DIT4TraM

Distributed Intelligence & Technology
for Traffic & Mobility Management

# D1.2: Methodologies based on network science for evaluating and improving urban traffic

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 953783.

Distributed Intelligence & Technology
for Traffic & Mobility Management

| Project | DIT4TraM |
|---|---|
| **Grant Agreement No.** | 953783 |
| **Start Date of Project** | 01-09-2021 |
| **Duration of the Project** | 36 months |
| **Deliverable Number** | 1.2 |
| **Deliverable Title** | Methodologies based on network science for evaluating and improving urban traffic |
| **Dissemination Level** | Public |
| **Deliverable Leader** | BIU |
| **Submission Date** | |
| **Author** | Shlomo Havlin |
| **Co-author(s)** | Efrat Blumenfeld Lieberthal |

## Release Approval

| Name | Role | Date |
|---|---|---|
| Prof. Eleni I. Vlahogianni | Reviewer | 27/2/23 |
| Prof. Ludovic Leclercq | Reviewer | 23/2/23 |
| Prof. Dirk Helbing | Reviewer | 27/2/23 |
| | | |
| | | |
| | | |

## Document History

| Version | Description | Name |
|---|---|---|
| V0.1 | Draft version | D1.2 report v0.docx |
| V0.2 | Draft version | D1.2 report v2.docx |
| V0.3 | Final version | D1.2 report |

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 953783.

## Legal Disclaimer

# Contents

1. Introduction. ....5
1.1. State of the Arts....5
2. Methodology....9
2.1. Jam Tree – Basic method....9
2.2 Economic Cost - Prioritization Strategy....11
2.3 Extended Definition of Jam Trees....12
2.4 Spatio-Temporal Evolution of Jam Trees....14
3. Results....15
3.1. Macro-stability....16
3.2. Meso-dynamics....17
3.3. Urban Fingerprints....18
3.4. Asymmetry in the congestion dynamics....23
3.5 Correlation between spatial and temporal evolution....26
4. Summary of the methodology....29
5. An example - Utrecht....31
6. An algorithmic approach for the adaptation / self-calibration of computer agents....33
7. Publications for D1.2....37
8. Code repository....38
9. Bibliography....39

# 1. Introduction

One of the challenges in assessing urban traffic is developing universal tools that apply to different types of urban forms, transportation systems, drivers' habits, etc. In this report, we address the assessment methodology for qualitative and quantitative data-driven analyses which is based on the percolation methodology.

Here, we present a novel methodology for identification, cost evaluation, and thus, prioritisation of congestion origins, i.e., their bottlenecks. The presented work is based on network analysis of the entire road network from a global point of view. We identify and prioritise traffic bottlenecks based on big data of traffic speed retrieved in near-real-time. Our approach highlights the bottlenecks that have the most significant effect on the global urban traffic flow. We follow the evolution of every traffic congestion in the entire urban network and rank all the congestions, based on the cost they cause (in Vehicle Hours units). We show that the macro-stability that represents the seeming regularity of traffic load both in time and space, overshadows the existence of meso-dynamics, where the bottlenecks that create these congestions usually do not reappear on different days or hours. Thus, our method enables to identify in near-real-time both recurrent and nonrecurrent congestions and their sources.

## 1.1. State of the art

The 21st century can be characterized as the century of the cities. Since 2008, more than 50% of the world's population lives in urban areas. The increasing urbanization process (with urban population annual growth rate of about 1.8, based on the world bank estimations) is accompanied by the growing usage of vehicles, which leads to a significant increase in traffic congestion in cities around the world (1–4). The price of congestion is the enormous time spent on roads (5–8), as well as the increasing fuel consumption, air pollution, and Carbone Dioxide emission (6–8). Current technological development gave hope that autonomous cars will solve congestion problems as they were expected to reduce the number of private cars by increasing car-sharing. Recent studies, however, suggest that this is not the case (9–12). There exists extensive work in various disciplines, e.g. urban planning (2), traffic (13–15), complexity, and networks (16–22) that aims at reducing traffic congestion generally, and in urban areas in particular.

