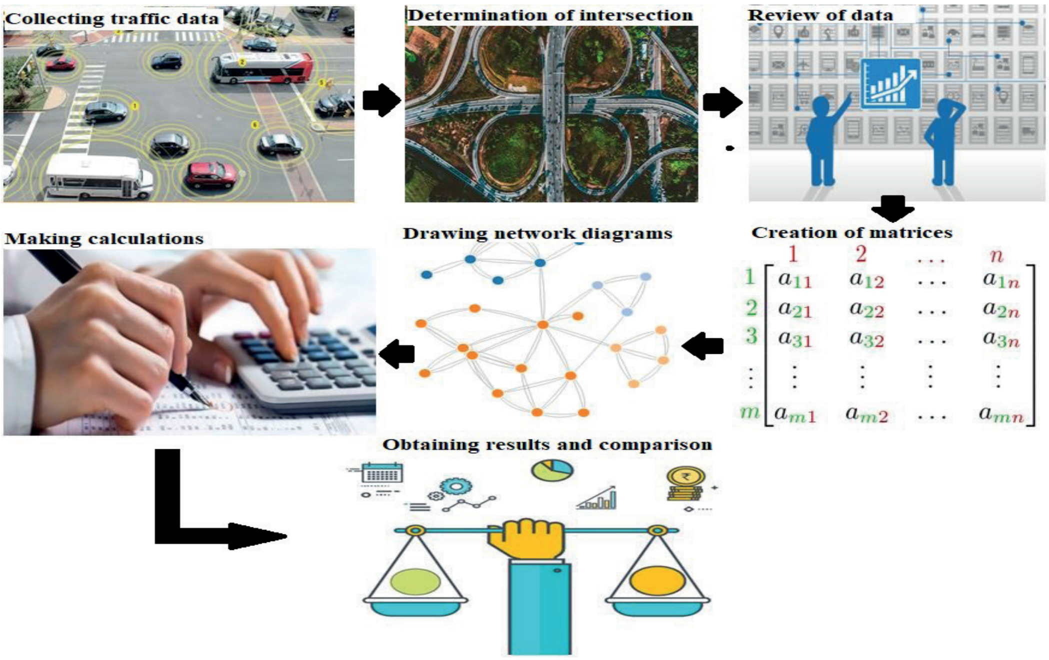

Figure 1 *Flow chart of SNA application in transportation planning*

3 Method and network model

The SNA examines the relationships between individuals or objects in terms of political, formal-informal, familial, geographical, or any other factors. SNA, which also means the digitization and scientific making of interpersonal relations, is used to explore the relations between various organizations, or the networks formed by these organizations and important events [19]. Therefore, the network analysis allows for examining how the elements, organizations, or systems within a network affect the way the network functions. The advantage of SNA is that, unlike many other methods, it focuses on interactions (rather than individual behavior); it is also less time-consuming and more practical than many other methods. In addition, it is an interdisciplinary analysis method that can be applied in many areas, such as social networks, political networks, electricity networks and transportation networks. The SNA can be especially important to a country's policies when it is applied to transportation networks because it is an indisputable fact that innovative strategies for transportation policies contribute to sustainable economic growth [20]. Implementation of the SNA in transportation networks is provided, as shown in Figure 1.

3.1 Centrality

Centrality focuses on the following question: "Who is the most important and centralized actor (node) in the network?" Measures of local centrality are those that examine the very close environment of the individual; however, to understand an individual's position in the social network, it is necessary to examine its place within the general structure of the network. In the SNA, the node point with the highest centrality value is called the most central (important) node [21]. Similarly, the most central intersection points in the road network structure are called the most central intersections. When determining the most central intersections, not only the traffic volume values of the intersections are considered. At the same time, the location of the intersection points in the network structure and the neighbor status of the intersections are also examined. Therefore, each concept of centrality examines the network in a different way.

3.1.1 Closeness centrality

This centrality measures the degree of closeness (or distance) of a node in the network to (or from) other nodes, either directly or indirectly. Closeness is a sum of individual's shortest distances to other individuals in the network and the ability to access information reflects how quickly a node can connect to other nodes in the network. Additionally, the closeness measure can calculate the weakness or strength of the connections between the nodes [22]. It is calculated using equations (1) and (2). In this centrality, the concept is not necessary for a node to have a high degree and closeness. Closeness centrality defined as $C_C(i)$ in Equation (1). Normalized closeness centrality defined as $C_c'(i)$ in Equation (2). Where d (i, j) represents the distance between vertices i and j; N represents the number of nodes in the network.

$$C_C(i)=\left[\sum_{j=1}^{N} d(i,j)\right]^{-1}, \tag{1}$$

$$C_c'(i) = (C_C(i))/(N-1). \tag{2}$$

3.1.2 Degree centrality

Degree centrality is equal to the number of ties (links) of a node to other nodes. It not only reflects the connectivity of each node to other nodes, but it also relies on the network size. In other words, the larger the network, the higher the maximum degree of centralization [23]. Thus, a certain degree of centrality indicates that a node either has many connections in a small network or has only a small number of ties in a large network. When the degree of centrality of any node is calculated, the sum of the tie values to that node is calculated as shown in Equation (3). X is the factor, i and j are the network nodes.

$$C_d(i) = \sum_j X_{ij}. \tag{3}$$

3.1.3 Betweenness centrality

This centrality addresses how the relationships between the binary elements that are not directly connected are controlled or directed by other nodes. Betweenness centrality is an important indicator of excessive information exchange within a network or control of the flow of resources. This measurement is based on the shortest number of ties passing through a node [24]. It is calculated using Equations (4) and (5).

$$C_b(i) = \sum_{j<k} g_{jk}(i)/g_{jk}, \tag{4}$$

usually normalized by:

$$C_b'(i) = C_b(i)/[(n-1)(n-2)/2], \tag{5}$$

where $g_{jk}g_{jk}$ is the shortest number of ties connecting jk and g_{jk} is the number that node i is on.

