

<https://doi.org/10.1038/s42005-025-02049-6>

# Unveiling city jam-prints of urban traffic based on jam patterns



Guanwen Zeng<sup>1,2,3</sup>, Nimrod Serok<sup>4</sup>, Efrat Blumenfeld Lieberthal<sup>4</sup>, Jinxiao Duan<sup>5,6</sup>, Shiyan Liu<sup>3</sup>, Shaobo Sui<sup>3</sup>, Daqing Li<sup>3</sup> & Shlomo Havlin<sup>2</sup>✉

The complexity of traffic jam is expected to increase significantly as cities expand. While existing studies have examined local bottlenecks and global traffic patterns, the daily evolution of traffic congestion is unclear yet crucial for understanding network-scale jam formation. We analyze the daily patterns of traffic jams in typical urban networks using real-world data based on a recently developed jam tree model. We extend the model by integrating additional realistic jam components into the model and find that, while the locations of traffic jams can vary significantly, the daily distribution of the costs associated with these jams follows a consistent pattern, i.e., a power law with similar exponents. This consistent pattern persists across different days within individual cities while exhibiting distinct variations between different cities, forming a unique signature we term “jam-prints”. Our findings are useful for assessing the quality of urban traffic and for establishing new traffic management goals.

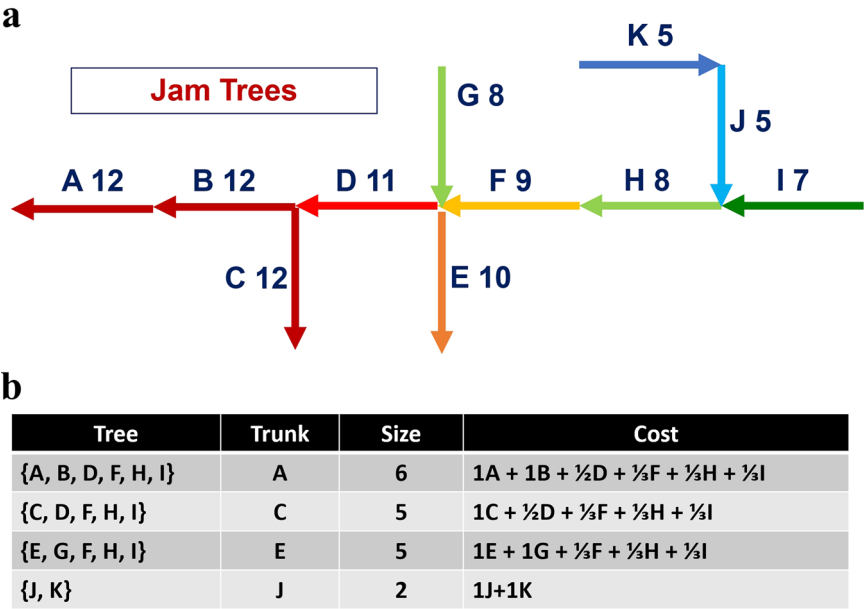
Traffic jams have been a long-term worldwide painful problem for travelers and urban traffic administrators. They are often considered as the result of the decision-making processes of individual travelers (i.e., agents) who display bounded rationality in a system with limited transportation resources. The complexity of traffic jam is expected to significantly increase as urban populations and the size of cities grow. This is particularly true in megacities and urban agglomerations, where traffic jams can be challenging to address. With the emergence of new technologies, i.e., Internet of Things, cloud computing, big data, and artificial intelligence, modern transportation systems could be improved and become more intelligent and efficient, known as the Intelligent Transportation Systems (i.e., ITS). The ITS integrates various types of networks, including cyber networks (i.e., sensing and communication networks), physical networks (i.e., traffic network) and social networks (i.e., traveling demand and traveling trajectory networks). On one hand, ITS is aiming at reducing traffic jams by using comprehensive sensing, communicating, and monitoring technologies. On the other hand, the introduction of these new technologies could also enhance the inter-dependent interactions among different elements in the transportation system (i.e., travelers, cars, roads, devices, etc.), potentially creating new risks. For instance, it has been found that as the strength of these inter-dependent interactions grows, cascading failures appear and lead to abrupt collapses of the system of systems even from a small region of initial failures<sup>1,2</sup>.

A key challenge in addressing traffic congestion is detecting and mitigating bottlenecks, i.e., jam sources, in urban traffic. If traffic bottlenecks are not alleviated at their early stages, they may evolve to regional congestion or even gridlocks<sup>3,4</sup> in the traffic network. Most existing studies have focused on the cause of traffic bottlenecks from a road-level perspective<sup>5–9</sup>, which have laid solid foundation for research on traffic congestion mechanism. Meanwhile, many related solutions have also been suggested and applied to bottleneck mitigation, such as diversion and lane restriction<sup>10</sup>, signal timing<sup>11</sup>, speed restriction<sup>12</sup>, and connected vehicles system<sup>13,14</sup>. Besides road-level perspective, some scholars have also sought for the emergence of network-level bottlenecks<sup>15</sup>. These studies often focused on uncovering the key elements that influence the spreading process in the network<sup>16</sup> to identify the potential bottlenecks. In recent years, studies based on percolation theory have also attempted to uncover the potential bottlenecks of urban traffic network by considering their effects on the global traffic organization in the city. For example, Li et al. manifested that bottlenecks in different locations can play different roles in the organization process of traffic networks, and addressing these network-level bottlenecks can significantly improve global traffic<sup>17</sup>; Hamedmoghadam et al. uncovered traffic bottlenecks by designing a percolation-based framework based on flow heterogeneity of the traffic network<sup>18</sup>. However, the practical mitigation of traffic bottlenecks is still complex due to the complexity behaviors of the travelers<sup>19,20</sup> and cascading jams in traffic networks<sup>21,22</sup>.

<sup>1</sup>School of Systems Science, Beijing Jiaotong University, 100044 Beijing, China. <sup>2</sup>Department of Physics, Bar-Ilan University, Ramat Gan, 52900, Israel. <sup>3</sup>School of Reliability and Systems Engineering, Beihang University, 100191 Beijing, China. <sup>4</sup>Azrieli School of Architecture, Tel Aviv University, Tel Aviv, 6997801, Israel.

<sup>5</sup>School of Traffic and Transportation, Beijing Jiaotong University, 100044 Beijing, China. <sup>6</sup>School of Economics and Management, Beihang University, 100191 Beijing, China. ✉e-mail: [havlins@gmail.com](mailto:havlins@gmail.com)

**Fig. 1 | Definition of a jam tree.** **a** Examples of several jam trees. Each directed link represents a road segment (indicated by an upper case letter), with an arrow indicating the direction of traffic flow on it. The number above each link indicates its jam duration, i.e., number of successive time intervals of traffic jam. Here, each time interval represents 10 min. If the difference in jam duration between two adjacent segments is less than a threshold (in the present study 2 intervals, i.e., 20 min), they are considered part of the same jam tree<sup>31</sup>. **b** Key information about the presented jam trees in (a). The table presents information of each jam tree, including its trunk, size, and cost. Each jam tree has only a single specific trunk based on the above definition. The size of a jam tree is determined by the number of segments that belong to it, including both the main trunk and associated branches, while the temporal cost of each jam tree is defined as the weighted sum cost of its trunk and branches (See Methods).



