Hindawi
Journal of Advanced Transportation
Volume 2023, Article ID 2910454, 14 pages
https://doi.org/10.1155/2023/2910454

*Research Article*

# Expressway Usage Pattern Analysis Based on Tollgate Data: A Case Study of the Shandong Province, China

**Chengcheng Wang**,[1] **Xiaoe Ding**,[1] **Chao Wang**,[1] **Mengqi Lv**,[1] **Run Xu**,[1] **Yufeng Bi**,[1] **and Zhiren Huang**[2]

[1]*Shandong Provincial Communications Planning and Design Institute Group Co., Ltd, Jinan 250000, China*
[2]*Department of Computer Science, Aalto University, Espoo 00076, Finland*

Correspondence should be addressed to Zhiren Huang; zhiren.huang.cn@gmail.com

Received 7 March 2023; Revised 3 August 2023; Accepted 4 September 2023; Published 29 September 2023

Academic Editor: Francesco Galante



Expressway transportation is an essential part of regional development. An efficient expressway system can enhance cities' connectivity and coordinate long-distance trips between urban areas. Understanding how travel demand affects the flow of expressways is crucial to designing an efficient expressway management system. However, congested expressways are substantial obstacles to unimpeded expressway travel. Here, we explore the relationship between demand origin locations and congested expressways. We extract the time-varying OD demand matrix from empirical tollgate data collected in Shandong province, China. The incremental traffic assignment method is introduced to obtain the traffic flow of expressway road segments. It was found that congested expressways were generated due to only a few origin locations. In addition, expressway congestion during peak hours could be effectively alleviated by controlling the travel demand of these origin locations. Therefore, the proposed method can provide a novel perspective for expressway management.

## 1. Introduction

Expressway transportation is an indispensable part of the construction of the national economy. In recent years, with the continuous increase in traffic demand, expressway congestion conditions have worsened, leading to more fuel consumption, polluting gas emissions, and affecting travellers' experience [1, 2]. To reduce expressway traffic congestion, in the early stage, a new expressway was constructed to meet the growth of traffic demand. However, with the dense expressway network, intelligent management measures play a more important role in alleviating congestions [3, 4]. To alleviate the expressway congestion problem, the widely used traffic management strategies are ramp control [5, 6], variable speed-limit control [7, 8], mainline and ramp integrated control [9], dynamic route guidance [10], time-of-day pricing [11–13], and temporary opening of hard shoulders [14].

Tollgate station control is an easy-implement control strategy in the expressway system. Hong-jun et al. [15] proposed a traffic control system for toll stations to reduce fuel consumption and reduce congestion for vehicles going through toll stations. In recent years, the emergence of data mining technology [16] has provided more information about strategies for reducing congestion [17, 18]. In particular, by accumulating massive toll data generated by tollgate systems, accurate OD information and road traffic volumes can be calculated [19], which provides data support for research on tollgate station control strategies [20, 21]. Yang et al. [22] extracted travel data from the toll collection system of the freeway network in Shandong province of China to investigate the weather effect on intercity travel demand. Zheng et al. [23] proposed a new approach of inferring traffic flow on expressway networks using toll ticket data to obtain driving time and its variation, dwell time and its variation, and link flow. Fu et al. [24] developed a hybrid neural network for large-scale expressway network OD prediction based on toll data and delivered a better prediction performance. Chen et al. [25] presented a method to predict the exit station's traffic flow with different three scenes that are ETC, MTC and the mix of ETC, and MTC combining spatial-temporal matrix and long short-term memory model. Petrovic et al. [26] proposed a methodology based on a combination of recurrent neural networks, queuing theory, and metaheuristics to predict the optimal number of active modules in toll stations for continuous-time optimal control of expressway tolls. In terms of toll station management and control, some scholars [27] presented nonparametric regression models to predict the traffic volume of all stations periodically based on the analysis of both spatial and temporal business characteristics, while others [28] attempted to analyze the relevant factors affecting toll station safety through vehicle collision risk analysis. These research results have practical importance for the risk control and optimal management of toll stations.

Existing studies on expressway traffic management and control mainly focus on traffic flow analysis and strategies for solving single-point toll station congestion and mainline congestion [29]. There is a lack of analysis of highway road usage patterns, especially analyzing the balance between traffic demand and road supply and detecting driver sources [30]. It can combine traffic flow analysis and complex network theory to explore the causes of congestion, provide a theoretical basis for the traffic control strategy, and improve the effectiveness and flexibility of the strategy [31]. To explore road usage patterns, some studies have focused on traffic demand analysis and the equilibrium of road supply and demand. Wang et al. [32] used large-scale mobile phone data to obtain traffic demand and explore road usage patterns, and they found that there are several "driver sources" on road segments that provide most vehicles. Wang et al. [33] followed this research and proposed the concept of "dynamic driver sources," locating trip origin districts instead of residential districts as driver sources to support traffic control. These "driver sources" are built on residential districts and dynamic trip origins, both based on traffic flow analysis. However, Gong et al. [34] estimated the destination spatial distribution of vehicles in the research area and analyzed the characteristics of the "main traffic destinations." To analyze the main area causing the most congestion of the whole city, Li et al. [35] proposed an algorithm to identify congested road segments and construct congestion propagation graphs to model congestion propagation in urban road networks. Wang et al. [36] used radio frequency identification (RFID) data collected in Nanjing to estimate dynamic travel demand and developed an RFID data-based vehicle routing model that can be applied to a group of targeted vehicles only, providing more adaptive, efficient, and feasible routing strategies to mitigate traffic congestion. Toole et al. [37] combined multiple algorithms to generate a traffic demand matrix and constructed interactive visual web pages for the main sources and destinations of travel, showing the usage patterns of residential areas and roads. Chu et al. [38] extracted the OD tensor and then used a new deep learning model, Multi-Conv-LSTM, to predict future dynamic travel demand. Based on a large-scale New York real traffic dataset, the prediction model outperforms existing forecasting methods, and the study can help balance the supply and demand of roads and optimize on-demand transportation services. Demissie et al. [39] presented a methodology to estimate passenger demand for public transport services using cell phone data. Substantial origins and destinations of inhabitants were extracted and used to build origin-destination matrices that resemble travel demand. The outcome of this study can be useful for the development of policies that can potentially help fulfill the mobility needs of city inhabitants. From the perspective of complex network theory, Wang et al. [40] explored the spatial structure characteristics of intercity travel patterns during National Day, identifying aspects such as the "small world" phenomenon and a core-periphery radial structure. In addition, Kiashemshaki et al. [41] utilized ego-centric networks to model the travel patterns of Finnish cities, studying shifts in mobility patterns during the COVID-19 pandemic. Furthermore, Mimar et al. [42] applied network structure indicators, including centrality and PageRank, to evaluate welfare levels based on intercity mobility data.

This paper estimates the traffic demand of expressway systems based on empirical tollgate data, explores the relationship between traffic demand and the traffic flow of expressways, and constructs a bipartite network to analyze the driver source distribution. This paper calls this bipartite network expressway usage patterns. By identifying the sources of vehicles on high-volume expressways, a congestion mitigation strategy is proposed. Finally, a visualization system is developed to support transport management. The main contributions of this paper are summarized as follows:

(1) Based on real-world expressway tollgate data, we propose a road usage pattern analysis model to locate dynamic expressway driver sources for each segment, and a visualization system for transport management has been developed. The driver source analysis in this paper is based on accurate tollgate data, while previous studies were mainly based on mobile phone data [32, 33]. Therefore, our approach improves the accuracy and dynamics of driver source identification, making tollgate control strategies based on driver sources easier to implement in the closed expressway system.

(2) Based on the driver source information, a simple and effective congestion mitigation strategy is proposed. By controlling a few driver sources, efficient congestion mitigation is achieved. The results from this paper indicate that the major driver sources may be far from the target segment, which is crucial for improving tollgate control that is often close to the target segment.

## 2. Data and Methodology

In the present study, vehicle tollgate record data were used to obtain travel demand information. To identify driver sources contributing to traffic congestion, traffic demand should be assigned to road segments to construct the association between tollgate stations and road segments and to explore road usage patterns. The flowchart of the proposed method is shown in Figure 1, and there are three key steps in the main methodology (OD estimation, traffic assignment, and driver sources detection).

*2.1. Road Network Data.* The road network data in this paper were collected from the expressway road network in Shandong province, which is a coastal province in the Eastern China region with a population of more than 100 million. The total length of Shandong province expressway road network was more than 8,000 kilometres in 2022 [43], and the expressways were charged for different vehicle types. The road network data used in this paper are composed of 814 road nodes (i.e., grade separation and tollgate station) and 1766 road segments (Figure 2(a)). Due to the dense geographical distribution of interchanges and tollgates (where traffic volumes changed), the average length of road segments in Shandong province expressway road network was around 8.5 kilometers (Figure 2(b)). The properties of the segments contain road name, length, design speed, and number of lanes. The capacity $C$ of the road segment is calculated by the lane capacity of the road segment and the number of lanes. The lane capacity of road segment is obtained here from the Technical Standard of Highway Engineering, which is dedicated to the design and construction of expressways in China (Table 1). The free travel time $t^{(\text{free})}$ of each road segment is calculated by dividing the road length by the designed speed limit.