3.1.4 Eigenvector centrality

Eigenvector centrality shows the importance of a node in the network. The eigenvector center is not equal to all the connections in the network; it assumes that efficient nodes act on less effective nodes to which they are associated. It depends on the number of connections and the quality of the ties. If a node has a small number of high-quality connections, it can cross a node with a large number of medium-quality connections. For network G = (V, E), with |V| as the number of nodes, let the adjacency matrix A = ($a_{v,t}$). M(v) is a set of neighbors of v, $a_{v,t}$ = 1 if node v is tied to node t and $a_{v,t}$ = 0 otherwise, λ is a constant [25]. The eigenvector centrality value of v is obtained by equation (6):

$$x_v = \frac{1}{\lambda}\sum_{t\in M(V)} x_t = \frac{1}{\lambda}\sum_{t\in G} a_{v,t}x_t. \tag{6}$$

3.1.5 Bonacich power centrality

Bonacich power is a measure that determines the node centrality based on the degree centrality of adjacent nodes, where a node's degree centrality is its summed connections to others, weighted by their centralities. This centrality depends not only on how many connections a node has but also on how many connections its neighbors have (and on how many connections its neighbors' neighbors have and so on). The highest value found here represents the most powerful node on the network, as shown in equation (7) [26].

$$C(\alpha,\beta) = \alpha(I-\beta R)^{-1}R1, \tag{7}$$

where R is the adjacency matrix (can be valued), β reflects the extent to which you weight the centrality of nodes tied to the studied node, I is the identity matrix, α is a scaling vector, 1 is a vector of all ones.

3.2 Ucinet and netdraw

Ucinet is a comprehensive package for analysis of the social networks. It can read and write in several formatted text and Excel files. In addition, it can create a network model (i.e. using Netdraw), constitute centrality measures, conduct a matrix analysis and conduct a statistical analysis based on permutation. The Netdraw Software is also included in this package to integrate with Ucinet and allow for drawing diagrams of social networks. No additional program is needed to create network diagrams from data in the Ucinet Software as well as in the Netdraw Software. Since this program is free, it is frequently preferred in the SNA [27].

3.3 Support vector machine (SVM)

In machine learning, the SVM is a supervised learning algorithm. It can be used for classification, as well as for regression. It is very efficient for many applications in science and engineering, especially for classification problems (pattern recognition). The main objective is to identify a hyper-plane that can classify the data points distinctively. The SVM classifier can be made non-linear by mapping the input variables to a higher dimension space and then applying the algorithm to the processed parameters [28]. This mapping is done using the SVM kernels. Two kernels are evaluated, radial basis function (RBF) and polynomial. Expressions for radial basis function (RBF) and polynomial in Scikit-learn are shown in equations (8) and (9), respectively [29].

$$K(x,z) = \exp(-y\| x - z \|^2), \tag{8}$$

$$K(x,z) = (y(x^T z) + c)^d. \tag{9}$$

In equations (8) and (9), K is the kernel function, x and z are vectors in the input space, c is the coefficient, d is the order of the polynomial and γ is the kernel parameter. The kernel parameter γ gives the influence of a single training data point [30].

It is necessary to compare the accuracy values of the order of the central intersections determined by the concepts of centrality. Performance analysis can be performed for each concept of centrality using the Kappa statistics, mean absolute error (MAE) and root mean squared error (RMSE) values. MAE and RMSE represent error scales and Kappa statistic expresses the accuracy value [31]. These calculations can be made easily with Weka (Waikato Environment for Knowledge Analysis).

The Kappa statistics is frequently used to test interrater reliability. The importance of rater reliability lies in the fact that it represents the extent to which the data collected in the study are correct representations of the variables measured. One may compare two or more tests or examinations to measure their agreement beyond that caused by chance. The Kappa statistic is formulated as :

$$K = \frac{p_o - p_e}{1 - p_e}, \tag{10}$$

where p_o and p_e are expectation and observation, respectively. The meaning of this calculation has come into question and ranges for the measure vary. However, an example would include the following: $K < 0.20$ = poor agreement; K = 0.21 to 0.40 is fair; K = 0.41 to 0.60 is moderate; K = 0.61 to 0.80 is substantial; and $K > 0.81$ is good [32].

The MAE is a measure of errors between paired observations expressing the same phenomenon. To put it shortly, it is the average of all the absolute errors and a model evaluation metric used with regression models. MAE error of a model with respect to a test set is the mean of the absolute values of the individual prediction errors on over all the instances in the test set. Each prediction error is the difference between the true value and the predicted value for the instance. It also measures closeness of the predictions to the eventual outcomes. Examples of Y versus X include comparisons of predicted versus observed, subsequent time versus initial time and one technique of measurement versus an alternative one. It can be expressed by equation (11), where $x_{f,i}$ and $x_{o,i}$ are the i-th expectation and observation, respectively:

$$MAE = \frac{1}{N}\sum_{1=1}^{N} |x_{f,i} - x_{o,i}|. \tag{11}$$

The RMSE is the square root of mean squared error; it measures the differences between values predicted by a hypothetical model and the observed values. In other words, it measures the quality of the fit between the actual data and the predicted model. The RMSE is one of the most frequently used measures of the goodness of fit of generalized regression models. It can be expressed as: $x_{f,i} x_{o,i}$

$$RMSE = \sqrt{\frac{1}{N}\sum_{1=1}^{N}(x_{f,i} - x_{o,i})^2}. \tag{12}$$

3.4 Case studies

Traffic counts represent the volume of traffic in both directions in a 24-hour period. Traffic counting systems are categorized into two main groups: road systems and roadside systems. Roadside systems, such as pneumatic hose systems, magnetic loop systems and systems with piezoelectric sensors are used by the General Directorate of Highways [33]. The most commonly used roadside systems are manual counting systems. In recent years, advances in traffic management applications within the scope of intelligent transportation systems have fostered roadside counting systems using technologies such as video visual systems, passive and active infrared, ultrasonic, radar, laser and microwave. Traffic volume values are obtained using these counting systems. For the traffic volume, the most common method on wide roads is the counting method, using hoses laid on the roads. The calculated traffic volume is obtained using these counting methods, as well as the calculation of average estimated value. These data are derived from the annual average daily traffic values. In this study, the compound of the state roads and ring roads in Istanbul traffic volume maps were used. Traffic volume values were obtained from the General Directorate of Highways and Transport Management Center of Istanbul Metropolitan Municipality [34-35]. The left side of Istanbul is located on the European continent and the right side is located on the Asian continent.

