Understanding the topological characteristics and flow complexity of urban traffic congestion

Article *in* Physica A: Statistical Mechanics and its Applications · January 2017

DOI: 10.1016/j.physa.2017.01.035

3 authors:


Tzai-Hung Wen
National Taiwan University


Benny Chin
Singapore University of Technology and Design


Pei-Chun Lai
National Taiwan University

Physica A 473 (2017) 166–177

Contents lists available at ScienceDirect

Physica A

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

Understanding the topological characteristics and flow complexity of urban traffic congestion

Tzai-Hung Wen *, Wei-Chien-Benny Chin, Pei-Chun Lai

Department of Geography, College of Science, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

HIGHLIGHTS

- A ranking algorithm is proposed to investigate traffic demands and flow complexity.
- Traffic demand reflects the concentration of human movements.
- Flow complexity reflects the street network topology and turning traffic.
- The ranking algorithm captures the potential urban congestion zones.
- This study provides a network topological insight into street-scale urban mobility.

ARTICLE INFO

Article history:
Received 19 September 2016
Received in revised form 16 December 2016
Available online 6 January 2017

Keywords:
Traffic congestion
PageRank algorithm
Network topology
Mobility

ABSTRACT

For a growing number of developing cities, the capacities of streets cannot meet the rapidly growing demand of cars, causing traffic congestion. Understanding the spatial–temporal process of traffic flow and detecting traffic congestion are important issues associated with developing sustainable urban policies to resolve congestion. Therefore, the objective of this study is to propose a flow-based ranking algorithm for investigating traffic demands in terms of the attractiveness of street segments and flow complexity of the street network based on turning probability. Our results show that, by analyzing the topological characteristics of streets and volume data for a small fraction of street segments in Taipei City, the most congested segments of the city were identified successfully. The identified congested segments are significantly close to the potential congestion zones, including the officially announced most congested streets, the segments with slow moving speeds at rush hours, and the areas near significant landmarks. The identified congested segments also captured congestion-prone areas concentrated in the business districts and industrial areas of the city. Identifying the topological characteristics and flow complexity of traffic congestion provides network topological insights for sustainable urban planning, and these characteristics can be used to further understand congestion propagation.

1. Introduction

For a growing number of developing cities, the capacities of streets cannot meet the growing demand of cars, causing traffic congestion. Some cities have complex street structures that also cause traffic congestion. Congestion causes global concerns, such as increased commuting times and fuel usage as well as environmental deterioration. Urban congestion can be categorized into phenomena at the local and global scales. Local congestion, such as single interactions, only decreases

* Corresponding author.
E-mail address: wenthung@ntu.edu.tw (T.-H. Wen).

http://dx.doi.org/10.1016/j.physa.2017.01.035
0378-4371/

the velocity of individual vehicles, whereas global congestion often decreases the velocity of the overall street network and requires additional traffic control [1]. Therefore, understanding the process of traffic flow and detecting traffic congestion are important issues associated with developing urban policies to resolve congestion.

Traffic congestion is a complex spatial–temporal process [2]. It is often triggered by recurring factors, such as inadequate capacity or recurrent weather [3], and unexpected factors, such as accidents, lane closures, and special events [3]. Understanding the factors that cause congestion has become an important issue in traffic management. The causes of traffic congestion can be measured by physical and psychological factors. Physical causes measure the traffic volume, speed, and street density. Psychological congestion is more difficult to measure. Some people are accepting of slight congestion, whereas others are not; thus, fuzzy logic has been used for the evaluation of congestion [4,5]. Some studies have focused on developing methods to detect congestion, whether it is psychological or physical. In early studies, researchers developed point-based detectors that examined the traffic volume over a unit distance [6,7]. In recent studies, development of vehicle-to-vehicle systems has increased [8–10]. A vehicle-to-vehicle system is a dynamic, non-location-based set of detectors. A vehicle-to-vehicle system utilizes nodes on the sides of the street and vehicles and uses short-range communication equipment to inform closed nodes of the current state of movement. The vehicle is a moving node, and the street-side nodes are fixed and connected to a management center. By analyzing the dynamic characteristics of traffic movement, unexpected congestion can be detected, and the management center can make better decisions for traffic control.

The spatial–temporal process of traffic congestion needs to be examined simultaneously from the perspectives of transportation infrastructure and traffic demand. Regarding the transportation infrastructure, traffic congestion could be the result of the topological complexity of traffic flow in terms of turning probability [11,12]. If the street structure is more complex, moving vehicles may have more opportunities for turning from one street to another. Thus, the complexity of turning directions at a crossroads could influence the moving speeds of automobiles and cause traffic congestion. In contrast, if the street structure is relatively simple, moving vehicles might choose the same direction; thus, these vehicles would maintain constant moving speeds through a crossroads. In other words, the topological structure of a street network influences the turning probabilities between streets and the moving speeds of automobiles on those streets. Therefore, if more vehicles are moving between street segments at lower moving speeds, the street becomes congested.

Analysis of traffic demand considers the functionalities of destinations and connectivities of street segments. The functionality of a destination includes the types of buildings, facilities, or land use at a specific location [13]. In cities, most people work in the commercial districts and areas where most office buildings are concentrated. Therefore, high commuting needs are expected in those areas and nearby areas. The traffic conditions near central business districts (CBD) and financial districts would be heavier than those in residential areas [14]. Moreover, the connectivity of street segments reflects the degree to which the street system facilitates people moving toward their destinations. Assuming that people tend to choose the shortest or fastest routes for moving toward their destinations [15,16], the route selection of each moving person is dependent on the connectivity of the street system. For example, a street segment with a higher connectivity to the city center means that it is easier to travel to the city center using that street. Therefore, higher connectivity in a street network could reflect a higher traffic flow concentration and traffic demand. Thus, higher connectivity could have a high potential to cause traffic congestion.

In summary, traffic congestion can be attributed simultaneously to the topological structure of the street network and to traffic demand. Therefore, the objective of this study is to propose an innovative flow-based ranking procedure for investigating the traffic demand in terms of traffic flow concentration and the flow complexity of the street network based on the turning probability. By analyzing street network structures and traffic flow, we attempted to calibrate the turning probabilities of street segments and to measure the complexity of traffic flow for identifying congestion-prone street segments. Identifying the topological characteristics and flow complexity of traffic congestion can provide network-structure insights for sustainable urban planning, and these characteristics can be used to further understand congestion propagation.

2. Methods

2.1. Study framework

The study framework is composed of the three major procedures illustrated in Fig. 1. In the first procedure, we proposed a flow-based ranking algorithm (termed the Flow-based PageRank, FBPR), which contains two components: one is to develop the iterative process to analyze the attractiveness of a node for capturing traffic demands (see Section 2.4), and the other is to use a small fraction of segments with an actual traffic volume of moving vehicles to calibrate the attractiveness of all the nodes in the graph (see Section 2.5). In the second procedure, we measured the topological complexity of traffic flow in terms of outgoing entropy and calculated the traffic flow concentration based on the calibration results. The congested segments were defined as the street segments that are prone to traffic congestion. In the third procedure, congested segments were identified by combining the high traffic demand in terms of FBPR scores and the high traffic flow complexity in terms of outgoing entropy. Taipei City, which is one of the major metropolitan areas in East Asia, was used as a case study to demonstrate the feasibility of the proposed framework.