*2.2. Vehicle Tollgate Record Data.* This dataset was collected by the toll collection system (TCS). Tollgate stations are located at the junctions of urban road networks and expressway road networks, and the tollgate stations in this study are densely distributed (Figure 3(a)). A record is generated when a vehicle enters or exits the expressway road network through the tollgate station. Records in the provincial boundary were recorded by the gantry station provincial boundary where tollgates canceled. For data cleaning, records with abnormal time information and records without entrance or exit information were deleted. Around 1.5 million trips generated by one million vehicles were obtained in Shandong expressway system. Figure 3(b) illustrates the temporal trends in travel demand across the Shandong expressway system.

*2.3. Traffic Assignment.* Based on estimated travel demand, different methods of assigning trips to road networks can be utilized. The incremental traffic assignment (ITA) [44] method, which is suitable for large-scale road networks, can solve this problem by updating travel costs that consider travel congestion. As shown in Table 2, in the ITA method, original trips are usually first split into different subtrips that contain different percentages of the original trips. First, one of the subtrips is assigned to the network using the free travel time. Then, the actual travel time in a road segment is updated using the Bureau of Public Roads (BPR) function. Then, the next subtrips are assigned using the updated travel time, and the process is continued until all subtrips are assigned to the road network. Incremental traffic assignment method is an improvement of all-or-nothing assignment methods, and because of its convenience and the fact that it considers additional travel costs, it is often used in complex road networks with a large number of trips [31, 32].

*2.4. Driver Source Detection.* In Table4, to quantify the traffic flow contribution from tollgates to road segments, an array variable $S[x]$ ($x$ is the ID of a tollgate) is defined. If the trip route passes through road segment $x$, $S[x]$ plus one, then the traffic flow contribution from each tollgate to the segment is calculated by counting $S[x]$ from all paths in the OD matrices. Sorting and analyzing the traffic contributions of each tollgate to each segment can reflect the different congestion impacts of various tollgates. We define the top-ranked driver sources as driver sources providing the most traffic flow on a road segment as its major driver sources (the major driver sources in total produce 80% of a segment's traffic flow).

In Figure 4, the road usage bipartite network is generated in this paper to explore the relationship between major driver sources and road segments. In the road usage bipartite network, road segments and major driver sources are two types of nodes and links only exist in different types of nodes, i.e., each road segment relates to its major driver sources. The weight of the link is the traffic flow contribution from the driver source to the segment. In the modeling framework of a road usage bipartite network, the degree of a tollgate and the degree of a road segment were proposed [31]. The degree of a tollgate is the number of road segments for which the driver source is the major driver source, and the degree of a road segment is the number of major driver sources of the road segment [32].

*2.5. Comparison with Complex Network Indicators.* For a more thorough understanding of the system's behavior, we opted for complex network indicators to make a comparison to interpret and visualize the relationships and dependencies within the traffic network data [46, 47]. Specifically, closeness centrality [48] measures the average length of the shortest path from a node to all other nodes in the network, providing an indication of how long it will take to spread from that node to others; betweenness centrality [49] indicates how much a node acts as a bridge' between other nodes. Both closeness centrality and betweenness centrality can thus give us information about the most critical points in the traffic network, which could be of interest when planning interventions. The definition for driver source and other complex network indices of key nodes is as follows:

(a) Main driver source $\text{MD}(v)$:

If node $v$ is one major driver source of the segment, $\text{MD}(v)$ plus one.

(b) Production $\text{Pr}(v)$:

If node $v$ is origin of a trip, $\text{Pr}(v)$ plus one.

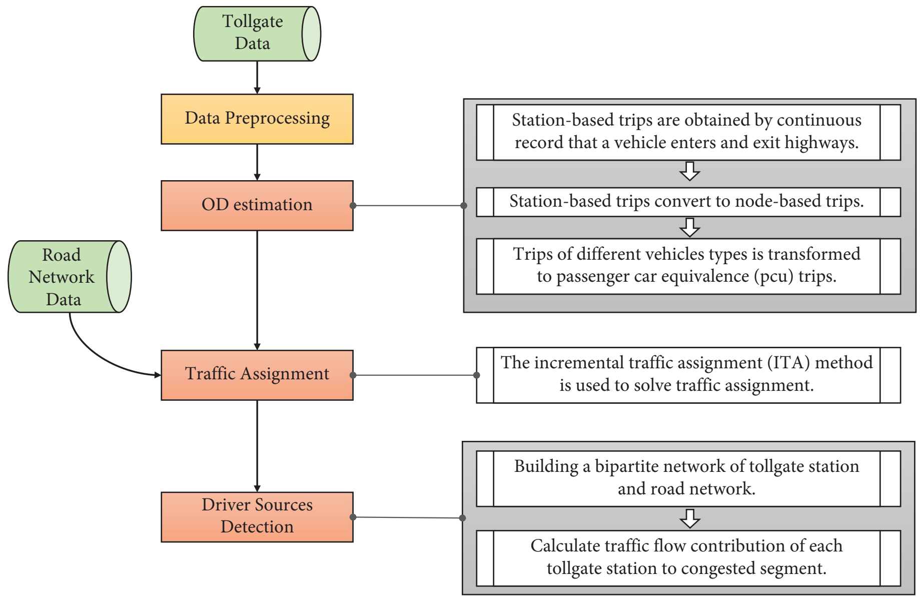
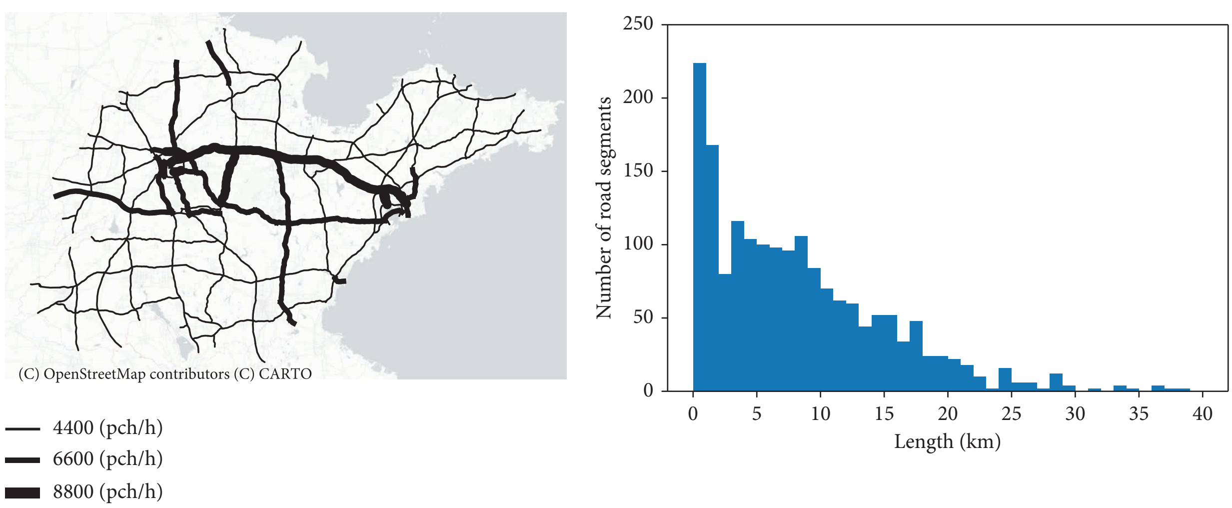



FIGURE 1: Flowchart of the proposed methodology.




(a) (b)

FIGURE 2: Statistics of the expressway network. (a) Map of expressway network. The width of black lines represents the capacity of road segments; (b) the length distribution of road segments.