Since it can be important yet difficult to control network bottlenecks, researchers have also analyzed traffic jams based on their global patterns. Geroliminis and Daganzo<sup>23</sup> found a fundamental relationship called Macroscopic Fundamental Diagram (MFD) in global traffic (i.e., the relationship between average flow, average density, and average speed of all roads) in a region where the distribution of cars is relatively homogeneous. Based on MFD, many corresponding traffic control strategies, e.g., perimeter control, have been designed<sup>24–27</sup>. Zeng et al. found that traffic dynamics could represent multiple network percolating states based on percolation theory<sup>28,29</sup>. In addition, another study focused on the recovery behavior of traffic jams, which proposed a measure for evaluating the spatio-temporal resilience of urban traffic congestion<sup>30</sup>. They discovered a scaling law for the distribution of resilience, suggesting that there is an inherent universal behavior behind traffic resilience which is independent of details at the microscopic scale. A recent study of Serok et al. developed a method for identifying “traffic jam trees” based on the spatio-temporal relationships between the sources of the congestion (defined as bottlenecks) and the streets it affects<sup>31</sup>. While the study analyzed the overall daily impact (or “jam cost”) of traffic congestion in specific cities, the evolution of traffic congestion during the day is unclear yet critical for network jam formation.

In the present study, we extend the jam tree model of Serok et al.<sup>31</sup> and study congestion patterns in urban traffic of five different cities and an urban agglomeration region in China. We analyzed the evolution of jam tree cost (in vehicle hours), and particularly the characteristics of jam trees in different scenarios (i.e., rush hours and non-rush hours). We find that although the locations of the jam trees vary significantly from day to day, the distribution of their costs follow a similar pattern every day, i.e., a power law distribution with similar exponents. We also find that the exponents for the same city in different days are similar not only for the entire day, but also for the same hours in different days. On the other hand, they are different for different cities. This suggests that the daily pattern of jam tree exponents can be used as a fingerprint of the city, i.e., as a unique characteristic of a city’s traffic during a 24-h period. Our results suggest a new way to classify and compare the evolution of traffic in different cities, which can be useful for evaluating traffic reliability as well as for strategies for improving traffic management and control by observing the changes in the fingerprint.

Results

Modeling and measurements

The traffic jam tree model<sup>31</sup> relates the propagation of traffic congestion to the growth of a tree: the downstream road that becomes congested first

represents the trunk of a tree, while the upstream roads that become later congested represent the branches which sprout from the trunk. Here, we show based on the tree-like structure, the characteristics of the emergence, growth, and dissipation of the urban traffic congestion. By analyzing the jam trees, one can identify the trunks of trees at a given time (see Fig. 1). These can be regarded as the bottlenecks representing the origins of large-scale urban congestion.

In this paper, we extend the definition of jam trees to include additional types of jam trees and analyze their statistical properties using big real-world data of high resolution (i.e., 1 min). Specifically, we analyze here the traffic patterns of typical regions in China over the course of a full day. Our generalization of the jam tree model enables us to identify additional types of jam trees beyond those identified in ref. 31, and to determine their jam trunks, as well as their corresponding duration, sizes and costs. To identify the jam trees in urban traffic, our first step is to calculate the initial time and duration of the congestion (until a given time) for each road segment in the urban traffic network. The duration of the congestion is defined as the successive time that a road segment has been congested (i.e., the relative velocity of the road is below a predefined threshold<sup>17</sup>, see also Methods). The basic assumption is that the traffic congestion is initiated on a downstream road(s) and propagates over time to its neighboring upstream roads and then to its next nearest neighbors and so on. Therefore, we can identify and follow the evolution of each jam tree in the traffic network by analyzing the congestion on different road segments and the connectivity of the roads. The methodology of identifying the jam trees in urban traffic is in principle based on ref. 31. However, in this work, we extend the definition of a jam tree to include further general cases; in particular, cases where several jam trees may overlap and share the same trunk or branches. The core idea of our extended approach is to associate the influence of a jam to its potential trunk(s) in a reasonable way (see Fig. 1), that is, to share the influences from parallel jammed downstream roads or to determine the original influence from successive downstream roads. This approach is highly necessary as these cases are common, particularly in megacities with complex traffic patterns; however, they have not been discussed in the previous study. Further details are given in the Methods section.

As can be seen in Fig. 1, each letter (A to K) in the figure represents a road segment, with an arrow pointing in the direction of the traffic flow on that segment. The number above each link is the jam past duration (in time interval units) of the link at a given time. The table in Fig. 1b demonstrates how we calculate the tree size and cost. The size of a tree is defined as the number of road segments that this jam tree contains, while its cost is defined

**Table 1 | The number of trunks (trees) according to their sizes (daily average)**

| Size threshold | Beijing (52,968 links) | Shenzhen (22,248 links) |
|----------------|------------------------|-------------------------|
| $\geq 1$       | 34,647 (65.4%)         | 13,341 (60.0%)          |
| $\geq 5$       | 15,483 (29.2%)         | 6371 (28.6%)            |
| $\geq 10$      | 3850 (7.3%)            | 1905 (8.6%)             |
| $\geq 20$      | 416 (0.8%)             | 218 (1.0%)              |

The percentage in the parenthesis is derived considering the number of trunks out of the number of total links in each city (e.g., 65.4% = 34,647/52,968). These percentages represent the relative jam scale of a city during the day.

as the weighted sum cost of its related road segments. Following ref. 31, the cost of a single road segment is calculated by the vehicle hours (VH) spent in traffic jams at this road segment, which represents the total extra travel time taken by drivers to cross the road segment above the optimal travel time under normal traffic conditions. Thus, the cost of a link at a given time  $t$ ,  $C(t)$ , can be calculated by the equation<sup>31</sup>:

$$C(t) = d^* \left( \frac{1}{v(t)} - \frac{1}{v_{op}} \right) * \frac{q(t) * l}{\frac{60}{T}}, \quad (1)$$

where  $d$  is the length of the link,  $v(t)$  is the current velocity on the link at time  $t$ ,  $v_{op}$  is the optimal velocity on the link at which the flow in this link can be maximal according to the flow-density diagram (details can be found in the Supplementary Fig. S1 in the Supplementary Information),  $q(t)$  is the current flow on the link at time  $t$ ,  $l$  is the number of lanes in the link,  $T$  is the time interval that we set here as 10 min. Based on Eq. (1), the cost of congestion highly depends on the number of agents (i.e., vehicles or people) affected by the congestion, which reflects the severity of the congestion impact better than simply looking at speed. Note that in our dataset we do not have information on the number of lanes in each link. Therefore, we estimate this value based on the official rank of each link (details can be found in Methods). According to Eq. (1), it is necessary to initially evaluate traffic flow on each link. The flow  $q(t)$  on the link can be calculated by:

$$q(t) = v(t) * k(t), \quad (2)$$

where  $k(t)$  is the current car density on the link in time  $t$ . Besides, the relationship between the velocity and density on the link is based on the generalized car-following model assumed as<sup>32</sup>:

$$\left( \frac{v(t)}{v_f} \right)^{1-m} = 1 - \left( \frac{k(t)}{k_j} \right)^{l-1}, \quad (3)$$

where  $v_f$  is the free-flow velocity on the link and we assume it to be its velocity limit,  $k_j$  is the jam density on the link which is set as 150 vehicles per kilometer as suggested in earlier studies<sup>31,33</sup>. Different combinations of distance headway exponent  $m$  and speed exponent  $l$  represent utilizing different car-following models to describe the traffic flow characteristics in road segments. Therefore, the values of parameters  $m$  and  $l$  are used to determine the shape of speed-density relation curve, which are usually derived from empirical data and yield  $m = 0.8$  and  $l = 2.8$  in this case<sup>34</sup>. We have also checked the fundamental diagram of typical links under the above parameters, and find that the specific flow-density diagram can be well described (see Fig. S1 in the Supplementary Information for more details). Based on Eqs. (2) and (3), we can infer the relations between flow and velocity of each link (details can be found in the Supplementary Fig. S1 in the Supplementary Information), and therefore calculate the flow on each link at any specific time with our velocity dataset.

