



# Comparison of traffic reliability index with real traffic data

Limiao Zhang<sup>1,2</sup>, Guanwen Zeng<sup>1,2</sup>, Shengmin Guo<sup>3,4</sup>, Daqing Li<sup>1,2\*</sup> and Ziyou Gao<sup>5</sup>

\*Correspondence:

daqingl@buaa.edu.cn

<sup>1</sup>School of Reliability and Systems Engineering, Beihang University, Beijing, China

<sup>2</sup>Science and Technology on Reliability and Environmental Engineering Laboratory, Beijing, China

Full list of author information is available at the end of the article

## Abstract

Existing studies have developed different indices based on various approaches including network connectivity, delay time and flow capacity, estimating the traffic reliability states from different angles. However, these indices mainly estimate traffic reliability from single view and rarely consider the combined effect of city traffic dynamics and underlying network structure. Based on percolation theory, Li et al. has developed a traffic reliability index to address this issue (Proc. Natl. Acad. Sci. USA 112(3):669-672, 2015) [1]. Here we compare this percolation-based index with one of the well-known index - congestion delay index (CDI). Using real traffic data of Beijing and Shenzhen (two large cities in China), we compare the two indices in the macroscopic trends and microscopic extreme values. The two indices are found to indicate the state of real-time traffic reliability in different consideration. Our results can be used for better evaluation of traffic system reliability and mitigation measures of traffic jams.

**Keywords:** traffic reliability; reliability index; percolation theory; traffic data

## 1 Introduction

Given the rapid urbanization process and the sharp growth in travel demand, people spend more time on road, which has led much economic and environmental loss. A 2015 Texas Transportation Institute report found that U.S. commuters spend about 42 hours a year stuck in traffic congestion. The total nationwide price tag: \$160 billion, or \$960 per commuter [2]. Behind the staggering number is the increasing concern about traffic reliability.

Traffic reliability is a critical measure to assess the performance of transportation systems, especially under unexpected events [3]. Researchers have developed different types of traffic reliability indices with different considerations. Existing traffic reliability indicators include connectivity reliability, travel time reliability, capacity reliability, travel cost reliability, traffic flow recession reliability, traffic demand satisfaction reliability, user satisfaction reliability etc.

Connectivity reliability was firstly defined by Mine and Kawai [4] in 1982, which mainly reflects the connection probability between a random pair of nodes in road network. A given road segment in the road network are classified into two states: connected or disconnected. Further works supplement the theory by extending the definition of connectivity from two nodes to  $k$  nodes [5]. However, this measure of traffic reliability neglects the limitation of real-time flow, and mainly quantifies the ability of road static capacity.

Anthony Chen et al. [6, 7] proposed the concept of capacity reliability. Capacity reliability deals with the probability of a road network to meet traffic demand under certain service level. Lindley [8] developed an index based on peak hour traffic volume of urban highways. The index is calculated by comparing volume to capacity ( $V/C$ ), and roads with  $V/C$  higher than 0.77 are regarded as congested. Research at the Texas Transportation Institute [9] led to the development of the roadway congestion index (RCI) methodology to quantify the relative congestion levels in urban areas, which combines the indicator of urban area daily vehicle kilometers of travel (DVKT) per lane kilometer of roadway for both freeways and principal arterial streets.

Travel time reliability (TTR) is widely used to estimate the temporal damage caused by daily traffic congestion, which not only affects the daily travel of the public but also causes frustration among drivers [10–13]. TTR is defined as the probability of trips completed within a specified time between a given origin and destination (OD) at a certain level of service (LOS) [14]. TomTom International B.V. proposed a congestion index by comparing travel time in peak hours with travel time during non-congested periods (free flow) [15]. The difference is expressed as a percentage increase in travel time. Higher index indicates a longer delay in real-time compared with that in free-flow periods. The deformation of this index, congestion delay index (CDI) [15] that reflects the average delay time of real travel trajectories, is well-applied especially in China.

With the development of urban traffic and intelligence technology, there is a pressing need to estimate the real-time traffic performance from the system operator's viewpoint [16, 17]. However, existing reliability studies may not be sufficient for a comprehensive network performance measurement [7]. Most of traditional reliability analyses mainly focus on the effect of single fact on the performance of the network, neglecting the combined effect of traffic dynamics and network structure. Here we use a traffic reliability index based on percolation theory [1] to measure real-time traffic reliability in a comprehensive way. We compare this traffic reliability index  $q_c$  with congestion delay index, and analyze the performance of these two indices in large cities of China: Beijing and Shenzhen.

By constructing a traffic dynamical network, Li et al. found that the organization of city traffic could be considered as a percolation-like transition [1]. Percolation theory [18–20] is a useful tool to study network transition, providing a possibility to overcome the limitations mentioned above. In percolation process, different clusters form as failed nodes/edges are removed (due to congestion) from original network, during which the transition can be clearly identified between a well-connected global giant cluster and isolated local clusters. Percolation theory can present a systematic viewpoint to analyze the influence of localized jam on system. The transition of traffic network phase can be quantified by the probability threshold  $q_c$ , which can be taken as a statistical indicator of the operational limits of a network [21–23]. Therefore, the index  $q_c$  has three main advantages that we will discuss in the sections below: (a)  $q_c$  is the threshold distinguishing the network dynamics from connected global scale to isolated local scale; (b)  $q_c$  is less influenced by the trip sampling, faced with possible extreme local conditions; (c)  $q_c$  measures the relative traffic reliability from network operator's viewpoint, which is scalable for comparing different cities.

In Section 2, we describe the dataset of real traffic. In Section 3, we explain the construction of dynamical traffic network and definition of index. In Section 4, we compare the two

different indices in different aspects. The application in different cities is also illustrated in this part. Conclusions and discussion are presented in Section 5.

## 2 Data description

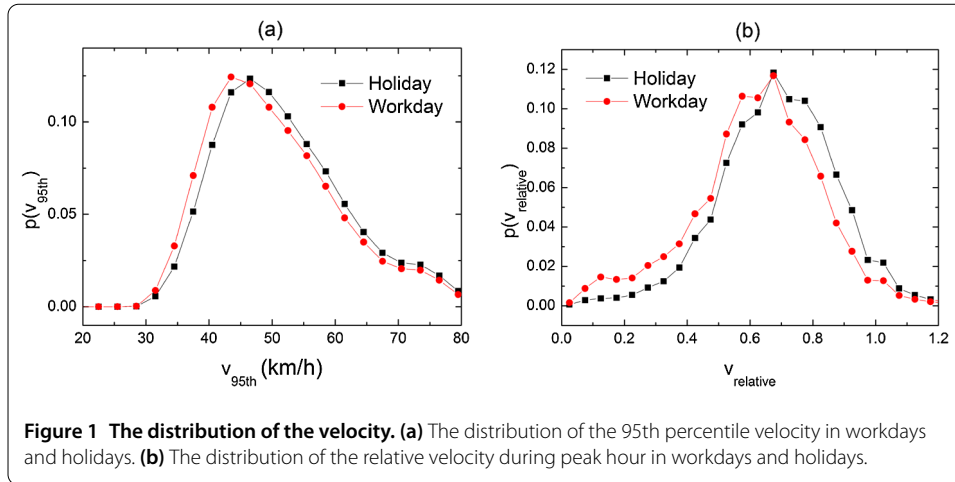
For the road network, intersections are represented by nodes and road segments between two intersections are represented by links. The road network of Beijing includes over 52,000 road segments (links) and 27,000 intersections (nodes). The road network of Shenzhen includes over 22,000 road segments (links) and 12,000 intersections (nodes). For each link, the velocity  $v_{ij}$  ( $i$  and  $j$  stands for the node on each end of the link respectively) is recorded according to real-time traffic. Here we consider a directed traffic network, because  $v_{ij}$  is in general different from  $v_{ji}$ . The dataset covers velocity records of roads in Beijing and Shenzhen for 30 days in October 2015, including a representative holiday period in China, the National Day, from Oct. 1st to Oct. 7th. Velocity (km/h) is recorded through floating cars, with a resolution of minute.

