# Percolation transition in dynamical traffic network with evolving critical bottlenecks


**Daqing Li[a,b,1], Bowen Fu[a,b], Yunpeng Wang[c,1], Guangquan Lu[c], Yehiel Berezin[d], H. Eugene Stanley[e,1], and Shlomo Havlin[d]**

[a]School of Reliability and Systems Engineering and [c]Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems and Safety Control, Beihang University, Beijing 100191, China; [b]Science and Technology on Reliability and Environmental Engineering Laboratory, Beijing 100191, China; [d]Department of Physics, Bar Ilan University, Ramat Gan 52900, Israel; and [e]Center for Polymer Studies and Physics Department, Boston University, Boston, MA 02215



Contributed by H. Eugene Stanley, December 2, 2014 (sent for review June 19, 2014)



**A critical phenomenon is an intrinsic feature of traffic dynamics, during which transition between isolated local flows and global flows occurs. However, very little attention has been given to the question of how the local flows in the roads are organized collectively into a global city flow. Here we characterize this organization process of traffic as "traffic percolation," where the giant cluster of local flows disintegrates when the second largest cluster reaches its maximum. We find in real-time data of city road traffic that global traffic is dynamically composed of clusters of local flows, which are connected by bottleneck links. This organization evolves during a day with different bottleneck links appearing in different hours, but similar in the same hours in different days. A small improvement of critical bottleneck roads is found to benefit significantly the global traffic, providing a method to improve city traffic with low cost. Our results may provide insights on the relation between traffic dynamics and percolation, which can be useful for efficient transportation, epidemic control, and emergency evacuation.**



emergence | percolation | traffic


Traffic, as a large-scale and complex dynamical system, has attracted much attention, especially on its dynamical transition between free flow and congestion (1–3). The dynamics of traffic have been studied using many types of models (4–11), ranging from models in macroscopic scales based on the kinetic gas theory or fluid dynamics to approaches in microscopic scales with equations for each car in the network. However, there is still a gap between the microscopic behavior of individual vehicles and the emergence of macroscopic city traffic. Indeed, a fundamental question has rarely been addressed: how the local flows in roads interact and organize collectively into global flow throughout the city network. This knowledge is not only necessary to bridge the gap between local traffic and global traffic, but also essential for developing efficient traffic control strategies.

There are mainly two obstacles in studying how the collective network dynamics of real traffic emerge from local flows. The first obstacle is the lack of valid methods to quantify the dynamical organization of traffic in the road network. The second is the lack of data on traffic dynamics in a network scale. To overcome the first obstacle, we develop here a quantitative framework based on percolation theory, which combines evolving traffic dynamics with network structure. In this framework, instead of the commonly used structural topology, only roads in the network with speed larger than a variable threshold are considered functionally connected. In this way, we can characterize and understand the formation process of traffic dynamics.

To overcome the second obstacle of missing data on a network scale and understand the organization processes of real traffic in a network, we collected and analyzed velocities of more than 1,000 roads with 5-min segments records measured in a road network in a central area of Beijing (Fig. 1*A*). This area of more than 22 km$^2$ contains the largest train station in Beijing and is considered a typical region showing transition between free flow and congestions. The data cover a time span of 2 wk in 2013. For the road network, nodes represent the intersections and edges represent the road segments between two intersections. For each road, the velocity $v_{ij}(t)$ varies during a day according to real-time traffic. For each road $e_{ij}$, we set the 95th percentile of its velocity in each day as its limited maximal velocity and define $r_{ij}(t)$ as the ratio between its current velocity and its limited maximal velocity measured for that day (Fig. 1*B* and *SI Appendix*, Fig. S1). For a given threshold $q$, the road $e_{ij}$ can be classified into two categories: functional when $r_{ij} > q$ and dysfunctional for $r_{ij} < q$,

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

In this way, a functional traffic network can be constructed for a given $q$ value from the traffic dynamics of the original road network, which becomes more diluted as the value of $q$ increases.

## Results

To observe the emergence of global city traffic in the network scale at a given time, we can vary the value of $q$ and study the formation process of the dynamical traffic network. For $q = 0$, the traffic network is the same as the original road network and for $q = 1$ it becomes completely fragmented. For a certain value of $q$, the hierarchical organization of traffic in different scales emerges, where only clusters of roads with $r_{ij}$ higher than $q$ appear (clusters in Fig. 2 *A*–*C*). These clusters represent functional modules composed of connected roads with speed higher than $q$. For example, during a typical lunchtime instant, for $q = 0.69$, as

### Significance

The transition between free flow and congestions in traffic can be observed in our daily life. Although this traffic phenomenon is well studied in highways, traffic in a network scale (representing a city) is far from being understood. A fundamental unsolved question is how the global flow in a city is being integrated from local flows. Here, we identify a fundamental mechanism of traffic organization in a network scale as a percolation process, and we show how global traffic breaks down when identified bottlenecks are congested. These bottlenecks evolve with time according to traffic dynamics and are different from structural bottleneck links found by traditional network analysis. Improvement of traffic on these bottlenecks can significantly improve the global traffic.


Author contributions: D.L., B.F., Y.W., G.L., Y.B., H.E.S., and S.H. designed research; D.L., B.F., Y.W., G.L., Y.B., H.E.S., and S.H. performed research; D.L., B.F., Y.W., G.L., Y.B., H.E.S., and S.H. analyzed data; and D.L., Y.B., H.E.S., and S.H. wrote the paper.

The authors declare no conflict of interest.



[1]To whom correspondence may be addressed. Email: hes@bu.edu, ypwang@buaa.edu.cn, or daqingl@buaa.edu.cn.



This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419185112/-/DCSupplemental.

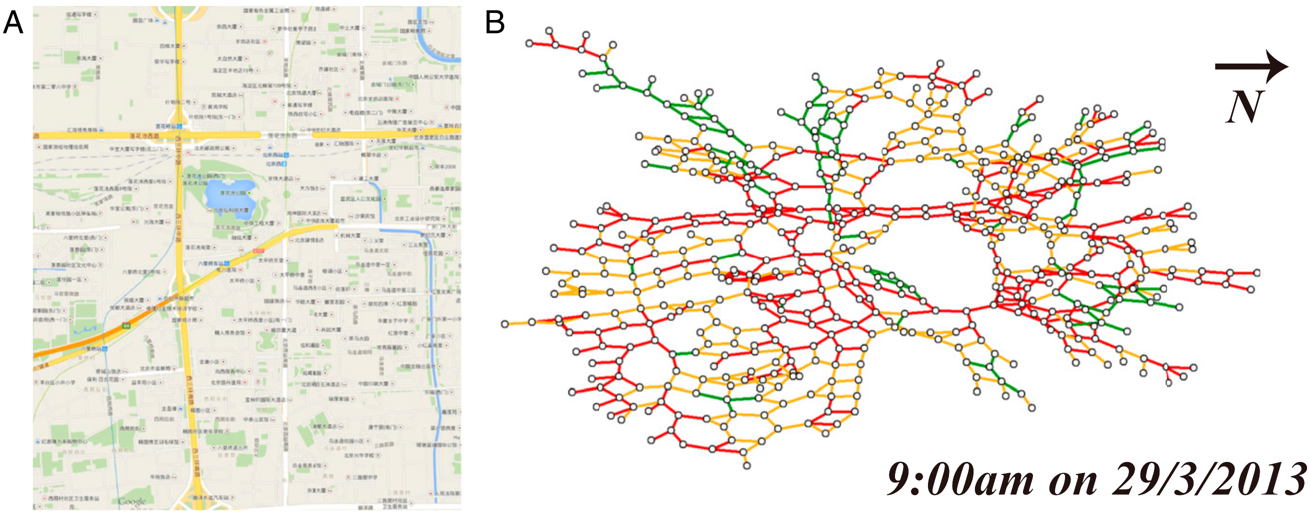


**Fig. 1.** Road network of the observed district. (*A*) Map of the investigated district. (*B*) Road network of the investigated district. Road network at 9:00 AM on March 29, 2013 is shown, where links are classified into three categories according to their velocity ratio $r_{ij}$: velocity ratio below 0.4 (red), between 0.4 and 0.7 (yellow), and above 0.7 (green). Note the clustering of each color.

shown in Fig. 2*A*, only small clusters of connected roads with high velocity emerge, which cannot maintain the global network traffic. As the value of $q$ decreases to 0.19 in Fig. 2*C*, these small clusters merge together and a giant cluster is formed, where the functional network (with lower velocity) extends to almost the full scale of original road network. For $q = 0.38$ (Fig. 2*B*), the size of second-largest cluster becomes maximal, which signifies the phase transition point for network connectivity of a functional traffic network, according to percolation theory (12, 13). This percolation-like process can be better understood in Fig. 2*D* (more examples in *SI Appendix*, Fig. S4). As $q$ increases, the size of the giant component decreases, and the second-largest cluster reaches a maximum at the critical threshold ($q_c$) separating the fragmented phase from the connected phase of the traffic network.