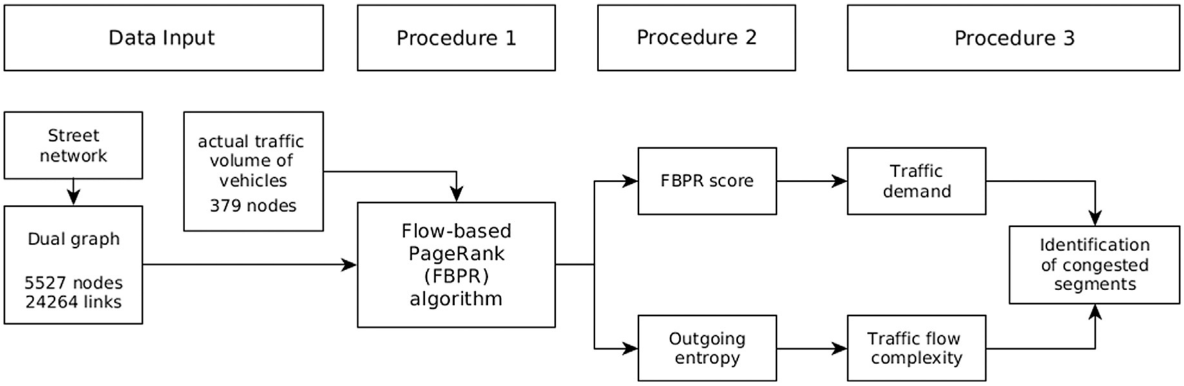


Fig. 1. The study framework.

2.2. The study area

Taipei City, a major metropolitan area and the commercial/political center of Taiwan, has an area of approximately 71.8 km^2, and the overall population density of Taipei City is approximately 9.9 thousand people/km^2. The village-level population densities range from below 9 thousand people/km^2 to above 30 thousand people/km^2 (Fig. 2(a)). The physical landscape of Taipei City (Fig. 2(b)) includes the main basin areas at the western boundary. The Keelung River flows from east to northwest and converges with the Tamsui River at the western boundary, which separates the population in the city center from that in the corridor areas of northwest Taipei City. Mountainous areas are located to the northeast and southeast of Taipei City. Fig. 3 shows important facilities and infrastructure in Taipei City. These include the three major railway stations (Songshan, Wanhua, and Taipei Main stations), areas with high densities of offices (the highest commercial building in Taiwan (Taipei 101) and Neihu Technology Park), important markets and leisure areas (Beitou Hot spring resorts, Shihlin market, Guanghua digital plaza, Nanmen market, and Ximen market), and three facilities (Rong Zong hospital, NTU hospital, and National Taiwan University). The locations of these facilities and infrastructures were used for validating the congestion segments identified by the proposed method.

2.3. Data

2.3.1. Street network and its dual representation

The street network data were collected by the Institute of Transportation, Ministry of Transportation and Communication of Taiwan. The street network data contain the street names, types, and locations. The street types include national streets (freeways), elevated streets (viaducts), county streets, normal streets, country streets and lanes. This study focused on traffic conditions in the major planar street network; therefore, we filtered out the lanes, which have little influence due to low traffic volumes, and elevated streets.

This study uses a dual graph (or dual representation) of the street topology [17–20] to analyze the relationships between the street segments. In the dual graph, the nodes represent the street segments and the links represent the connections between the street segments (Fig. 4). The weight of a link can be represented as the turning probability from one segment to another. Using this dual representation of the street network, the structure of the connectivity between street segments and the vehicle turning relationships could be captured intuitively [17]. The dual graph in this study includes 5527 nodes and 24,264 links.

2.3.2. Traffic volume

Traffic volume data were provided by the Taipei Traffic Control Center and measured by vehicle detectors (VD). A VD installed in a street segment automatically detects the traffic volume and average speed of automobiles crossing the street segment. The traffic information was updated every five minutes, and the data items included exchange time, device name, the name of the location, traffic volume, average speed, and the longitude and latitude coordinates of the VD. A total of 379 VDs were used in this study, and traffic volume data were collected from July 1st to August 31st in 2012. The VDs were scattered only on major street segments across the city. Therefore, in the dual graph of the street network, there are 379 nodes (street segments) that have traffic volume data from VDs, which are only 6.8% of the total nodes in the city.

2.3.3. Land use types

The land uses of Taipei City were grouped into seven types, including residential areas, business districts, industrial areas, and land used for schools, leisure activities, transport facilities, and others. Residential areas are often places where travel originates. Business districts, industrial areas, and land for schools are common destinations of commuter trips. Land used

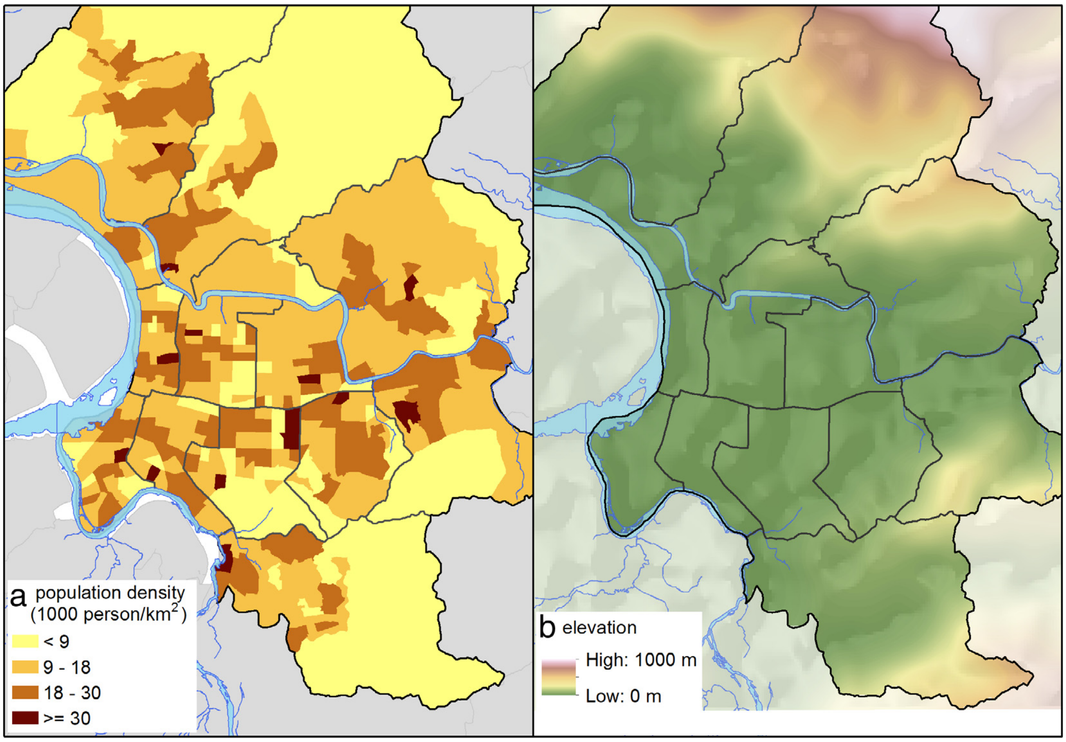


Fig. 2. The spatial distribution of population density (a) and elevation (b) of Taipei City.

for leisure activities is a destination for weekends or holiday trips. Transport facilities include airports, railway stations, and elevated roads. The type of land use might influence the traffic demands, destination attractiveness and moving speeds of the street segments. Therefore, each street segment was assigned one of these land use types. We further investigated the association of the flow concentration and the complexity with land use types.