TABLE 1: Capacity of expressway road lanes.

| Capacity of lane (pcu/h) | Design speed (km/h) |
|---|---|
| 2,200 | 120 |
| 2,100 | 100 |
| 2,000 | 80 |

(c) Closeness centrality:

$$\mathrm{CC}(v) = \frac{n-1}{\sum_u l(v,u)}, \tag{1}$$

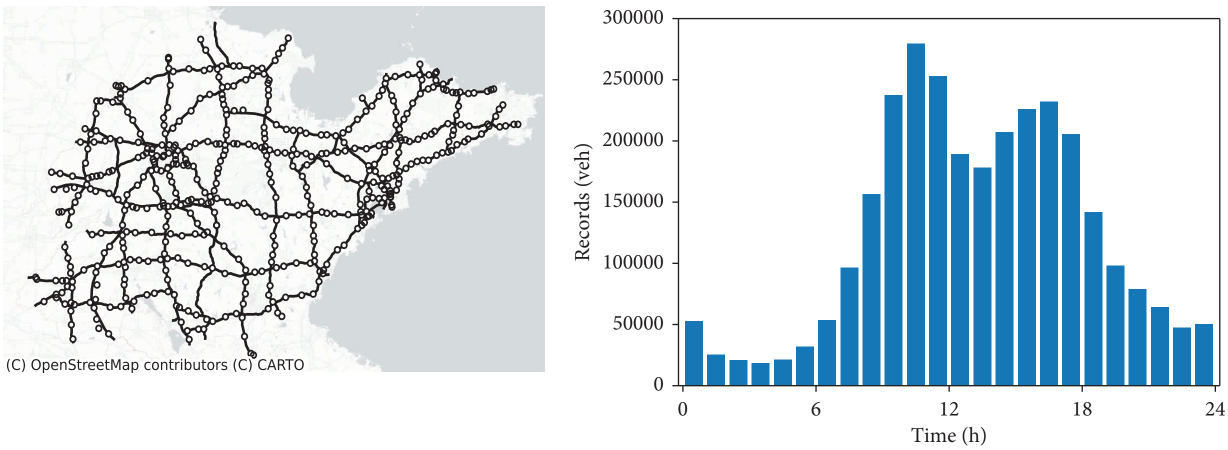
where $l(v,u)$ is the distance between node $u$ and node $v$, and $n$ is the number of nodes in the network.

(d) Betweenness centrality:

$$\mathrm{BC}(v) = \frac{\sum_{i,j\neq v}\left[P_v(i,j)/P(i,j)\right]}{(n-1)(n-2)/2}, \tag{2}$$

where $P(i,j)$ presents the number of shortest paths between node $i$ and node $j$. $P_v(i,j)$ presents the number of shortest paths between node $i$ and node $j$, but those shortest paths must pass through node $v$.

*2.6. Congestion Mitigation Strategy Based on Driver Sources.* As shown in Table 5, the road segments with the V/C greater than the threshold (0.75) are selected as the congested road segments. The driver sources of congested road segments are




(a) (b)

Figure 3: Statistics of tollgate records. (a) Tollgate locations map; (b) hourly recordings.

Table 2: Incremental traffic assignment algorithm. (The parameters are explained in Table 3)

*Input*:
Traffic network $G = (V, A)$, $\forall a \in A$;
Capacity of road segment $\{C_a\}$;
Free travel time of road segments $\{t_a^{\text{free}}\}$;
Traffic demand of OD pair $\{q_{\text{rs}}\}$, $(r, s) \in \text{RS}$;
Number of iterations $N$
*Output*:
Traffic flow of road segments $f_a, f_a^n (n = N)$;
Routes of each trip $L_{\text{rs}} = \{a_1, a_2, \ldots, a_i\}$;
Travel time of each road segments $t_a, t_a^n (n = N)$
*Steps*:
(1) Initialization
Set current iteration $n = 1$, the traffic flow of each road segment $f_a^0 = 0$, and each OD pair is divided into $N$ equal parts $q_{\text{rs}}^n = q_{\text{rs}}/N, n = 1 \ldots N$
(2) Update travel cost
Recalculate the travel time of road segment by
$$t_a^n = t_a^{\text{free}} (1 + \alpha (f_a^{n-1}/C_a)^{\beta}), \forall a$$
where $\alpha = 0.15$, $\beta = 4$ according to the Bureau of public roads (BPR) function [45]
(3) Incremental assignment
Assign $n$ th part OD by using all-or-nothing assignment based on $\{t_a^n\}$, the yields a flow pattern $\{w_a^n\}$. Update the traffic flow $f_a^n = f_a^{n-1} + w_a^n, \forall a$
(4) Stop condition
If $n = N$, stop the iteration; otherwise $n = n + 1$, go to step 1)

Table 3: Definition of variable.

| Variables | Definition | Unit |
|---|---|---|
| *Road network* | | |
| $C_a$ | The capacity of road segment $a$, calculated by the lane capacity of the road segment and the number of lanes | pcu/h |
| $t_a^{(\text{free})}$ | The free travel time of road segment $a$, calculated by dividing the road length by the designed speed limit | mins |
| $t_a$ | Travel time of road segment $a$ | mins |
| $f_a$ | Traffic flow of road segment $a$ | pcu |
| *Trips* | | |
| $q_{\text{rs}}$ | Traffic demand of OD pair from $r$ to $s$ | — |
| $L_{\text{rs}}$ | Routes of each trip from $r$ to $s$, $L_{\text{rs}} = \{a_1, a_2, \ldots, a_i\}$, $\{a_1, a_2, \ldots, a_i\}$ are road segments in the route | — |

Table 3: Continued.

| Variables | Definition | Unit |
|---|---|---|
| *Driver source* | | |
| $S_a^{x_m}$ | Contribution of driver source to segment $a$ | pcu |
| $\mathrm{MD}_a$ | Major driver sources of segment $a$, $D_a = \{x_1, x_2, \ldots, x_m\}$ , $\{x_1, x_2, \ldots, x_m\}$ are tollgates in major driver sources of segment $a$ | — |
| $\mathrm{MD}(v)$ | Main driver source in the road network, if node $v$ is one major driver source of a segment, $\mathrm{MD}(v)$ plus one, node $v$ is the road network node corresponding to a tollgate | — |
| *Complex network indicators* | | |
| $\mathrm{Pr}(v)$ | If node $v$ is origin of a trip, Pr $(v)$ plus one | — |
| $\mathrm{CC}(v)$ | Closeness centrality, $\mathrm{CC}(v) = (n-1)/(\sum_u l\ (v,u))$, where $l\ (v,u)$ is the distance between node $u$ and node $v$ | |
| $\mathrm{BC}(v)$ | Betweenness centrality, $\mathrm{BC}(v) = (\sum_{i,j\neq v}[P_v(i,j)/P(v,u)])/((n-1)(n-2)/2)$, where $P(v,u)$ presents the number of shortest paths between node $i$ and node $j$. $P_v(i,j)$ presents the number of shortest paths between node $i$ and node $j$, but those shortest paths must pass through node $v$ | |

Table 4: Driver source detection algorithm (the parameters are explained in Table 3).

*Input*:
Traffic network $G = (V, A)$, $\forall a \in A$;
Traffic demand of OD pair $\{q_{rs}\}$, $(r, s) \in RS$;
Trip number $N$
Tollgate $\{x_i\}$;
Routes of each trip $L_n = L_{rs} = \{a_1, a_2, \ldots, a_i\}$, $\forall n \in N$;
Traffic flow of road segments $f_a$;
*Output*:
Contribution of driver source $x_m$ to segment $a$, $S_a^{x_m}$;
Major driver sources of each segment $a$, $\mathrm{MD}_a = \{x_1, x_2, \ldots, x_m\}$
*Steps*:
(1) Contribution calculation
Set $S_a^{x_m} = 0$, $n = 1$;
Cycle the set of route segment of trip $n$, if trip $n$ goes through segment $a$, the contribution of trip origin $S_a^{x_m} = S_a^{x_m} + 1$, until $n = N$
(2) Contribution accumulation
Sort contribution from tollgate $x_m$ to segment $a$, $S_a^{x_m}$
Accumulate the contribution from the driver source with the largest contribution to the road segment $a$, until $\sum_{i=1,2,\ldots,j} S_a^{x_i}/f_a^N \geq 0.8$

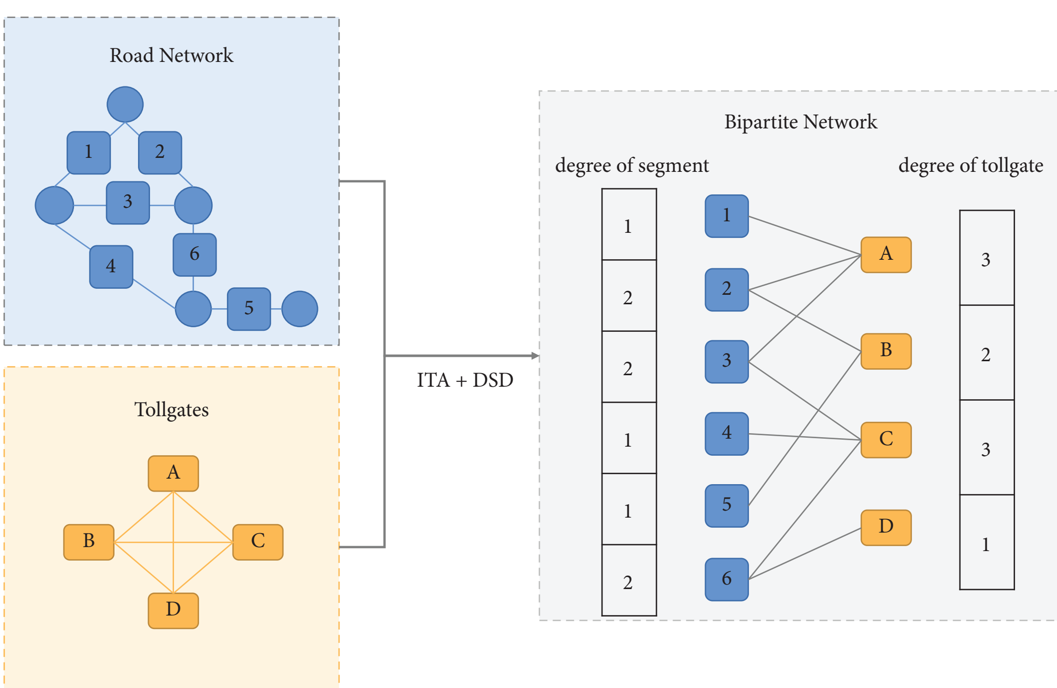



Figure 4: Bipartite network and degree of segment/tollgate.