As an indicator of the robustness characteristics of network connectivity (14–20), the critical threshold $q_c$ in this percolation-like process here quantifies the organization efficiency of real traffic. An individual car can travel most of the city (giant component of traffic network) only with velocity below $q_c$, whereas this car will be trapped in small isolated clusters when it drives with velocity above $q_c$. Hence, $q_c$ measures effectively the maximal relative velocity one can travel over the main part of a network, which reflects the global efficiency of traffic in a network view.

Due to the traffic evolution, $q_c$ is found to change dramatically during the day as seen in Fig. 2*E* (details in *SI Appendix*, Fig. S5). In a typical working day, $q_c$ is found to be maximal from about midnight until 5:30 AM, indicating that the whole road network can function with high velocity. Close to 6:00 AM, $q_c$ begins to drop abruptly and shows a minimum around 8:00 AM corresponding to morning rush hours in Beijing. There are usually two local minima during a typical working day, which are around 8:00 AM and 6:00 PM. Note that $q_c$ reaches an intermediate level around noon, 12:00 PM, which might correspond to a possible third phase between free phase and congested phase. Due to the diverse commuting habit during weekends, only one local minimum appears in weekends around 2:00 PM (Fig. 2*E*).

The network at percolation criticality has a very dilute structure and behaves as the "backbone" of the original network (21). In the backbone of the traffic network, we find some links (called "red bonds" in percolation) that play a critical role in bridging different functional clusters of traffic. Therefore, these bridging links can be considered as bottlenecks because their velocities are lowest with respect to the whole backbone and $q_c$ is determined according to their value. We identify the bottleneck links of the traffic network by comparing the functional network just below and immediately above the criticality threshold. Fig. 3 *A* and *B* demonstrates the links removed at criticality, $q_c$, showing that they can disintegrate the giant cluster and result in a maximal second-largest cluster. Some of these links connect different traffic clusters and are thus considered bottlenecks. Because the roads in the real data are directed, we define the connected component as the "strongly connected component" (22, 23), in which all pairs of nodes are mutually reachable from each other along a directed path. Therefore, removal of two roads in Fig. 3*A* will lead to loss of directed paths bridging different clusters and disintegration of the giant strongly connected component.

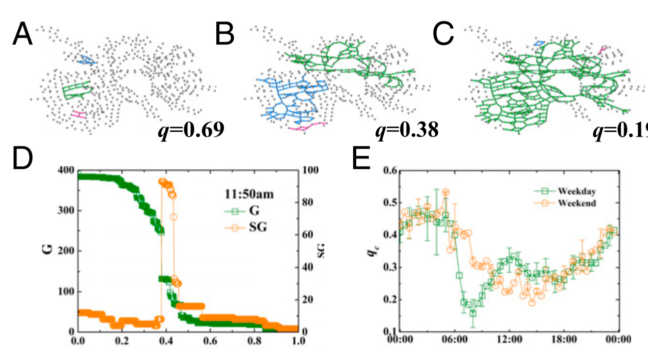


**Fig. 2.** Percolation of traffic networks: Traffic networks during the noon period (at 11:50 AM on March 27) for three $q$ values corresponding to different connectivity states. *A*, *B*, and *C* exhibit the traffic networks under different $q$ values with 0.69, 0.38, and 0.19 representing the states of high-, medium-, and low-velocity thresholds, respectively. For clarity, only the largest three clusters are plotted, which are marked in green (largest cluster), blue (second-largest cluster), and strawberry (third-largest cluster). Here the clusters are strongly connected components, considering road direction (more details in *SI Appendix*). (*D*) Size of the largest cluster (G) and the second-largest cluster (SG) of traffic networks as a function of $q$ (more examples in *SI Appendix*). Critical value, $q_c$, is determined as the $q$ value when SG becomes maximal. (*E*) $q_c$ as a function of time, averaged separately over nine weekdays and two weekends.

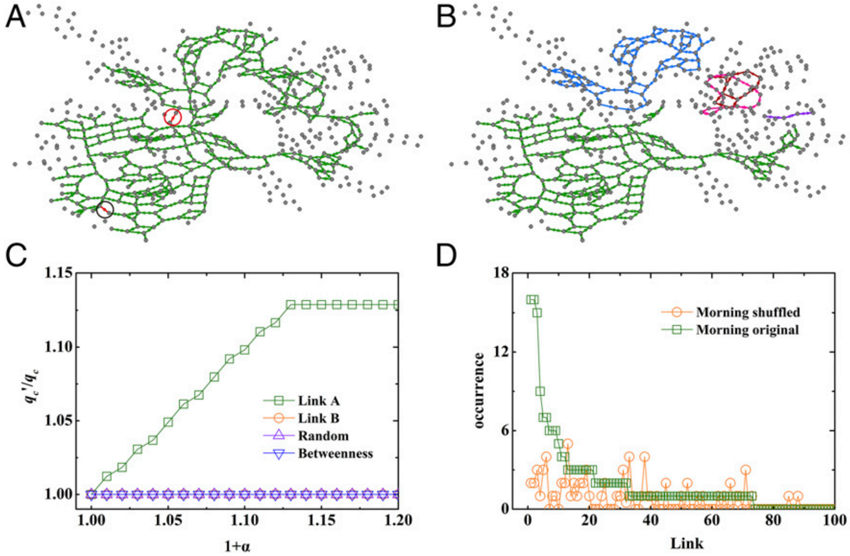


**Fig. 3.** Bottleneck links of a traffic network. (*A*) A typical example of a traffic network just below criticality, where two links (in red within red or black circles) are removed at criticality. Removal of them will disintegrate the giant functional network. (*B*) Same traffic network after removal of the two links, where the giant functional cluster is disintegrated into five clusters. We find all strongly connected clusters of the traffic network for each $q$ and identify the links removed at threshold $q_c$ when the second-largest strongly connected cluster reaches a maximum. Although some of these links are removed by chance, a few links do play a critical role of bridging different traffic clusters of higher velocities. These bridging links are identified as bottleneck links, because when increasing their velocity large clusters can join together to become the largest component (more details in *SI Appendix*, Fig. S7). (*C*) The improvement of $q_c$ by increasing separately the ratio ($r'_{ij} = r_{ij}(1+\alpha)$) of two links marked in *A*, within which improvement of $q_c$ can be achieved only with one (marked with red circle) of them. This link is considered a bottleneck link for global traffic. This is compared with the improvement of one link randomly chosen and the link with highest betweenness. (*D*) Zipf plot of occurrence times of links as bottlenecks during morning rush hours. It is compared with occurrence times of bottlenecks in the same network with shuffled values of $r_{ij}$ during morning rush hours. For the shuffled case, we shuffle the $r_{ij}$ values 100,000 times at each instant and find the bottleneck links with the same method.

To observe the impact of bottleneck links on the global traffic, we increase the road velocity ($r_{ij}$) of bottleneck links by a factor of $1+\alpha$ ($\alpha > 0$) and measure the new $q_c$ of the modified traffic network. It can be seen in Fig. 3*C* that $q_c$ of the traffic network is significantly increased only through increasing the velocity of the bottleneck link (more examples in *SI Appendix*, Fig. S7). This improvement of global traffic is dramatically higher compared with the improvement when increasing the velocity of a single link chosen randomly. Surprisingly, the improvement of $q_c$ is negligible when the velocity of the link with the highest hop-count–based betweenness is increased, although such a link is usually considered a bottleneck link because it bridges different topological communities (24, 25). This suggests that the bottleneck links found in our dynamical network are unique and different from results of network analysis based only on structural information. In addition, the links found based on weighted betweenness (26, 27) are also compared and found to be different from bottlenecks found by our method (*SI Appendix*, Fig. S12).

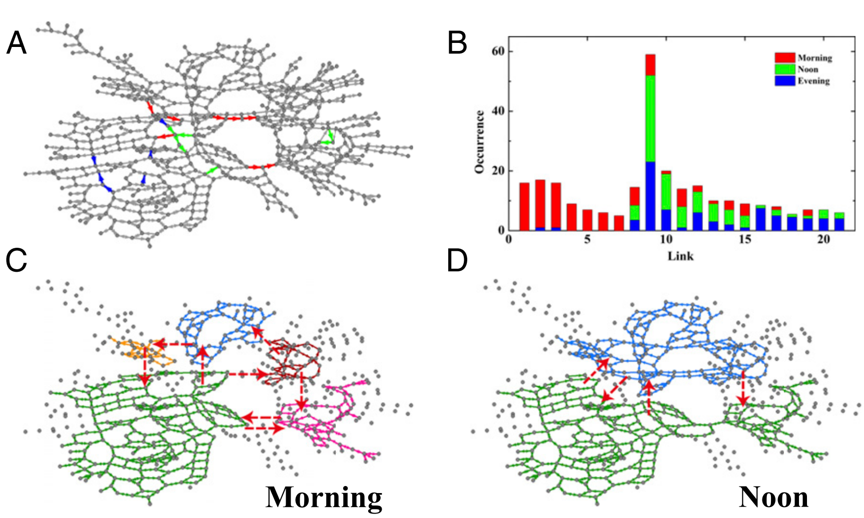


**Fig. 4.** Evolving bottlenecks in different periods in 1 d. (*A*) Bottleneck links with high occurrence in different periods are marked: morning (red), noon (green), and evening (blue). (*B*) The occurrence times of links (marked in *A*) as bottlenecks in different periods are plotted: morning (red), noon (green), and evening (blue). (*C* and *D*) The network breaks into several clusters after removal of bottlenecks with highest occurrence (top 10 in the morning in *C* or top 8 at noon in *D*). Red arrows in *C* and *D* are paths bridging different clusters, which are fragmented by the removal of bottleneck links.