There are three Bosphorus Bridges connecting the two continents. The 1st and 2nd Bosphorus Bridges (i.e. the 1st is 15 Temmuz Sehitler Bridge and the 2nd is Fatih Sultan Mehmet Bridge) are old structures, while the 3rd Bosphorus Bridge (Yavuz Sultan Selim Bridge) was constructed recently. Accordingly, two different situations were investigated to determine the effect of the newly constructed 3rd Bosphorus Bridge on the traffic of Istanbul. The traffic volumes from before and after the construction of the 3rd Bosphorus Bridge and the Northern Ring Motorway are illustrated in Figures 2 and 3 as the first and second case studies. In the first case, 13 intersections have formed, belonging to the transportation network consisting of ring roads and state roads. In the second case, 14 intersections (added I node) have formed because new roads have been added to the network. The 3rd Bosphorus Bridge added to

***Figure 2** The first case study; traffic volume map from before the construction of the 3rd Bosphorus Bridge and the Northern Ring Motorway*

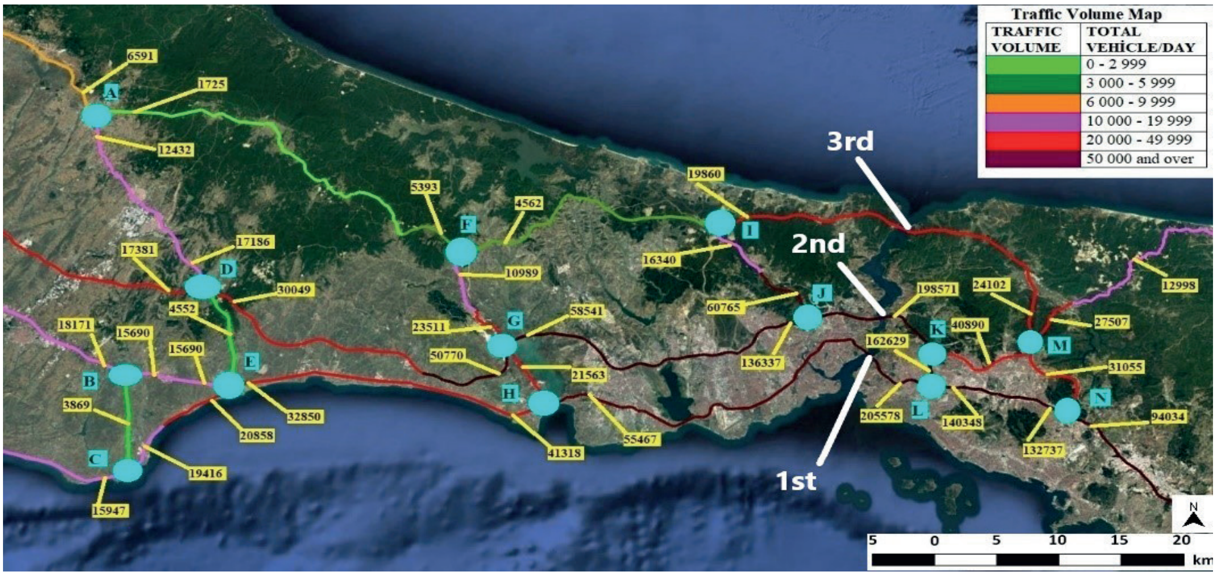

***Figure 3** The second case study; traffic volume map from after the construction of the 3rd Bosphorus Bridge and the Northern Ring Motorway (Connection of I and M intersections)*

the highway network structure has reduced the traffic load of the other two Bosphorus Bridges. Traffic flow has accelerated in the 1st and 2nd Bosphorus Bridges, especially due to the obligation to cross the 3rd Bosphorus Bridge for heavy land vehicles. With this situation, the 3rd Bosphorus Bridge has become an important transit point as it facilitates access to settlements in the regions near A, F, I and M intersections. Thus, the highway network structure gained momentum in the north direction.

The reason for investigating their situation before and after is to determine whether the new road added to the network structure has changed the central intersection. If the central intersection has changed, it has to move a little closer to the route newly added to the network structure, since the centrality of the network structure is expected to shift in that direction. At the same time, the new central intersection is expected not to go far from the existing central intersection. In addition to this situation, it is expected that the central intersection will be close to the middle points in the network structure and the traffic volume value will be high. The application method giving results suitable to these conditions increases the accuracy of the method.

It is very important to determine the central intersection in a highway network structure. Thanks to the improvement of the central intersection determined for the network structure, the entire network structure is optimally affected. For this reason, the central intersections can be prioritized in order to obtain maximum efficiency from investments in the road network. In order to detect these central intersections, the SNA method was used in the study. Current transportation analysis tools are expensive and time consuming and require rigorous data

***Table 1** Adjacency matrices of the first case study*

	A	B	C	D	E	F	G	H	J	K	L	M	N
A	0	0	0	1	0	1	0	0	0	0	0	0	0
B	0	0	1	0	1	0	0	0	0	0	0	0	0
C	0	1	0	0	1	0	0	0	0	0	0	0	0
D	1	0	0	0	1	0	1	0	0	0	0	0	0
E	0	1	1	1	0	0	0	1	0	0	0	0	0
F	1	0	0	0	0	0	1	0	1	0	0	0	0
G	0	0	0	1	0	1	0	1	1	0	0	0	0
H	0	0	0	0	1	0	1	0	0	0	1	0	0
J	0	0	0	0	0	1	1	0	0	1	0	0	0
K	0	0	0	0	0	0	0	0	1	0	1	1	0
L	0	0	0	0	0	0	0	1	0	1	0	0	1
M	0	0	0	0	0	0	0	0	0	1	0	0	1
N	0	0	0	0	0	0	0	0	0	0	1	1	0

***Table 2** Adjacency matrices of the second case study*

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	0	0	0	1	0	1	0	0	0	0	0	0	0	0
B	0	0	1	0	1	0	0	0	0	0	0	0	0	0
C	0	1	0	0	1	0	0	0	0	0	0	0	0	0
D	1	0	0	0	1	0	1	0	0	0	0	0	0	0
E	0	1	1	1	0	0	0	1	0	0	0	0	0	0
F	1	0	0	0	0	0	1	0	1	0	0	0	0	0
G	0	0	0	1	0	1	0	1	0	1	0	0	0	0
H	0	0	0	0	1	0	1	0	0	0	0	1	0	0
I	0	0	0	0	0	1	0	0	0	1	0	0	1	0
J	0	0	0	0	0	0	1	0	1	0	1	0	0	0
K	0	0	0	0	0	0	0	0	0	1	0	1	1	0
L	0	0	0	0	0	0	0	1	0	0	1	0	0	1
M	0	0	0	0	0	0	0	0	1	0	1	0	0	1
N	0	0	0	0	0	0	0	0	0	0	0	1	1	0

for reliable results. Accordingly, a quick and inexpensive methodology to preliminarily analyze traffic networks is beneficial to better direct more detailed transportation analyses. Due to its ability to grasp the full complexity and connectivity of networks in a timely and cost effective manner, SNA can fulfill this requirement. The SNA approach has been able to easily and quickly determine the most central intersections in the investigated transportation networks. Previous research reveals that the road network structures, examined using the SNA, give correct results and provide high efficiency in improving traffic flow. Accordingly, SNA is believed to be an effective and innovative tool in transportation analysis [36].