Table 5: Congestion mitigation strategy based on driver sources (the parameters are explained in Table 3).

*Input*:
OD matrix;
Driver sources of each road segment;
Productions of tollgate $i$, $\mathrm{PR}(i)$;
V/C threshold $\mathrm{thr}_{V/C}$;
Number of controlled tollgates $N_n$
*Output*:
Strategy, controlled nodes with responding number of controlled OD $\{i\colon r_i \text{ for } i \text{ in } N_n\}$
*Steps*:
(1) Congested road segments selection:
Sort road segments according to V/C in descending order. Select the road segments with V/C larger than $\mathrm{thr}_{V/C}$;
(2) Sourcing:
Merge the driver sources of congested road segments into one set, the contributions accumulated;
(3) Selecting controlled nodes:
Sort the sourcing set according to V/C in descending order, and keep top $N_n$ sources
(4) Allocating deductions:
$w_i = (\mathrm{PR}(i)/\sum_i^{N_n}\mathrm{PR}(i))$
(5) Simulating OD deduction:
Remove $w_i$ trips randomly from all trips from source $i$
(6) Selecting controlled nodes:
Recalculated the flow of the road network

analyzed and extracted, which are sorted according to the V/C of the corresponding road segment to identify the driver sources that need to be controlled.

## 3. Results

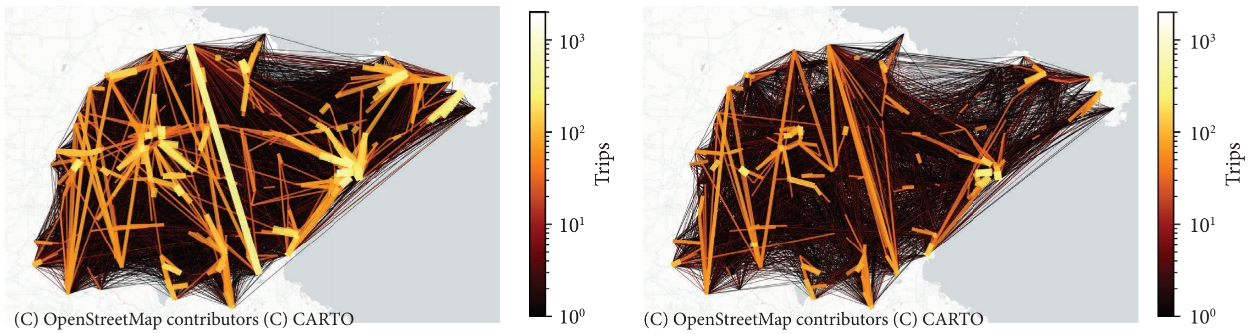
*3.1. Time-Varying OD Distributions.* The origin-destination pairs during a peak hour (10:00–11:00), and an off-peak hour (18:00 p.m.–19:00) were analyzed. As shown in Figure 5(a) and 5(b), the color and width of a link illustrate volume over trip volumes. The same spatial distribution patterns but different volume quantities in various time windows can be discovered. During peak hours, expressway trips are mainly concentrated around the city, and the trips are usually short to medium distances. Especially around Jinan, Qingdao, and Yantai, there is a large amount of short and medium distance travel. In long-distance trips, the amount of transit traffic passing through north of Binzhou to the south of Linyi is relatively large, which might be large transit trips. There is a similar geographical distribution during the off-peak hour, but the numbers are considerably lower. The total number of trips during the peak hour and the off-peak hour is 162,571 and 77,933, respectively, where the number of OD pairs with the highest number of trips is 1,025 and 593, respectively.

*3.2. Traffic Flow Distributions.* Traffic flows show different spatial distribution patterns during the two studied time windows. In Figure 6(a) and 6(b), traffic flows with the maximum flow reach 5,084 pcu/h in the peak hour and 2,631 pcu/h in the off-peak hour. During the peak hour, the roads with large traffic flows are mainly concentrated in the Jiguang expressway, Qingyin expressway, Qingxin expressway, Changshen expressway, Beijing-Shanghai expressway, and other expressways around Jinan, Qingdao, and Yantai. During the off-peak hour, the overall traffic flows are reduced, especially on the expressways around cities, where the traffic flows are most considerably reduced. However, there are still large traffic flows on the Jiguang expressway, Qingyin expressway, and Changshen expressway.

In this paper, saturation (V/C) and extra travel time are used as indices to measure the degree of road network congestion. The ratio of traffic flow $f$ and road capacity C was defined here as volume over capacity (V/C). A road segment with $\mathrm{V/C} > 0.75$ (service level is four) [50] was defined as a congested road segment. As Figure 7 demonstrates, our results show that few road segments (21 in the peak hour time window) had a V/C that was slightly larger than 0.75, with a maximum V/C of 0.874. For a road segment, the extra travel time (the difference between the actual travel time and free travel time) could also be used to measure its level of congestion [31]. If a driver travels through congested roads, he or she will experience a large amount of extra travel time. When comparing Figure 7(b) with Figure 7(a), the results show that the extra travel time caused by congestion increases substantially, and it is most obvious between 10:00 and 12:00.

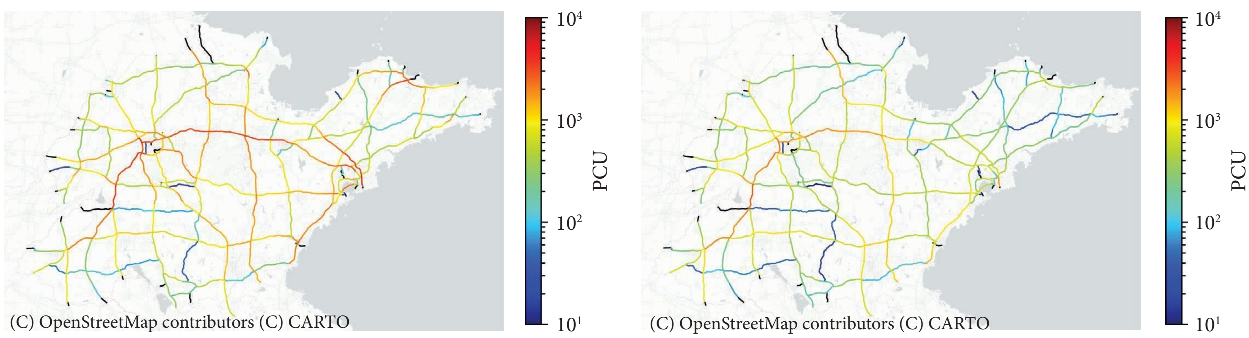
*3.3. Driver Source Analysis.* To further analyze the usage patterns of road sections, three types of road sections are selected as cases, i.e., congested road sections ($\mathrm{V/C} > 0.75$), high-traffic-flow sections (both flow and capacity are large), and ordinary cases (V/C~0.25), to analyze the distribution of driver sources under different road conditions. Figure 8 shows that the distribution of driver sources during peak hours is more concentrated.

By analyzing the difference between major driver sources and driver sources, we found that production sources are usually numerous and unstable (Figure 9), so our analysis of major driver sources is essential in practice. Because the number of major driver sources is relatively stable and small, it is convenient to implement control strategies based on these locations.



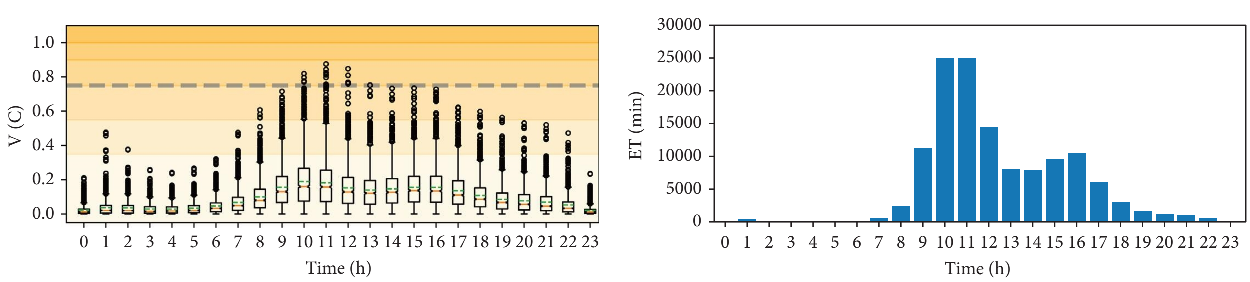

(a) (b)

Figure 5: Spatial distribution of trip volumes during a peak hour (10:00 a.m.–11:00 a.m.) and an off-peak hour (12:00 p.m.–1:00 p.m.). Color and width of a link illustrate volume over trip volumes.