Further discussion can be found in *SI Appendix*. The appearance of the bottleneck links is not accidental. As shown in Fig. 3*D* (*SI Appendix*, Fig. S8), some bottleneck links appear to be much more frequent than the random case. The high occurrence of bottleneck links demonstrates the dynamical percolation feature of real traffic and shows that the approach could be useful for significantly improving city traffic. Note that these bottleneck links with high occurrence are different from those found in the shuffled case, which reflect only the structural feature of the road network.

Static bottleneck links of a network are identified usually based on structural information (28–33), by considering links that are critical for network connectivity. However, traffic is a dynamical nonequilibrium system, which evolves with time as a result of collective individual competition. Therefore, we expect that bottlenecks of global traffic will also evolve accordingly, different from those found by structural methods. From the bottleneck links identified in different hours during a typical day (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S10), one can conclude that the bottleneck roads are essentially different in the morning, lunchtime, and evening rush hours. This is due to the different individual travel habits and interactions, which result in different global traffic patterns during different rush hours. As seen in Fig. 4*A*, in the morning, red bonds are distributed along a central path (city highway), whose congestion disintegrates the whole network into isolated clusters; however, in the evening hours, red bonds are distributed in less central roads, whose congestion influences only local areas, and the main part of the network stays functional. Indeed, as shown in Fig. 4*B*, the occurrence of links as bottlenecks changes dramatically from morning to evening rush hours; however, they appear repeatedly in different days in the same hours (*SI Appendix*, Fig. S11).

Bottleneck links result from the interactions among local functional clusters. Different bottleneck links signify distinct organization of global traffic in different hours. As shown in Fig. 4 *C* and *D*, traffic networks become disintegrated in different ways when different bottleneck links are removed. In Fig. 4*C*, removal of bottleneck links in the morning causes the giant cluster to break into one large cluster and four smaller clusters. In Fig. 4*D*, however, removal of bottleneck links at noon breaks the giant cluster into two clusters of similar size.

## Conclusion

As we reveal the percolation feature in organization of real traffic, the percolation threshold can be considered a measure for traffic efficiency, which takes into account the interaction between roads' network structure and flow. This proposed framework enables us to identify instantaneously those roads bridging different traffic clusters of higher velocity (with respect to the bottleneck). These bottleneck links identified at $q_c$ can provide opportunities to improve significantly the global network traffic with minor cost (e.g., improving a single road). Understanding the congestion formation and dissipation mechanisms in a network view through our framework can serve to predict and control traffic, in particular in the future realization of the "smart city." Particularly, our study can be useful in mitigating congestion (34) or traffic-driven epidemics (35) through certain self-healing algorithms (36) based on real-time information on traffic dynamics in the network.

**ACKNOWLEDGMENTS.** This work is supported by The National Basic Research Program of China (2012CB725404). D.L. acknowledges support from the National Natural Science Foundation of China (Grant 61104144). S.H. thanks Defense Threat Reduction Agency (DTRA), Office of Naval Research (ONR) (N62909-14-1-N019), United States-Israel Binational Science Foundation, the LINC (Grant 289447) and the Multiplex (Grant 317532) European projects, the German Research Foundation, and the Israel Science Foundation for support. The Boston University work was supported by ONR (Grant N00014-14-1-0738), DTRA (Grant HDTRA1-14-1-0017), and National Science Foundation (Grant CMMI 1125290).

1. Chowdhury D, Santen L, Schadschneider A (2000) Statistical physics of vehicular traffic and some related systems. *Phys Rep* 329(4):199–329.
2. Helbing D (2001) Traffic and related self-driven many-particle systems. *Rev Mod Phys* 73(4):1067–1141.
3. Kerner BS (2004) *The Physics of Traffic* (Springer, Berlin).
4. Lighthill M, Whitham G (1955) On kinematic waves. I. Flood movement in long rivers. *Proc R Soc Lond A* 229(1178):281–316.
5. Prigogine I, Herman R (1971) *Kinetic Theory of Vehicular Traffic* (Elsevier, New York).
6. Newell GF (1993) A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transp Res Part B* 27(4):281–287.
7. Bando M, Hasebe K, Nakayama A, Shibata A, Sugiyama Y (1995) Dynamical model of traffic congestion and numerical simulation. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 51(2):1035–1042.
8. Treiber M, Hennecke A, Helbing D (2000) Congested traffic states in empirical observations and microscopic simulations. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 62(2 Pt A):1805–1824.
9. Nagel K, Schreckenberg M (1992) A cellular automaton model for freeway traffic. *J Phys I* 2(12):2221–2229.
10. Kerner BS (1998) Experimental features of self-organization in traffic flow. *Phys Rev Lett* 81:3797–3800.
11. Helbing D, Huberman BA (1998) Coherent moving states in highway traffic. *Nature* 396(6713):738–740.
12. Bollobás B (1985) *Random Graphs* (Academic, London).
13. Bunde A, Havlin S, eds (1991) *Fractals and Disordered Systems* (Springer, Berlin).
14. Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74(1):47–97.
15. Newman ME (2003) The structure and function of complex networks. *SIAM Rev* 45(2): 167–256.
16. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U (2006) Complex networks: Structure and dynamics. *Phys Rep* 424(4):175–308.
17. Dorogovtsev SN, Goltsev AV, Mendes JF (2008) Critical phenomena in complex networks. *Rev Mod Phys* 80(4):1275–1335.
18. Barrat A, Barthelemy M, Vespignani A (2008) *Dynamical Processes on Complex Networks* (Cambridge Univ Press, Cambridge, UK).
19. Cohen R, Havlin S (2010) *Complex Networks: Structure, Robustness and Function* (Cambridge Univ Press, Cambridge, UK).
20. Barthélemy M (2011) Spatial networks. *Phys Rep* 499(1):1–101.
21. Wu Z, Braunstein LA, Havlin S, Stanley HE (2006) Transport in weighted networks: Partition into superhighways and roads. *Phys Rev Lett* 96(14):148702.
22. Broder A, et al. (2000) Graph structure in the web. *Comput Netw* 33(1):309–320.
23. Newman ME, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E Stat Nonlin Soft Matter Phys* 64(2 Pt 2): 026118.
24. Goh K-I, Kahng B, Kim D (2001) Universal behavior of load distribution in scale-free networks. *Phys Rev Lett* 87(27 Pt 1):278701.
25. Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12):7821–7826.
26. Ercsey-Ravasz M, Toroczkai Z (2010) Centrality scaling in large networks. *Phys Rev Lett* 105(3):038701.
27. Ercsey-Ravasz M, Lichtenwalter RN, Chawla NV, Toroczkai Z (2012) Range-limited centrality measures in complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 85(6 Pt 2):066103.
28. Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry* 40(1):35–41.
29. Costa LdaF (2004) The hierarchical backbone of complex networks. *Phys Rev Lett* 93(9):098702.
30. Guimerà R, Mossa S, Turtschi A, Amaral LA (2005) The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci USA* 102(22):7794–7799.
31. Borgatti SP, Everett MG (2006) A graph-theoretic perspective on centrality. *Soc Networks* 28(4):466–484.
32. Serrano MÁ, Boguñá M, Vespignani A (2009) Extracting the multiscale backbone of complex weighted networks. *Proc Natl Acad Sci USA* 106(16):6483–6488.
33. Grady D, Thiemann C, Brockmann D (2012) Robust classification of salient links in complex networks. *Nat Commun* 3:864.
34. Toroczkai Z, Bassler KE (2004) Network dynamics: Jamming is limited in scale-free systems. *Nature* 428(6984):716.
35. Meloni S, Arenas A, Moreno Y (2009) Traffic-driven epidemic spreading in finite-size scale-free networks. *Proc Natl Acad Sci USA* 106(40):16897–16902.
36. Quattrociocchi W, Caldarelli G, Scala A (2014) Self-healing networks: Redundancy and structure. *PLoS ONE* 9(2):e87986.