2.4. Modeling traffic demand using a flow-based ranking algorithm

A new flow-based ranking algorithm was proposed to capture the conceptual traffic demand on street segments. The algorithm is called Flow-based PageRank (FBPR) and was based on PageRank (PR). PR is an algorithm used by the Google search engine [21]. PR uses weighting scores to rank web pages and identify the significant web pages on the enormous and complicated World Wide Web. The calculation process of PR is similar to the flow of vehicles into and out of a street network. The streets with higher volumes lead to higher volumes on neighboring streets, and streets with more connections will have higher volumes. However, some settings in PageRank do not correspond to the structures of street networks. Human movement is influenced by both network structure and individual options [22]. Puzis et al. [23] proposed that network analyses should consider the characteristics of human movements. Socioeconomic structures and the purposes of trips are often used as variables when simulating movement patterns [24,25].

Therefore, the FBPR uses a different weighting scheme to emphasize the attractiveness of street segments and determine the turning probabilities between streets (weights of link relationships) [22,26]. The link weights reflect location attractiveness, which is associated with land use patterns and destination attractiveness. Real traffic volumes from VDs were used to calibrate the attractiveness and link relationships to make the FBPR scores fit the distributions of real traffic volumes. A genetic algorithm (GA) was used to determine the optimal weights of the link relationships.

The procedure of FBPR is illustrated in Fig. 5. Link weights were determined by the attractiveness of street segments, which is a standardized number from one to 100. Eq. (1) indicates the transfer proportion, based on attractiveness, from street segment v to u. If $Weight(v, u)$ equals 0.5, then half of the traffic volume on street segment v moves to street segment u. The FBPR score of street segment u is defined by its attractiveness proportion among other outgoing street segments from street segment v, which is multiplied by the FBPR score of street segment v in the previous iteration (Eq. (2)). After an iteration of this procedure is performed, the current score of each node is used in the next iteration. This process is repeated until the scores of all nodes reach equilibrium. For this study, we tested the number of iterations from one to 1000 and found that the mean of the absolute difference of the scores for each node between iteration 500 and 501 was lower than

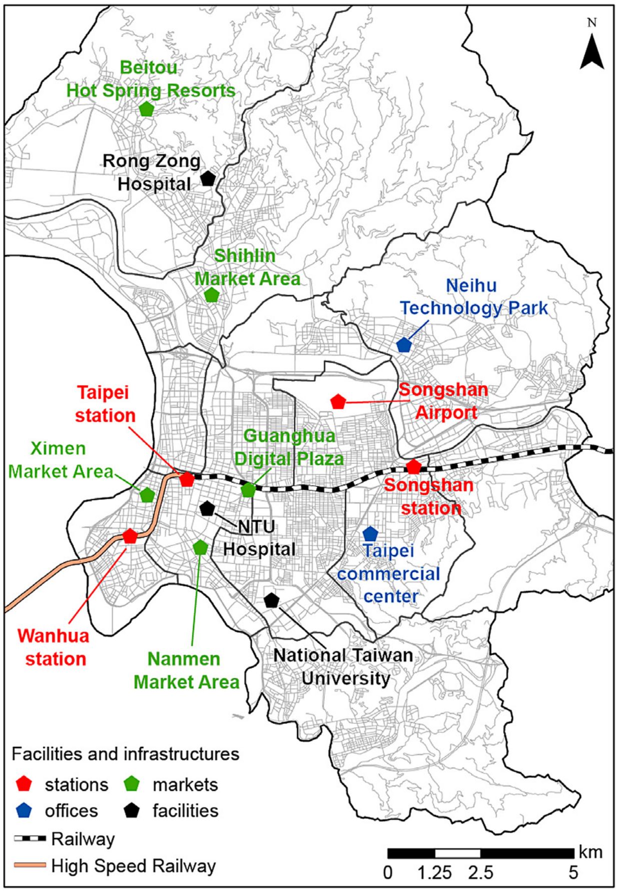


Fig. 3. Spatial distribution of important facilities and infrastructures in Taipei City.

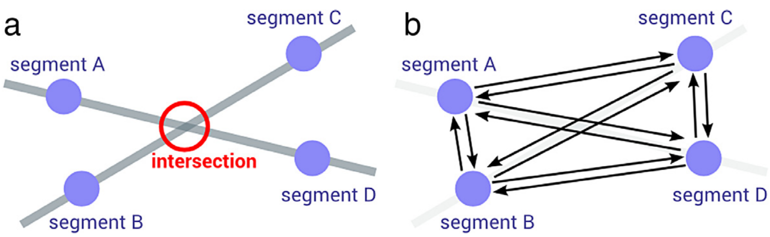


Fig. 4. An illustration of transforming from street network to dual graph. (a) Two streets intersect and separated into four segments by the intersection; (b) the four segments are connected to each other.

1.1×10^{-8}. Spearman's rank correlation between iterations 500 and 501 was almost equal to one. Therefore, the nodes reached equilibrium in 500 iterations. Then, the final FBPR score was used to represent the traffic demand of vehicle flows.

$$Weight(v, u) = \frac{Attr(u)}{\sum_{p \in R(v)} Attr(p)} \tag{1}$$

$$FBPR_t(u) = \sum_{v \in B(u)} FBPR_{t-1}(v) \times Weight(v, u) \tag{2}$$

Here, $Weight(v, u)$ is the link weight from street segment v to u; $Attr(u)$ and $Attr(p)$ are the values of the attractiveness of street segments u and p, respectively; $R(v)$ is the set of outgoing nodes of street segment v, which includes street segment u; $B(u)$ is the set of incoming nodes of street segment u; and $FBPR_t(u)$ and $FBPR_{t-1}(v)$ are the FBPR scores of street segment u at iteration t and street segment v at iteration $t - 1$, respectively.