The work on identifying traffic bottlenecks has been developed from studying freeway bottlenecks, through urban active bottlenecks, and lately, with the availability of big data — to near real-time identification of traffic bottlenecks. Many studies on freeway identification of traffic bottlenecks suggested evaluating traffic attributes such as flow, speed, or the differences between the travel duration in the road upstream and downstream (23–25, 26). These methods, however, cannot be applied directly to urban areas due to the different patterns of the road network (e.g., freeways have no intersections of traffic lights) and the travel behavior on it. Hence, other methods were proposed to identify urban bottlenecks (27-34). For example, Lee et al. (27) implemented a mining model which defined spatiotemporal traffic bottleneck (STB), and thereafter developed three methods to identify STBs in urban networks. Tao et al. (29), used the Cell Transmission Model (CTM) theory, where the Average Journey Velocity was selected as the measurement of congestion. The availability of big data, retrieved from traffic flow sensors, intrigued new methodologies that use data-driven techniques to identify urban bottlenecks. These works propose new methodologies that may be developed into tools and implemented in real traffic control systems to relieve congestion and enhance the network performance. Such works employed correlation tests and the implementation of a Dynamic Bayesian network to overcome the lack of data for the entire urban street network (e.g., 26, 30–34). Ma et al. (33) combined complex network theory with a user equilibrium model to analyse the evaluation process of traffic bottleneck. Chen et al. (26) proposed a method to identify traffic bottlenecks by modelling causal relationships between traffic flow sensors located in urban areas. For that, they estimated transfer entropy among these sensors, and constructed causality graphs to identify traffic bottlenecks and discover congestion propagation patterns.

Existing traffic-management solutions that optimise traffic lights address each intersection individually and use bottom-up solutions such as synchronisation and slotting to mitigate local congestion (35). Currently, there is a lacuna in providing an approach that prioritizes specific bottlenecks over the others, in order to optimize the entire road network in near-real-time, as well as to provide a dynamic road pricing that charges each vehicle according to its unique effect on the entire system. As explained by Hamilton: “When a holistic view of traffic management is taken, individual junction efficiencies can suffer to improve the state of the network as a whole... A strategic view of the entire urban network, with improved detection and communication technologies, is required to enter the next evolution of urban traffic control" (36). Recent work has tried to address the optimisation of traffic management solutions. For example, Backfrieder et al. (37) developed a forecasting algorithm that identifies expected bottlenecks before a traffic jam emerges. It is based on origin–destination data of the vehicles and assumes utilisation of vehicle-to-X communication for transmission of contemporary vehicle data, such as route source and destination or current position, as well as for the provision of the routing advice for vehicles. Zhao et al. (38) also focused on urban bottlenecks. They divided the urban road network into a uniform orthogonal grid and identified sources of traffic jams in specific cells. Li et al. (39) developed a method to identify traffic jams bottlenecks based on the percolation process while using big data, retrieved in real-time, of traffic speeds. They address the issue of how local traffic flows organize collectively into a global urban flow and refer to this process as "traffic percolation". Hamedmoghadam et al. (40) studied the way heterogeneity of flow demand affects the network flow dynamics under congestion. They used a percolation approach to identify the bottlenecks with the highest impact on the network flows.

In this present work, we developed a methodology to follow in near-real-time and simultaneously the evolution of every traffic congestion in the entire urban network, and rank all the traffic congestions, based on their cost (in vehicle hours (VH) units). We find that non-recurrent traffic congestion incidents dominate the urban traffic and, therefore, an efficient real-time identification of traffic congestion is critically needed. Our method is innovative as it uses a new strategy, which overcomes the challenges that the near-real-time identification problem poses. Specifically, our method is innovative in two main aspects: (1) It does not aim at predicting the location of future traffic bottlenecks, but identifies them as they emerge. Thus, it allows to accurately follow simultaneously all bottlenecks' dynamics and evolution in near real-time even during intervention in the system, for example, by using an adaptive traffic light control system. Moreover, as our method is not based on the identification of historical patterns, it considers all types of bottlenecks – recurrent as well as non-recurrent; and (2) By identifying and prioritising simultaneously all the bottlenecks in the network, at different times, it highlights which bottlenecks have the most significant effect on the urban traffic flow. These advantages can be implemented in planning transportation systems and reduce urban traffic congestion.