The first important step in SNA is to create an adjacency matrix. In both case studies, after creating a map where the nodes and the traffic volume values were processed, an adjacency matrix was created for the nodes. During the creation of the adjacency matrix, if there was a transition state between the two nodes, a value of 1 was entered in the matrix and if there was no transition between two nodes, a value of 0 was entered. At this stage, a point needs to be written in the field 0 corresponding to each node. Tables 1 and 2 show the adjacency matrices for both case studies.

The traffic volume values for each node were entered in the generated adjacency matrix. Once the generated adjacency matrix had been loaded into the Ucinet Software, a graph of the road network was obtained from the Netdraw Software, a tab of Ucinet [37]. This process was performed for both case studies and the two different road networks are presented in Figures 4 and 5. In the first case, 13 nodes were formed in the road network; in the second case, with the addition of the I node, the road network consisted of 14 nodes. The I node was located between the F, J and M nodes. In addition, the first case consisted of 18 ties, while 20 ties were formed in the second case.

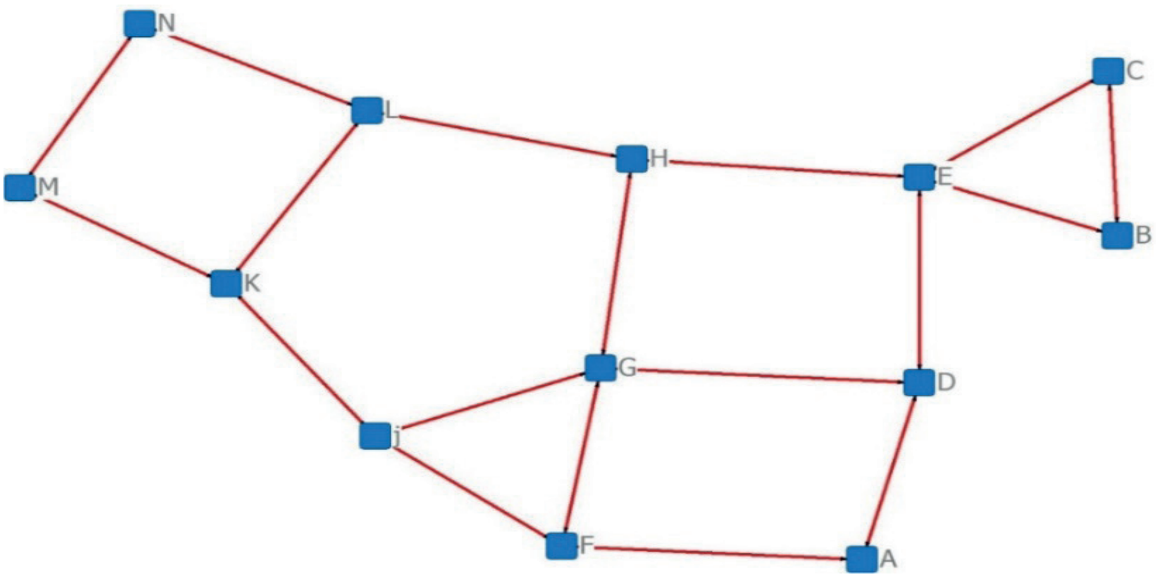

***Figure 4** The first case study network structure*

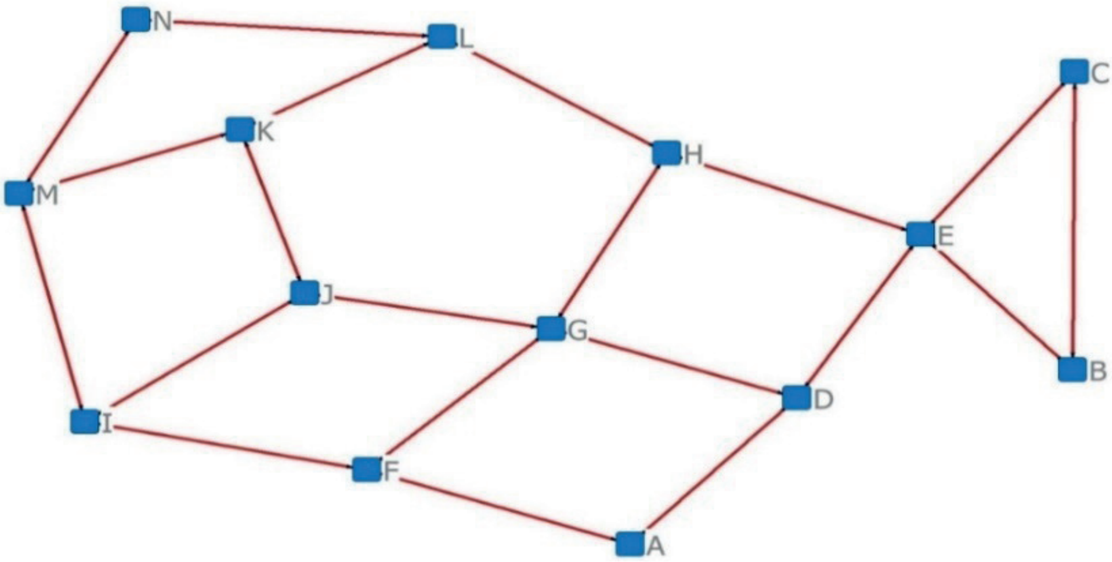

***Figure 5** The second case study network structure*

4 Results and discussion

Once the road network has been established, it was available to examine the situations of centralization for SNA. After the centrality values has been compared, the first 10 nodes with the highest values in the network have been determined. One of the most basic concepts of centrality is the closeness centrality. In this centrality, the tie values in the network are not important; rather, the structure of the network is important. The geodesic distances between nodes need to be determined correctly when calculating this centrality. It is easy to calculate the closeness centrality values when the closeness matrix with the shortest distances has been determined. Tables 3 and 4 present the closeness matrices, which were formed according to the location of the nodes in the network.