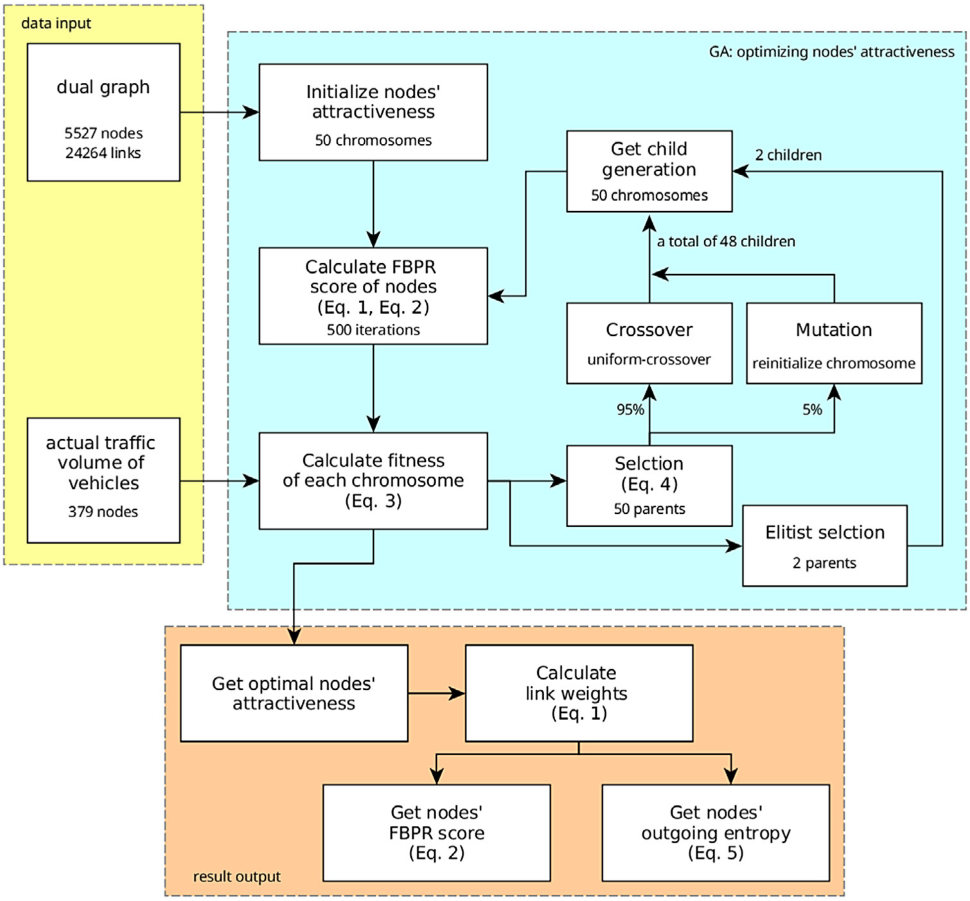


Fig. 5. The procedure of Flow-based PageRank (FBPR) algorithm.

2.5. Calibrating the attractiveness of street segments using the genetic algorithm

We used the GA to calibrate the attractiveness of street segments based on network characteristics and the traffic volumes of nodes. In the first step, the initial generation of 50 chromosomes (solutions) with lengths equal to the total number of nodes in the graph were created by randomly choosing an integer from one to 100 for each node. In the second step, each chromosome was used to calculate the weights (Eq. (1)) based on the network connectivity. Then, the iteration process in the FBPR calculation was repeated until equilibrium (500 iterations), and the final FBPR score (Eq. (2)) of each node was determined. Spearman's rank correlation between the FBPR scores and the actual traffic volume of the nodes from VD data was used as the fitness function (Eq. (3)). The objective of the GA was to maximize the rank correlation to calibrate the attractiveness of the nodes.

$$fitness(s) = \frac{\sum_i ((R_{FBPR}(i) - \overline{R_{FBPR}})(R_{Vol}(i) - \overline{R_{Vol}}))}{\sigma_{rFBPR} \times \sigma_{rVol}} \tag{3}$$

Here, $R_{FBPR}(i)$ and $R_{Vol}(i)$ are the ranks of node i based on the FBPR scores and traffic volumes, respectively; $\overline{R_{FBPR}}$ and $\overline{R_{Vol}}$ are the average ranks of the FBPR scores and traffic volumes, respectively; and σ_{rFBPR} and σ_{rVol} are the standard deviations of the ranks of the FBPR scores and traffic volumes, respectively.

When a generation of the GA process had been completed, the fitnesses of all chromosomes were calculated. Selection, crossover, and mutation were then performed, based on the fitness of each chromosome, to create the offspring generation. For the selection process, the fitness was transformed into a relative probability (Eq. (4)) and was used as the selection criterion. The higher the relative probability, the more likely the solution would be selected into the mating pool. After the selection process, the crossover procedure randomly chose two chromosomes from the mating pool as the parent solutions, cut the same parts of the chromosomes for exchange, and created a pair of offspring. With a low probability (set to 5% in this study), the crossover procedure between two chosen chromosomes was replaced by a mutation procedure, which would create two random solutions as the offspring. After the crossover and mutation processes, 48 chromosomes would be

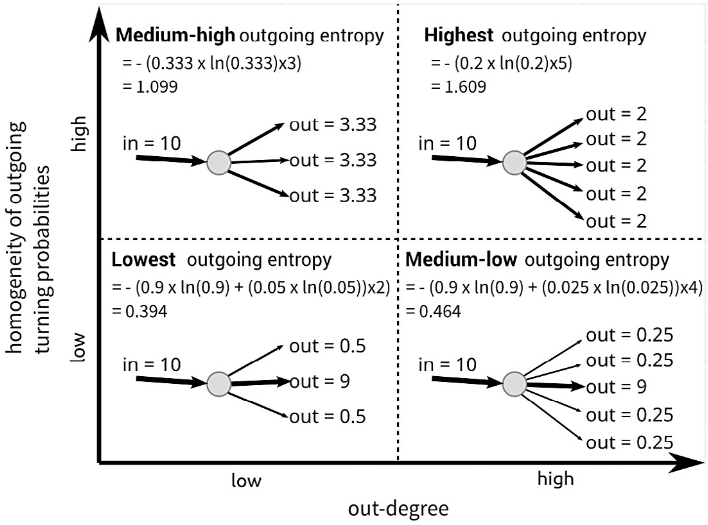


Fig. 6. Illustration of the patterns of traffic flow complexity in terms of outgoing entropy for different combinations of homogeneities of outgoing turning probabilities and out-degrees of nodes.

created. The two chromosomes with the highest and second-highest fitnesses from the parent generation would be cloned into the child generation (elitist selection). Thus, 50 chromosomes were created as the offspring. The procedures would then be repeated with the offspring solutions, starting from the second step. In this study, the GA process was repeated for 2000 generations before termination. The optimal solution in the last generation was used in the following analysis and discussion.

$$P(s) = \frac{fitness(s)}{\sum_{r=1}^{N} fitness(r)} \tag{4}$$

Here, $P(s)$ is the probability of the solution s being selected; $fitness(s)$ and $fitness(r)$ are the fitnesses of solution s and r, respectively; and N is the total number of solutions in one generation.

2.6. Determining the homogeneity of turning directions based on the outgoing entropy

The link weight reflects the turning probability between street segments, which can be calibrated using the real traffic volume data. If the outgoing link weights of a node are equal, the node's outgoing flow is moving equally toward its outgoing nodes. Thus, every vehicle moving along the street segment might turn in any direction at a junction, which can potentially lead to a lower moving speed. We used the concept of Shannon entropy [27] to measure the homogeneity of outgoing turning probabilities onto each street segment (Eq. (5)). The outgoing entropy in this study is defined in Eq. (5), which captures the complexity of network topology and turning traffic simultaneously.

$$Entropy(v) = -\sum_{p \in R(v)} (Weight(v, p) \times \ln(Weight(v, p))) \tag{5}$$

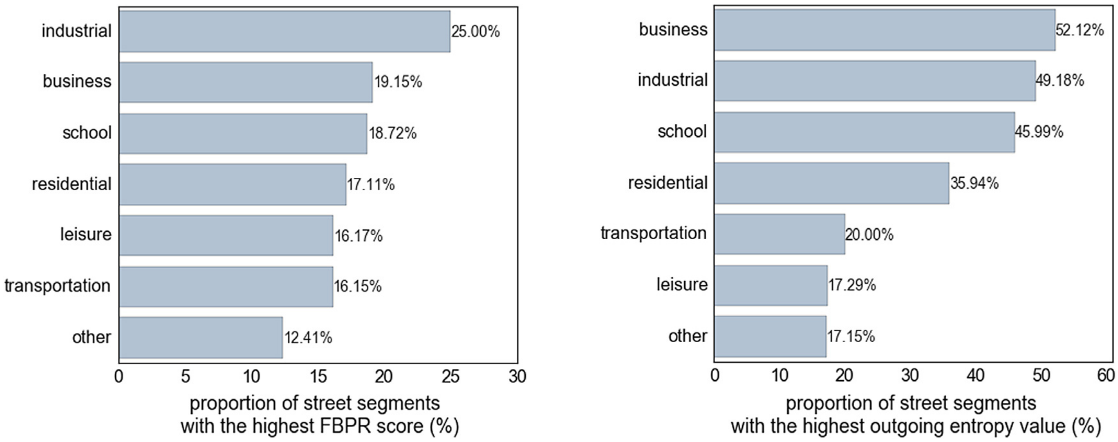
Here, $Entropy(v)$ is the outgoing entropy of a node (street segment v), $Weight(v, p)$ is the link weight from street segment v to p, and $R(v)$ is the set of outgoing nodes of street segment v.