Similar to (30, 31, 33, 39, 40) we address the traffic urban flow as a directed weighted network. We suggest identifying traffic bottlenecks based on the definition coined by (41): "*The main feature of a bottleneck is that its downstream is in free flow and its upstream is jammed*". Thus, our method is based on the idea that if a bottleneck causes its upstream to be congested, the bottleneck must have been congested prior to it. Hence, for the definition of a bottleneck, time is as important as space.

# 2. Methodology

In this chapter, we present the methodology for identifying, evaluating, and thus, prioritizing urban traffic bottlenecks. Our methodology is based on complex network theory. We developed a three-stage methodology to identify functional clusters and use them to identify and evaluate traffic congestion. In the following sections, we present these stages.

## 2.1. Jam Tree — Basic method

To identify traffic bottlenecks, we converted datasets of urban areas to dynamic, directed traffic networks, where each node represents a junction, and each link represents a street segment between two junctions. The direction of the links represents the allowed traffic on that street segment, and the weight of the link at time segment $t$, $W(t)$, represents the temporal traffic relative speed, i.e., the ratio between the temporal speed and the speed at maximal flow for the link. We defined a street segment as currently congested if $(W(t)) < 0.5$. Next, we construct for a given time $t$ a new dynamic weighted network, where $W'(t)$ is the sum of $W(t)$ of all times each link has been considered as congested up to time $t$ (see figure 1) and used the following process to create tree-shaped clusters of congested links:

1. At each time $t$, we identify the links with the highest weight $W'$ (i.e. those that have been congested for the longest time) and define them as potential trunks of a jam-tree (JT). Next, we identify the branches of the JT by adding links or other trunks, connected to each trunk, with $W' \leq W'_{trunk}$. By doing so, we identify links that became congested no more than a predefined parameter $\theta$, in this case — defined as 2 measurement units, after the trunk or after a neighboring road. The value of $\theta$ is only used to limit the connections of new branches to a JT; in other words, it reflects the maximal duration that a congested street segment is considered as the cause for the congestion in its upstream. High values of $\theta$ allow a street segment to connect to its downstream longer times after its downstream became congested. This leads to larger JTs on one hand but reduces the probability of causality on the other. In other words, in our analysis, if a street segment became congested no more than 30

minutes after its trunk we can assume that the traffic load in these links resulted from the trunk of the JT. To test this assumption, we compared the result of the analyses of the real data to those of a controlled random model. The results of this comparison present a qualitative difference, which strengthens our assumption of causality. By using this definition, we consider the street segment that acts as the trunk as a bottleneck of the JT. Note, that we chose $\theta = 2$ as our datasets had 15 minutes time-intervals and thus, our analysis considered the macro-dynamics of urban traffic. Not that for other datasets with higher resolution of shorter time intervals, lower values of $\theta$ haev been used.

2. We continue assigning connected links to these JTs in the same iterative process until no more connected links (roads) with $W' \leq \theta$ for the last added branches are found.
3. We start again at stage 1, but now we look for the link with the highest weight $W'$, that has not been assigned to an existing (JT).
4. We continue this process until there are no more congested links that are not assigned to any JTs.

The resulted clusters represent JTs and the time each of their links was loaded. Examples of JTs are shown in figure 1.