In order to estimate traffic state of Beijing, at least the information of 31,000 floating vehicles is needed [24]. Now about 100,000 floating cars were monitored. The number of sampled floating cars varies with time. There are around 64.38% of high-level roads having more than 5 vehicles records every 5 minutes, while the percent for low-level roads is 15.21%. We use an interactive-voting based map matching algorithm to associate the measured velocity to a given road, which is introduced by reference [25]. Our GPS data includes multiple types of floating cars with different resolutions. For taxis, the resolution of GPS data is 1 min or 30 s. For private cars, the resolution of GPS data is 1 s. The vehicle position error is less than 30 m. We compute the road velocity based on Dempster-Shafer theory [26], which includes a voting process. Each road is classified into one of three categories according to a pair of thresholds  $(v_1, v_2, v_1 < v_2)$  based on road levels. We compare the instantaneous velocities  $v$  of vehicles on a given road with its velocity thresholds. For  $v \in (0, v_1)$ , we vote for congested state. For  $v \in (v_1, v_2)$ , we vote for intermediate state. For  $v \in (v_2, \infty)$ , we vote for free-flow state. We regard the state with the highest votes as the real-time traffic state of this road. Then for this road, we smooth velocities within the voted category and calculate the road velocity. All sampled vehicles are pre-filtered to ensure their representativeness of road condition properly. The accuracy of data is greatly influenced by traffic lights. Our test results suggest that the accuracy of data is more than 85% on closed roads, while it is more than 70% on open roads.

The dataset is incomplete, with some velocities missing. We compensate the missing velocities by considering road network topology [1], where the missing velocity of road equals to the average velocity of its neighboring roads with the same direction. The data availability is constrained by our agreement with data provider of company.

## 3 Model

A traffic dynamical network is constructed based on both road network topology and traffic velocity data. Instead of directly using the absolute velocity of road segment, we take the relative velocity to illustrate real-time traffic road performance. For each road segment, we rank its velocity over one day in an increasing order and regard the 95th percentile of that as the limited maximal velocity for this road. The distribution of the 95th percentile velocity of all roads is shown in Figure 1(a). It can be observed that the distribution the 95th percentile velocity has a characteristic value around 45 (km/h) in workdays and holidays.

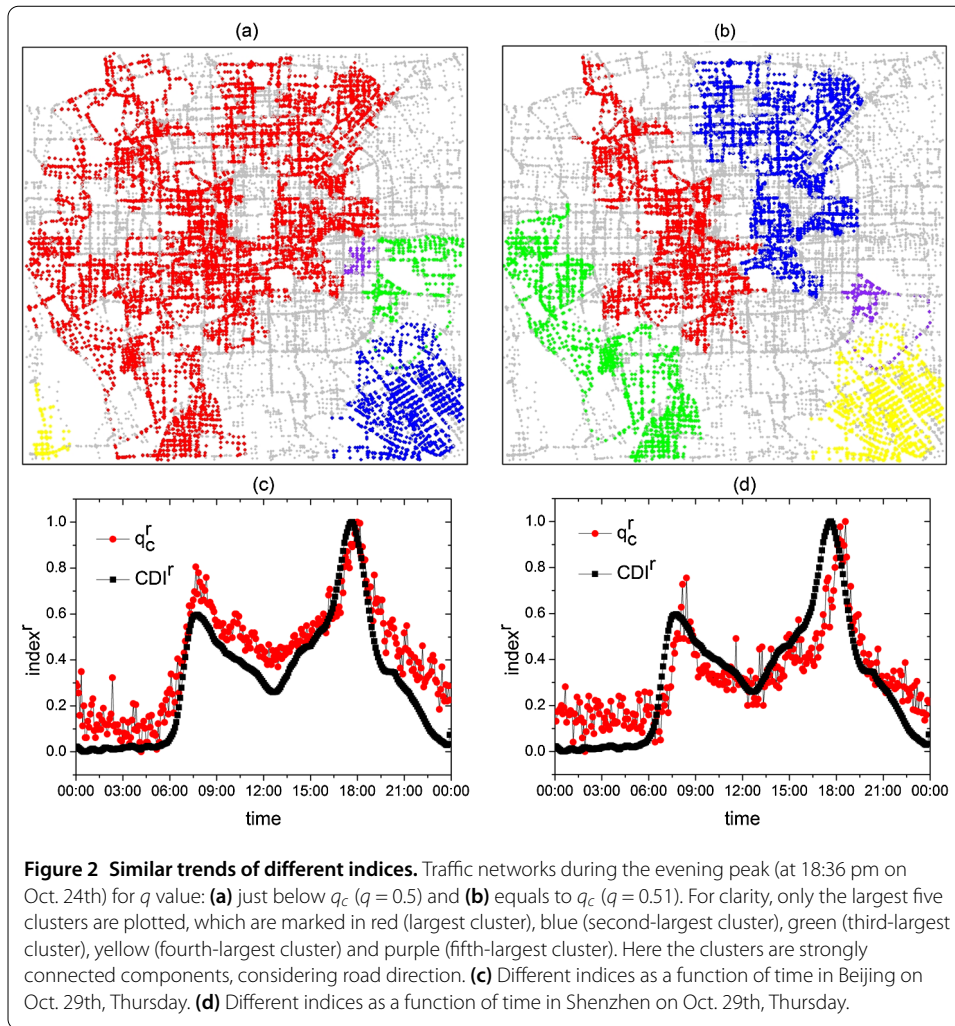


For each link, the ratio between its instantaneous velocity and the limited maximal velocity,  $r_{ij}$ , is its relative velocity. The distribution of the relative velocity during peak hour is shown in Figure 1(b), with a characteristic value around 0.65. Then a tunable percolation parameter  $q$  is defined to determine the state of road segments [1]. The state of each road segment  $e_{ij}$  with relative velocity  $r_{ij}$  will be classed into two cases: functional state for  $r_{ij} \geq q$  and congested for  $r_{ij} < q$ , i.e.

$$e_{ij} = \begin{cases} 1, & r_{ij} \geq q, \\ 0, & r_{ij} < q. \end{cases} \tag{1}$$

Then we remove all links with congested state and calculate strongly connected clusters in the rest of the network. A strongly connected cluster is a set of nodes, where there is a path in both directions between each pair of nodes [27]. In our calculation, we use Tarjan algorithm [28] to identify strongly connected clusters in the functional network.