Fig. 6 illustrates the patterns of traffic flow complexity in terms of outgoing entropy for different combinations of homogeneities of outgoing turning probabilities and out-degrees of nodes. The out-degree reflects the complexity of network topology, while the homogeneity of outgoing turning probabilities reflects the complexity of turning traffic at crossroads. A node with the highest outgoing entropy has the highest out-degree and equal outgoing turning probabilities simultaneously. Therefore, a street segment with a higher outgoing entropy suggests that it has more outgoing links and the outgoing flow is equally divided among them. This higher entropy could result in automobiles moving slowly and the street segment being prone to traffic congestion.

3. Results

3.1. The FBPR scores of traffic demand

The FBPR scores reflect the traffic demands of street segments. The higher the FBPR score is, the higher the traffic demand on a street segment. We categorized the traffic demand into three categories, high, medium and low using the Jenks natural



(a) FBPR score.

(b) Outgoing entropy.

Fig. 7. The comparison between the proportions of land-use types and the street segments with (a) high traffic demand (FBPR score); (b) high complexity of turning traffic (outgoing entropy value).

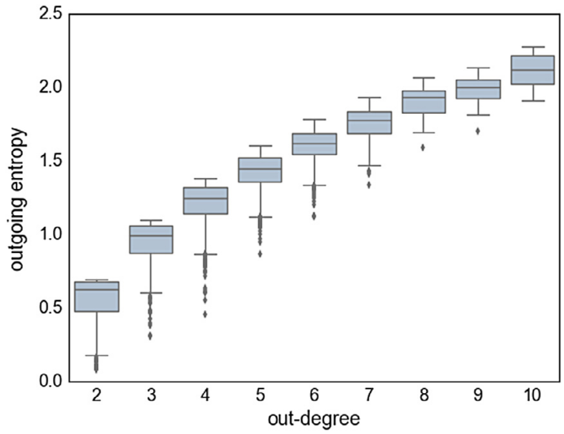


Fig. 8. The relationship between out-degree and outgoing entropy of street segments.

break classification method [28]. The Jenks natural break classification is a data clustering method designed to identify groups of data whose variances within the groups are minimized and between the groups are maximized. The proportions of land-use types with high traffic demand are compared in Fig. 7(a). The results show that industrial areas, business districts and land for schools have higher proportions of street segments with high traffic demand than residential areas and land used for leisure. This finding indicates that traffic flow is likely to concentrate in commercial and industrial areas, which could be a characteristic of commuting flow for urban mobilities.

3.2. The outgoing entropy as the topological complexity of traffic flow

The outgoing entropy captures the homogeneity of the turning probabilities of a street segment's outgoing links. This indicator reflects the complexity of the turning traffic. The streets were classified into three levels according to their outgoing entropy by using the Jenks natural break method [28]. As shown in Fig. 7(b), the industrial areas, businesses districts and land for schools have higher proportions of street segments with more complex turning traffic.

Fig. 8 shows the relationship between the out-degree and outgoing entropy of street segments. The figure shows that the outgoing entropy rises with the values of out-degree. However, the street segments with lower out-degree have larger variations of outgoing entropy. This finding indicates that the street segments with lower out-degree are more sensitive to the homogeneity of the turning directions of the outgoing traffic. The figure demonstrates the complex interactions between the number of outgoing links and directions of outgoing turning traffic at a crossroads.

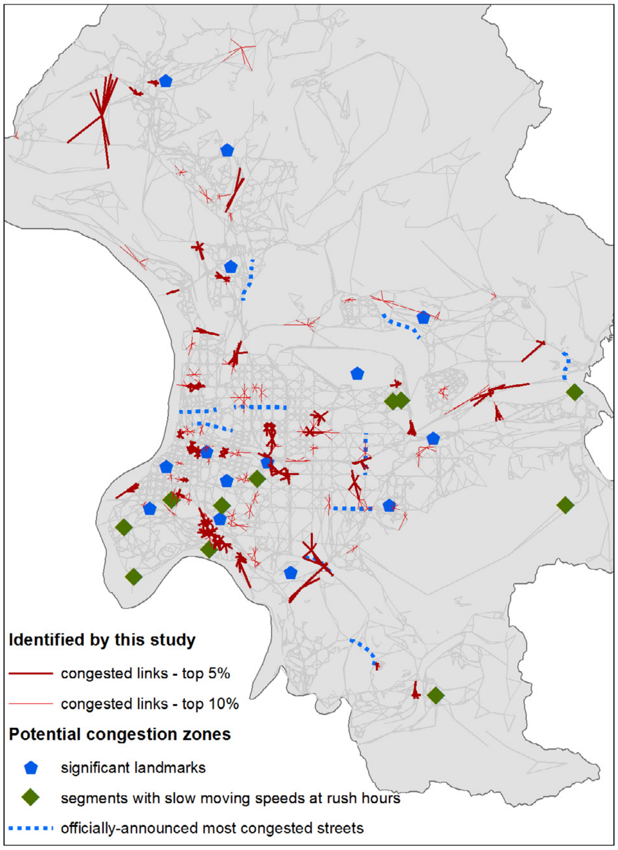


Fig. 9. The street segments prone to traffic congestion are identified as the street segments with both the highest (5% and 10% of) FBPR scores (high traffic demand) and the highest (5% and 10% of) outgoing entropy values (high flow complexity of turning traffic). The figure compared the identified congested segments with three kinds of potential congestion zones: 1. The most congested streets announced officially by the Taipei City Government (dashed lines); 2. Street segments with slow moving speeds at rush hours (average speed is less than 25 km/h) (diamond markers); and 3. The important facilities and infrastructure in Taipei City (pentagons markers).

3.3. Identifying the congested segments

Traffic congestion can occur on the street segments with high traffic demands and high complexity of turning traffic. Therefore, the street segments prone to traffic congestion can be identified as the street segments with both the highest (5% and 10% of) FBPR scores (high traffic demand) and the highest (5% and 10% of) outgoing entropy values (high flow complexity of turning traffic), highlighting the congested segments (Fig. 9).

We used three data sources as the potential congestion zones to validate the congested segments identified in this study. These potential congestion zones include the locations of the most congested streets announced officially by the Taipei City Government, street segments with slow moving speeds at rush hours (average speed is less than 25 km/h), and the locations of main rail stations or significant landmarks. Fig. 9 shows that most of the congested segments that were identified in this study successfully capture the potential congestion zones, including the officially announced most congested streets, segments with slow moving speeds or locations of significant facilities and landmarks. These identified congested segments also appeared around the main train or mass rapid transit (MRT) stations and significant landmarks.