Closeness centrality values calculated for both cases are given in Table 5. In the first case, the value of node H was slightly higher than that of node G, but both values were the same in the second case. As mentioned previously, the closeness centrality is higher in the regions at the midpoints of the network. Another main centrality measure is the degree centrality. In this centrality, the sum of the tie values of each node in the network is calculated. Unlike in the case of the closeness centrality in the case of this centrality, the location of the node in the network structure is not important. All that matters is that the tie values connect the nodes. When the degree of centrality values was examined, it was observed that the value of the node L was higher in both cases, since in both cases the traffic volume values on the road network reached their maximum levels at node L. At the same time, the traffic volume

Table 3 Closeness centrality matrix of the first case study

	A	B	C	D	E	F	G	H	J	K	L	M	N
A	0	3	3	1	2	1	2	3	2	3	4	4	5
B	3	0	1	2	1	4	3	2	4	5	3	5	4
C	3	1	0	2	1	4	3	2	4	4	3	5	4
D	1	2	2	0	1	2	1	2	2	3	3	4	4
E	2	1	1	1	0	3	2	1	3	3	2	4	3
F	1	4	4	2	3	0	1	2	1	2	3	3	4
G	2	3	3	1	2	1	0	1	1	2	2	3	3
H	3	2	2	2	1	2	1	0	2	2	1	3	2
J	2	4	4	2	3	1	1	2	0	1	2	2	3
K	3	5	4	3	3	2	2	2	1	0	1	1	2
L	4	3	3	3	2	3	2	1	2	1	0	2	1
M	4	5	5	4	4	3	3	3	2	1	2	0	1
N	5	4	4	4	3	4	3	2	3	2	1	1	0

Table 4 Closeness centrality matrix of the second case study

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	0	3	3	1	2	1	2	3	2	3	4	4	3	4
B	3	0	1	2	1	4	3	2	5	4	5	3	5	4
C	3	1	0	2	1	4	3	2	5	4	4	3	5	4
D	1	2	2	0	1	2	1	2	3	2	3	3	4	4
E	2	1	1	1	0	3	2	1	4	3	3	2	4	3
F	1	4	4	2	3	0	1	2	1	2	3	3	2	3
G	2	3	3	1	2	1	0	1	2	1	2	2	3	3
H	3	2	2	2	1	2	1	0	3	2	2	1	3	2
I	2	5	5	3	4	1	2	3	0	1	2	3	1	2
J	3	4	4	2	3	2	1	2	1	0	1	2	2	3
K	4	5	4	3	3	3	2	2	2	1	0	1	1	2
L	4	3	3	3	2	3	2	1	3	2	1	0	2	1
M	3	5	5	4	4	2	3	3	1	2	1	2	0	1
N	4	4	4	4	3	3	3	2	2	3	2	1	1	0

Table 5 Closeness and degree centrality values of case studies

Degree centrality				Closeness Centrality			
First Case		Second Case		First Case		Second Case	
L	516549	L	508555	H	0.52	G	0.50
J	421119	K	402090	G	0.5	H	0.50
K	401575	J	395673	E	0.46	L	0.43
N	166094	N	163792	L	0.44	J	0.43
G	142994	G	154385	J	0.44	D	0.43
H	106491	H	118348	D	0.44	E	0.433
E	65959	M	96047	K	0.41	F	0.42
M	45230	E	73950	F	0.40	K	0.39
D	45077	D	51787	A	0.36	I	0.38
C	20812	I	40762	C	0.33	A	0.37

***Table 5** Closeness and degree centrality values of case studies*

Degree centrality				Closeness Centrality			
First Case		Second Case		First Case		Second Case	
L	516549	L	508555	H	0.52	G	0.50
J	421119	K	402090	G	0.5	H	0.50
K	401575	J	395673	E	0.46	L	0.43
N	166094	N	163792	L	0.44	J	0.43
G	142994	G	154385	J	0.44	D	0.43
H	106491	H	118348	D	0.44	E	0.433
E	65959	M	96047	K	0.41	F	0.42
M	45230	E	73950	F	0.40	K	0.39
D	45077	D	51787	A	0.36	I	0.38
C	20812	I	40762	C	0.33	A	0.37

***Table 6** Betweenness and eigenvector centrality values of case studies*

Betweenness centrality				Eigenvector centrality			
First Case		Second Case		First Case		Second Case	
E	33.76	E	30.98	L	0.55	L	0.57
H	33.61	H	29.74	K	0.53	K	0.52
L	24.95	G	25.83	J	0.43	J	0.39
G	21.82	L	19.02	H	0.32	H	0.38
K	17.15	D	15.66	N	0.25	N	0.25
D	15.25	F	12.67	G	0.23	G	0.19
J	14.87	I	10.58	F	0.09	M	0.09
F	7.48	J	10.19	M	0.06	I	0.08
M	3.79	K	7.59	D	0.04	E	0.05
N	3.79	M	7.59	E	0.04	D	0.03

values reached their second highest level at node J in the first case study, but reached their second highest level at node K in the second case study. It was observed that the tie values attached to the I node added in the network structure were not high. Therefore, the I node in the second case study did not much affect the other nodes in terms of the degree centrality.

Like closeness centrality, betweenness centrality is a shortest-path-based centrality concept. The difference is that a node is located based on the extent to which it takes part in the shortest paths between the other nodes rather than based on its closeness to the other nodes. For example, when calculating the betweenness centrality of a node, the shortest distances between the other nodes and how many of them pass through that node are determined. The more the node takes part in the shortest paths, the higher its betweenness centrality value. The most important factor to consider while calculating the betweenness centrality of the nodes is that there may be more than one shortest path between the nodes. It can be seen that the E node had the highest betweenness centrality value in Table 4. This is because there was no alternative to the E node for the transition to the C and B nodes. At the same time, like the closeness centrality values of nodes G, L and H, the betweenness centrality values of nodes G, L and H seemed to be high. In the second case, the I node attached to the network induced a change in the order of the other nodes, especially since it facilitated the transition to the M and N nodes.

Eigenvector centrality is a measurement of centrality that takes degree centrality a step further. In this concept, the importance of a node depends not only on the number of ties it has but on the importance of its neighbors, as well. In this way, eigenvector centrality can yield results that reflect the network structure more accurately than degree centrality. The nodes with the highest eigenvector centrality, as revealed in Table 6, appeared to be ranked in the same way as the nodes with the highest degree centrality. Since the concept of eigenvector centrality is similar to the concept of degree centrality, it was seen in the results that the rankings were similar. The L node ranked first in both working positions because it had both high degree centrality and neighbors with high degree centrality. In both situations, the K node had a higher eigenvector centrality value than the J node. This is because the K node was a neighbor to the L node, which had the high

__Table 7__ Bonacich power centrality values of case studies

First Case		Second Case	
L	78160776	K	80128328
K	77112864	L	68472496
N	57223936	J	60620028
J	53638424	H	50518176
H	49054532	N	33966540
G	31645060	G	32196792
F	11707991	M	15508986
M	11152561	I	13719325
E	5692315	E	7213459
D	5371831	D	5668435