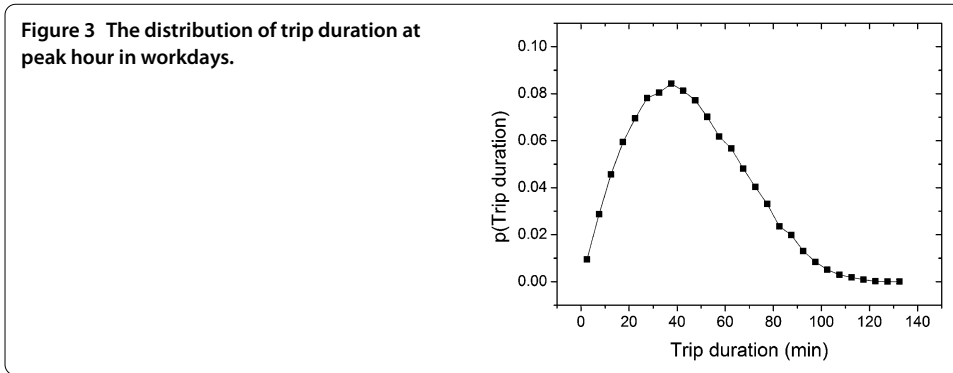
We increase  $q$  from 0 to 1 with an interval  $\Delta q = 0.01$ , representing users' increasing requirement of service level. With the increasing of  $q$ , more links are considered congested and higher proportion of links is removed. This leads to a decrease of the size of the giant component  $G$ . Meanwhile, the second-largest component,  $SG$ , increases and reaches its maximum when  $q$  equals to the critical threshold  $q_c$ , according to percolation theory. The critical threshold  $q_c$  (shown in Figure 2(a) and (b)) represents the robustness characteristics of the functional network connectivity [21], which also signifies the phase transition from free flow phase to congestion phase in the functional traffic network [1]. In the percolation theory, the failure of a node/edge of network is modeled by removal, and network connectivity transition is observed during the component failure process. Here in the traffic network, different from only considering structural information as before, we remove the roads whose velocity is under certain threshold ( $q$ ). In this way, the traffic dynamics is incorporated into the percolation framework. With the increase of the tunable percolation parameter  $q$ , which represents users' increasing requirement of service level, the removal of roads increases, and the network undergoes a transition from the phase of connectivity (free flow roads connected as a whole network) to the phase of disintegration. This transition is significantly influenced by the traffic dynamics and corresponding road velocity configuration. The probability threshold,  $q_c$ , is obtained by identifying this critical transi-



tion as a result of random process of road removal. Thus  $q_c$  can be different from a given sample, like the sample at the same time of different days.

Here we use  $q_c$  as an index [1] to measure the reliability level of city traffic: only cars with relative velocity below  $q_c$  can travel the main part of the city, i.e. the giant component of traffic network; otherwise, cars with relative velocity above  $q_c$  will be trapped in local isolated clusters. Therefore,  $q_c$  indicates the maximal relative velocity that allows one to travel the main part of the city, which reflects the global efficiency of traffic in a network view [1]. For comparison, we also take a widely applied index - congestion delay index (CDI) [15]. In practice, with advanced information processing technology, the trajectory can be precisely positioned on the map based on the data obtained from floating cars and GPS navigation. And the travel time of users can be obtained from GPS time stamp records [29]. In this paper, we use the same data set for both indices. Due to the lack of trip details in our data, we have to sample the origin and destination of trip from the traffic network. For simplicity and generality, 120,000 pairs of nodes are randomly selected as origin and destination (OD) separately.

We reconstruct the trips according to life experience: people usually choose the shortest way to the destination in real life. Then the shortest path between each pair of OD is



calculated. The shortest path between two nodes in directed network is a directed path, where the sum of the weights along its constituent edges is minimized. In traffic network, the weight of each edge is the length of corresponding road segment. We use Dijkstra algorithm to identify the shortest path between each pair of OD, and take each pair of nodes as sample trips to calculate CDI. Each trip has a different duration, which depends on the real-time traffic situation. The distribution of trip duration at peak hour is shown in Figure 3, most of which is distributed around 40 (min).

The travel time  $T_r$  for each trip is calculated according to recorded velocity. We select the reference moment  $t_f$  when the average speed of all links in the network is maximal. We then calculate free flow travel time  $T_f$  at reference moment  $t_f$ . Since our dataset covers velocity records of roads with resolution of minute,  $v_i(t)$  may change every minute. Each trip corresponds to a ratio between actual travel time  $T_r$  and free flow travel time  $T_f$ , and the average ratio of 120,000 trips is regarded as the CDI. We assigned the CDI index to the starting time of each trip, at which we compute corresponding  $q_c$ . Therefore, the CDI at  $t_r$  can be defined as:

$$CDI(t_r) = \frac{1}{S} \sum_S \frac{T_r}{T_f}, \tag{2}$$

where  $S$  is the total number of trips. For this index, larger value means people spend longer time on the way to the destination compared with free flow condition.

Because the two indices have different value ranges, in order to make a comparison between the two indices, we use relative values of indices:  $q_c^r$  and  $CDI^r$ . The relative value for each index can be obtained by:

$$q_c^r = \frac{q_c^{\max} - q_c}{q_c^{\max} - q_c^{\min}}, \tag{3}$$

$$CDI^r = \frac{CDI - CDI^{\min}}{CDI^{\max} - CDI^{\min}}, \tag{4}$$

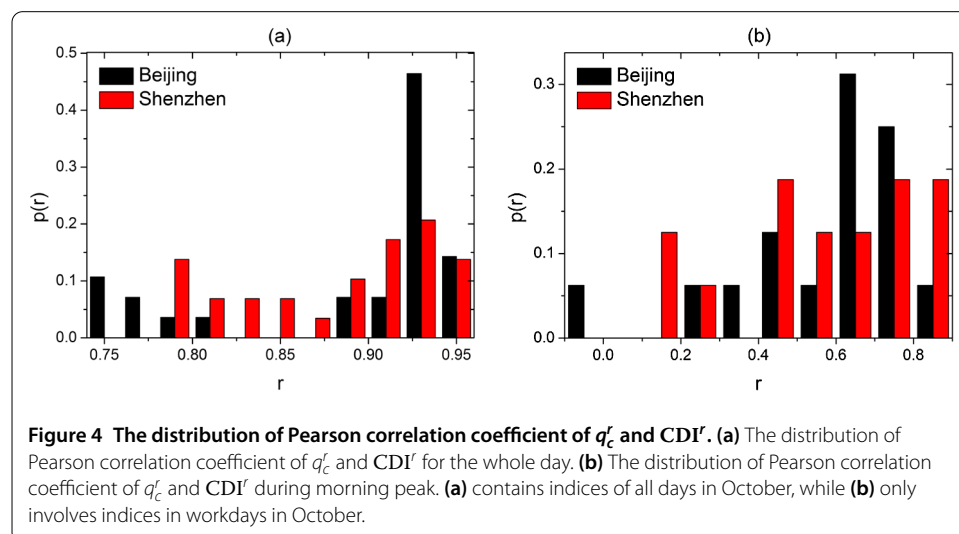
where  $q_c^{\max}$  and  $q_c^{\min}$  is the maximum and minimum of  $q_c$  from 00:00 to 24:00 in each day. The definition of  $CDI^{\max}$  and  $CDI^{\min}$  is similar to  $q_c^{\max}$  and  $q_c^{\min}$ . In this case, both indices range from 0 to 1. In the following sections, we will use these two relative indices:  $q_c^r$  and  $CDI^r$ , unless noted elsewhere.

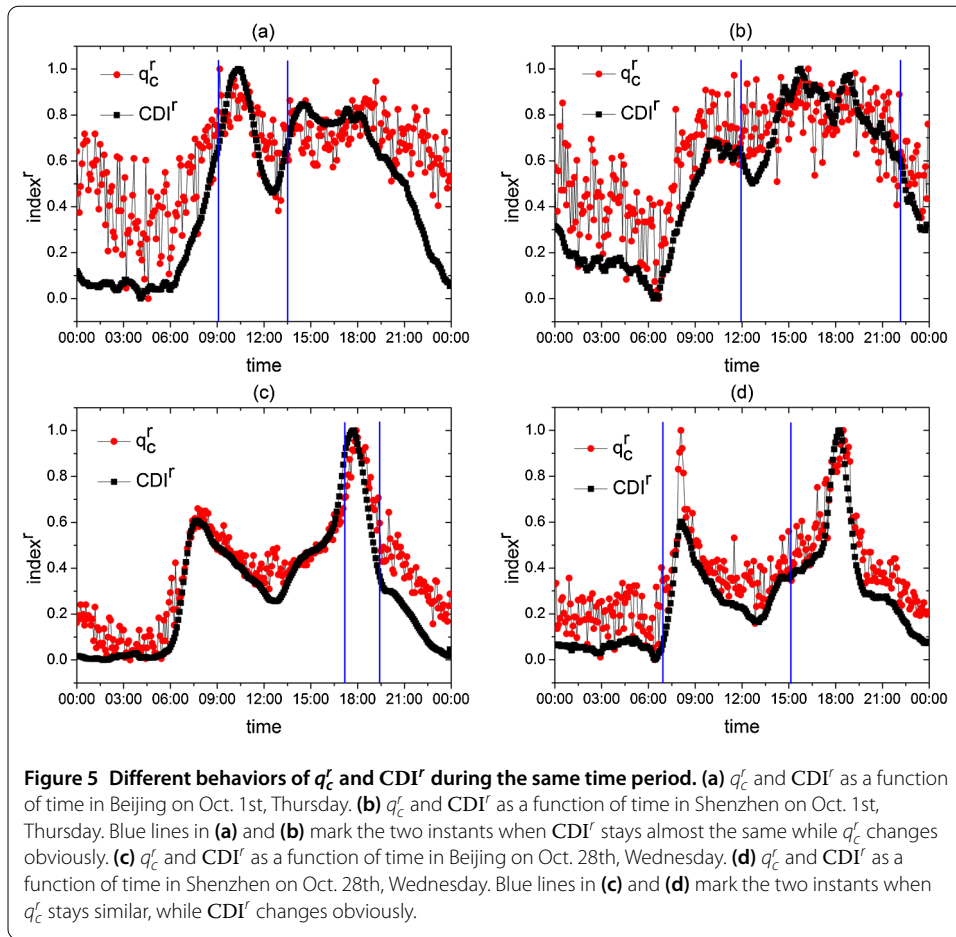
#### 4 Results

We calculate the two indices in Beijing and Shenzhen for a month and compare trends of different indices. We find the trends of the two indices are similar, while the degree of congestion reflected by different indices is different from each other. Figure 2(c) and (d) show examples that the trends of the two indices are similar. For Beijing on Oct. 29th (see Figure 2(c)), both of the two indices increase in the morning and reach to morning peaks around 7:30. Then indices decrease to relative low values, which correspond to a better traffic condition during noon. Around 18:00, evening peaks appear and afterwards both indices begin to decrease. The same curve trends can be also found in Shenzhen, as shown in Figure 2(d). Since CDI has already been widely used to measure traffic state, similar trends reflected by the two indices better illustrate the basic indicative function of  $q_c^r$ .

This similar trend is also confirmed by the correlation analysis between the two indices. We calculate the Pearson correlation coefficient  $r$  of the two indices for the whole day or during morning peak (morning peak  $\pm 0.5$  hour), respectively. The distribution of  $r$  is shown in Figure 4. Pearson correlation coefficient  $r$  is a measure of linear correlation between two processes.  $r$  ranges from  $-1$  to  $1$ . If  $r > 0$ , two variables are positively correlated; if  $r < 0$ , two variables are negatively correlated. The larger the absolute value of  $r$  is, the stronger correlation between two variables is. As shown in Figure 4(a), a strong positive correlation between the two indices has been observed. This further illustrates their similarity in indicative function. Although  $CDI^r$  and  $q_c^r$  shows a strong correlation for the whole day, they seems less correlated with each other during morning peak in workdays. As shown in Figure 4(b), although most of  $r$  distributes close to  $0.8$ , there are values even range from  $-0.25$  to  $0.5$ . In Beijing, the average Pearson correlation  $\langle r \rangle$  is  $0.576$  during morning peak, while  $\langle r \rangle$  is  $0.887$  for the whole day. In Shenzhen,  $\langle r \rangle$  is  $0.559$  during morning peak and  $0.882$  for the whole day. This also proves that the two indices display different behaviors during peak hours, while sharing similar trends in the whole day. The differences are important because people care about the specific traffic condition especially during peak hour.