# Supplementary Materials for

## Percolation transition in dynamical traffic network with evolving critical bottlenecks

Daqing Li, Bowen Fu, Yunpeng Wang, Guangquan Lu, Yehiel Berezin, H. Eugene Stanley, Shlomo Havlin

## TABLE OF CONTENTS

## Data and Methods

### Data description

The dataset analyzed in the manuscript is the sampling velocity data of vehicles offered by Beijing Municipal Commission of Transport, with resolution of 5-minute segments over a district of about 22.5 $km^2$ around the largest train station in Beijing, Beijing West Railway Station. This area is one of the typical examples illustrating the daily traffic evolution in Beijing, where traffic congestions occur frequently. Road network in the district are composed of 530 nodes and 1,002 directed links.

### Data preprocessing

To build the dynamic network according to traffic data at each instant, interpolation from original data is necessary due to the lack of velocity data on some roads. For a road without velocity data at certain instant, we consider its velocity as the average velocity over its entire neighboring roads. By repeating this process, we complete the velocity data of all roads at each instant.

Since roads can be classified into different categories according to their capacity, it is not appropriate to define their traffic states based on the absolute value of velocity. To define an objective criterion, we consider the velocity limit for each road as the 95 percentile of its velocity values during each day (see Fig. S1). Then the ratio $r_{ij}$ between the current velocity and velocity limit of each road is used to quantify its relative velocity.

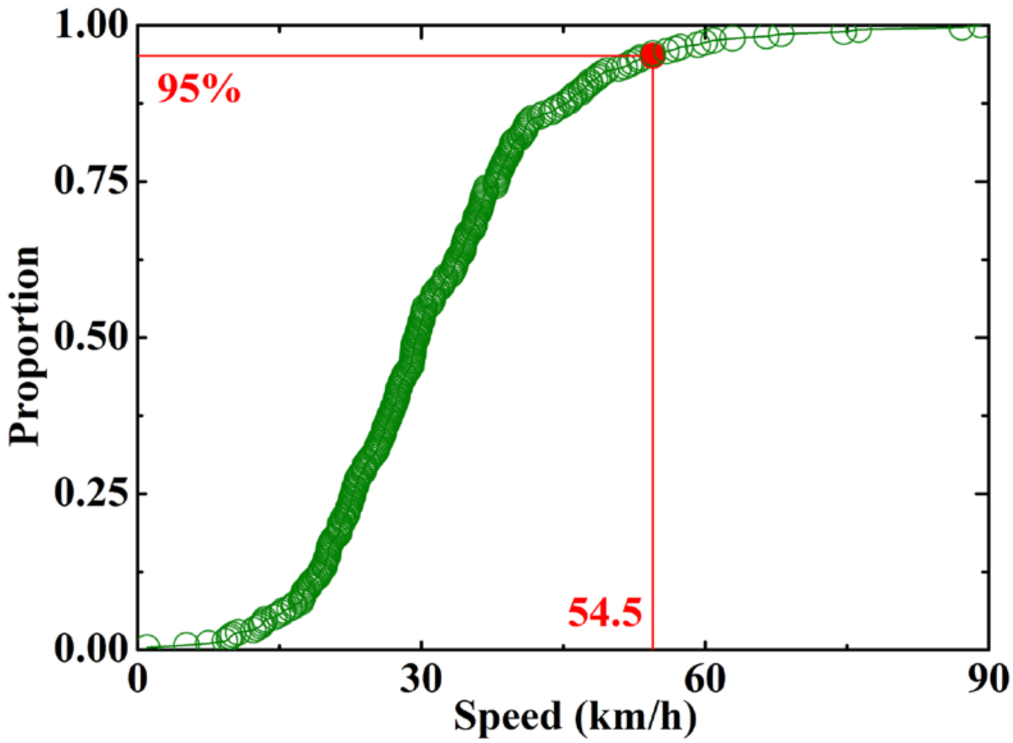


**Fig. S1.**

Illustration of finding velocity limit for an individual link along a whole day.

To test the significance of variance, we calculate the coefficient of variation (CV), which is the ratio between the standard deviation and the average of speed. As Fig. S2 shows, the CV is relatively small (below 0.3) for all instants during a typical day. Thus, the average speed represents well the sampling characteristics. Furthermore, the scatter plot of CV and the average of speed shown in Fig. S3, indicates that a higher CV is correlated with a lower average speed, which generates relatively smaller standard deviation.

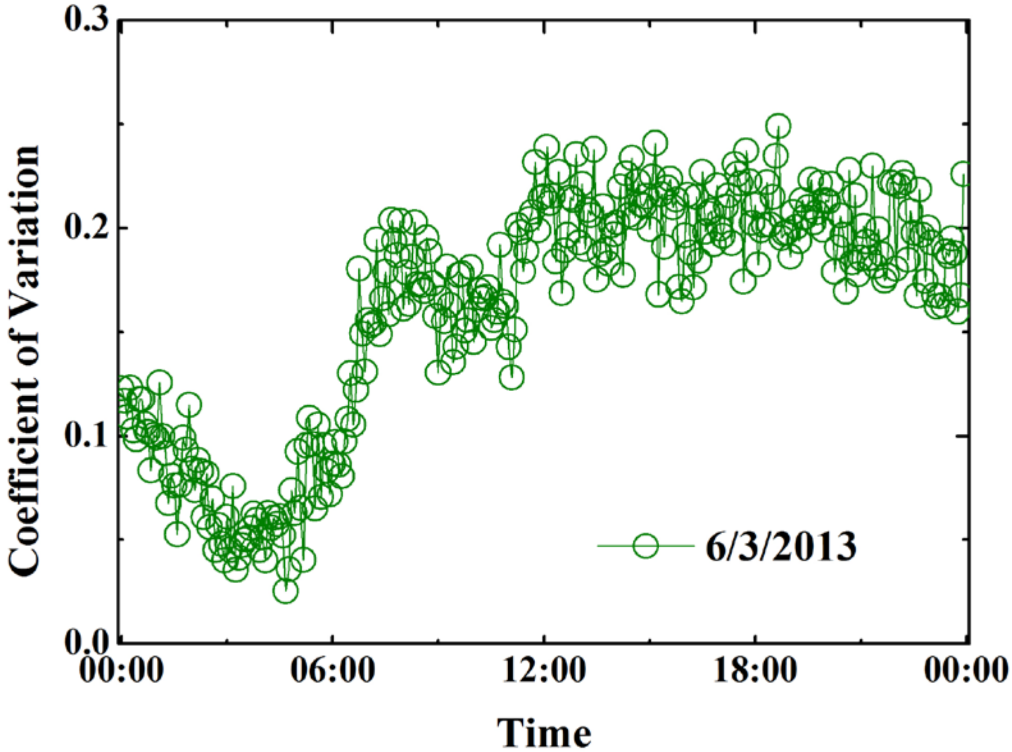


**Fig. S2.**

Coefficient of variation (CV) as a function of time for a typical day. Similar results have been found in other days.

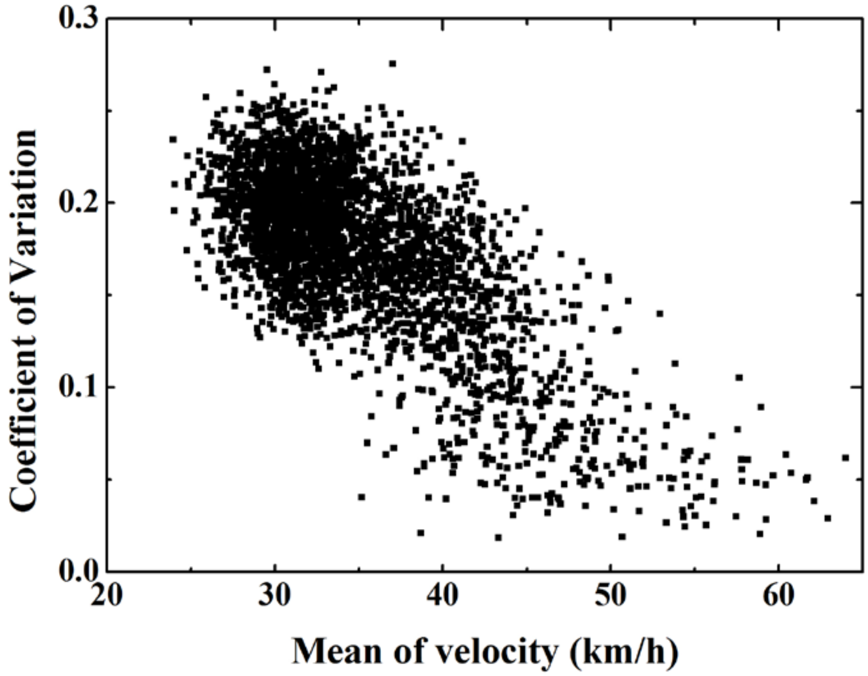


**Fig. S3.**

Scatter plot of coefficient of variation (CV) and the average speed. The plot includes the data of all the days analyzed in this study.