(a) (b)

Figure 6: Spatial distribution of traffic flows during (a) a peak hour and (b) an off-peak hour.




(a) (b)

Figure 7: Traffic condition over time. (a) Hourly V/C; (b) hourly extra travel time (the total additional time spent by all drivers on the road network).

The degrees of driver sources (tollgate) and road segments can be approximated by an exponential distribution (Figure 10). The distributions of the degree of a driver source and the degree of a road segment reflect the road usage internal relations. First and useful for congestion mitigation, only a few driver sources provide the major usage of a road segment. Second, a similar and small number (5–15) of road segments were used by drivers from each driver source, which indicates that there is less influence on the tollgate control strategy.

Compared with other complex network indices, the spatial distribution of key nodes is very different, as shown in Figure 11. Distributing the key nodes in the road network from the network structure and trip distribution is more conducive to the implementation of emergency security and control strategies.

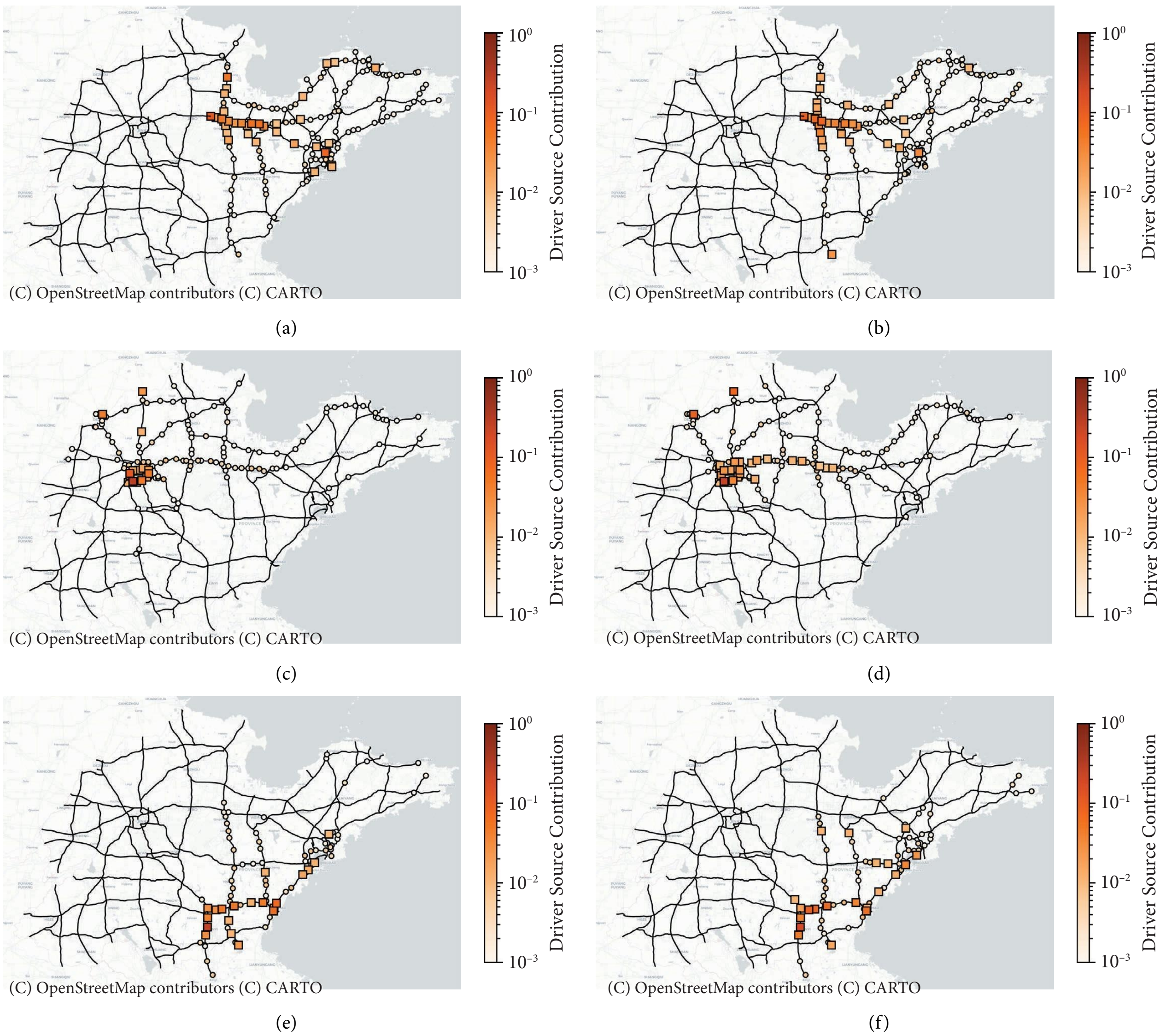



Figure 8: Driver source distributions. The major driver sources of the road segment are denoted by squares, the driver sources of the road segment are denoted by circles, and the colors represent the ratio of flow contribution. (a) Road 1 (10:00-11:00). (b) Road 1 (18:00-19:00). (c) Road 2 (10:00-11:00). (d) Road 2 (18:00-19:00). (e) Road 3 (10:00-11:00). (f) Road 3 (18:00-19:00).

## 4. Discussion

Based on controlling tollgates, the changes in extra travel time and congested road sections after reducing different percentages of OD volumes while controlling various percentages of tollgates are shown in Figure 11. The lighter color of the heatmap means a more obvious optimization effect, and the darker color means a less obvious optimization effect. The specific calculation of the heatmap parameters is as follows:

$$\frac{\left(C_{\text{after}} - C_{\text{before}}\right)}{C_{\text{before}}}, \tag{3}$$

where $C_{\text{after}}$ represents the extra travel time or congested road segments after the implementation of the strategy and $C_{\text{before}}$ represents the extra travel time or congested road segments before the strategy implementation.

Figure 12 shows that reducing OD volumes and controlling tollgates will effectively reduce expressway congestion. In general, the optimization effect becomes more obvious as the OD volume decreases, but the optimization effect does not become obvious as the proportion of controlled nodes increases, which is more effective at 1.5% of controlled tollgates. The best optimization is achieved with 1.5% controlled tollgates and a 3% reduction in OD volumes, which reduces the additional travel time by 35% and the number of congested road segments by 100%. The worst optimization is achieved with 2.5% of controlled tollgates and a reduction of 3% of OD volumes, which reduces the additional travel time by 5% and the number of congested road segments by 20%.

To compare the implementation effects of various strategies, Figures 13 and 14 show the changes in the total extra travel time and the number of congested road segments after

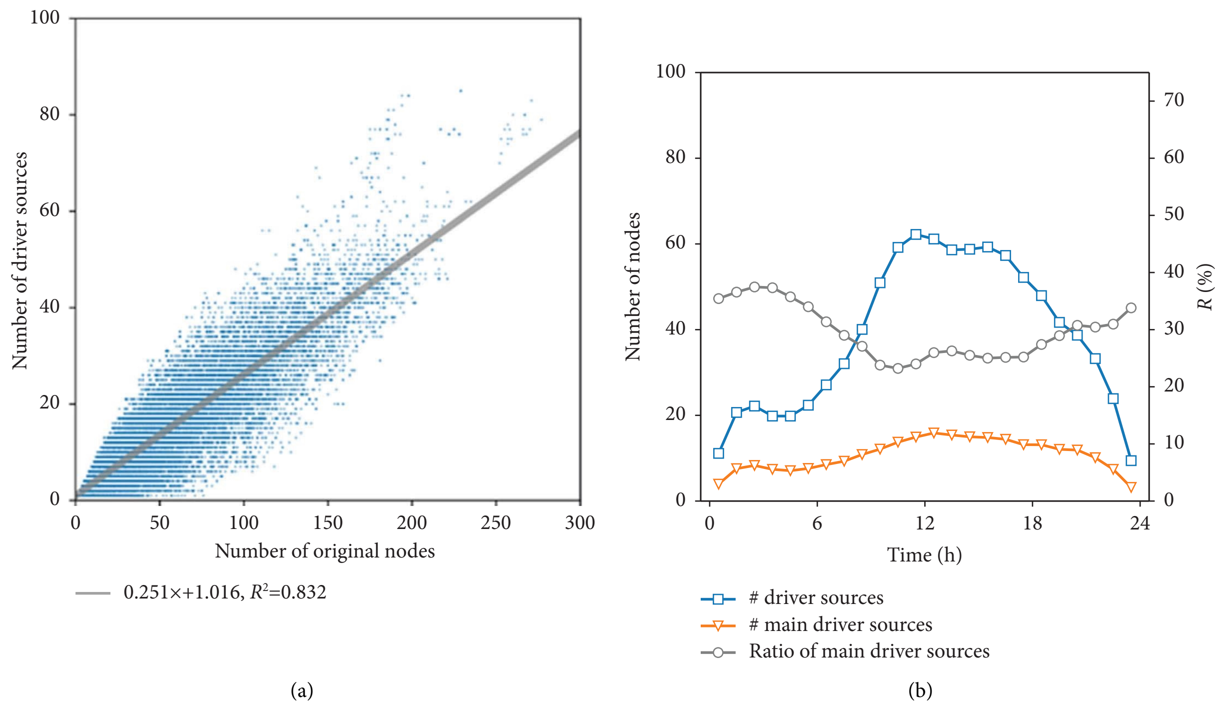



FIGURE 9: (a) The number of major driver sources and driver sources; (b) the change in the number of major driver sources and driver sources over time.