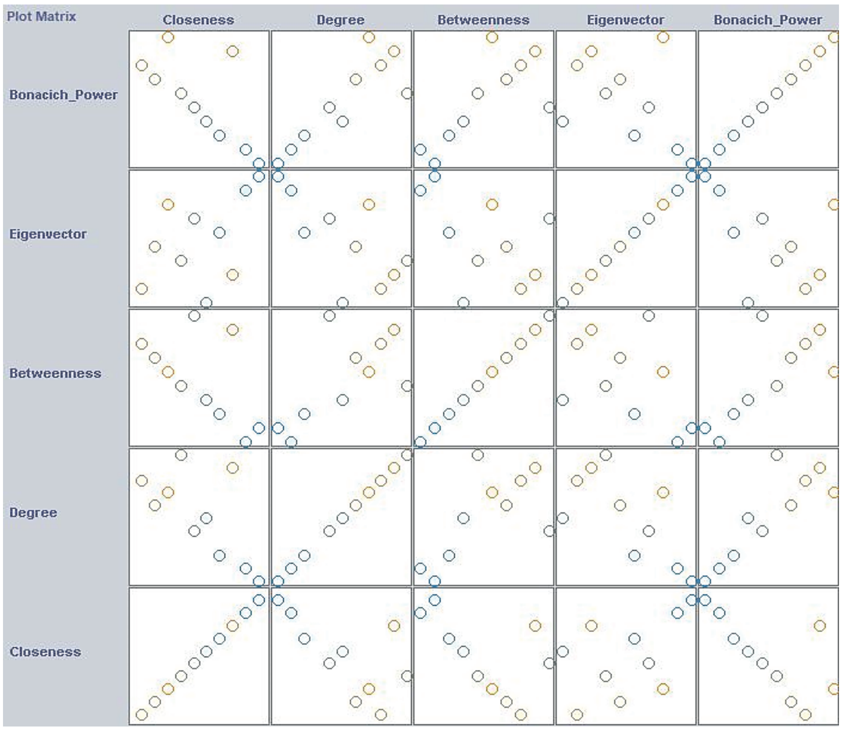

__Figure 6__ Plot matrix of centrality concepts

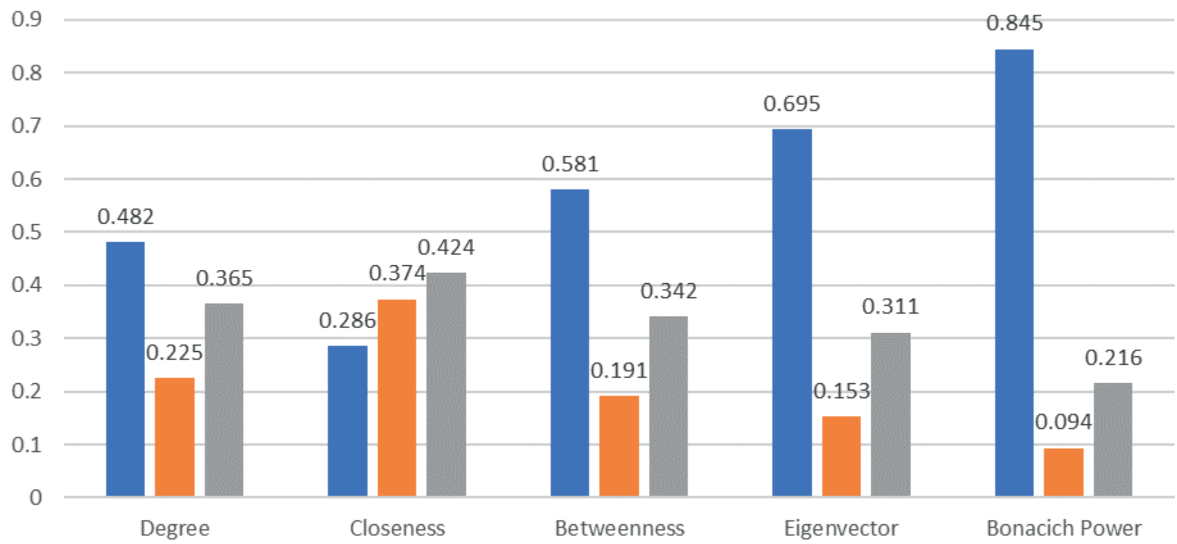

__Figure 7__ Error scales and accuracy value

degree centrality. In addition, it was observed that the addition of the I node to the network structure barely affected this centrality measure because the order of the nodes with the six highest eigenvector centrality values did not change.

Bonacich power was proposed by Phillip Bonacich as an important method in SNA because it considers the tie values and location of the nodes in the network structure. The node in the network structure is a comprehensive method that considers the neighbors of the nodes in the whole network, not merely their neighbors [38]. When the Bonacich power centrality values of this study, as presented in Table 7, were examined, it was observed that the L node had the highest value in the first case. Although the L node, with the highest volume value, ranked first in eigenvector centrality and degree centrality in both case studies, the K node had the highest Bonacich power centrality value in the second case study. It was observed that the highest Bonacich power centrality value passed from the L node to the K node with the addition of the I node in the second case study. In other words, the node moved to a more northern point. These results reflected the reality that the new bridge and the node that was added to the network structure are located in the northern part of the city of Istanbul.

Each centrality measure examines the network structure from different approaches. For this reason, it should be determined which centrality criterion gives better results by using the results obtained. The concept of centrality, which gives the best results, will be more suitable for the road network structure. The SVM algorithm was used to perform this operation. First of all, the relationship between this algorithm and the concepts of centrality was examined. The relationship of each centrality concept with each other is shown in Figure 6. After this process, Kappa statistical value and error scales for each centrality criterion, were calculated using the SVM. These results are shown in the graph in Figure 7. The column in dark blue represents the kappa statistical value, the column in orange represents the MAE value and the column in gray represents the RMSE value.

The Bonacich power centrality has values of 0.094 MAE, 0.216 RMSE. Accordingly, the centrality value with the lowest error rate was the Bonacich power centrality. Similarly, the Kappa statistic value is 0.845. This value is quite high compared to other centrality concepts. Bonacich power centrality has the best results in terms of both error scales and accuracy.

As a result of these analyses and comparisons, it was decided that the centrality indicator that provided the most appropriate results for this transportation network structure is the Bonacich power centrality, which examined the network structure in terms of both centrality and degree. Furthermore, this concept of centrality examined the network structure completely rather than locally. For these reasons, it was observed that the Bonacich power centrality values provided more realistic results than the other centrality measurements in determining the most central intersection of the transportation road network.