Fig. 10 shows a histogram of distances between the identified congested segments and the potential congestion zones. The results show that 87.3% and 80.6% of identified congested segments (highest 5% and 10% of segments with both FBPR scores and outgoing entropy values, respectively) are within one kilometer of the potential congestion zones. In other words, the segments identified as congested segments in this study are significantly close to the potential congestion zones in Taipei City.

4. Discussion

Urban street topology and flow complexity can reflect the spatial flow of urban activities, which can cause traffic congestion. Relying only on street network topology and a small fraction of segments with traffic volume, this study developed an iterative voting (sending scores) process in the FBPR algorithm to capture the topological connectivity structure

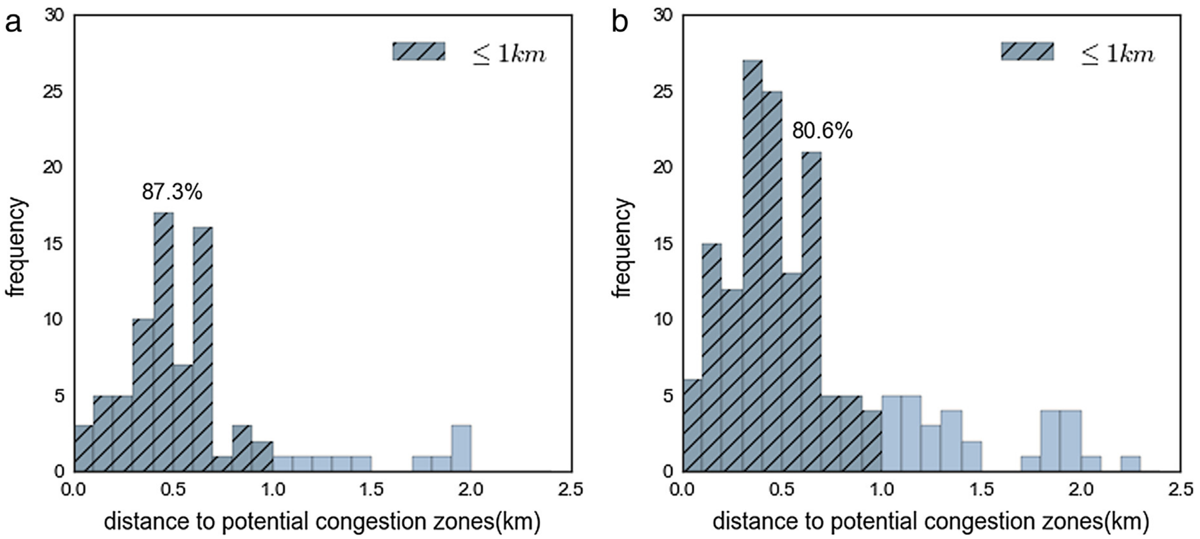


Fig. 10. Histogram of distances between the potential congestion zones and the identified congested segments: (a) the highest 5%; (b) the highest 10%, of segments with both FBPR scores and outgoing entropy values.

and determine the optimal link weights. The mechanism of the FBPR reflects the process of traffic flow; therefore, the FBPR scores represent the traffic demands of street segments. As shown in Fig. 9, our findings suggest that the spatial clustering areas with both high FBPR scores and high outgoing entropy are more concentrated at the city centers, which are near areas with the highest population densities (Fig. 2(a)). This finding is consistent with the findings from previous studies, which indicated that population size is a key indicator that explains the extent of urban congestion [29,30]. Moreover, the street segments with high FBPR scores tended to be concentrated in areas near important facilities, including major railway stations, MRT stations, significant landmarks and CBDs.

Previous studies used the network connectivity structures of street networks to investigate traffic congestion [11,23,31,32], suggesting that betweenness centrality is an indicator of the traffic situation of a street segment. The betweenness centrality of a node measures the number of shortest paths between all pairs of nodes that pass through the node. A node with high betweenness can be considered a bridge node within the network; thus, the betweenness centrality can capture bottlenecks of traffic flow. However, the betweenness centrality only measures the position of a node (street segment) in the network, and simplification of the street network restricts its ability to capture the patterns of traffic flow [23,31]. In this study, we have used the turning probability to quantify the strength of the tie between two street segments, which is based on the attractiveness of the target node relative to the other outgoing targets of the original node. Based on the concept of turning probability, the outgoing entropy is used to measure the variation in the turning probability from a street segment onto the outgoing links [33]. The higher the outgoing entropy, the more diverse the outgoing direction is. This finding implies that vehicles have similar probabilities for turning onto any of the outgoing segments. Because vehicles will slow down before making a turn, street segments with higher outgoing entropies will have slower moving speeds than segments with lower outgoing entropies. Hence, the street segments with high outgoing entropies might become bottlenecks of the traffic flow.

CBDs usually attract many people and vehicles, which can result in higher traffic demand and flow complexity. Previous studies suggested that vehicles tend to turn onto streets that provide more destination options [34]. Our FBPR algorithm can capture the flow concentration and identify the locations of places with multiple functionalities, showing that the turning probabilities of outgoing links in the CBDs are more homogeneous than those in the land used for leisure (Fig. 7(b)). This result suggests that the outgoing street segments could provide similar destination options in CBD areas. On the other hand, the turning probability is more concentrated on one or two outgoing links in the land used for leisure and transport. This situation appears in the outer ring of Taipei City, where the streets are designed for high moving speeds and for connecting the central city and external areas.

Predicting traffic flow and congestion has been an important topic for decades [35–44]. In this study, we argue that congested segments result from high traffic demands and the topological complexity of traffic flow. Therefore, the congested segments can be identified as those segments with both high FBPR scores and high outgoing entropy values. Although the results in Fig. 9 show that the congested segments might not be the slowest street segments, most are connected to segments with slow moving speeds, including the officially announced most congested streets and the historical records from VDs. Vehicles may be turning from one segment onto another, decreasing the moving speed and triggering a chain reaction of congestion. Hence, congested segments can be considered as possible sources of street congestion. In addition, the results also indicate that slow movement situations often occur at locations where the congested segments are spatially clustered. If vehicles keep turning onto slow moving street segments in the network and the number of vehicles in an area continues to increase, traffic congestion may occur. Therefore, slow movement scenarios become more serious if the slow segments are spatially close. In summary, only relying on the topological characteristics of streets, the proposed framework can identify congested segments that may be sources of traffic congestion.

Previous studies have found that transportation and commercial facilities with multiple functionalities, such as union stations and shopping districts, could attract large numbers of people and influence nearby traffic mobility [45,46]. Our results indicated that most of the identified congested segments are close to potential congestion zones, including important transportation facilities and landmarks, such as rail stations, shopping districts, hospitals, and night markets. Our flow-based ranking algorithm relies on analyzing street network topology and a small fraction of segments with traffic volume, and the algorithm successfully captured the characteristics of traffic demand and identified urban congestion-prone segments.