In this work, we extend the methodology to encompass additional scenarios involving the aggregation of segment costs linked to a specific tree.

Such scenarios were not addressed in the previous model. If a road segment can be associated to multiple trunks, e.g., link F in Fig. 1b, which is a shared branch of trunks A, C, and E, the cost of this segment is divided equally among all the trunks that the link can be associated with. In this case, 1/3 of the cost would be assigned to each trunk. This is since if a congested link (branch) is connected to more than one trunk, then each trunk could be the possible origin of this jam. We therefore consider each associated trunk to have the same probability to induce this congested link. Based on this definition, we identify the jam trees in two typical megacities in China (i.e., Beijing and Shenzhen) on different days. We summarize the daily number of the jam trees (i.e., number of trunks) according to their sizes in Table 1. In the following, we present results for additional three large cities as well as an urban agglomeration region.

### Jam patterns analysis

During its lifespan, a jam tree will grow and then dissipate. It is important to understand for each time during the day the information about jam trees, each connected to a specific trunk, i.e., what is their impact on the traffic system. In Fig. 2a, b, we demonstrate a typical jam tree of large influence during the day of October 26 in Beijing and Shenzhen, respectively. Next, to quantitatively address this issue, we begin by analyzing the evolving number of jam trees throughout the day. By calculating the number of trunks (each represents a jam tree) during the day, we find two distinct peaks: one between 7:00-9:00 and the other between 17:00-19:00 in both Beijing and Shenzhen, which represent the rush-hour windows in both cities (Fig. 2c, d). Note that, in both cities, we find more jam trees during the afternoon rush hours compared to the morning rush hours. This is probably because in the afternoon, individuals are heading to a wider variety of locations (such as their homes), as opposed to the morning when the travel is primarily directed towards central workplaces. The number of jam trees in each time unit (Fig. 2c, d) depicts the frequency of existing jam trees but it does not give us information about their actual importance (i.e., cost). For the cost of a jam tree, there are several important definitions covering different time and space scale, including the momentary cost, the total momentary cost and daily cost. The momentary cost of a jam tree is defined as the sum of the costs of all links that belong to its tree trunk at a specific time:

$$\text{MomentaryCost}(t) = \sum_{l \in \{\text{tree}\}} w \cdot C(t), \quad (4)$$

where  $\{\text{tree}\}$  is the set of all members of this jam tree (including a trunk and its associated branches),  $C(t)$  is the cost of a specific link,  $w = 1/m$  is the weight of the related link which is shared by  $m$  different trunks (Note that  $w$  is regarded as 1 for a trunk). The total momentary cost is therefore determined by the summed momentary cost of all the jam trees at a given time, which can be evaluated as:

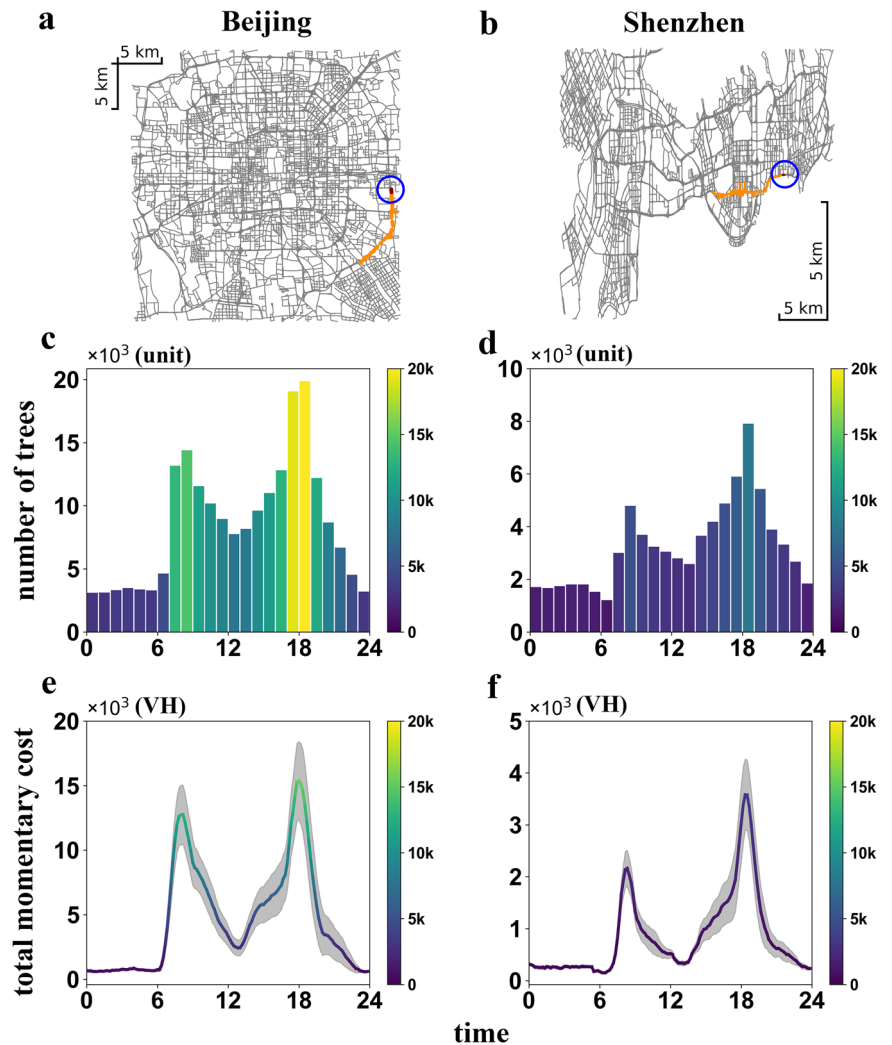
$$\text{TotalMomentaryCost}(t) = \sum_1^n \text{MomentaryCost}(t), \quad (5)$$

where  $n$  is the number of jam trees having tree size larger than 1 at a given time. The daily cost of a jam tree is defined as its cumulative momentary cost in a successive time window from the moment it appears until the moment it recovers.

$$\text{DailyCost}(t) = \sum_{t=T_0}^{t=T_1} \text{MomentaryCost}(t), \quad (6)$$

where  $T_0$  is the moment when its trunk appears and  $T_1$  is the moment when its trunk recovers. Thus, next we calculate the total momentary cost by summing up the momentary cost of all jam trees in the urban traffic network at a given time, and analyze how this value evolves during the day. As seen in Fig. 2e, f, the evolution of the total momentary cost also presents, for both cities, two peaks during rush hours. These results not only indicate, as