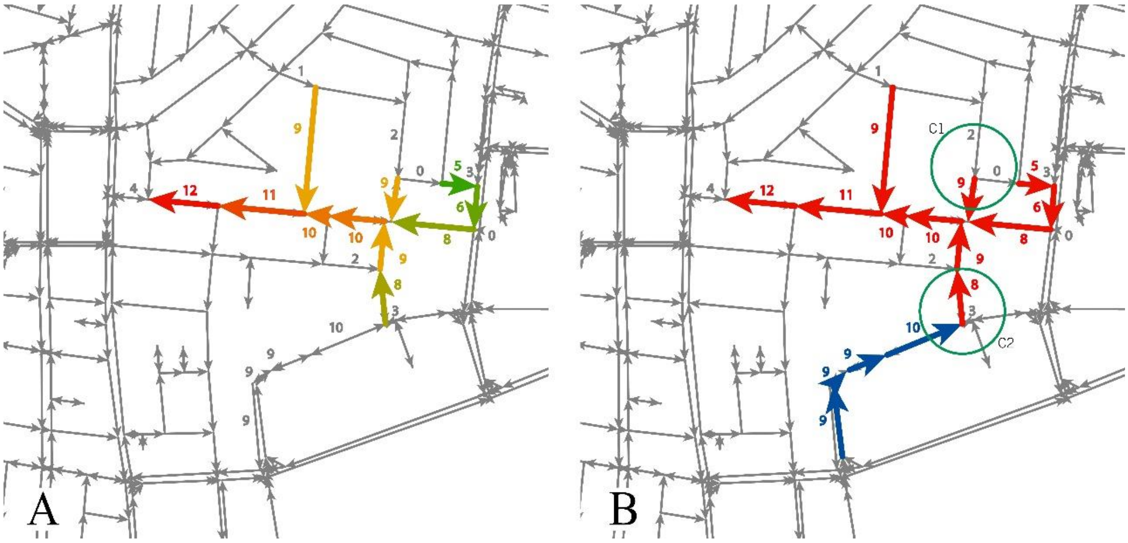



*Figure 1: Clusters of JTs. The numbers represent the time (in 15 minutes units) each street segment was congested. (A) All the coloured streets are part of one JT where the red street represents its trunk: its duration (12 successive measurements that represent 3 hours) is the longest, which indicates it was the first street with traffic load in this JT. (B) Two JTs (represented by red and blue colours). The red JT does not include the street that has been loaded for 2 measurements, as the time gap between this street and its adjacent one is larger than the pre-defined threshold $\theta$ (see the upper green circle). The blue JT cannot be considered as part of the red JT, as the duration of its trunk is longer than that of its adjacent street in the red JT (see the lower green circle). When a bottleneck is released but the JT that follows it remains congested, the next street segment with the longest duration becomes the new trunk of the JT and carries the cost of the remaining branches of the JT.*

## 2.2. Economic Cost — Prioritisation Strategy

While some traffic congestion incidents can last many hours, their economic cost might be marginal, if, for example, they occur in peripheral small streets. To assign prioritisation for traffic congestions, we measure their cost in vehicle hours (VH). In this section, we introduce four formulas used to calculate the cost at different times of the JTs and the links they include.

The cost of a link $C_{ij}(t)$ is calculated for every measurement unit — 15 minutes in this case (this duration can be changed according to the required accuracy. Shorter periods will reflect more accurate calculations while longer periods will reflect more general estimations) — relative to its cost $U_f$, free-flow speed. This measurement unit demonstrates the meso-dynamics of urban traffic. Indeed, using shorter periods of time will allow following the micro-dynamics of urban traffic. This cost represents the time it takes to cross a road (link) in comparison to the time it takes to cross this road in its maximal flow (calculated for each link), multiplied by the number of drivers who crossed the endpoint of this link at a specific time:

$$C_{ij}(t) = dist_{ij} * \left(\frac{1}{u_{ij}(t)} - \frac{1}{u_{qmax_{ij}}}\right) * \frac{q_{ij}(t) * l_{ij}}{\frac{60}{T}} \qquad (1)$$

Here, $\boldsymbol{dist_{ij}}$ is the length of the link in km, $\boldsymbol{q_{ij}(t)}$ is the current flow on the link, $\boldsymbol{u_{ij}(t)}$ is the current speed on the link, $\boldsymbol{u_{qo_{ij}}}$ is the speed when the flow is optimal, $\boldsymbol{T}$ represents a measurement unit which corresponds to 15 minutes (in the present case) and $\boldsymbol{l_{ij}}$ is the number of lanes in the link (i.e. the number of lanes in each street segment of the JT).