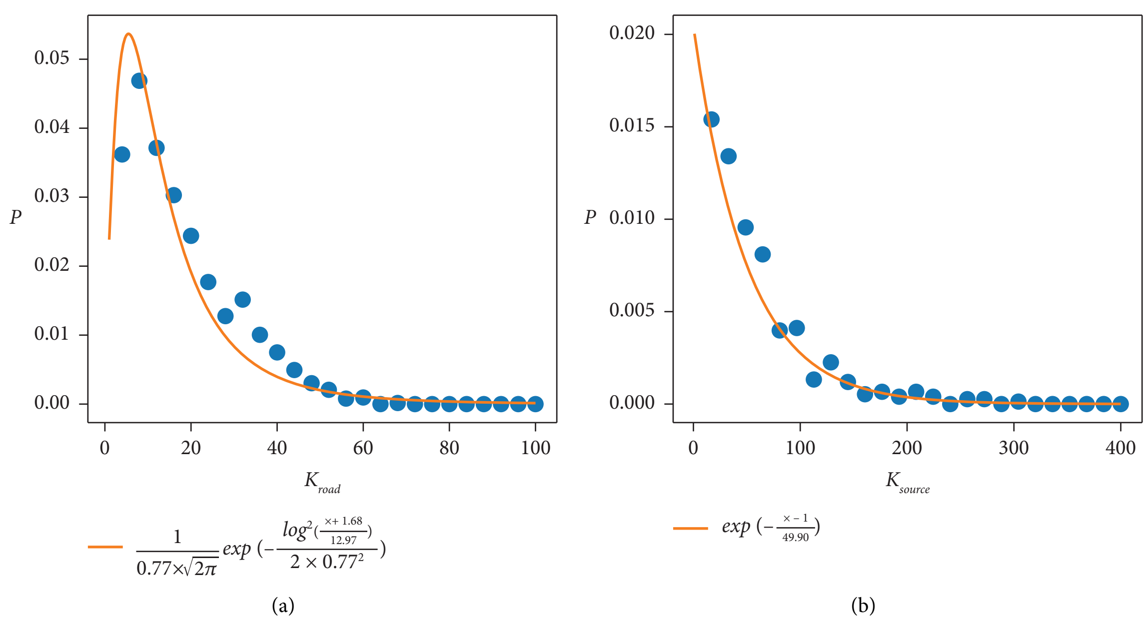



FIGURE 10: The degree distribution of (a) tollgates and (b) road segments.

the implementation of different strategies in two cases. Results show that controlling driver sources are much more effective than other strategies. When reducing OD volumes by 3% and controlling nodes by 1.5%, controlling driver sources can completely eliminate congested road segments. Even though the combination of control ratios is not effective, it can reduce the expressway congestion to some extent. Controlling production sources is better than the effect of other strategies, such as controlling closeness centrality. Note that the effect of controlling production, closeness, and betweenness is worse than randomly selected node control, indicating that finding the driver sources on congested road segments is essential for managing expressway congestion. Because the production sources are usually numerous and unstable, controlling driver sources can better reduce the flow of congested road segments to effectively decrease congestion.

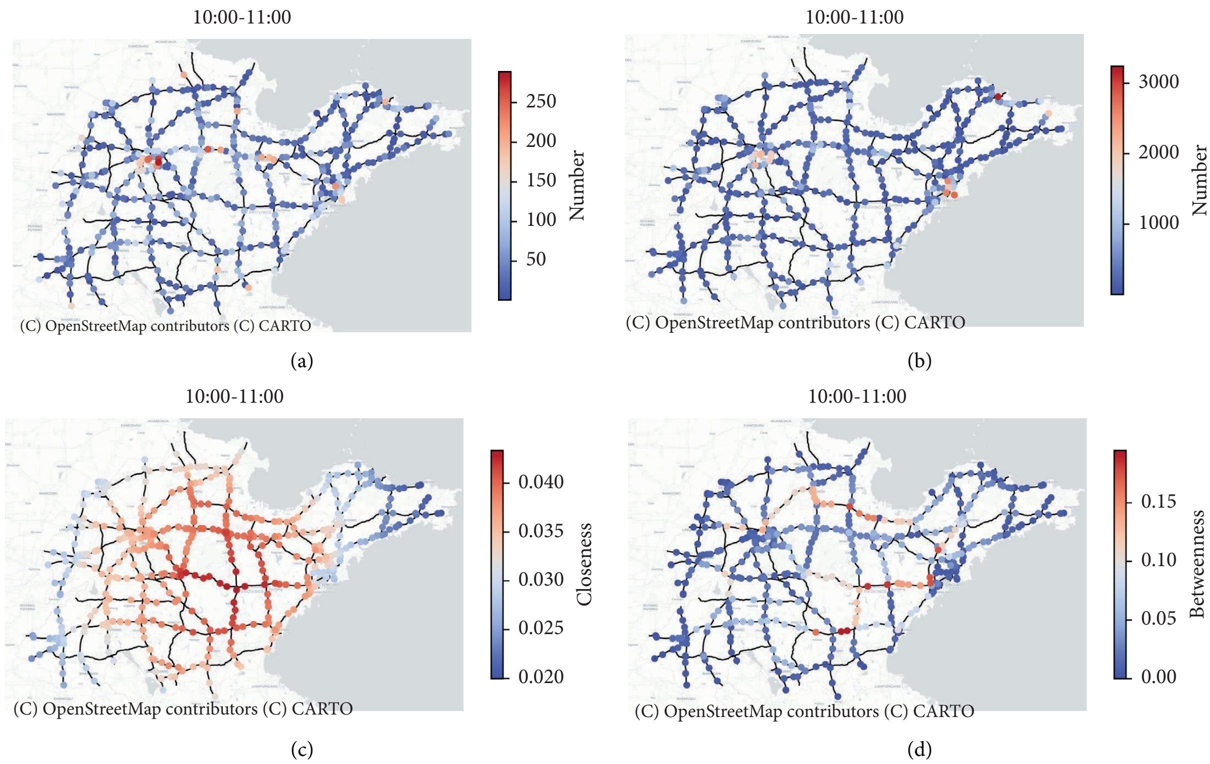



Figure 11: The spatial distribution of driver sources and key nodes for other complex network indices. (a) Main driver source. (b) Production. (c) Closeness centrality. (d) Betweenness centrality.

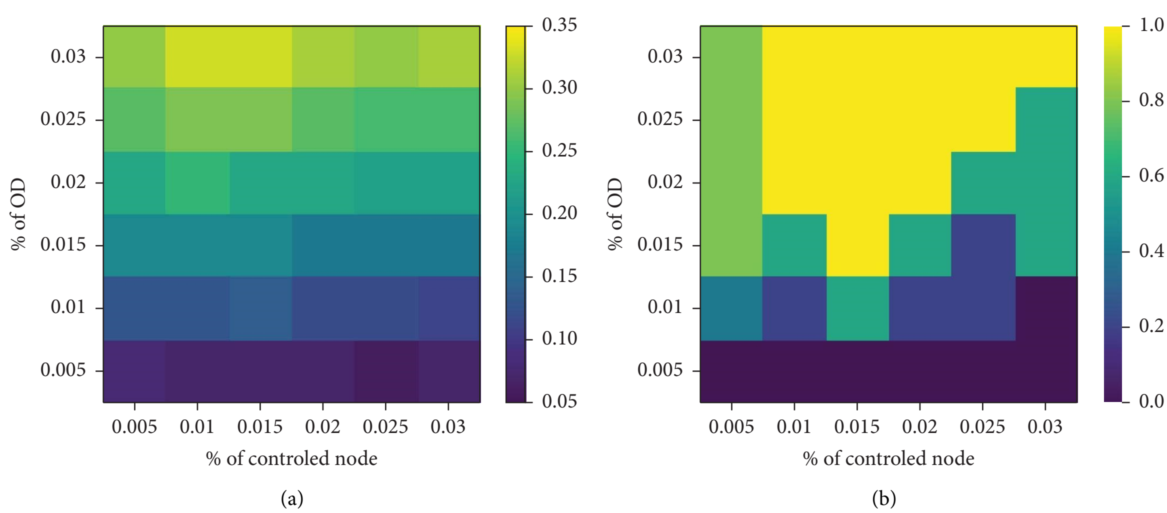



Figure 12: (a) The reduced proportion of extra travel time changes; (b) the proportion of congested road segments (V/C > 0.75) become uncongested road segments.