**Fig. 2 | Temporal evolution of jam trees during workdays.** Demonstration of a typical jam tree with large influence during the day in October 26 in **a** Beijing and **b** Shenzhen, where the dark red link represents the trunk (surrounded by blue circle) while the orange links represent the branches. The scale bars are also presented in both maps. The number of existing jam trees during the day (average on 17 working days with the resolution of 1 h) in **c** Beijing and **d** Shenzhen. The total momentary cost of jam trees during the day (with the resolution of 10 min) in **e** Beijing and **f** Shenzhen. The error-bars (gray) which are surprisingly small are the variances of the total momentary cost in 17 working days. It can be seen from (c) to (f) that although the number of jam trees at rush hours in Beijing is 2–3 times compared to Shenzhen, their corresponding cost is however 4–5 times higher, indicated by the color bars.



expected, that during rush hours the urban transportation network has significantly more and larger traffic jams, but also quantify their momentary cost. Moreover, when comparing Beijing and Shenzhen, we find that the total momentary cost during rush hours in Beijing is 4–5 times as that in Shenzhen, which is most probably because Beijing is a larger city. It can be also seen that in both cities the evening rush hour is more congested than the morning rush hour, with a higher total momentary cost during the evening rush hours of over 15,000 VH in Beijing and over 3000 VH in Shenzhen. However, it is worth noting, as seen in Supplementary Fig. S2 in the Supplementary Information, that when it comes to the average momentary cost at each moment, the situation can be rather different. In Shenzhen, the average cost during the two rush-hour periods is almost the same, while in Beijing the average momentary cost during the morning rush hours is larger than that during the evening rush hours.

Next, we focus on the distribution of tree cost of all jam trees during the day. Specifically, for a given day, we sum up the cost of jam trees associated with each specific trunk during the entire day and consider it as the daily cost associated with this trunk. The finding of highly cost trunks is important for identifying, managing, and improving the traffic in these locations which could help improve global urban traffic. As can be seen in Fig. 3a, b, for a typical working day (Oct. 26, 2015), we find that, for both cities, the probability density of daily cost of jam trees follows a power-law distribution:

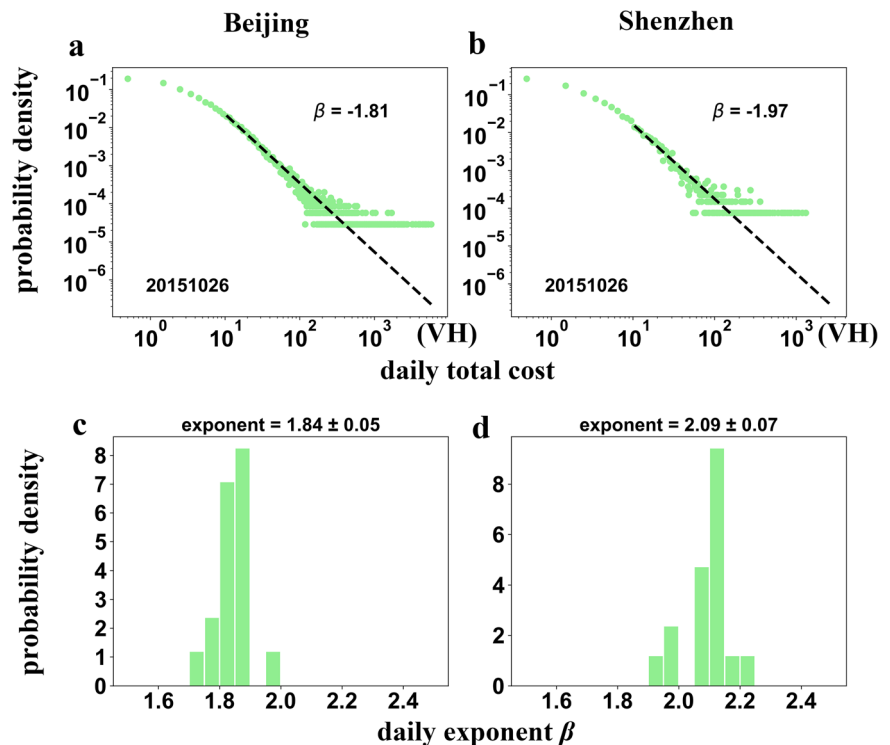
$$P(c) \sim c^{-\beta} \quad (7)$$

where  $c$  represents the cost,  $P(c)$  is probability density function of the cost distribution, and  $\beta$  is the corresponding distribution exponent, with the value of 1.81 and 1.97 for Beijing and Shenzhen respectively in this case. Furthermore, when considering the working days only (over a month), we also find that the exponents of all cost distributions of jam trees for individual days have very similar characteristics with similar exponents. As can be seen in Fig. 3c, d, the cost distribution exponent in Beijing is  $1.84 \pm 0.05$ , while in Shenzhen the value is  $2.09 \pm 0.07$ . We have also tested and verified the goodness-of-fitting of the power-law distributions (details can be found in the Supplementary Fig. S5 in the Supplementary Information). Considering the heavy-tail characteristic of a power-law distribution, the smaller the exponent is, the higher heterogeneity the distribution will be, which means a higher fraction of large-cost jam trees (i.e., the tail values) exists. Thus, the smaller exponent in Beijing suggests that the costs of jams in Beijing are higher compared to Shenzhen. We also calculate the exponents of three other cities including Shanghai, Hangzhou and Jinan, respectively. Results of above cities are shown in Table 2. It is vital to be convinced that the above conclusions will be valid for different configuration of model parameters. To test the sensitivity of different model parameters, we have also used other combinations of parameters, including both uniform parameters and non-uniform parameters, to calculate the jam cost. We find that the phenomenon unveiled by our proposed approach is still robust regardless of the specific microscopic model parameters (see Supplementary Figs. S6, S8 in the Supplementary Information).

An important finding of Serok et al.<sup>31</sup> is that the specific bottlenecks of traffic jams, i.e., the jam tree trunks, are poorly predictable in London and



**Fig. 3 | Distribution of jam tree costs in urban traffic.** The distribution of jam costs associated with a given trunk in a typical working day for all trunks in **a** Beijing and **b** Shenzhen. Results for other working days are similar and can be seen in Supplementary Fig. S3 (Beijing) and Supplementary Fig. S4 (Shenzhen) in the Supplementary Information. Probability density function of the jam cost exponent  $\beta$  for 17 working days for **c** Beijing and **d** Shenzhen. We also apply the two-sample Kolmogorov–Smirnov test (KS test) to evaluate if the distributions of daily exponents in two cities are indeed different and obtain a  $p$ -value of nearly 0 (i.e.,  $3 \times 10^{-8}$ ), which is much lower than the common accepted significance threshold of 0.05.