The momentary cost of a JT represents the sum of the costs (eq. 2) of all the links that are included in it at a specific measured time:

$$MomentaryCost(t)_{JT} = \sum_{b_{ij}}^{n} \left(C_{ij}(t)\right) \qquad (2)$$

And the cumulative cost of a JT is the cost of the JT from the moment it was created until the time (t) which is calculated as:

$$CumulativeCost(t)_{JT} = \sum_{b_{ij}} \left(\sum_{t_I \leq t}^{t} C_{ij}(t_I)\right) \qquad (3)$$

Here, $b_{ij}$ is a branch (i.e. link) in the JT and $t_I$ is the time each branch $b_{ij}$ was a part of the JT (in 15 minutes units).

Lastly, to follow the spatio-temporal dynamics of the system, we combine all the different JTs that had the same street as their trunk throughout the entire examined week and refer to them as Repetitive Jam Trees (RJT). The cumulative cost of the RJTs represents the sum of all the JTs they contain at a specific time window (e.g., day or week):

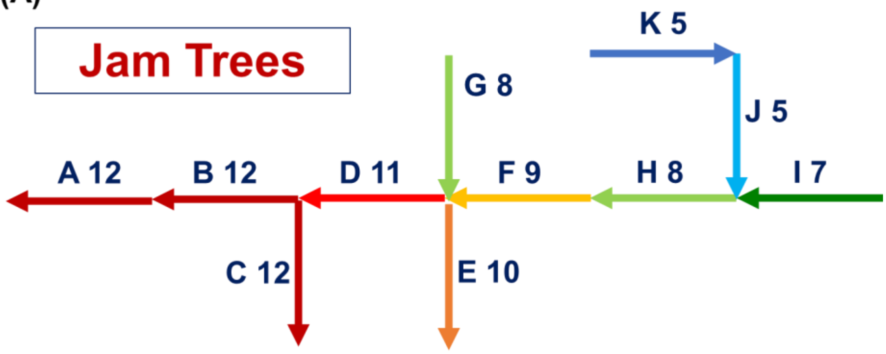
$$TotalCost_{RJT} = \sum TotalCost_{JT} \qquad (4)$$

Equations (2)-(4) allow calculating not only the cost of each JT from the moment it became congested until it was dissolved but also its dynamics and temporal costs at different times.

## 2.3. Extended Definition of Jam Trees

We extend the definition of a jam-tree to include more general cases where several jam trees may overlap and share the same trunk or branches. This is highly significant as these cases are very common, particularly in megacities with complex traffic patterns (Fig. 2).

(A)




(B)

| Tree | Trunk | Size | Cost |
|---|---|---|---|
| {A, B, D, F, H, I} | A | 6 | 1A + 1B + ½D + ⅓F + ⅓H + ⅓I |
| {C, D, F, H, I} | C | 5 | 1C + ½D + ⅓F + ⅓H + ⅓I |
| {E, G, F, H, I} | E | 5 | 1E + 1G + ⅓F + ⅓H + ⅓I |
| {J, K} | J | 2 | 1J+1K |

*Fig. 2. Definition of a jam tree. (A) An example of several potential jam trees . Each directed link represents a road segment, with an arrow indicating the direction of traffic flow on it. The number above each link indicates the jam duration i.e., successive time intervals of traffic jam. Here, each time interval represents 10 minutes. If the difference in jam duration between two nearby segments is less than a threshold, (in this work 2 intervals that represent 20 minutes), they are considered part of the same "jam tree." This is based on the assumption that there is a causal relationship between upstream and downstream traffic flow. (B) Key information about the presented jam trees. The Table shows information on the tree trunk, size, and cost of each specific jam tree in (A). Each jam tree has only one 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 any branches., while the temporal cost of each jam tree is defined as the weighted sum cost of its trunk and branches. (See Methods).*

As can be seen in Fig.2, 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 duration (in time units) of the link at a given time. The Table in Fig. 2b shows how we calculate the tree size and cost. The tree size is defined as the number of road segments that a jam tree contains. To calculate the cost of each road segment and sum them up, we followed stage #1. If a road segment can be associated with 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 it is associated with, in this case, 1/3 of the cost would be assigned to each trunk. This is because each trunk has a the same probability to cause the congestion.