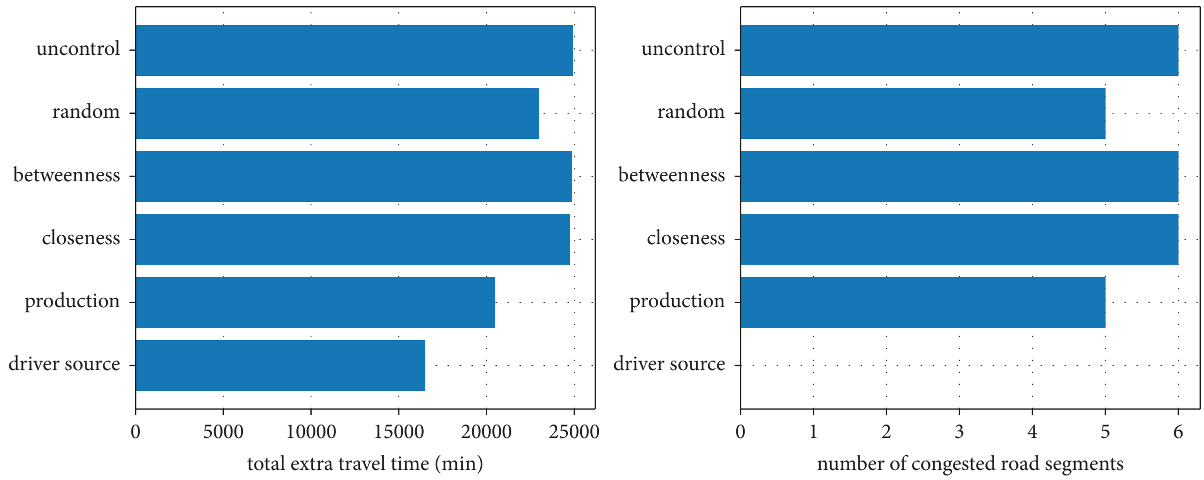



Figure 13: 3% of OD and 1.5% of controlled nodes.

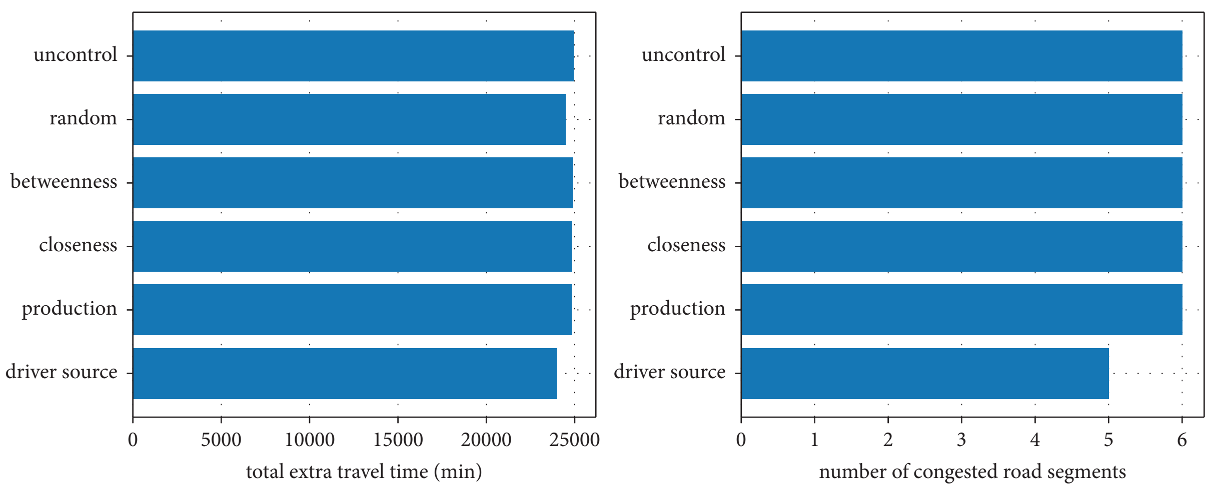



Figure 14: 0.5% of OD and 2.5% of controlled nodes.

## 5. Conclusions

This study provides valuable insights into the relationship between demand origins and congested expressways. By analyzing empirical data from Shandong province, China, this study can identify the specific origins that contribute to congestion on expressways. The findings suggest that only a small portion of driver sources are responsible for congestion and controlling the demand from these driver sources during peak hours can help resolve the problem. These results highlight the importance of considering road usage analysis when addressing congestion on expressways and can inform strategies for reducing congestion in the future. However, the current study only focuses on the expressway system. In the future, we can consider lower hierarchy roads to provide alternative routes [30, 51, 52] for congested driver sources. In practice, guidance information can be published via variable message signs [53]. The current simulated congestion schema did not consider fairness for expressway drivers. Expressway authorities can apply incentive strategies [54] to manage travel demand. Hence, this approach can provide a novel perspective for expressway management.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank Dr. Bao Guo for valuable discussions. This research was funded by the Natural Science Youth Foundation of Shandong (Grant no. ZR202103040494/

ZR2021QF137), Science and technology plan project of Shandong Provincial Department of Transportation (Grant no. 2022B62), and Key Science and Technology Project in Transportation Industry (Grant no. 2022-ZD6-077). The work of Z. Huang was supported by the Guangdong Science and Technology Strategic Innovation Fund (the Guangdong–Hong Kong-Macau Joint Laboratory Program), under Project no.: 2020B1212030009.

## References

[1] M. H. Askariyeh, M. Venugopal, H. Khreis, A. Birt, and J. Zietsman, "Near-road traffic-related air pollution: resuspended PM2.5 from highways and arterials," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, p. 2851, 2020.

[2] H. Yu, S. Koga, T. R. Oliveira, and M. Krstic, "Extremum seeking for traffic congestion control with a downstream bottleneck," *Journal of Dynamic Systems, Measurement, and Control*, vol. 143, no. 3, 2021.

[3] Y. Ren, M. Ercsey-Ravasz, P. Wang, M. C. González, and Z. Toroczkai, "Predicting commuter flows in spatial networks using a radiation model based on temporal ranges," *Nature Communications*, vol. 5, no. 1, pp. 5347–5349, 2014.

[4] Z. B. Li, P. Liu, C. C. Xu, and W. Wang, "Development of analytical procedure for selection of control measures to reduce congestions at various freeway bottlenecks," *Journal of Intelligent Transportation Systems*, vol. 22, no. 1, pp. 65–85, 2018.

[5] Y. S. Ci, H. L. Wu, Y. C. Sun, and L. N. Wu, "A prediction model with wavelet neural network optimized by the chicken swarm optimization for on-ramps metering of the urban expressway," *Journal of Intelligent Transportation Systems*, vol. 26, no. 3, pp. 356–365, May 2022.

[6] J. W. Zeng, Y. Qian, Z. Lv et al., "Expressway traffic flow under the combined bottleneck of accident and on-ramp in framework of Kerner's three-phase traffic theory," *Physica A: Statistical Mechanics and Its Applications*, vol. 574, Article ID 125918, 2021.

[7] J. R. D Frejo, I. Papamichail, M. Papageorgiou, and B. De Schutter, "Macroscopic modeling of variable speed limits on freeways," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 15–33, Mar 2019.

[8] P. Wang, Y. Zhang, S. Wang, L. Li, and X. Li, "Forecasting travel speed in the rainfall days to develop suitable variable speed limits control strategy for less driving risk," *Journal of Advanced Transportation*, vol. 2021, Article ID 6639559, 13 pages, 2021.

[9] X. Chen, T. Li, Z. Ma, and J. Sun, "Integrated mainline and ramp signal control for expressway on-ramp bottleneck with unequal lane-setting," *Journal of Intelligent Transportation Systems*, vol. 26, no. 1, pp. 100–115, 2022.

[10] J. Q. Wang and H. M. Niu, "A distributed dynamic route guidance approach based on short-term forecasts in cooperative infrastructure-vehicle systems," *Transportation Research Part D: Transport and Environment*, vol. 66, pp. 23–34, 2019.

[11] A. F. Lentzakis, R. Seshadri, A. Akkinepally, V. A. Vu, and M. Ben-Akiva, "Hierarchical density-based clustering methods for tolling zone definition and their impact on distance-based toll optimization," *Transportation Research Part C: Emerging Technologies*, vol. 118, Article ID 102685, 2020.