Tel Aviv. This finding is also verified here in our large datasets of Beijing and Shenzhen. As can be seen in Supplementary Fig. S9 in the Supplementary Information, it is found that the most influential jam trees (i.e., the trunks with cost above 20 VH) that appear repeatably on all five workdays only account for less than 20% of the total large trees. This suggests that, although the impact of the heavily costly trunks (i.e., cost distribution) can be similar from day to day, most of the specific heavily costly trunks (i.e., spatial locations), about 80%, do not usually recur in different days. Thus, it is difficult to forecast their spatio-temporal location based on historic data. We also calculate the Jaccard Index, often applied to analyze the similarity between two sample sets, to evaluate the overlap of tree trunks between every pair of working days in our dataset. We find that large traffic jams are usually induced by different influential tree trunks on different days (Supplementary Fig. S10 in the Supplementary Information). Particularly, during specific periods (i.e., rush hours or non-rush hours), the overlap of tree trunks with large costs is even lower. Based on our definition of a jam tree, a specific trunk is considered as a specific bottleneck at a fixed location. Therefore, the low result of overlapping trunks on different days suggests that the locations of these bottlenecks, especially the influential ones, are different on different days.

Next, we wish to analyze the evolution of cost during a day in a higher resolution of temporal scale, since the overall macroscopic pattern of traffic congestion in a city (i.e., the distribution of costs throughout the day as in Fig. 3) remains stable and consistent from day to day. To this end, we first tested if in a more specific temporal scale, i.e., in the scale of 20 min, the jam cost still follows the power-law distribution suggested in Eq. (7). We find that the momentary cost during these periods can still be well-fitted in different typical periods, as can be seen in the Supplementary Fig. S11 in the Supplementary Information. Next, we fit and calculate the exponent  $\beta$  for every 20 min during a day. In this way, we obtain one curve indicating the

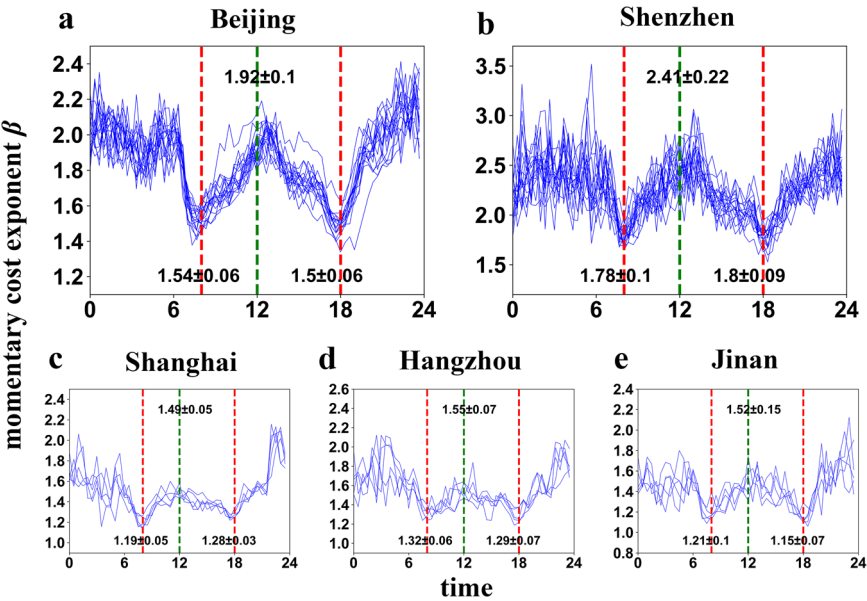
evolution of exponent  $\beta$  during the entire day. We then plot the evolution of  $\beta$  for every working day. We find a similar evolution pattern of  $\beta$  in different days. It is seen that the evolution curve for different days is highly concentrated to form a “W-like” shape curve, where the values of  $\beta$  are typically lower during the morning and evening rush hours. We also find that this “W-like” shape is universal for different cities, as seen in Fig. 4. It is worth noting that the patterns and the  $\beta$  values are similar for the same city on different working days, but are with significantly different  $\beta$  values for different cities. Thus, the daily pattern of all exponents of momentary jam costs during the day for traffic jams can be used as a fingerprint for urban traffic, i.e., jam-prints. The “W-like” shape curves for the momentary cost exponent in each day are also consistent with the results found in Fig. 2, indicating that traffic jams are significantly heavier during morning and evening rush hours. Focusing on the evolution of the exponent of the tree cost distribution in different cities, it is interesting to observe that in Shanghai, the traffic situation is typically worse during the morning rush hours than during the evening rush hours. This is represented by smaller momentary cost exponents in the morning rush hour period. However, for the other cities, the momentary cost exponents during the two periods are similar. Overall, the momentary cost exponent for a given time is typically larger in Shenzhen than that in the other cities, suggesting a better traffic flow in Shenzhen compared to other cities.

To test if the congestion pattern of a specific city is unique and characterizes the city, we calculate for different cities, the  $p$ -values for the cost distribution exponents during rush hours. As shown in Table 3, by comparing the distribution of the momentary cost during the morning and evening rush hours in each pair of cities (the results in non-rush hours can also be seen in Supplementary Table S3 in the Supplementary Information), we can conclude that generally, different cities have different characteristic values of cost distribution exponents. Although some cities may have similar jam features over the whole day (see Supplementary Table S2 in the Supplementary Information) or during a certain period, for instance, morning rush hours in Hangzhou and Jinan or evening rush hours in Shanghai and Hangzhou, we did not find a case that a pair of cities have similar distributions throughout both rush hour periods. These results suggest that the congestion pattern in each city during the day, as shown in Fig. 4, is unique for a city, and can also serve as a jam-print for a city.

**Table 2 | Daily cost distribution exponent  $\beta$  of five different cities**

| City    | Beijing         | Shenzhen        | Shanghai        | Hangzhou        | Jinan           |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| $\beta$ | $1.84 \pm 0.05$ | $2.09 \pm 0.07$ | $1.80 \pm 0.05$ | $1.78 \pm 0.07$ | $1.75 \pm 0.05$ |

**Fig. 4 |** Jam patterns of tree cost during the day can be considered as the city jam-prints of urban traffic. The temporal evolution of the exponent of the cost distribution of jam trees during the day is shown, in 5 different cities **a** Beijing (over 17 days), **b** Shenzhen (over 17 days), **c** Shanghai (over 5 days), **d** Hangzhou (over 5 days) and **e** Jinan (over 5 days). The numbers in the figures are the mean and standard deviation of daily exponents at three typical time windows (20 min for each) indicated by dash lines, at 8:00AM and 18:00PM for rush hours (red) and at 12:00 for non-rush hours (green).



**Table 3 |** The *p*-values for cost distribution exponents in morning/evening rush hours of pairs of cities

|          | Beijing | Shenzhen | Shanghai | Hangzhou | Jinan  |
|----------|---------|----------|----------|----------|--------|
| Beijing  | 1.000   | <0.001   | <0.001   | <0.001   | <0.001 |
| Shenzhen | <0.001  | 1.000    | <0.001   | <0.001   | <0.001 |
| Shanghai | <0.001  | <0.001   | 1.000    | 0.008    | 0.873  |
| Hangzhou | <0.001  | <0.001   | 0.079    | 1.000    | 0.008  |
| Jinan    | <0.001  | <0.001   | 0.008    | 0.008    | 1.000  |

**Note:** The yellow color represents results comparing the morning rush hours (i.e., 7:00-9:00AM) while green represents results comparing the evening rush hours (i.e., 17:00-19:00PM). If the *p*-value is lower than threshold of 0.05, the two distributions are regarded as well separated.