More differences lie in the degree of congestion in details. According to  $CDI^r$ , as shown in Figure 5(a) and Table 1, the road condition in Beijing at 9:10 and 13:35 are similar. How-



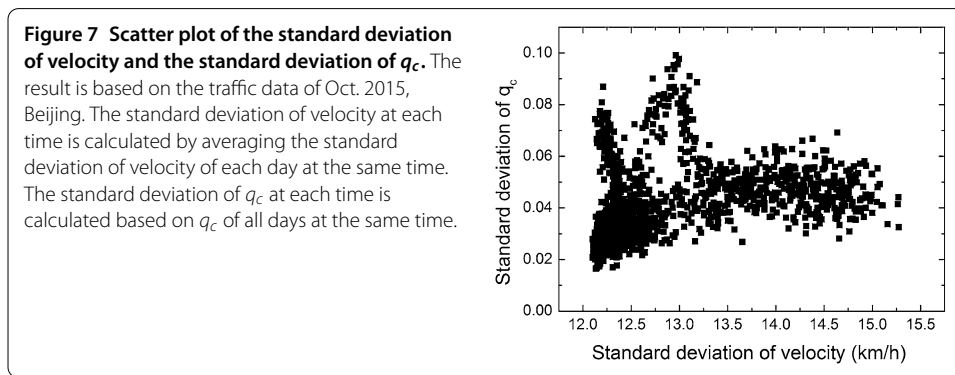
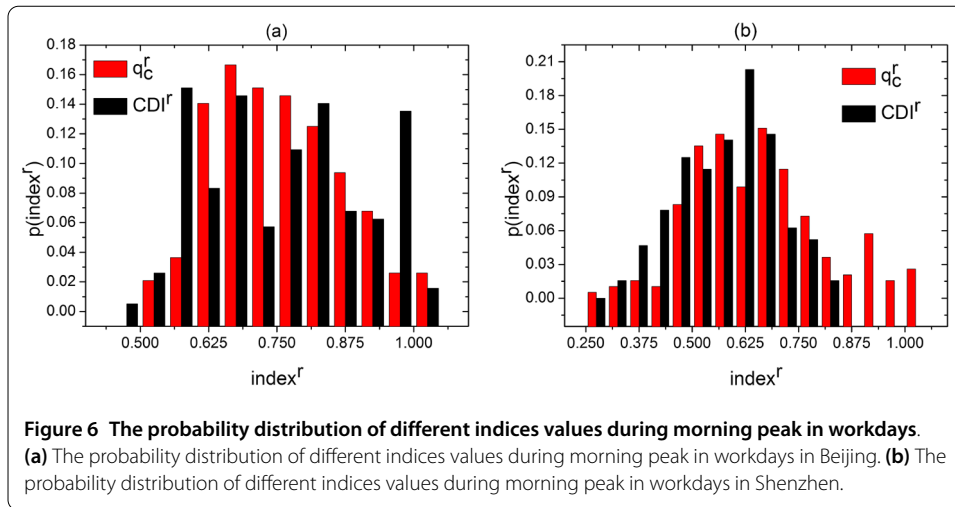


**Table 1** Example of differences in indices values

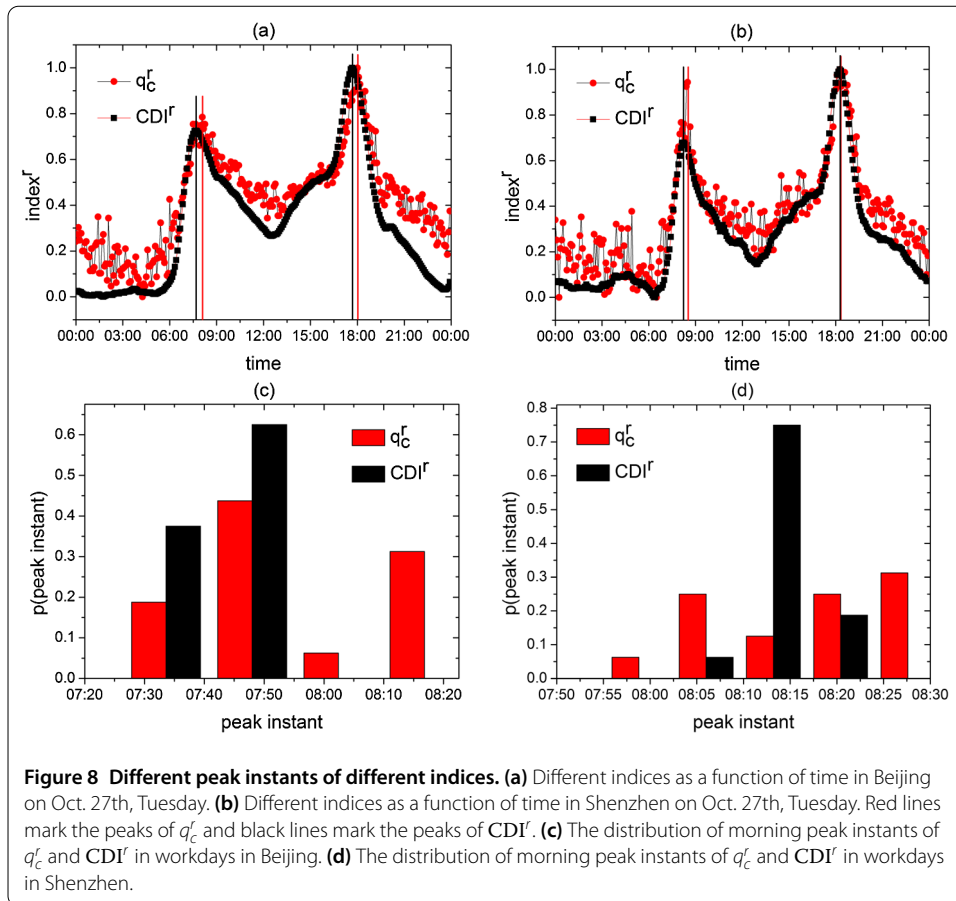
Date	Index	Time1	Value1	Time2	Value2	Change	
Beijing	$CDI^r$	9:10	0.69	13:35	0.69	0	
		$q_c^r$	1		0.6	-40%	
	Oct. 28th	$CDI^r$	17:10	0.89	19:20	0.39	-56.2%
		$q_c^r$		0.71		0.71	0
Shenzhen	Oct. 1st	$CDI^r$	11:50	0.65	22:05	0.61	6.2%
		$q_c^r$		0.58		0.89	53.4%
	Oct. 28th	$CDI^r$	6:55	0.07	15:10	0.37	427.6%
		$q_c^r$		0.34		0.34	0

ever,  $q_c^r$  changes from 1 to 0.6 during this period, indicating a 40% relief of traffic congestion. The same situation can be observed in Shenzhen, as shown in Figure 5(b). At 11:50 and 22:05,  $CDI^r$  shows that the traffic reliability levels at the two instants are similar, while  $q_c^r$  increases more than 53%. There are also situations that  $q_c^r$  stays stable, while  $CDI^r$  changes obviously between two instants. Figure 5(c) and (d) show this difference in Beijing and Shenzhen, separately. In Figure 5(c),  $q_c^r$  is the same at 17:10 and 19:20, however,  $CDI^r$  decreases from 0.89 to about 0.39. The same situation can be observed from examples in Shenzhen, as shown in Figure 5(d). Table 1 shows detailed indices values mentioned above.