[12] L. Zheng, P. Liu, H. Huang, B. Ran, and Z. He, "Time-of-day pricing for toll roads under traffic demand uncertainties: a distributionally robust simulation-based optimization method," *Transportation Research Part C: Emerging Technologies*, vol. 144, Article ID 103894, 2022.

[13] L. H. Han, C. J. Zhu, D. Z. W. Wang, H. J. Sun, Z. J. Tan, and M. Meng, "Discrete-time dynamic road congestion pricing under stochastic user optimal principle," *Transportation Research Part E: Logistics and Transportation Review*, vol. 131, pp. 24–36, Nov 2019.

[14] K. Arora and L. Kattan, "Operational and safety impacts of integrated variable speed limit with dynamic hard shoulder running," *Journal of Intelligent Transportation Systems*, pp. 1–30, 2022.

[15] C. Hong-jun, Z. Zhi-lei, L. Xia, and Z. Min-qing, "Congestion flow control technology based on energy conservation and emission reduction at expressway toll stations," *China Journal of Highway and Transport*, vol. 28, no. 5, p. 125, 2015.

[16] W. Chen, F. Guo, and F.-Y. Wang, "A survey of traffic data visualization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2970–2984, 2015.

[17] Q. Shi and M. Abdel-Aty, "Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 380–394, 2015.

[18] G. Boquet, A. Morell, J. Serrano, and J. L. Vicario, "A variational autoencoder solution for road traffic forecasting systems: missing data imputation, dimension reduction, model selection and anomaly detection," *Transportation Research Part C: Emerging Technologies*, vol. 115, Article ID 102622, 2020.

[19] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions On Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2014.

[20] L. W. Chen and D. E. Chen, "Exploring spatiotemporal mobilities of highway traffic flows for precise travel time estimation and prediction based on electronic toll collection data," *Vehicular Communications*, vol. 30, Article ID 100356, 2021.

[21] S. Kim, H. Yu, and H. Yeo, "A study on travel time estimation of diverging traffic stream on highways based on timestamp data," *Journal of Advanced Transportation*, vol. 2021, Article ID 8846634, 13 pages, 2021.

[22] X. Yang, X. Yue, H. Sun, Z. Gao, and W. Wang, "Impact of weather on freeway origin-destination volume in China," *Transportation Research Part A: Policy and Practice*, vol. 143, pp. 30–47, 2021.

[23] K. Zheng, E. Yao, J. Zhang, and Y. Zhang, "Traffic flow estimation on the expressway network using toll ticket data," *IET Intelligent Transport Systems*, vol. 13, no. 5, pp. 886–895, 2019.

[24] X. Fu, H. Yang, C. Liu, J. Wang, and Y. Wang, "A hybrid neural network for large-scale expressway network OD prediction based on toll data," *PLoS One*, vol. 14, no. 5, Article ID e0217241, 2019.

[25] Z. Chen, B. Y. Wu, B. Li, and H. S. Ruan, "Expressway exit traffic flow prediction for ETC and MTC charging system based on entry traffic flows and LSTM model," *IEEE Access*, vol. 9, pp. 54613–54624, 2021.

[26] A. Petrovic, M. Nikolic, U. Bugaric, B. Delibasic, and P. Lio, "Controlling highway toll stations using deep learning, queuing theory, and differential evolution," *Engineering*

*Applications of Artificial Intelligence*, vol. 119, Article ID 105683, 2023.
[27] W. L. Ding, X. F. Wang, and Z. F. Zhao, "CO-STAR: a collaborative prediction service for short-term trends on continuous spatio-temporal data," *Future Generation Computer Systems*, vol. 102, pp. 481–493, 2020.
[28] L. Xing, J. He, M. Abdel-Aty, Q. Cai, Y. Li, and O. Zheng, "Examining traffic conflicts of up stream toll plaza area using vehicles' trajectory data," *Accident Analysis and Prevention*, vol. 125, pp. 174–187, 2019.
[29] H. Gan and X. Ye, "Whether to enter expressway or not? The impact of new variable message sign information," *Journal of Advanced Transportation*, vol. 49, no. 2, pp. 267–278, 2015.
[30] S. Çolak, A. Lima, and M. C. González, "Understanding congested travel in urban areas," *Nature Communications*, vol. 7, no. 1, pp. 10793–10798, 2016.
[31] C. Wang and P. Wang, "Data, methods, and applications of traffic source prediction," in *Transportation Analytics in the Era of Big*, S. Ukkusuri and C. Yang, Eds., pp. 105–120, Springer, Berlin, Germany, 2019.
[32] P. Wang, T. Hunter, A. M. Bayen, K. Schechtner, and M. C. González, "Understanding road usage patterns in urban areas," *Scientific Reports*, vol. 2, no. 1, pp. 1001–1006, 2012.
[33] J. Wang, D. Wei, K. He, H. Gong, and P. Wang, "Encapsulating urban traffic rhythms into road networks," *Scientific Reports*, vol. 4, no. 1, pp. 4141–4147, 2014.
[34] H. Gong, C. Wang, Y. Qu, K. He, and P. Wang, "Locating traffic information sources in urban areas," in *Proceedings of the 16th COTA International Conference of Transportation Professionals*, pp. 141–151, Shanghai, China, July 2016.
[35] C. L. Li, W. W. Yue, G. Q. Mao, and Z. G. Xu, "Congestion propagation based bottleneck identification in urban road networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4827–4841, 2020.
[36] C. Wang, Z. Xu, R. Du, H. Li, and P. Wang, "A vehicle routing model based on large-scale radio frequency identification data," *Journal of Intelligent Transportation Systems*, vol. 24, no. 2, pp. 142–155, 2020.
[37] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, "The path most traveled: travel demand estimation using big data resources," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 162–177, 2015.
[38] K.-F. Chu, A. Y. Lam, and V. O. Li, "Deep multi-scale convolutional LSTM network for travel demand and origin-destination predictions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3219–3232, 2020.
[39] M. G. Demissie, S. Phithakkitnukoon, T. Sukhvibul, F. Antunes, R. Gomes, and C. Bento, "Inferring passenger travel demand to improve urban mobility in developing countries using cell phone data: a case study of Senegal," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2466–2478, 2016.
[40] Z. Wang, W. Nie, and T. Lin, "Analysis of the spatio-temporal characteristics of intercity travel based on SVD and complex network: take Bohai Rim City Group as an example," in *Proceedings of the International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2021)*, pp. 357–365, Zhengzhou, China, November 2021.
[41] M. Kiashemshaki, Z. Huang, and J. Saramäki, "Mobility signatures: a tool for characterizing cities using intercity mobility flows," *Frontiers in big data*, vol. 5, p. 18, 2022.
[42] S. Mimar, D. Soriano-Paños, A. Kirkley et al., "Connecting intercity mobility with urban welfare," *PNAS Nexus*, vol. 1, no. 4, Article ID pgac178, 2022.
[43] Shandong, "Art," 2022, http://jtt.shandong.gov.cn/art/2022/12/31/art_15690_10307566.html.
[44] M. Chen and A. S. Alfa, "A network design algorithm using a stochastic incremental traffic assignment approach," *Transportation Science*, vol. 25, no. 3, pp. 215–224, 1991.
[45] J. de Dios Ortúzar and L. G. Willumsen, *Modelling Transport*, John Wiley and Sons, Hoboken, NJ, USA, 2011.
[46] A. Jayasinghe, K. Sano, and H. Nishiuchi, "Explaining traffic flow patterns using centrality measures," *International Journal of Traffic and Transportation Engineering*, vol. 5, no. 2, pp. 134–149, 2015.
[47] Y. Zhang, W. Xuesong, Z. Peng, and C. Xiaohong, "Centrality characteristics of road network patterns of traffic analysis zones," *Transportation Research Record*, vol. 2256, no. 1, pp. 16–24, 2011.
[48] A. Bavelas, "Communication patterns in task-oriented groups," *Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 725–730, 1950.
[49] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, pp. 35–41, 1977.
[50] L. Z. R. Zhou, *China Highway Capacity Manual*, China Communications Press Co., Ltd., China, 2017.
[51] K. He, Z. Xu, P. Wang, L. Deng, and L. Tu, "Congestion avoidance routing based on large-scale social signals," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2613–2626, 2016.
[52] P. Wang, C. Wang, J. Lai, Z. Huang, J. Ma, and Y. Mao, "Traffic Control Approach based on multi-source data fusion," *IET Intelligent Transport Systems*, vol. 13, no. 5, pp. 764–772, 2019.
[53] W. Zhao, M. Quddus, H. Huang, J. Lee, and Z. Ma, "Analyzing drivers' preferences and choices for the content and format of variable message signs (VMS)," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 1–14, 2019.
[54] N. Dogterom, D. Ettema, and M. Dijst, "Tradable credits for managing car travel: a review of empirical research and relevant behavioural approaches," *Transport Reviews*, vol. 37, no. 3, pp. 322–343, 2017.