5 Conclusions

With the global increase in traffic congestion, modeling and analyzing the road networks have become very important. Various modeling and analysis methods have been developed for this purpose. In this study, the traffic networks were modeled and analyzed using Netdraw and Ucinet, which are SNA Software with a practical and innovative structure. The results were compared using support vector machine with Weka Software, a machine learning program. When the measures of centrality were examined, it was observed that the most appropriate concept of centrality for the transportation network analyzed in this study was Bonacich power centrality, which has examined the nodes in a network structure both in terms of their positions and degrees. This situation is very suitable for the road network structure. Since in determining the most central intersection in road network structures, both the locations of the intersections and the traffic volume values are taken as basis. For the first case study, the L intersection on the south of the Anatolian side was identified as the most central intersection using this centrality measure. This intersection is at the intersection point of the ongoing ring road of 2nd Bosphorus Bridge and the extension of the 1st Bosphorus Bridge. In the second case study, the K intersection was identified as the most central intersection. This intersection is also on the Anatolian side and closer to the midpoints, compared to the L intersection. It is located at an important point in the intersection zone of the 2nd Bosphorus Bridge. Both nodes are in the transition zone to the bridges and they are at intersections with high traffic volumes. The 3rd bridge and new road added to the road traffic in Istanbul caused the most central intersection to change. All these situations support the idea that these nodes are the central points.

Overall, improvements made at the most central intersection in the road network structure can reduce traffic congestion and delays in the entire network structure. The results of this study can guide the local governments. Decision makers can prioritize the idea of practical implementation and useful methods such as SNA when planning the improvement of any transport network. Additionally, the future line of research can measure how improvements could occur at central intersections and the overall impact of these improvements on traffic. This article proposes application of the SNA method in engineering studies involving the relationship between transport networks and nodes the need to identify the most central nodes in the network structure.

References

[1] LIU, X., WANG, J., LIU, T., XU, J. Forecasting spatiotemporal boundary of emergency-event-based traffic congestion in expressway network considering highway node acceptance capacity. *Sustainability* [online]. 2021, **13**(21), 12195. eISSN 2071-1050. Available from: https://doi.org/10.3390/su132112195

[2] BECKMANN, M. J. Traffic congestion and what to do about it. *Transportmetrica B: Transport Dynamics* [online]. 2013, **1**(1), p. 103-109. ISSN 2168-0566, eISSN 2168-0582. Available from: https://doi.org/10.1080/21680566.2013.780988

[3] MATHEW, T. V., BOMBAY, I. T. Automated traffic measurement. Transportatin System Engineering [online] [accessed 2021-09-11]. Lecture notes and presentations. Available from: https://www.civil.iitb.ac.in/tvm/1100_LnTse/ceTseLn/ceTseLn.html

[4] KIM, J., SARKAR, S., VENKATESH, S. S., RYERSON, M. S., STAROBINSKI, D. An epidemiological diffusion framework for vehicular messaging in general transportation networks. *Transportation Research Part B: Methodological* [online]. 2020, **131**, p. 160-190. ISSN 0191-2615. Available from: https://doi.org/10.1016/j.trb.2019.11.004

[5] WANG, D. Z. W., XIE, D. F. Reliable transportation network design considering uncertain demand variability. *Journal of Sustainable Transportation* [online]. 2016, **10**(8), p. 752-763. ISSN 1556-8318, eISSN 1556-8334. Available from: https://doi.org/10.1080/15568318.2016.1149644

[6] Traffic counting systems - General Directorate of Highways [online] [accessed 2022-03-21]. Available from: http://www.kgm.gov.tr/SiteCollection Documents /KGMdocuments/Istatistikler/TrafikveUlasimBilgileri/17TrafikUlasimBilgileri.pdf

[7] KOC, E., CETINER, B., ROSE, A. SOIBELMAN, L., TACIROGLU, E., WEI, D. CRAFT: Comprehensive resilience assessment framework for transportation systems in urban areas. *Advanced Engineering Informatics* [online]. 2020, **46**, 101159. ISSN 1474-0346. Available from: https://doi.org/10.1016/j.aei.2020.101159

[8] GAUTHIER, P., FURNO, A., EL FAOUZI, N. E. Road network resilience: how to identify critical links subject to day-to-day disruptions. *Transportation Research Record* [online]. 2018, **2672**, p. 54-65. ISSN 0361-1981, eISSN 2169-4052. Available from: https://doi.org/10.1177/03611981187921

[9] WEN, H., SUN, J., ZENG, Q., ZHANG, X., YUAN, Q. The effects of traffic composition on freeway crash frequency by injury severity: a Bayesian multivariate spatial modeling approach. *Journal of Advanced Transportation* [online]. 2018, **2018**, 6964828. ISSN 0197-6729, eISSN 2042-3195. Available from: https://doi.org/10.1155/2018/6964828

[10] HADAS, Y., GNECCO, G., SANGUINETI, M. An approach to transportation network analysis via transferable utility games. *Transportation Research Part B: Methodological* [online]. 2017, **105**, p. 120-143. ISSN 0191-2615. Available from: https://doi.org/10.1016/j.trb.2017.08.029

[11] ZHANG, X., ZHANG, H. M., LI, L. Analysis of user equilibrium traffic patterns on bottlenecks with time-varying capacities and their applications. *International Journal of Sustainable Transportation* [online]. 2010, **4**(1), p. 56-74. ISSN 1556-8318, eISSN 1556-8334. Available from: https://doi.org/10.1080/15568310601060036

[12] FREEMAN, L. C. *The development of social network analysis*. Vancouver, BC Canada: Empirical Press, 2014. ISBN 1-59457-714-5.

[13] PRITCHARD, J. P., MOURA, F., ABREU E SILVA, J. Incorporating social network data in mobility studies: benefits and takeaways from an applied survey methodology. *Case Studies on Transport Policy* [online]. 2016, **4**(4), p. 279-293. Available from: https://doi.org/10.1016/j.cstp.2016.09.002

[14] CHENG, Y. Y., LEE, R. K. W., LIM, E. P., ZHU, F. Measuring centralities for transportation networks beyond structures [online]. In: *Applications of Social Media and Social Network Analysis*. KAZIENKO, P., CHAWLA, N. (eds.). Lecture Notes in Social Networks. Cham: Springer, 2015. ISBN 978-3-319-19002-0, eISBN 978-3-319-19003-7, p. 23-39. Available from: https://doi.org/10.1007/978-3-319-19003-7_2

[15] COOPER, C. H. V. Predictive spatial network analysis for high-resolution transport modeling applied to cyclist flows, mode choice and targeting investment. *International Journal of Sustainable Transportation* [online], 2018, **12**(10), p. 714-724. ISSN 1556-8318, eISSN 1556-8334. Available from: https://doi.org/10.1080/15568318.2018.1432730

[16] WATTS, R., WITHAM, A. Social network analysis of sustainable transportation organizations [online]. Available from: http://www.uvm.edu/~transctr/research/trc_reports/UVM-TRC-12008.pdf 2012

[17] EL-ADAWAY, I. Analyzing traffic layout using dynamic social network analysis. NCITEC Project, 2014.