## Supplementary Text

Identifying criticality of traffic percolation

Since the road network is directed, clusters are defined as strongly connected components (1, 2), where all pairs of nodes are mutually reachable from each other along a directed path. For each $q$ value, roads with ratio $r_{ij} < q$ in the network are removed, and different strongly connected clusters are formed, including the largest cluster G and the second largest cluster SG. According to percolation theory (3, 4), the critical threshold $q_c$ of percolation transition is reached when the size of SG reaches maximal (see Fig. S4).

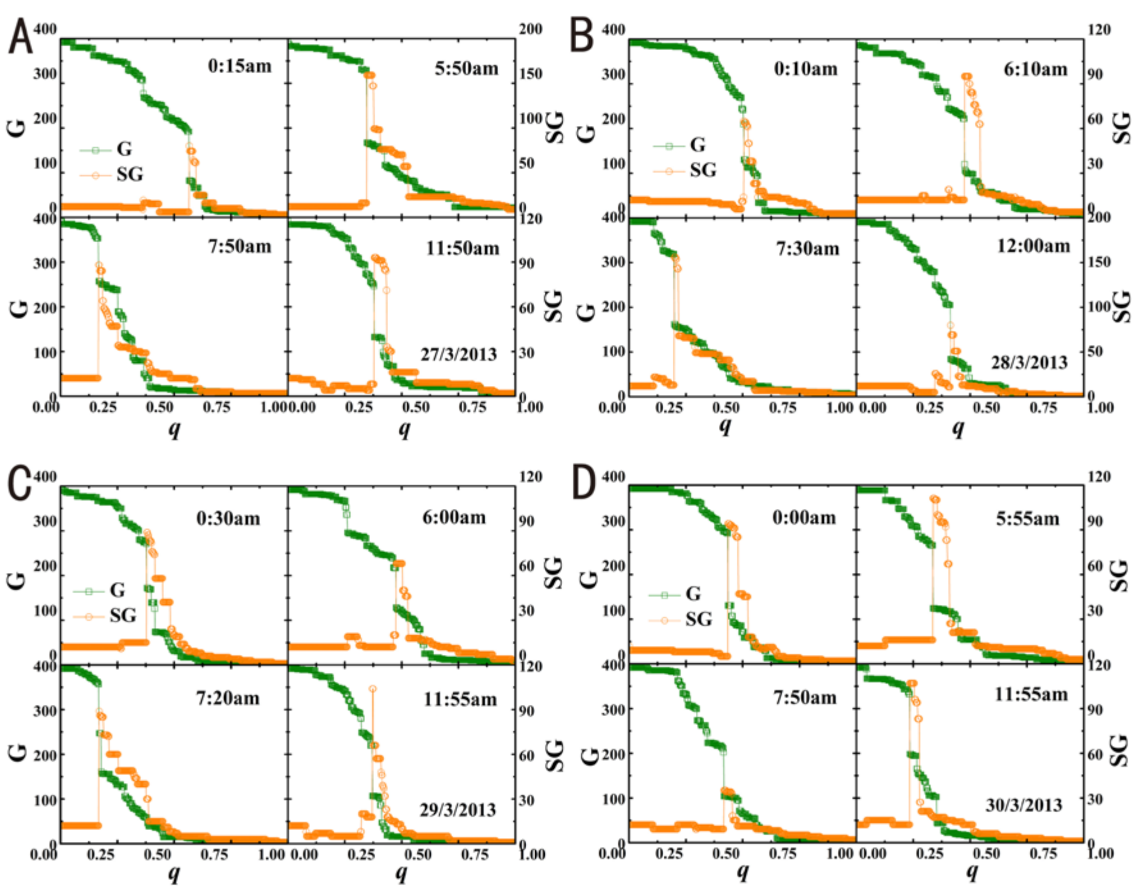

**Fig. S4.**

Giant component (G) and second giant component (SG) as a function of $q$ at different instants during a day.

Critical threshold of percolation in dynamical traffic network is calculated at different instants in all available days (Fig. S5). The figure (S5) demonstrates the evolving efficiency of global traffic during the day. In each day $q_c$ is calculated for all 288 instants (every 5 minutes during 24 hours in a day) along time, where different patterns between working day and weekend can be observed (see Fig. S5). During working days, $q_c$ has 2 local minima corresponding to rush hours in the morning and evening. Note that $q_c$ is smaller during morning than that of evening showing traffic congestion is heavier in the morning.

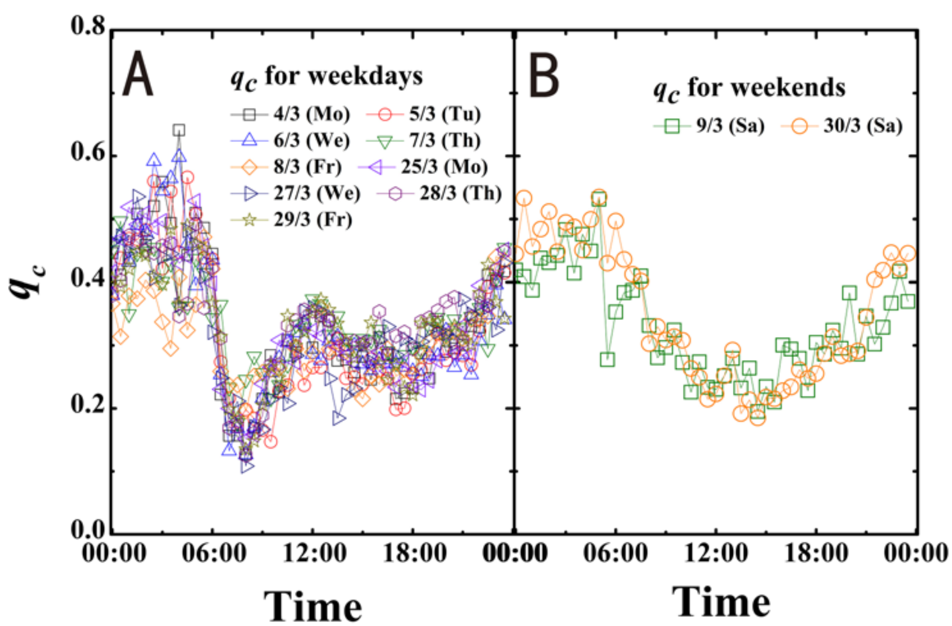

**Fig. S5.**

The global traffic efficiency, $q_c$, as a function of time during a day. The value of $q_c$ is smoothed by averaging over a time window of 30 minutes.

We also calculate (Fig. S6) the distribution of original $q_c$ for weekdays and weekends, during morning and afternoon rush hours. Note that while in the morning hours $q_c$ are much smaller in working days, the opposite behavior is seen during lunch time.

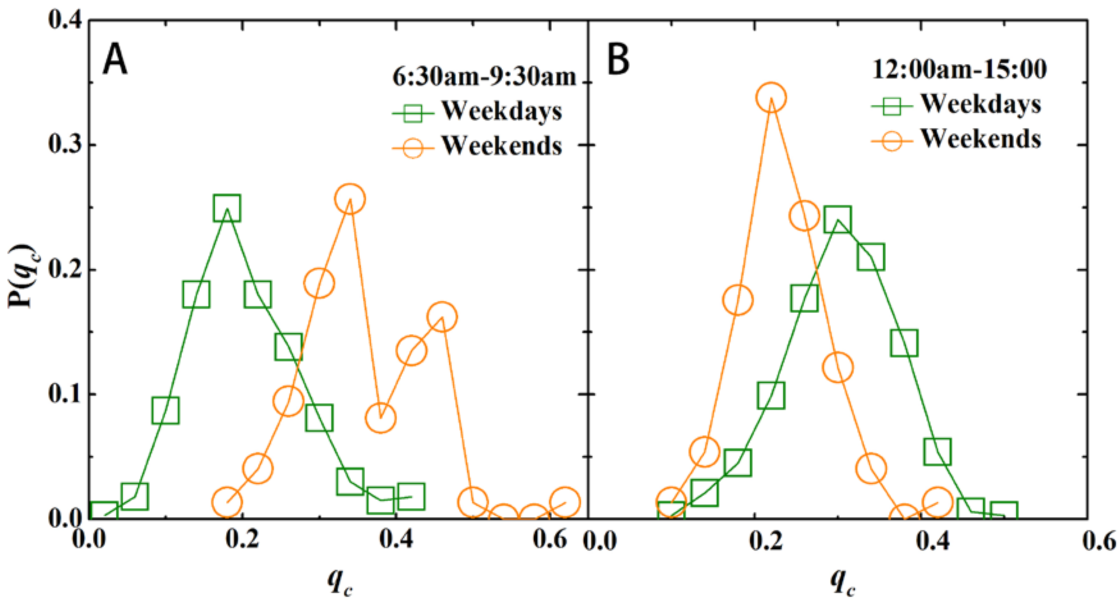


**Fig. S6.**

Distribution of $q_c$ during 6:30am-9:30am (A) and 12:00am-15:00 (B) in weekdays and weekends.