This study has several limitations. First, the street network was simplified as a planar network. Elevated roads and underground tunnels, which cross ground-level streets but are not connected to them, exist in Taipei City. The intersections between elevated roads or underground tunnels are farther apart; thus, these paths may connect distant street segments. The mechanisms and effects of these types of connections require further analysis. Second, the edge effect was observed in this study. Taipei City is geographically and topologically connected to the surrounding area (the New Taipei City). There are several bridges at the western boundary that cross rivers, immediately connecting Taipei City to a neighboring city with a high population density and different street network. On the other side of the city, several highways and tunnels also connect Taipei City to surrounding areas. Future studies may analyze the street networks and data in external areas to avoid the edge effect. Third, our methodology does not incorporate temporal variations in traffic flow. We did not specify whether congestion on segments occurred during rush or non-rush hours or on weekdays or weekends. This limitation may be resolved in future studies by aggregating temporal data in different temporal ranges (rush or non-rush hour periods) to identify the congested segments during rush or non-rush hours.

5. Conclusions

By analyzing the street network topological structure and calibrating the street segments based on the partial traffic volume data, this study proposed the FBPR algorithm to capture traffic demands and street complexity simultaneously. Traffic demands can be determined by FBPR scores, which capture human movements, and street complexity can be measured by the outgoing entropy, which represents the topological complexity in terms of turning probability. We also examined the association of urban land use types with traffic demand and street complexity. This study provides network topological insight to identify street-scale urban mobilities, which could be used for future planning and development of sustainable urban environments.

Acknowledgments

The research was supported by the grants of the Ministry of Science and Technology in Taiwan (MOST 105-2627-M-002-018; MOST 105-2410-H-002 -150-MY3). The funder had no role in the study design, data collection and analysis or in the preparation of the manuscript.

References

[1] W.S. Vickrey, Congestion theory and transport investment, Amer. Econ. Rev. 59 (2) (1969) 251–260. URL http://www.jstor.org/stable/1823678.

[2] B. Kerner, Congested traffic flow: Observations and theory, Transp. Res. Rec.: J. Transp. Res. Board 1678 (1999) 160–167. http://dx.doi.org/10.3141/1678-20. URL http://trrjournalonline.trb.org/doi/abs/10.3141/1678-20.

[3] J. Kwon, M. Mauch, P. Varaiya, Components of congestion: Delay from incidents, special events, lane closures, weather, potential ramp metering gain, and excess demand, Transp. Res. Rec.: J. Transp. Res. Board 1959 (2006) 84–91. http://dx.doi.org/10.3141/1959-10. URL http://trrjournalonline.trb.org/doi/abs/10.3141/1959-10.

[4] J. Lu, L. Cao, Congestion evaluation from traffic flow information based on fuzzy logic, in: 2003 IEEE Intelligent Transportation Systems, 2003. Proceedings, vol. 1, 2003, pp. 50–53. http://dx.doi.org/10.1109/ITSC.2003.1251919.

[5] P. Pongpaibool, P. Tangamchit, K. Noodwong, Evaluation of road traffic congestion using fuzzy techniques, in: TENCON 2007 - 2007 IEEE Region 10 Conference, 2007, pp. 1–4. http://dx.doi.org/10.1109/TENCON.2007.4429119.

[6] B. Coifman, Identifying the onset of congestion rapidly with existing traffic detectors, Transp. Res. A 37 (3) (2003) 277–291. http://dx.doi.org/10.1016/S0965-8564(02)00016-2. URL http://www.sciencedirect.com/science/article/pii/S0965856402000162.

[7] A.I. Gall, F.L. Hall, Distinguishing between incident congestion and recurrent congestion: A proposed logic, Transp. Res. Rec. (1232). https://trid.trb.org/view.aspx?id=308664.

[8] R. Bauza, J. Gozalvez, Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications, J. Netw. Comput. Appl. 36 (5) (2013) 1295–1307. http://dx.doi.org/10.1016/j.jnca.2012.02.007. URL http://www.sciencedirect.com/science/article/pii/S1084804512000628.

[9] A. Lakas, M. Shaqfa, Geocache: Sharing and exchanging road traffic information using peer-to-peer vehicular communication, in: Vehicular Technology Conference, VTC Spring, 2011 IEEE 73rd, 2011, pp. 1–7. http://dx.doi.org/10.1109/VETECS.2011.5956785.

[10] X. Yang, L. Liu, N.H. Vaidya, F. Zhao, A vehicle-to-vehicle communication protocol for cooperative collision warning, in: The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2004, pp. 114–123. http://dx.doi.org/10.1109/MOBIQ.2004.1331717.

[11] P. Holme, Congestion and centrality in traffic flow on complex networks, Adv. Complex Syst. 06 (02) (2003) 163–176. http://dx.doi.org/10.1142/S0219525903000803. URL http://www.worldscientific.com/doi/abs/10.1142/S0219525903000803.

[12] A. Mazloumian, N. Geroliminis, D. Helbing, The spatial variability of vehicle densities as determinant of urban network capacity, Phil. Trans. R. Soc. A 368 (1928) (2010) 4627–4647. http://dx.doi.org/10.1098/rsta.2010.0099. URL http://rsta.royalsocietypublishing.org/content/368/1928/4627.

[13] J.-C. Foltête, A. Piombini, Urban layout, landscape features and pedestrian usage, Landsc. Urban Plan. 81 (3) (2007) 225–234. http://dx.doi.org/10.1016/j.landurbplan.2006.12.001. URL http://www.sciencedirect.com/science/article/pii/S0169204606002696.

[14] M.K. Reilly, M.P. O'Mara, K.C. Seto, From Bangalore to the Bay Area: Comparing transportation and activity accessibility as drivers of urban growth, Landsc. Urban Plan. 92 (1) (2009) 24–33. http://dx.doi.org/10.1016/j.landurbplan.2009.02.001. URL http://www.sciencedirect.com/science/article/pii/S0169204609000188.

[15] S. Pallottino, M.G. Scutellà, Shortest path algorithms in transportation models: Classical and innovative aspects, in: P. Marcotte, S. Nguyen (Eds.), Equilibrium and Advanced Transportation Modelling, Centre for Research on Transportation, Springer, US, 1998, pp. 245–281. http://dx.doi.org/10.1007/978-1-4615-5757-9_11. URL http://link.springer.com/chapter/10.1007/978-1-4615-5757-9_11.

[16] W.-B. Du, Z.-X. Wu, K.-Q. Cai, Effective usage of shortest paths promotes transportation efficiency on scale-free networks, Physica A 392 (17) (2013) 3505–3512. http://dx.doi.org/10.1016/j.physa.2013.03.032. URL http://www.sciencedirect.com/science/article/pii/S0378437113002653.

[17] J. Aez, T. De La Barra, B. Pérez, Dual graph representation of transport networks, Transp. Res. B 30 (3) (1996) 209–216. http://dx.doi.org/10.1016/0191-2615(95)00024-0. URL http://www.sciencedirect.com/science/article/pii/0191261595000240.

[18] M.B. Hu, R. Jiang, Y.H. Wu, W.X. Wang, Q.S. Wu, Urban traffic from the perspective of dual graph, Eur. Phys. J. B 63 (1) (2008) 127–133. http://dx.doi.org/10.1140/epjb/e2008-00219-5. URL http://link.springer.com/article/10.1140/epjb/e2008-00219-5.

[19] B. Jiang, Ranking spaces for predicting human movement in an urban environment, Int. J. Geogr. Inf. Sci. 23 (7) (2009) 823–837. http://dx.doi.org/10.1080/13658810802022822.

[20] J.-P. Rodrigue, C. Comtois, B. Slack, The Geography of Transport Systems, third ed., Routledge, New York, 2013. URL https://people.hofstra.edu/geotrans/.