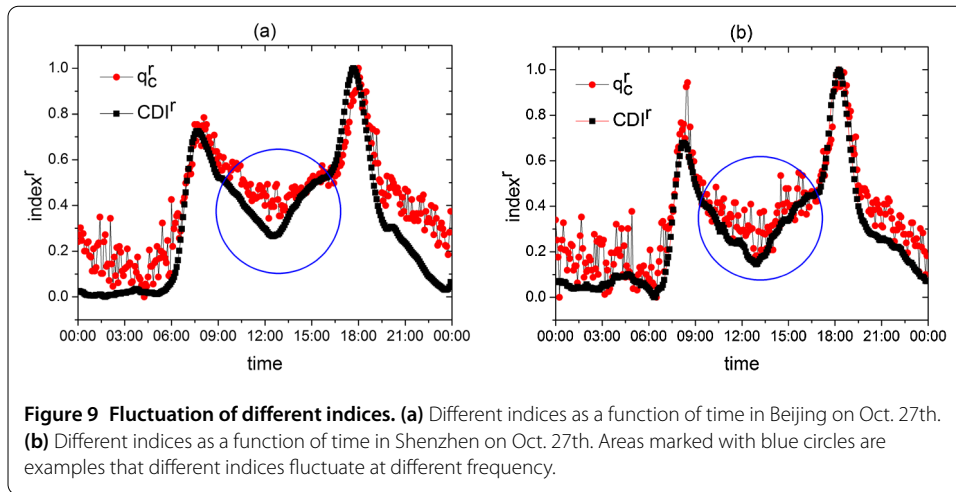


The distribution of index values during morning peak (morning peak  $\pm 0.5$  hour) in workdays is illustrated in Figure 6. In Beijing,  $q_c^r$  shows a bell shaped distribution and the maximum of  $p(q_c^r)$  is at  $q_c^r = 0.675$ . The distribution of  $CDI^r$  is more uniform with no obvious peak. This suggests that the evaluation of congestion reflected by the two indices is different, and  $q_c^r$  performs more stable since it focuses on the whole network performance. Performing stably is a critical property of an index because people may use it to compare traffic condition in different days. In Shenzhen, both indices show a bell shaped distribution. Indices values with higher probability are approximately equal:  $q_c^r = 0.675$  and  $CDI^r = 0.625$ . Moreover, we find the average  $q_c^r$  value of Beijing (0.75) is generally larger than those of Shenzhen (0.65). For  $CDI^r$ , the average value is 0.76 and 0.58 separately. It is shown that Beijing, as the capital city of China, has a worse traffic condition than Shenzhen, especially during morning peak. The differences between  $CDI^r$  and  $q_c^r$  result from index concepts underlying their calculation methods. The variation of  $CDI^r$  depends on the number of sample trips.  $CDI^r$  is influenced by the travel demands of traffic network, and roads with larger travel demands are more likely to be sampled.  $q_c^r$  does not depend on trip sampling. Instead, it considers the network as a whole system. Traffic fluctuations can influence congestion formation [30]. The scatter plot of the standard deviation of velocity and the standard deviation of  $q_c$  is shown in Figure 7.  $q_c$  will not change significantly with the velocity fluctuations of a few roads since it depends on the global connectivity.

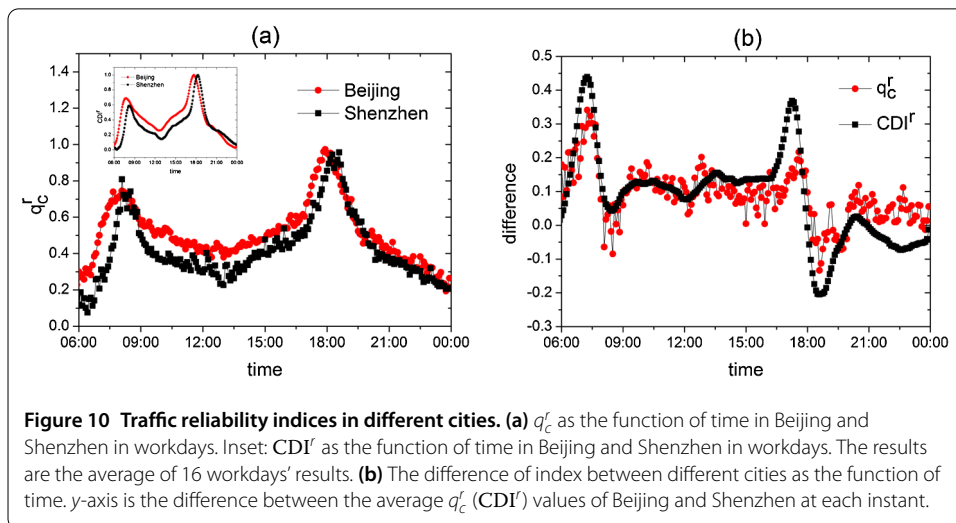


It is important to know when the traffic will get the most congested for congestion avoidance or traffic prediction. There are also differences in peak instants of the two indices, as shown in Figure 8. In Beijing on Oct. 27th, Tuesday, as shown in Figure 8(a), the time when  $q_c^r$  at peak in the morning is 8:05, while the peak instant of  $CDI^r$  is 7:40. In the evening, the peak instant of  $q_c^r$  is 18:00, slightly behind that of  $CDI^r$ , 17:40. The same can be observed in Shenzhen in the morning. Peak instants imply the worst traffic condition of the day. Furthermore, the distribution of peak instants of the two indices in workdays is illustrated in Figure 8(c) and (d). Obvious differences between the two indices can be observed. In Beijing, peak instants of  $CDI^r$  range from 7:33 to 7:48, while peak instants of  $q_c^r$  have a wider range from 7:33 to 8:16. In Shenzhen, peak instants of  $CDI^r$  distribute mainly around 8:13. However,  $q_c^r$  have a much wider distribution of peak instants ranging from 7:58 to 8:27. This means that from a time delay perspective, peak hours appear more centralized among different days than those when we focus on the phase transition of network.

The difference in peak instants may be due to the feature of  $CDI^r$  that extreme values of travel time along frequently visited trips will determine significantly peak instants of  $CDI^r$ . With the accumulation of traffic volume, the velocity will become slow and travel time through mostly congested area will become one of the determining factors for  $CDI^r$ . This also makes  $CDI^r$  more sensitive to the change of local traffic. During the morning rush hours, road conditions of other parts in the city may not change so sharply and global connectivity of whole functional city may stay stable.



**Figure 9** Fluctuation of different indices. (a) Different indices as a function of time in Beijing on Oct. 27th. (b) Different indices as a function of time in Shenzhen on Oct. 27th. Areas marked with blue circles are examples that different indices fluctuate at different frequency.



**Figure 10** Traffic reliability indices in different cities. (a)  $q_c^r$  as the function of time in Beijing and Shenzhen in workdays. Inset:  $CDI^r$  as the function of time in Beijing and Shenzhen in workdays. The results are the average of 16 workdays' results. (b) The difference of index between different cities as the function of time. y-axis is the difference between the average  $q_c^r$  ( $CDI^r$ ) values of Beijing and Shenzhen at each instant.