Bottlenecks: red bonds found by percolation

Based on our observation that traffic can be viewed as a percolation process, it is suggested that high velocity links tend to form clusters, which are bridged by low-velocity roads (red bonds). These bottleneck roads determine the global connectivity of dynamical traffic network at each instant, and can help to develop real-time control strategy and improvement of city traffic.

At the critical threshold $q_c$, several roads are removed when we tune the $q$ value slightly higher. While some links are removed by chance, there are some roads that play a critical role in connecting different local traffic clusters in traffic network. Only these roads, whose improvement will determine the critical threshold $q_c$, are considered as bottleneck roads (Fig. S7). Removal of these bottleneck links will result in disintegration of traffic network into several local clusters (Fig. S7). Notice that improvement of other links including links with highest betweenness usually will not benefit significantly the global traffic in terms of increasing $q_c$. Different from fixed static bottleneck links found by structural analysis, the bottleneck links found here evolve with time and are usually different in different hours.

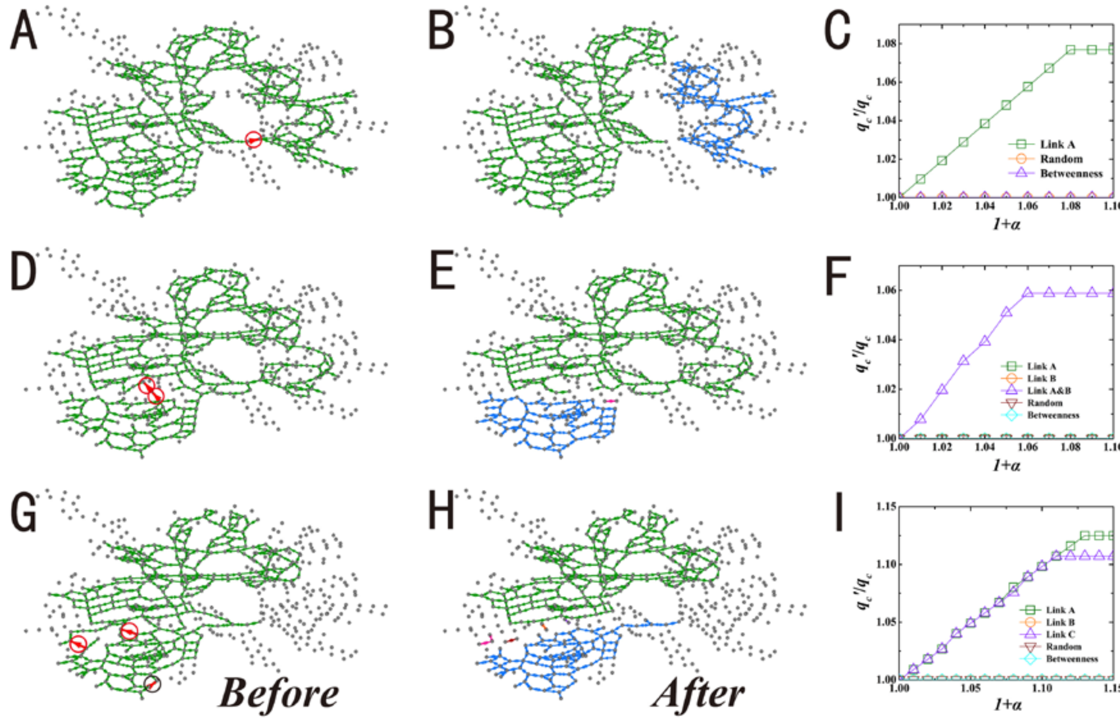


**Fig. S7.**

Examples of bottleneck links and effect of their improvement on global traffic. Links colored red in A (8:00am on 29/3/2013), D (8:55am on 7/3/2013), G (18:35 on 8/3/2013) are links that will be removed just above $q_c$. We find that only the improvement of links marked in red circles (bottleneck links) can significantly improve the global traffic in terms of $q_c$, while those marked in black circles are removed at criticality by chance. The largest component (G, in green) and the second largest component (SG, in blue) after the deletion of different red bonds are plotted in B, E, H. The effect on $q_c$ of improving velocity of bottleneck links is shown in C, F, I, together with a comparison with the link with highest betweenness and a randomly chosen link. Generally improvement of one bottleneck link will significantly improve $q_c$ (see C), but under some circumstances two or more links have to be improved simultaneously to gain an improvement of $q_c$ (see F), and in some other situations, improvement of either of several links will work (see I).

For the three cases mentioned above, the occurrence times (see next section) of links appearing as bottleneck are 1, 1 and $1/m$ respectively, where $m$ is the number of bottlenecks.

Occurrence times of bottlenecks

To further explore the temporal behavior of bottleneck links we found, we study also the occurrence of bottlenecks during two different periods (Fig. S8): 11:00-14:00 at noon, and 17:00-20:00 in the evening. It can be seen in Fig. S8 that some bottlenecks appear much more frequent than others. Identifying such highly occurrence links is critical for mitigating traffic congestions in different periods. When the link velocities ($r_{ij}$) are shuffled at a given instant, it is expected that structural bottleneck links will be mostly identified. Indeed, they have higher occurrence as bottleneck in the shuffled case, while they are not real bottlenecks if their improvement does not increase $q_c$ (Fig. 3C and S7). More surprisingly, as shown in Fig. S8, the bottleneck links found based on traffic information are different from those structural bottleneck links. This difference stems from the dynamical interactions between roads traffic.

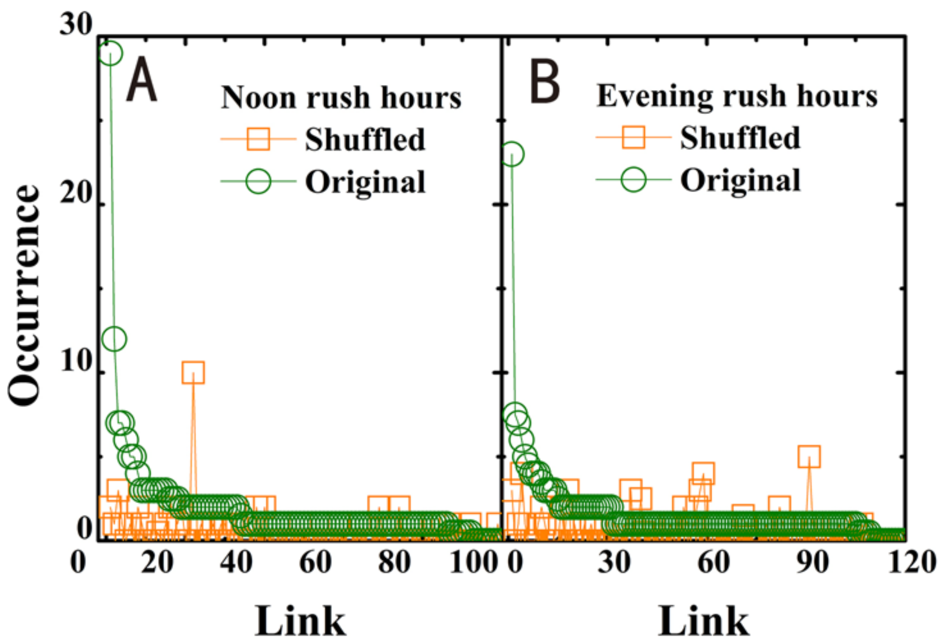

**Fig. S8.**

Occurrence (in a descending order) of links as bottlenecks. The number of times roads appearing as bottlenecks is plotted in different periods: (A) noon rush hours in 11:00-14:00; (B) evening rush hours in 17:00-20:00. Morning rush hours is shown in Fig. 3D. The results are compared to the shuffled velocities control case.

To explore the relationship between $q_c$ and $r_{ij}(t)$ of bottlenecks with highest occurrence times, we plot in Fig. S9 the $r_{ij}(t)$ of 3 top occurrence bottlenecks in the morning, as well as $q_c$ in the corresponding hours. As shown in Fig. S9, from 05:30 to 07:00, both $q_c$ and $r_{ij}(t)$ of bottlenecks are found to decrease, and they are overlapped significantly between 07:00 and 09:00, while after 09:00 they all tend to increase. Thus, the $r_{ij}(t)$ correlates with $q_c$. Note that for each time window different bottlenecks might dominate, some of which are not among top three.