[21] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, Comput. Netw. ISDN Syst. 30 (1), (1998) pp. 107–117. http://dx.doi.org/10.1016/S0169-7552(98)00110-X. URL http://www.sciencedirect.com/science/article/pii/S016975529800110X.

[22] W.C.B. Chin, T.H. Wen, Geographically modified pagerank algorithms: identifying the spatial concentration of human movement in a geospatial network, PLoS One 10 (10) (2015) e0139509. http://dx.doi.org/10.1371/journal.pone.0139509. URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0139509.

[23] R. Puzis, Y. Altshuler, Y. Elovici, S. Bekhor, Y. Shiftan, A. Pentland, Augmented betweenness centrality for environmentally aware traffic monitoring in transportation networks, J. Intell. Transp. Syst. 17 (1) (2013) 91–105. http://dx.doi.org/10.1080/15472450.2012.716663. URL http://www.tandfonline.com/doi/abs/10.1080/15472450.2012.716663.

[24] J. Jassbi, P. Makvandi, M. Ataei, P.A.C. Sousa, Soft system modeling in transportation planning: Modeling trip flows based on the fuzzy inference system approach, Afr. J. Bus. Manag. 5 (2) (2011) 505–514. URL https://www.researchgate.net/publication/228898293_Soft_system_modeling_in_transportation_planning_Modeling_trip_flows_based_on_the_fuzzy_inference_system_approach.

[25] L. Yao, H. Guan, H. Yan, Trip generation model based on destination attractiveness, Tsinghua Sci. Technol. 13 (5) (2008) 632–635. http://dx.doi.org/10.1016/S1007-0214(08)70104-4.

[26] W. Xing, A. Ghorbani, Weighted pagerank algorithm, in: Second Annual Conference on Communication Networks and Services Research, 2004. Proceedings, 2004, pp. 305–314. http://dx.doi.org/10.1109/DNSR.2004.1344743.

[27] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (1948) 379–423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.

[28] G. Jenks, The data model concept in statistical mapping, The Data Model Concept in Statistical Mapping. International Yearbook of Cartography 7 (1967) 186–190.

[29] M. Barthelemy, A global take on congestion in urban areas, Environ. Plann. B: Plann. Des. 43 (5) (2016) 800–804. http://dx.doi.org/10.1177/0265813516649955. URL http://epb.sagepub.com/content/early/2016/05/13/0265813516649955.

[30] R. Louf, M. Barthelemy, How congestion shapes cities: from mobility patterns to scaling, Sci. Rep. 4 (2014) 5561. http://dx.doi.org/10.1038/srep05561. URL http://www.nature.com/articles/srep05561.

[31] S. Gao, Y. Wang, Y. Gao, Y. Liu, Understanding urban traffic-flow characteristics: A rethinking of betweenness centrality, Environ. Plann. B: Plann. Des. 40 (1) (2013) 135–153. http://dx.doi.org/10.1068/b38141. URL http://epb.sagepub.com/content/40/1/135.

[32] C. Sarkar, C. Webster, M. Pryor, D. Tang, S. Melbourne, X. Zhang, L. Jianzheng, Exploring associations between urban green, street design and walking: Results from the Greater London boroughs, Landsc. Urban Plan. 143 (2015) 112–125. http://dx.doi.org/10.1016/j.landurbplan.2015.06.013. URL http://www.sciencedirect.com/science/article/pii/S0169204615001383.

[33] M. Gambhir, S. Guerin, S. Kauffman, D. Kunkle, Steps toward a possible theory of organization, in: Proceedings of International Conference on Complex Systems, 2004, p. 9 URL http://necsi.org/events/iccs/openconf/author/papers/370.doc.

[34] J. Krumm, Where will they turn: predicting turn proportions at intersections, Pers. Ubiquitous Comput. 14 (7) (2010) 591–599. http://dx.doi.org/10.1007/s00779-009-0248-1.

[35] E. Castillo, J.M. Menéndez, S. Sánchez-Cambronero, Predicting traffic flow using Bayesian networks, Transp. Res. B 42 (5) (2008) 482–509. http://dx.doi.org/10.1016/j.trb.2007.10.003. URL http://www.sciencedirect.com/science/article/pii/S0191261507001300.

[36] M. Lippi, M. Bertini, P. Frasconi, Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning, IEEE Transp. Intell. Transp. Syst. 14 (2) (2013) 871–882. http://dx.doi.org/10.1109/TITS.2013.2247040.

[37] Y. Lv, Y. Duan, W. Kang, Z. Li, F.Y. Wang, Traffic flow prediction with big data: A deep learning approach, IEEE Transp. Intell. Transp. Syst. 16 (2) (2015) 865–873. http://dx.doi.org/10.1109/TITS.2014.2345663.

[38] B.L. Smith, B.M. Williams, R. Keith Oswald, Comparison of parametric and nonparametric models for traffic flow forecasting, Transp. Res. C 10 (4) (2002) 303–321. http://dx.doi.org/10.1016/S0968-090X(02)00009-8. URL http://www.sciencedirect.com/science/article/pii/S0968090X02000098.

[39] B.L. Smith, M.J. Demetsky, Shorterm traffic flow prediction: Neural network approach, Transp. Res. Rec. (1453). https://trid.trb.org/view.aspx?id=424677.

[40] A. Stathopoulos, M.G. Karlaftis, A multivariate state space approach for urban traffic flow modeling and prediction, Transp. Res. C 11 (2) (2003) 121–135. http://dx.doi.org/10.1016/S0968-090X(03)00004-4. URL http://www.sciencedirect.com/science/article/pii/S0968090X03000044.

[41] T.T. Tchrakian, B. Basu, M. O'Mahony, Real-time traffic flow forecasting using spectral analysis, IEEE Transp. Intell. Transp. Syst. 13 (2) (2012) 519–526. http://dx.doi.org/10.1109/TITS.2011.2174634.

[42] M. Van Der Voort, M. Dougherty, S. Watson, Combining Kohonen maps with ARIMA time series models to forecast traffic flow, Transp. Res. C 4 (5) (1996) 307–318. http://dx.doi.org/10.1016/S0968-090X(97)82903-8. URL http://www.sciencedirect.com/science/article/pii/S0968090X97829038.

[43] E.I. Vlahogianni, M.G. Karlaftis, J.C. Golias, Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach, Transp. Res. C 13 (3) (2005) 211–234. http://dx.doi.org/10.1016/j.trc.2005.04.007. URL http://www.sciencedirect.com/science/article/pii/S0968090X05000276.

[44] J. Whittaker, S. Garside, K. Lindveld, Tracking and predicting a network traffic process, Int. J. Forecast. 13 (1) (1997) 51–61. http://dx.doi.org/10.1016/S0169-2070(96)00700-5. URL http://www.sciencedirect.com/science/article/pii/S0169207096007005.

[45] R. Dunphy, K. Fisher, Transportation, congestion, and density: new insights, Transportation Research Record: Journal of the Transportation Research Board 1552 (1996) 89–96. http://dx.doi.org/10.3141/1552-12. URL http://trrjournalonline.trb.org/doi/abs/10.3141/1552-12.

[46] R. Cervero, K. Kockelman, Travel demand and the 3ds: Density, diversity, and design, Transp. Res. D 2 (3) (1997) 199–219. http://dx.doi.org/10.1016/S1361-9209(97)00009-6. URL http://www.sciencedirect.com/science/article/pii/S1361920997000096.