**Generalization to larger scale regions**

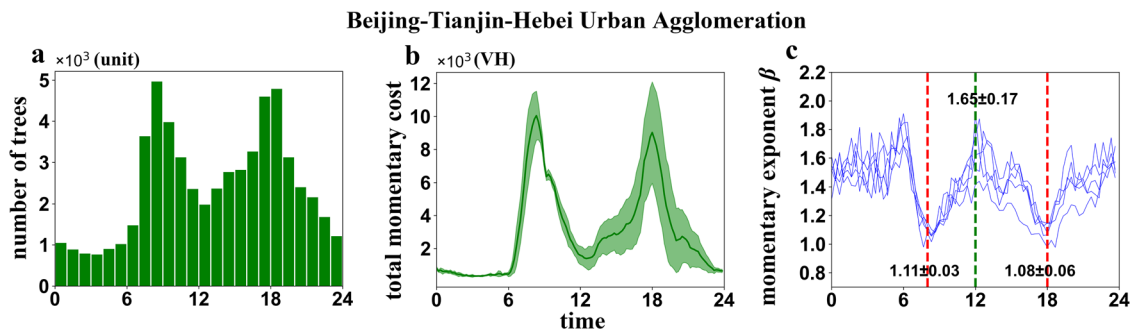
Besides single cities, we also consider if the “jam-print” can be generalized to larger scale regions, i.e., the urban agglomeration. Urban agglomeration is a highly developed spatial form of integrated cities where the cooperation among these cities is highly shifted, which renders the region one of the most important carriers for global economic development<sup>35</sup>. Here we focus on the traffic jams on inter- and intra-city highways of a typical urban agglomeration region, i.e., the Beijing-Tianjin-Hebei Urban Agglomeration of China. We first calculate the temporal evolution of jam trees during the workdays, as can be seen in Fig. 5a, b. Similar to single city results, one can observe two typical peaks during the rush hours representing heavier traffic jams during these two periods. However, it is interesting to note that the daily variation of total momentary cost during evening rush hours is much higher than that during morning rush hours. This may indicate that during morning rush hours, traffic flow tends to be more focused or intense, while in the evening traffic jams on highways might occur more sporadically. This could be related to the more dispersed nature of trips in the evening, as people travel to a variety of destinations rather than predominantly to work locations as in the morning.

Next, we analyze the evolution of exponent  $\beta$  during the entire day in urban agglomeration region. We mainly concentrate on jam propagation on the highways in this region. We still calculate the exponent  $\beta$  of momentary tree cost distribution every 20 min, and analyze how it evolves over time

during the day. As shown in Fig. 5c, the evolution of exponent  $\beta$  of this region also exhibit a similar “W-like” pattern for five different workdays. It can be noticed in this region that the average values of  $\beta$  during morning and evening rush hours are even smaller than any single city aforementioned, with values of approximately 1.11 and 1.08 respectively. This may represent the possible heavier jams in the highways than in the urban central areas. However, during other periods the highways would recover to a more efficient state (e.g., with exponent  $\beta$  around 1.65 at noon). Combination of these different exponents can therefore represent a new jam-print of this urban agglomeration region.

**Discussion**

In order to uncover the macroscopic patterns, as well as the spatio-temporal evolution, of traffic congestions in different cities, we study the congestion patterns using real-world data based on our extended jam tree model which analyzes the propagation of traffic jams associated with specific bottlenecks. Previous studies have focused more on the clusters in the road network and suggest also some universal critical behaviors in the urban traffic flows<sup>28–30</sup>. However, the jam tree model is vital since it applies stricter spatio-temporal relations between congestion links. Indeed, we find that the number of jam trees and the number of connected jam clusters are not always positively correlated, as seen in the Supplementary Note 9, as well as Supplementary Figs. S12 and S13 in the Supplementary Information. Compared to the



**Fig. 5 | Results of jam patterns in highways of an urban agglomeration region.** **a** The number of existing jam trees during the day (averaged over 5 working days with the resolution of 1 h). **b** The total momentary cost of jam trees evolving with time during the day (with the resolution of 10 min). The error-bars (lighter green) are the variances of the total momentary cost in 5 working days. **c** The temporal

evolution of the exponent  $\beta$  of the cost distribution of jam trees during the day. The numbers in the figures are the mean and standard deviation of different days at three typical time windows (20 min for each) indicated by dash lines, at 8:00AM and 18:00PM for rush hours (red) and at 12:00 for non-rush hours (green).

previous jam tree model in ref. 31, our extended model integrates additional realistic jam components, which appear frequently in traffic of large-scale urban area. Particularly, it addresses the cases which frequently appears in complex traffic systems such as traffic networks in megacities or even urban agglomerations, e.g., a trunk and its adjacent branches have the same length of jam duration, or several jam trees share the same trunk or branches (for demo see Fig. 1). As a demonstration, we calculate the number of branches as well as their total costs that are associated with multiple trunks in the road network of Beijing city. We find that the number of branches and their momentary costs recognized by our new methods account for over 10% and 25% of all jam trees, respectively (see Supplementary Note 10 and Supplementary Fig. S14 in the Supplementary Information). These results suggest that the previous model neglected the impact of significant jam trees that are typical in road networks of big cities. Meanwhile, combined with the extended model and massive real-time traffic data in typical districts, we find that the daily distribution of the costs associated with traffic jams follows a consistent pattern, i.e., a power law with similar exponents in different days, which suggest a possible universal characteristic of traffic congestion daily evolutions that has not been studied in earlier models. Based on the extended model, we study the time evolution of both, the number of existing jam trees and their cost during the day. We also calculate the distribution of the cost of jam trees during a day, and find that the cost distribution of the jam trees follows a power-law distribution based on Eq. (7), with a similar daily exponent for each of the analyzed cities, but different for different cities. Furthermore, we analyze the evolution of the cost distribution exponent during the entire day. We find that the patterns of the cost distribution exponent of the jam trees during the day, can be considered as a jam-print of the urban traffic in a city. This is since the patterns of the evolution of the exponent are similar each day for the same city, but different for different cities. To the best of our knowledge, our approach is unique in looking at exponents for distributions of momentary jam trees. This helps us to define a running parameter that characterize both the congestion characteristics in a given city and distinguish between different cities. As a comparison, earlier studies have neither applied statistical analysis to represent the jam patterns of urban traffic, nor focused on different times during the whole day of jam impacts. The analysis of the whole evolution process of traffic jams during the day is significant, since only the whole day exponent cannot capture the congestion pattern of a city (see Supplementary Table S2 in Supplementary Information). These unique patterns of traffic in urban areas which we suggest as a jam-print of a specific urban traffic may provide valuable insights for assessing the quality of traffic in the city, and for establishing new traffic management goals and assessing the effectiveness of various traffic improvement strategies. For example, these universal yet unique evolutionary behaviors can be used to immediately identify anomalies in traffic when a deviation from the universal curve occurs and alarm the traffic control of the city.

The consistent power-law scaling observed for the cost distributions of traffic jams across different days in each city may be linked to self-organized criticality<sup>36,37</sup> in the organization of urban traffic. On one hand, an increase in traffic demand during a specific period may lead to the spontaneous appearance of congestion; on the other hand, the individuals seeking the best possible route may redistribute the traffic and drive the system to an optimal state under current traffic conditions. The balance between these two forces could therefore push the traffic system towards its intrinsic operational limits.