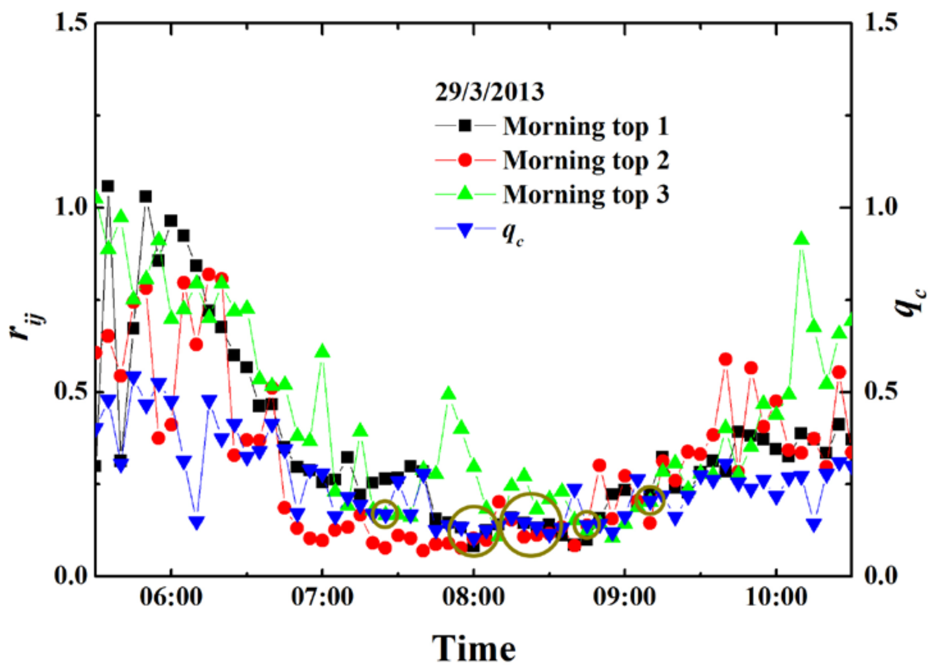

**Fig. S9.**

Velocity ratios $r_{ij}(t)$ of three bottleneck links with highest occurrence during morning rush hours, and $q_c$ as a function of time in the morning on 29/3/2013. The coincidence of $q_c$ and $r_{ij}(t)$ of the three bottleneck links are marked in brown circles. The fact that we see more coincidences inside the circles, demonstrates that the same link can act as a bottleneck many times.

To inspect the difference of bottleneck links found in three time periods of rush hours in a typical day, occurrence times of links as bottleneck in morning, noon and evening periods are plotted in Fig. S10. From this figure, we can see that the bottleneck links with highest occurrence in morning rush hour are different from those in the noon or evening rush-hours. Thus, according to our findings, different control strategies should be adopted in different rush hours to improve the global network traffic.

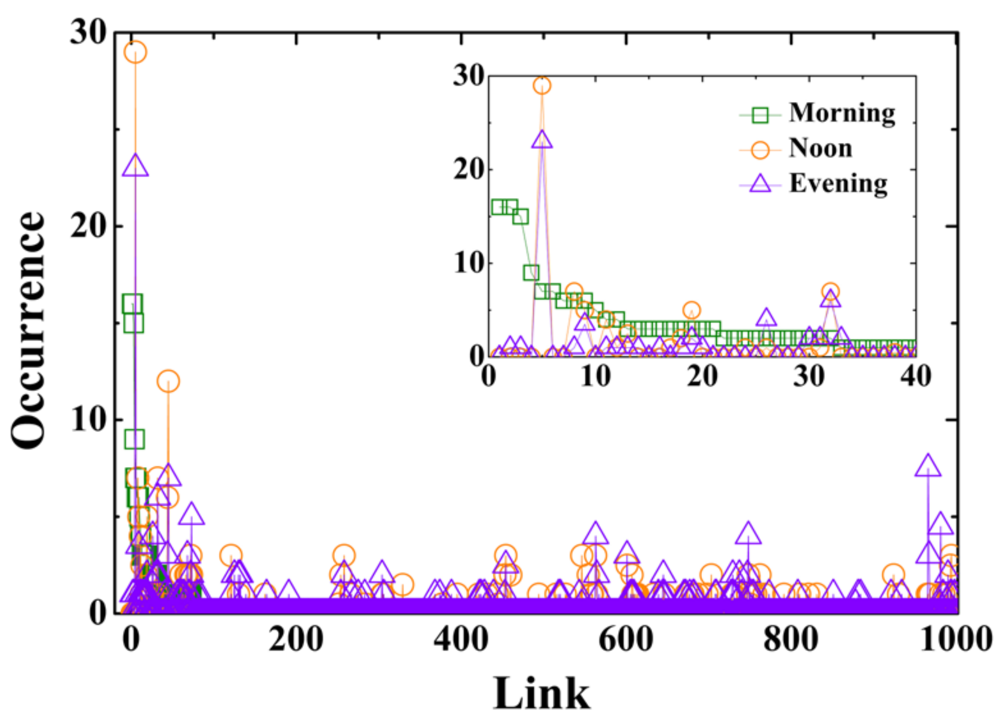

**Fig. S10.**

Occurrence times of bottleneck links during morning, noon and evening rush-hours. Links are ranked based on the occurrence times in the morning rush-hour. Inset is the same content of top 40 links in the morning rush hours.

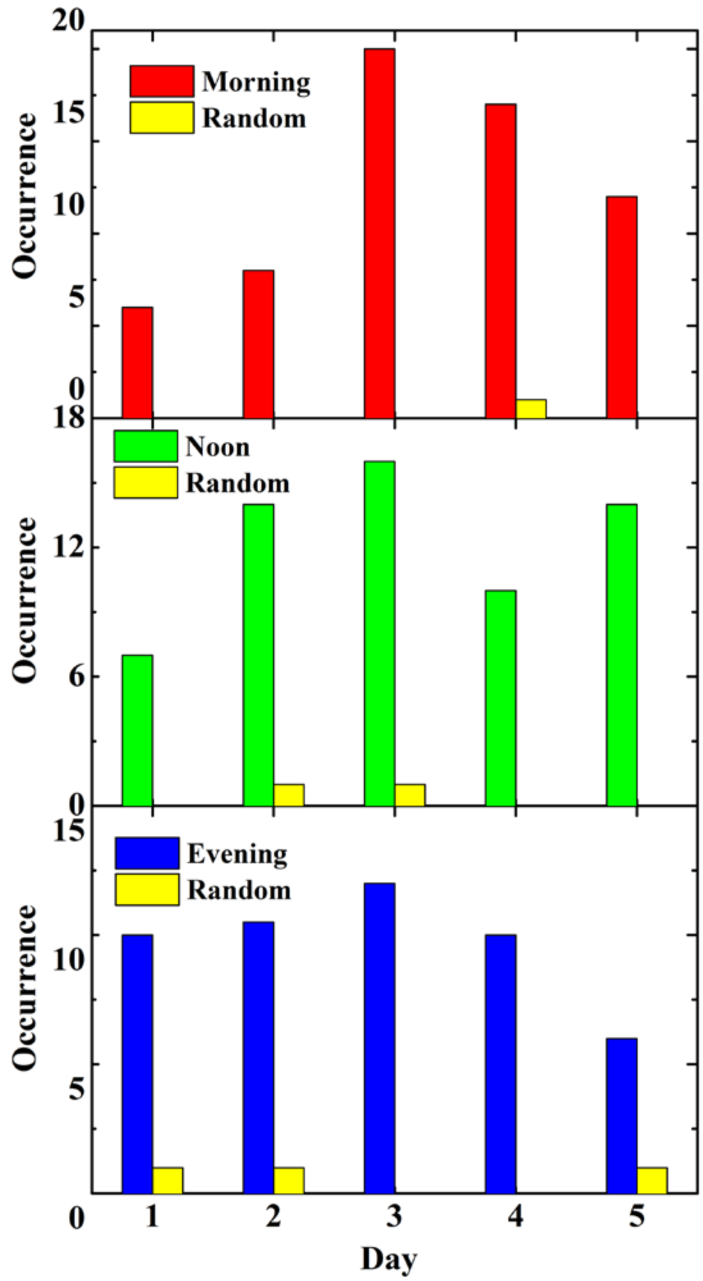


**Fig. S11.**

Sum of occurrence times of 5 bottleneck links with highest occurrence found in morning (red), noon (green) and evening (blue) in five different working days, compared with 5 links selected randomly.

Although the bottleneck links are different in different rush hours within a day, they appear stable and repeatedly in the same type of rush hours in different days. This frequent appearance of the same bottlenecks in the same hours in different days suggests a stable traffic pattern.

Comparison with weighted betweenness

To illustrate the bottleneck links found in our study, we calculated also the bottlenecks using the weighted betweenness (5, 6). We define the weight as instantaneous travel time of a given road and find the largest conductance path with minimal total travel time. Here instantaneous travel time of a given road is the ratio between length of the road and the instantaneous speed along this road. As shown in Fig. S12, links with high weighted betweenness are found to be different from those with high hop-count betweenness, and also different from dynamical bottlenecks found by our percolation approach. Links with high weighted betweenness have usually higher velocity, connecting different pairs in the road network. On the other hand, bottlenecks, found in our manuscript by percolation approach, are usually links with relatively lower speed bridging different high-speed clusters.