[18] EL-ADAWAY, I. H., ABOTALEB, I., VECHAN, E. Identifying the most critical transportation intersections using social network analysis. *Transportation Planning and Technology* [online]. 2018, **41**(4), p. 353-374. ISSN 0308-1060, eISSN 1029-0354. Available from: https://doi.org/10.1080/03081060.2018.1453456

[19] UYGUN, M. Stakeholders analysis of health tourism in Turkey: social network analysis approach practice in Ankara. Ankara: Yildirim Beyazit University, 2018.

[20] YAPICIOGLU, B., MOGBO, O. N., YITMEN, I. Innovative strategies for transport policies in infrastructure development: Nigerian stakeholders' perspective. *International Journal of Civil Engineering* [online]. 2017, **15**, p. 747-761. ISSN 1735-0522, eISSN 2383-3874. Available from: https://doi.org/10.1007/s40999-017-0172-0
[21] STORTO, C. An SNA-DEA prioritization framework to identify critical nodes of gas networks: the case of the US interstate gas infrastructure. *Energies* [online]. 2019, **12**, 4597. eISSN 1996-1073. Available from: https://doi.org/10.3390/en12234597
[22] KEMAL, Y. Increasing the effect of marketing from mouth to mouth with the social network analysis method: application on data sets. Ege University, 2011.
[23] OPSAHL, T., AGNEESSENS, F., SKVORETZ, J. Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks* [online]. 2010, **32**(3), p. 245-251. ISSN 0378-8733. Available from: https://doi.org/10.1016/j.socnet.2010.03.006
[24] FALZON, L., QUINTANE, E., DUNN, J., ROBINS, G. Embedding time in positions: temporal measures of centrality for social network analysis. *Social Networks* [online]. 2018, **54**, p. 168-178. ISSN 0378-8733. Available from: https://doi.org/10.1016/j.socnet.2018.02.002
[25] SABAH, L. Social network analysis and outbreak modeling. Duzce University, 2018.
[26] BONACICH, P. Some unique properties of eigenvector centrality. *Social Networks* [online]. 2007, **29**(4), p. 555-564. ISSN 0378-8733. Available from: https://doi.org/10.1016/j.socnet.2007.04.002
[27] HSIEH, M. Simple network analysis with UCINET - Engineering Systems Division [online]. Available from: https://ocw.mit.edu/courses/engineering-systems-division/esd-342-network-representations-of-complex-engineering-systems-spring-2010/tools/ucinet_slides.pdf. 2006
[28] LENDEL, V., PANCIKOVA, L., FALAT, L., MARCEK, D. Intelligent modelling with alternative approach: application of advanced artificial intelligence into traffic management. *Communications-Scientific Letters of the University of Zilina* [online]. 2017, **19**(4), p. 36-42. ISSN 1335-4205, eISSN 2585-7878. Available from: https://doi.org/10.26552/com.C.2017.4.36-42
[29] BUITINCK, L., LOUPPE, G., BLONDEL, M., PEDREGOSA, F., MUELLER, A., GRISEL, O., NICULAE, V., PRETTENHOFER, P., GRAMFORT, A., GROBLER, J., LAYTON, R., VANDERPLAS, J., JOLY, A., HOLT, B., VAROQUAUX, G. API design for machine learning software: experiences from the scikit-learn project. 2013. p. 1-15.
[30] KARANDIKAR, J. Machine learning classification for tool life modeling using production shop-floor tool wear data. *Procedia Manufacturing* [online]. 2019, **34**, p. 446-454. ISSN 2351-9789. Available from: https://doi.org/10.1016/j.promfg.2019.06.192
[31] GALDI, P., TAGLIAFERRI, R. Data mining: accuracy and error measures for classification and prediction [online]. In: *Encyclopedia of Bioinformatics and Computational Biology*. RANGANATHAN, S., GRIBSKOV, M., NAKAI, K., SCHONBACH, CH. (eds.). Vol. 1. Elsevier, 2018. ISBN 978-0-12-811432-2, p. 431-436. Available from: https://doi.org/10.1016/B978-0-12-809633-8.20474-3
[32] MCHUGH, M. L. Lessons in biostatistics interrater reliability : the kappa statistic. *Biochemica Medica* [online]. 2012, **22**(3), p. 276-282. ISSN 1330-0962, eISSN 1846-7482. Available from: https://doi.org/10.11613/BM.2012.031
[33] SAMMUT, C., WEBB, G. I. Mean absolute error [online]. In: *Encyclopedia of Machine Learning*. SAMMUT, C., WEBB, G. I. (eds.). Boston, MA: Springer US, 2010. ISBN 978-0-387-30768-8, eISBN 978-0-387-30164-8. Available from: https://doi.org/10.1007/978-0-387-30164-8_525
[34] Traffic density map - Istanbul Metropolitian Municipality [online] [accessed 2021-10-04]. Available from: https://uym.ibb.gov.tr/YHarita/ Harita_tr.aspx
[35] Traffic volume maps - General Directorate of Highways [online] [accessed 2021-09-26]. Available from: http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/ Trafik/TrafikHacimHaritasi.aspx
[36] EL-ADAWAY, I. H., ABOTALEB, I. S., VECHAN, E. Social network analysis approach for improved transportation planning. *Journal of Infrastructure Systems* [online]. 2017, **23**(2), p. 1-14. ISSN 1076-0342, eISSN 1943-555X. Available from: https://doi.org/10.1061/(ASCE)IS.1943-555X.0000331
[37] HART, M. G., YPMA, R. J. F., ROMERO-GARCIA, R., PRICE, S. J., SUCKLING, J. Graph theory analysis of complex brain networks: new concepts in brain mapping applied to neurosurgery. *Journal of Neurosurgery* [online]. 2016, **124**(6), p. 1665-1678. ISSN 0022-3085, eISSN 1933-0693. Available from: https://doi.org/10.3171/2015.4.JNS142683
[38] KUSKAPAN, E., CODUR, M. Y., TORTUM, A. Identifying the most critical intersections in transportation networks. *Technical Gazette* [online]. 2021, **28**(6), p. 1920-1926. ISSN 1330-3651, eISSN 1848-6339. Available from: https://doi.org/10.17559/TV-20201117162526