We find that each city has its own distinct macroscopic pattern of traffic congestion that is consistent on a daily basis at the same time of day, while the specific locations of traffic jams tend to vary. We hypothesize that the daily variations at the microscopic level might be associated with the fluctuation of human behavior and to the slight daily changes of origin-destination (OD) demand. This phenomenon is analogous to phase transition near criticality, where the distribution of finite clusters consistently follows the same power law while the spatial location of these finite clusters varies significantly in different realizations<sup>38</sup>. Note that power law distributions appear also in other urban features, e.g., in population of cities<sup>39</sup>. Future studies may attempt to focus on uncovering the topological and dynamical mechanisms that could explain the jam-prints of city patterns.

## Methods

### Data description

Our data covers five large cities (i.e., Beijing, Shenzhen, Shanghai, Hangzhou and Jinan) and an urban agglomeration (i.e., Beijing-Tianjin-Hebei Urban Agglomeration Region) in China. Our data includes the topology information of road networks and the operational information of traffic dynamics. We obtained information on each road segment which include the road identification information, direction, length, and rank. The identification information of each road segment includes the identification number of its segments, its source intersection, and its target intersection. The operational information is the real-time velocity records of each road segment with a resolution of 5 min up to 1 min, which is obtained from the global positioning system (GPS) data recorded by floating cars. Specifically, the scales of the road networks and the coverage period of velocity records are shown in Table 4.

### Dynamical traffic network

Based on the topology and operational information of the urban traffic system, we constructed the dynamic traffic network of each region at every moment. For each city, we regard the road segments as network links and the intersections as the network nodes. Then, we constructed the network representation based on their connecting relations. Based on the above network representation, we assign a real-time velocity of each road segment

**Table 4 | Key information of the dataset of different cities**

| Region   | Scale  | Duration   |
|--|--|--|
| Beijing  | road segments: 52968<br>intersections: 27877 | 17 working days in Oct. 2015 (resolution of 1 min) |
| Shenzhen   | road segments: 22248<br>intersections: 12337 | 17 working days in Oct. 2015 (resolution of 1 min) |
| Shanghai   | road segments: 50469<br>intersections: 25729 | 5 working days in Oct. 2015 (resolution of 1 min)  |
| Hangzhou   | road segments: 35815<br>intersections: 18591 | 5 working days in Oct. 2015 (resolution of 5 min)  |
| Jinan  | road segments: 22690<br>intersections: 12063 | 5 working days in Oct. 2015 (resolution of 5 min)  |
| Beijing-Tianjin-Hebei Urban Agglomeration (highways) | road segments: 66925<br>intersections: 64204 | 5 working days in Mar. 2023 (resolution of 1 min)  |

**Table 5 | Rank and estimated number of lanes for each link (road segment)**

| Rank | Road type           | Estimated lanes |
|------|---------------------|-----------------|
| 1    | Inter-city highways | 4               |
| 2    | Intra-city highways | 4               |
| 3    | National roads      | 3               |
| 4    | Provincial roads    | 2               |
| 5    | County roads        | 1               |
| 6    | Others              | 1               |

to its corresponding link, which can be regarded as the weight of the corresponding link. We used the relative velocity of each road segment as the link weight for normalization. The relative velocity of a given road segment is calculated by dividing its current velocity by its velocity limit. We chose the 95th percentile of the measured velocity during a day as this limit. This is to avoid the influence induced by extreme values. By this, we construct the weighted dynamical traffic network of each region at every moment, where the network topology is static while the network's dynamics is represented by the varying weight of each link.

#### Status of the link

In the dynamic traffic network, we consider the status of a link as either functional or congested based on its average weight within a certain time interval (here 10 min). Specifically, we regard the link weight of 0.5 as the relative velocity threshold - a link with an average weight lower than 0.5 is regarded as congested. It is worth noting that the velocity value of a given road segment could be vacant for some moments due to missing data. In this case, we assume that if such a road segment is located between two congested segments, and has the same traffic direction as the congested segments have, it is also highly likely to be congested; otherwise, we will consider the traffic flow on this road segment as fluent. Following this assumption, we can identify the status of every link in the dynamic traffic network for every time interval during the day.

#### Trunk and branches of a jam tree

After identifying the status of each link in the network, we analyzed the properties of jam trees in the dynamic traffic network. The propagation of traffic jams, from downstream roads to upstream roads, is similar to the growth process of a tree. The origin of the traffic jam (which we define as the trunk of the jam tree) is assumed to be the road segment in its downstream with the longest congestion duration compared to other upstream parts of the traffic congestion (which we define as the branches of the jam tree), as shown in Fig. 1. By following this principle, we can identify every jam tree in the dynamical traffic network and locate its trunk and branches by using the following steps:

(i) Calculating the jam duration of each link. The congestion duration of a link is the number of successive time intervals the link has been congested until the current moment.

(ii) Locating the trunk of a jam tree. For a congested link, if there is no neighboring link with a longer congestion duration along its downstream direction, it is assumed to be the trunk of a jam tree.

(iii) Identifying branches belonging to a trunk. Since the traffic congestion propagates from a downstream road to an upstream road, the jam duration of the upstream road should always be shorter than its downstream neighbors which belong to the same jam tree. Based on the spatio-temporal relations, the difference between two congestion durations of neighboring links must be below a reasonable time threshold  $\theta$ ; otherwise, it can be argued that the congestion of the upstream road may have not been caused by its downstream neighbor. Here we set this threshold  $\theta$  as 2 units of time intervals (i.e., 20 min)<sup>31</sup>. We identify the nearest neighbor branches of a given jam tree trunk by examining all links in the upstream direction and selecting those that are equal or below the duration threshold criteria to belong to that tree. We proceed to evaluate if any adjacent links connected to the identified branches are also part of the same jam tree, using the duration threshold criteria. We repeat this process iteratively, adding new branches to the jam tree until no further branches meet the criteria, and then go back to the first stage, looking for congested links that are not connected to any of the identified jam trees.

#### Estimation of number of lanes in each road

Since in our dataset we do not have information on the number of lanes in each link, we estimate this value based on the rank of each link as in Table 5. The number of lanes is estimated by technical rules of the highway engineering office in China. Typically, the number of lanes is 1–4, and the road with higher service level (i.e., which can be related to the rank in the data) are often designed with more lanes. From high to low service level, the roads in China are often divided into inter- and intra-city highways (with rank = 1 and 2 respectively), national roads (rank = 3), provincial roads (rank = 4), county roads (rank = 5), and other roads (rank = 6). Both inter- and intra-city highways are the skeleton of the transportation system and are of the highest service level, yet we assume them to be 4-lane roads; national roads should be of lower service level than highways, and they are assumed as 3-lane roads; the service level of provincial roads is often between that of national roads and county roads, yet they are assumed as 2-lane roads. County roads and other roads are therefore assumed as 1-lane roads.

#### Data availability

The necessary data generated in this study have been deposited in the Github database <https://github.com/GuanwenZeng/Jam-prints> to enable the reproducibility of the results in the paper. The raw velocity data are protected and unavailable due to data privacy laws.



## Code availability

We provide the source codes of our algorithms in GitHub <https://github.com/GuanwenZeng/Jam-prints> to enable the reproducibility of our findings.