We also find that the two indices show different sensibility to the change of road condition. Areas in Figure 9 marked with blue circle show one example of this difference. Although the two indices have similar trends - both decrease to minimum and then begin to increase -  $q_c^r$  fluctuates with a higher frequency, while  $CDI^r$  changes relatively smoothly. This shows that  $q_c^r$  is more sensible to real-time traffic variation. When calculating  $CDI^r$ , travel time cannot be obtained until travelers complete the whole trip. Thus there exists a smooth effect of traffic states for  $CDI^r$ . We also find the similar phenomenon in Shenzhen, as shown in Figure 9(b).

To further illustrate the application of the two indices in different cities, we compare the performance of indices in Beijing and Shenzhen. In workdays, as shown in Figure 10(a),  $q_c^r$  in different cities shows similar trends in general, with clear peaks in morning and evening. For  $CDI^r$ , the same conclusion can be observed. This means both cities experienced heavy traffic congestion during two peak instants and the traffic networks become less reliable. During daytime, values of both indices in Beijing are generally larger than that in Shenzhen. Then we calculate the difference of the same index between different cities - the average  $q_c^r$  ( $CDI^r$ ) values in Beijing are subtracted by the average  $q_c^r$  ( $CDI^r$ ) values in Shenzhen at each instant, the results are shown in Figure 10(b). Moreover, the average of abso-

lute difference values of  $CDI^r$  (0.127) is larger than that of  $q_c^r$  (0.1). This difference between  $CDI^r$  and  $q_c^r$  may be caused by following reasons: travel time usually scales with city size, which leads to much larger  $CDI^r$  value in Beijing than that in Shenzhen. However,  $q_c^r$  tends to evaluate the global management efficiency of cities based on phase transition and relative velocity, which stays stable among cities. Thus  $q_c^r$  is more scalable for comparing the traffic reliability among different cities.

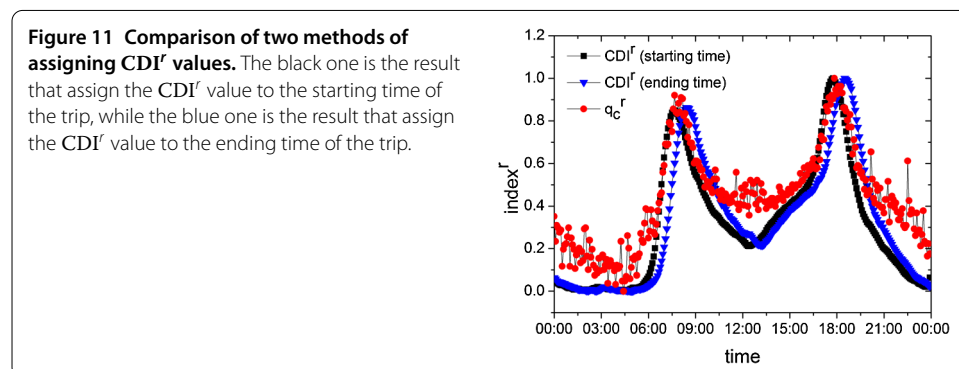
## 5 Conclusion

Traffic congestion has become increasingly frequent in many major cities around the globe. Congestions bring increasing extra economic and environmental costs to the whole society. It's critical to have an accurate estimation for traffic reliability for subsequent mitigation activities [31–33]. Here we compare a traffic reliability index based on percolation theory with congestion delay index (CDI), and calculate the results from different perspectives. When calculating the CDI, we assign the CDI index to the starting time of each trip. We also try to assign the CDI index to the ending time of each trip, as shown in Figure 11. There exists a clear delay (about 40 min) between two methods, which in accordance with Figure 3. It can be observed that  $q_c$  is more close to our original methods especially at peaks.

The percolation threshold naturally acts as a network reliability indicator, quantifying the operational limit of network traffic. Specifically, percolation theory focuses on connected clusters, which fills up the gap of other indices that rarely consider the macroscopic network congestion behaviors from a network view [21]. We find that  $q_c$  can reflect the transition of dynamical traffic network, faced with possible extreme local conditions. These features of  $q_c$  make it a useful tool under the variation and absence of complete traffic information, and provide supports for the congestion prediction and mitigation research.

Although we made a comparison between the two indices, we found that each index has its own advantages and limitations under certain situation. For example,  $q_c$  reflects traffic condition from managers' perspective, while is less intuitive for travelers; CDI is less scalable for comparing different cities, while has advantages of being more understandable and easier calculation. These important differences will decide the choice of a given traffic reliability index under specified requirement.

However, it should be noted that we just gave a brief analysis of indices in two cities during limited time span. More data and analyses are needed to summarize the travel



characteristics of different cities according to indices. In addition, the city is not homogeneous and the travel velocity depends on the trips length [34]. In our present work, we did not cut off the trip length. The influence of trips lengths on index calculation should be discussed in future research.

Our study focuses more on the quantitative comparison of two reliability indices, while the mechanisms behind these two indices are different. CDI assumes that the user trip information can reflect the overall performance of the traffic network, which incorporates the different OD information with their weights. Meanwhile, percolation concept suggests that the traffic organization over the whole network depends on the instantaneous connected clusters of high-speed roads with free flow, where the weight of each road is their velocity, instead of traffic flow. These underlying differences should be studied in the future, especially its relation with macroscopic fundamental diagram (MFD). Geroliminis and Daganzo [35] found that neighborhoods on the order of 10 km<sup>2</sup> in cities like Yokohama, Japan, should have a well-defined MFD. This MFD can be used to improve accessibility as measured by the city's trip completion rate. Both of their and our works discuss the index to measure the performance of traffic network. For  $q_c$  in our paper, it distinguishes the phase transition of the dynamical traffic network, where we divide based on the velocity the roads into two categories: free and congested.  $q_c$  reflects the real-time variation of traffic reliability. For the work of Geroliminis and Daganzo's, they used MFD for traffic state monitor, which reflects the relation of density, flow and velocity. Due to the lack of density and flow data, it is hard for us to explore the MFD and compare with percolation index in the current stage. Further analysis should be carried out when data are accessible.

Admittedly, only the comparison cannot reveal the underlying mechanism difference of these two indices. In our future work, we wish to gather the value of CDI from different sources including the mobile phone data, and further compare the fundamental relation between these two indices. Based on big data and other advanced technologies [36–39], we can perform a thorough cause-analysis for indices comparison in the future. Meanwhile, we can develop in the next step traffic optimization method based on percolation index and compare with other methods based on CDI.

#### **Acknowledgements**

This work is supported by National Natural Science Foundation of China (71621001, 71771009).

#### **Funding**

Not applicable.

#### **Abbreviations**

Not applicable.

#### **Availability of data and materials**

The traffic data used as a basis for the comparison cannot be shared because it relied on an anonymous travel dataset, which is protected by an internal agreement.

#### **Ethics approval and consent to participate**

Not applicable.

#### **Competing interests**

The authors declare no competing interests.