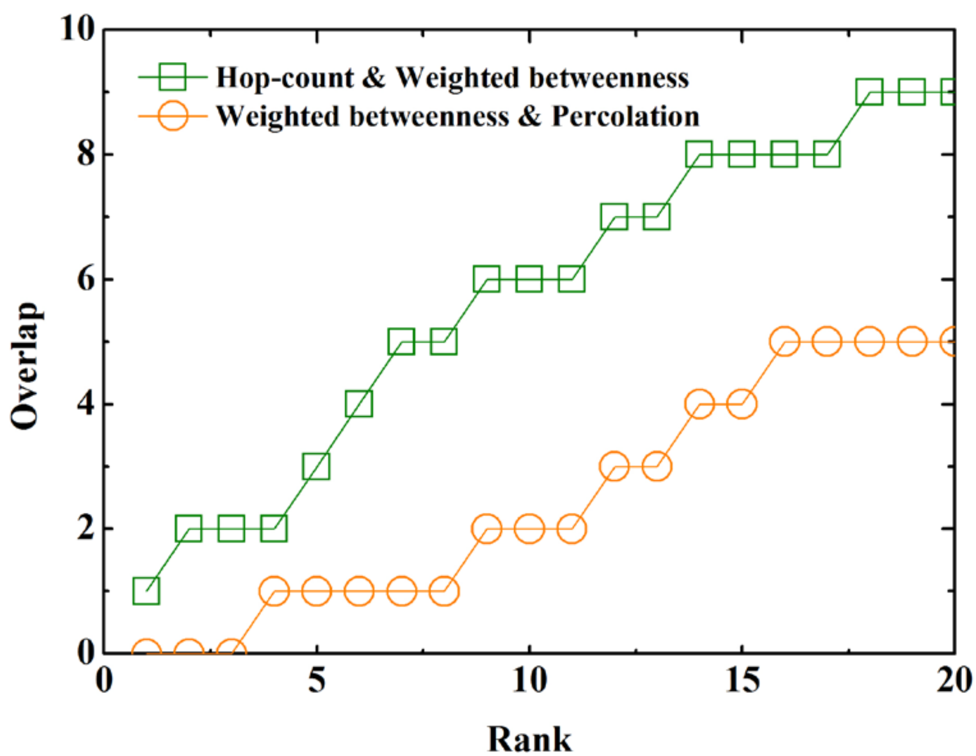

**Fig. S12.**

Overlap of top links with ranking according to hop-count betweenness, weighted betweenness or bottleneck occurrence based on percolation during 6:30am-9:30am. We rank the links by hop-count betweenness, weighted betweenness or bottleneck occurrence based on percolation, and compare the overlap of top link list. For example, for top 5 links obtained by weighted betweenness and percolation approaches, only one link among them is overlapped. For the weighted betweenness, we count the link with highest weighted betweenness at each instant, and rank the links by their occurrence times as highest weighted betweenness during 6:30am-9:30am.

Relation with previous theoretical results

To connect our results with theoretical results, first, we calculate in Fig. S13 the critical degree $k_c$ at criticality of traffic percolation as defined in the manuscript, and find that the critical degree fluctuates between 0.7 and 1.3 with time.

Second, we compare the critical degree of traffic percolation with that of random geometric graphs (RGG). The critical degree of RGG ($\approx$4.5) (7, 8) is much larger than that of traffic percolation (0.7-1.3). This is because in random geometric graphs, according to their construction rule, there are many local fully connected communities composed of sites within a given radius.

Third, we also compare our results with random directed percolation on square lattice (9, 10). In these systems, it is found that $\langle k \rangle_c = 2*0.6447 \approx 1.3$. Our result is smaller (0.7-1.3) in most cases possibly because the percolation of traffic network is not purely random as a result of the correlation in the traffic flow (11).

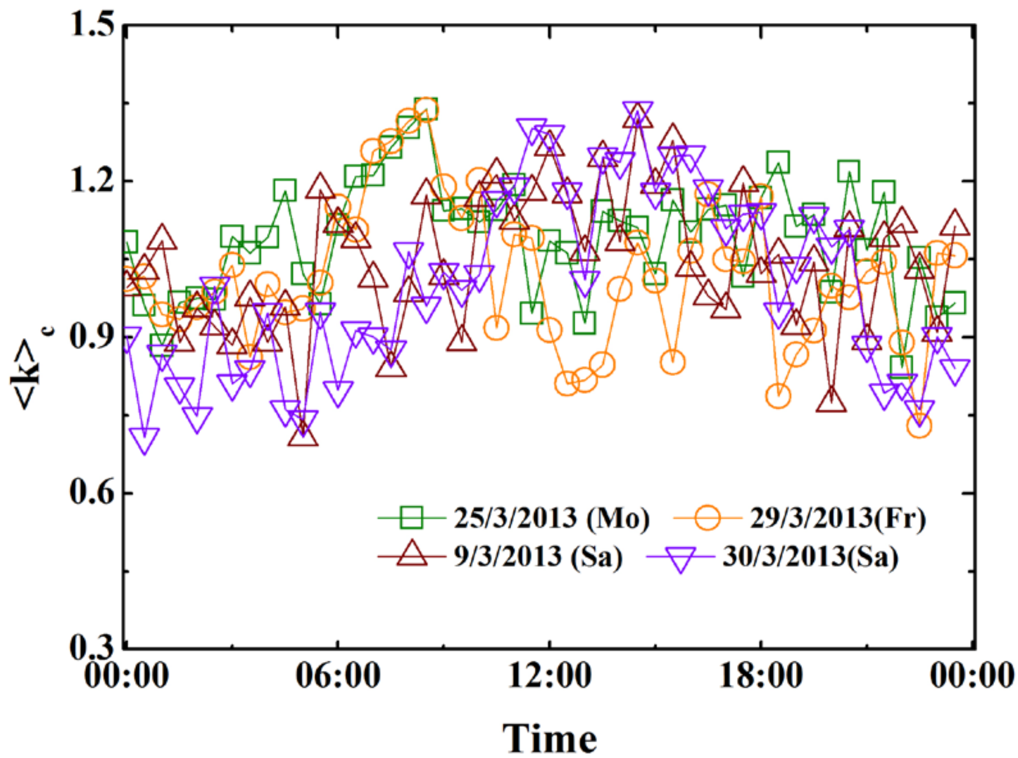


**Fig. S13.**

Threshold of degree, $<k>_c$, as a function of time on typical weekdays and weekends in traffic percolation, which is calculated as the average degree of traffic network at $q_c$ for each instant. Similar results are also found in other days.

In previous studies based on cellular automata model in 2d lattice (12, 13), the jamming transition to congestion has been studied as a function of car density. In our study, although we do not have data of car density, our approach based on velocity data, shows that realistic organization of local flows into global flows in a network scale is like a percolation process. With more available data including car density in the future, we believe that the relation between these theoretical results and our findings from real data could be further explored.

## References


1. Broder A *et al.* (2000) Graph structure in the web. *Comput Netw* 33(1-6):309-320.
2. Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E* 64(2):026118.
3. Bunde A, Havlin S (1991) *Fractals and Disordered Systems*, Ed. (Springer-Verlag, Berlin).
4. Stauffer D, Aharony A (1994) *Introduction to percolation theory*. (Taylor and Francis, London).
5. Ercsey-Ravasz M, Toroczkai Z (2010) Centrality Scaling in Large Networks. *Phys Rev Lett* 105(3):038701.
6. Ercsey-Ravasz M, Lichtenwalter RN, Chawla NV, Toroczkai Z (2012) Range-limited centrality measures in complex networks. *Phys Rev E* 85(6):066103.
7. Penrose M (2003) *Random geometric graphs* (Oxford University Press Oxford).
8. Dall J, Christensen M (2002) Random geometric graphs. *Phys Rev E* 66(1):016121.
9. Soares DJ, Andrade Jr JS, Herrmann HJ (2005) Precise calculation of the threshold of various directed percolation models on a square lattice. *J Phys A: Math Gen* 38(21):L413-L415.
10. Jensen I (1999) Low-density series expansions for directed percolation: I. A new efficient algorithm with applications to the square lattice. *J Phys A: Math Gen* 32(28):5233-5249.
11. Helbing D (2001) Traffic and related self-driven many-particle systems. *Rev Mod Phys* 73(4):1067-1141.
12. Biham O, Middleton AA, Levine D (1992) Self-organization and a dynamical transition in traffic-flow models. *Phys Rev A* 46(10):R6124-R6127.
13. Cuesta JA, Martínez FC, Molera JM, Sánchez A (1993) Phase transitions in two-dimensional

traffic-flow models. *Phys Rev E* 48(6):R4175-R4178.