Received: 7 April 2024; Accepted: 18 March 2025;

Published online: 28 March 2025

## References

- Buldryev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).
- Gao, J., Bashan, A., Shekhtman, L., & Havlin, S. *Introduction to Networks of Networks* (IOP Publishing, 2022).
- Daganzo, C. F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transp. Res. Part B Methodol.* **41**, 49–62 (2007).
- Shi, J. et al. Simulation and analysis of the carrying capacity for road networks using a grid-based approach. *J. Traffic Transp. Eng.* **7**, 498–506 (2020).
- Vickrey, W. S. Congestion theory and transport investment. *Am. Econ. Rev.* **59**, 251–260 (1969).
- Gazis, D. C. & Herman, R. The moving and “phantom” bottlenecks. *Transp. Sci.* **26**, 223–229 (1992).
- Arnott, R., De Palma, A. & Lindsey, R. A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *Am. Econ. Rev.* **83**, 161–179 (1993).
- Newell, G. F. A moving bottleneck. *Transp. Res. Part B Methodol.* **32**, 531–537 (1998).
- Helbing, D. Traffic and related self-driven many-particle systems. *Rev. Mod. Phys.* **73**, 1067 (2001).
- Davis, L. C. Mitigation of congestion at a traffic bottleneck with diversion and lane restrictions. *Phys. A Stat. Mech. Appl.* **391**, 1679–1691 (2012).
- Yuan, S., Zhao, X. & An, Y. Identification and optimization of traffic bottleneck with signal timing. *J. Traffic Transp. Eng.* **1**, 353–361 (2014).
- Strnad, I., Kramar Fijavž, M. & Žura, M. Numerical optimal control method for shockwaves reduction at stationary bottlenecks. *J. Adv. Transp.* **50**, 841–856 (2016).
- Piacentini, G., Goatin, P. & Ferrara, A. Traffic control via moving bottleneck of coordinated vehicles. *IFAC-PapersOnLine* **51**, 13–18 (2018).
- Grumert, E. F. & Tapani, A. Bottleneck mitigation through a variable speed limit system using connected vehicles. *Transp. A Transp. Sci.* **16**, 213–233 (2020).
- Duan, J. et al. Spatiotemporal dynamics of traffic bottlenecks yields an early signal of heavy congestions. *Nat. Commun.* **14**, 8002 (2023).
- Kitsak, M. et al. Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893 (2010).
- Li, D. et al. Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proc. Natl Acad. Sci.* **112**, 669–672 (2015).
- Hamedmoghadam, H., Jalili, M., Vu, H. L. & Stone, L. Percolation of heterogeneous flows uncovers the bottlenecks of infrastructure networks. *Nat. Commun.* **12**, 1254 (2021).
- Braess, D., Nagurney, A. & Wakolbinger, T. On a paradox of traffic planning. *Transp. Sci.* **39**, 446–450 (2005).
- Roughgarden, T. *Selfish routing*. Doctoral dissertation (Cornell University, 2002).
- Daqing, L., Yinan, J., Rui, K. & Havlin, S. Spatial correlation analysis of cascading failures: congestions and blackouts. *Sci. Rep.* **4**, 1–6 (2014).
- Duan, J., Daqing, L. & Hai-Jun, H. Reliability of the traffic network against cascading failures with individuals acting independently or collectively. *Transp. Res. Part C Emerg. Technol.* **147**, 104017 (2023).
- Geroliminis, N. & Daganzo, C. F. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transp. Res. Part B Methodol.* **42**, 759–770 (2008).
- Haddad, J. & Geroliminis, N. On the stability of traffic perimeter control in two-region urban cities. *Transp. Res. Part B Methodol.* **46**, 1159–1176 (2012).
- Aboudolas, K. & Geroliminis, N. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transp. Res. Part B Methodol.* **55**, 265–281 (2013).
- Hamedmoghadam, H., Zheng, N., Li, D. & Vu, H. L. Percolation-based dynamic perimeter control for mitigating congestion propagation in urban road networks. *Transp. Res. Part C Emerg. Technol.* **145**, 103922 (2022).
- Gao, S., Li, D., Zheng, N., Hu, R. & She, Z. Resilient perimeter control for hyper-congested two-region networks with MFD dynamics. *Transp. Res. Part B Methodol.* **156**, 50–75 (2022).
- Zeng, G. et al. Switch between critical percolation modes in city traffic dynamics. *Proc. Natl Acad. Sci.* **116**, 23–28 (2019).
- Zeng, G. et al. Multiple metastable network states in urban traffic. *Proc. Natl Acad. Sci.* **117**, 17528–17534 (2020).
- Zhang, L. et al. Scale-free resilience of real traffic jams. *Proc. Natl Acad. Sci.* **116**, 8673–8678 (2019).
- Serok, N., Havlin, S. & Blumenfeld Lieberthal, E. Identification, cost evaluation, and prioritization of urban traffic congestions and their origin. *Sci. Rep.* **12**, 1–11 (2022).
- May, A. D. & Keller, H. E. Non-integer car-following models. *Highw. Res. Rec.* **199**, 19–32 (1967).
- Long, J., Gao, Z., Zhao, X., Lian, A. & Orenstein, P. Urban traffic jam simulation based on the cell transmission model. *Netw. Spat. Econ.* **11**, 43–64 (2011).
- Brackstone, M. & McDonald, M. Car-following: a historical review. *Transp. Res. Part F Traffic Psychol. Behav.* **2**, 181–196 (1999).
- Fang, C. & Yu, D. Urban agglomeration: An evolving concept of an emerging phenomenon. *Landsc. Urban Plan.* **162**, 126–136 (2017).
- Bak, P., Tang, C. & Wiesenfeld, K. Self-organized criticality: An explanation of the 1/f noise. *Phys. Rev. Lett.* **59**, 381 (1987).
- Bak, P., Tang, C. & Wiesenfeld, K. Self-organized criticality. *Phys. Rev. A* **38**, 364 (1988).
- Bunde, A. & Havlin, S. *Fractals and Disordered Systems* (Springer, 1991).
- Batty, M. Rank clocks. *Nature* **444**, 592–596 (2006).

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grants 72225012, 72288101, 71890973/71890970), the Israel Science Foundation (Grant No. 189/19), the Binational Israel-China Science Foundation (Grant No. 3132/19), Science Minister-Smart Mobility (Grant No. 1001706769), and the European Union’s Horizon 2020 research and innovation programme (DIT4Tram, Grant Agreement 953783). G.Z. and J.D. are supported by the Talent Fund of Beijing Jiaotong University (Grants 2024XKRC065 and 2024XKRC067 respectively).

## Author contributions

G.Z. and N.S. contributed equally to this work. G.Z., N.S., E.B.L. and S.H. conceived and designed the research. G.Z., N.S., E.B.L., D.L. and S.H. implemented the methods. G.Z., N.S., J.D., S.L. and S.S. performed and checked the experiments. G.Z., N.S., E.B.L., J.D., S.L., S.S., D.L. and S.H. analyzed the data. G.Z., N.S., E.B.L., D.L. and S.H. wrote the paper with comments from other authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42005-025-02049-6>.

**Correspondence** and requests for materials should be addressed to Shlomo Havlin.

**Peer review information** *Communications Physics* thanks Sandro Sousa and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025