#### **Consent for publication**

Not applicable.

**Authors' contributions**

DL, LZ and GZ conceived and designed the research. SG acquired the data, and LZ, GZ analyzed the data. ZG and DL interpreted the results. All authors discussed, wrote, and approved the final version of the manuscript.

**Author details**

<sup>1</sup>School of Reliability and Systems Engineering, Beihang University, Beijing, China. <sup>2</sup>Science and Technology on Reliability and Environmental Engineering Laboratory, Beijing, China. <sup>3</sup>State Key Laboratory of Software Development Environment, Beihang University, Beijing, China. <sup>4</sup>Beijing PalmGo Infotech Co., Ltd, Beijing, China. <sup>5</sup>State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 2 November 2016 Accepted: 15 August 2017 Published online: 25 August 2017

**References**

- Li D, Fu B, Wang Y, Lu G, Berezin Y, Stanley HE, Havlin S (2015) Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proc Natl Acad Sci USA* 112(3):669-672
- Schrank D, Eisele B, Lomax T, Bak J (2015) 2015 urban mobility scorecard
- Bates J, Polak J, Jones P, Cook A (2001) The valuation of reliability for personal travel. *Transp Res, Part E, Logist Transp Rev* 37(2-3):191-229
- Mine H, Kawai H (1982) *Mathematics for reliability analysis*. Asakura, Tokyo
- Iida Y, Wakabayashi H (1989) An approximation method of terminal reliability of road network using partial minimal path and cut sets. In: *Transport policy, management & technology towards 2001: selected proceedings of the fifth world conference on transport research*
- Chen A, Yang H, Lo HK, Tang WH (1999) Capacity related reliability for transportation networks. *J Adv Transp* 33(2):183-200
- Chen A, Yang H, Hong KL, Tang WH (2002) Capacity reliability of a road network: an assessment methodology and numerical results. *Transp Res, Part B, Methodol* 36(3):225-252
- Lindley JA (1987) Urban freeway congestion: quantification of the problem and effectiveness of potential solutions. *ITE J* 57(1):27-32
- Schrank DL, Turner S, Lomax TJ (1994) Trends in urban roadway congestion - 1982 to 1991. *Statistics*
- Lei F, Wang Y, Lu G, Sun J (2014) A travel time reliability model of urban expressways with varying levels of service. *Transp Res, Part C, Emerg Technol* 48:453-467
- Hojati AT, Ferreira L, Washington S, Charles P, Shobeirinejad A (2014) Reprint of: modelling the impact of traffic incidents on travel time reliability. *Transp Res, Part C, Emerg Technol* 65:49-60
- Bhourri N, Haj-Salem H, Kauppila J (2013) Isolated versus coordinated ramp metering: field evaluation results of travel time reliability and traffic impact. *Transp Res, Part C, Emerg Technol* 28(3):155-167
- Sun L, Jin JG, Lee D-H, Axhausen KW (2015) Characterizing travel time reliability and passenger path choice in a metro network. In: *Transportation Research Board 94th annual meeting*
- Asakura Y, Kashiwadani M (1991) Road network reliability caused by daily fluctuation of traffic flow. In: *19th PTRC summer annual meeting, University of Sussex, UK*
- Cohn N, Kools E, Mieth P (2012) The TomTom congestion index. In: *19th ITS world congress*
- Min W, Wynter L (2011) Real-time road traffic prediction with spatio-temporal correlations. *Transp Res, Part C, Emerg Technol* 19(4):606-616
- Du W, Zhou X, Jusup M, Wang Z (2016) Physics of transportation: towards optimal capacity using the multilayer network framework. *Sci Rep* 6:19059
- Cohen R, Havlin S (2010) *Complex networks: structure, robustness and function*. Cambridge University Press, Cambridge
- Ben-Avraham D, Havlin S (2000) *Diffusion and reactions in fractals and disordered systems*. Cambridge University Press, Cambridge
- Bashan A, Parshani R, Havlin S (2011) Percolation in networks composed of connectivity and dependency links. *Phys Rev E* 83(5):051127
- Li D, Zhang Q, Zio E, Havlin S, Kang R (2015) Network reliability analysis based on percolation theory. *Reliab Eng Syst Saf* 93:556-562
- Wang F, Li D, Xu X, Wu R, Havlin S (2015) Percolation properties in a traffic model. *Europhys Lett* 112(3):38001
- Li D, Jiang Y, Kang R, Havlin S (2014) Spatial correlation analysis of cascading failures: congestions and blackouts. *Sci Rep* 4:5381
- Fushiki T, Yokota T, Kimita K, Kumagai M (2004) Study on density of probe cars sufficient for both level of area coverage and traffic information update cycle
- Yuan J, Zheng Y, Zhang C, Xie X, Sun GZ (2010) An interactive-voting based map matching algorithm. In: *Eleventh international conference on mobile data management*
- Dempster AP (2008) *Upper and lower probabilities induced by a multivalued mapping*. Springer, Berlin
- Diestel R (2011) *Graph theory*. *Math Gaz* 173(1):67-128
- Tarjan R (1971) Depth-first search and linear graph algorithms. *Sensors* 14(4):114-121
- Gong H, Chen C, Bialostozky E, Lawson CT (2012) A GPS/GIS method for travel mode detection in New York City. *Comput Environ Urban Syst* 36(2):131-139
- Andreotti E, Bazzani A, Rambaldi S, Guglielmi N, Freguglia P (2015) Modeling traffic fluctuations and congestion on a road network. *Adv Complex Syst* 18(03n04):1550009
- Liu C, Li D, Zio E, Kang R (2014) A modeling framework for system restoration from cascading failures. *PLoS ONE* 9:e112363

32. Liu C, Li D, Fu B, Yang S, Wang Y, Lu G (2014) Modeling of self-healing against cascading overload failures in complex networks. *Europhys Lett* 107(6):68003
33. Burkholz R, Garas A, Schweitzer F (2016) How damage diversification can reduce systemic risk. *Phys Rev E* 93:042313
34. Gallotti R, Bazzani A, Rambaldi S, Barthelemy M (2016) A stochastic model of randomly accelerated walkers for human mobility. *Nat Commun* 7:12600
35. Geroliminis N, Daganzo CF (2008) Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. *Transp Res, Part B, Methodol* 42(9):759-770
36. Sato AH, Sawai H (2014) Geographical risk assessment from tsunami run-up events based on socioeconomic-environmental data and its application to Japanese air transportation. *Proc CIRP* 19:27-32
37. Wang Q, Taylor JE (2015) Process map for urban-human mobility and civil infrastructure data collection using geosocial networking platforms. *J Comput Civ Eng* 30(2):04015004
38. Carbone A, Jensen M, Sato AH (2016) Challenges in data science: a complex systems perspective. *Chaos Solitons Fractals* 90:1-7
39. Lepri B, Antonelli F, Pianesi F, Pentland A (2015) Making big data work: smart, sustainable, and safe cities. *EPJ Data Sci* 4(1):16

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---