## 2.4. Spatio-Temporal Evolution of Jam Trees

We expand the method of jam tree to the spatiotemporal Evolution of Jam Tree (EJT) that describes the evolving features of MTJs (Momentary Jam Tree) originated from the same traffic bottlenecks $B$ to observe the evolution patterns of the traffic congestions. An EJT is expressed as

$$EJT = \{MJT(t_G), \cdots MJT(t_P), \cdots, MJT(t_R)\}, \tag{5}$$

where $MJT(t)$ is the set of the congested roads caused by a bottleneck $B$ at time $t$, where $t_G$, $t_P$, and $t_R$ are respectively the emergence time, peak time, and dissolving time of a jam tree. The size of a jam tree at the entire lifecycle can be given by

$$S = \{S(t_G), \cdots, S(t_P), \cdots, S(t_R)\} = \{|MJT(t_G)|, \cdots, |MJT(t_P)|, \cdots, |MJT(t_R)|\} \tag{6}$$

where the size $S(t)$ is the number of the congested roads included in the momentary jam tree at time $t$ originated from bottleneck $B$.

# 3. Results

Here, we present the results of the above methodology. Our methodology was tested on two datasets. The first includes three urban areas (London center, Tel Aviv center (including the Ayalon highway — the main road that crosses the city from North to South), and Tel Aviv Center (without the Ayalon highway); and the second covers five large cities in China (i.e., Beijing, Shenzhen, Shanghai, Hangzhou and Jinan). For the first dataset, we collected from Google Directions API the speeds of road sections in London center and Tel Aviv every 15 minutes over a week's time. We developed an algorithm that considers additional road segments (for which we did not have data) based on interpolating the data collected for their adjacent road segments. For each case, we analysed the data of 5 working days only (Mon–Fri in London and Sun–Thursday in Tel Aviv). This is because the results indicate that the dynamics of these systems are significantly different during the weekends. For the Chinese cities, the operational information is the real-time velocity records of each road segment with a resolution of up to 1 minute, which is derived from the global positioning system (GPS) data recorded by floating cars (e.g., taxies and private cars). The data has already been aggregated from numerous trajectories, and the velocity assigned to each link is the result of aggregation. The scale of the road network and the coverage period of velocity records are shown in Table 1.

*Table 1. Key information of the dataset of different cities.*

| City | Road segments | Period |
|---|---|---|
| London Center | 18,050 | 21–27/3/2018 |
| Tel Aviv | 5,425 | 12–18/2/2017 |
| Tel Aviv Center | 3,871 | 12–18/2/2017 |
| Beijing | 52,968 | Oct. 2015 |
| Shenzhen | 22,248 | Oct. 2015 |
| Shanghai | 50,469 | Oct. 2015 |
| Hangzhou | 35,815 | Oct. 2015 |
| Jinan | 22,690 | Oct. 2015 |

## 3.1. Macro-stability

Previous work found universal laws in urban traffic congestion (43–49). Some studies even identified a high degree of regularity in the measured speed of the street segments (42, 43). Others focused on the time evolution of urban congestion (46, 48), but not through the analysis of bottlenecks. At large scales, traffic dynamics and congestion have been found predictable (42, 43) and their weights follow power-law distributions (46, 48). With this in mind, we analysed the bottlenecks' dynamic at the macro-scale and found they indeed present such regularities. We analyzed three temporal scales: the largest scale is a work week (that includes all the examined working days), the intermediate scale is a 24-hour day, and the microscale is the different hours of the day.

We explored the behavior of the traffic congestion incidents of the different datasets for London and Tel Aviv. Throughout the entire weekdays and examined several attributes of the systems: the duration of the traffic congestions, their size (in terms of the number of road segments), and their cost (in Vehicle Hours units). Figure 3 shows that the distributions present similar behavior for all three datasets. For all datasets, the probability density functions (PDFs) present well approximated power-law distributions. This implies that despite the different infrastructure and transportation facilities in these two cities, there may be common characteristics in London and Tel Aviv in terms of their traffic macro-dynamics.

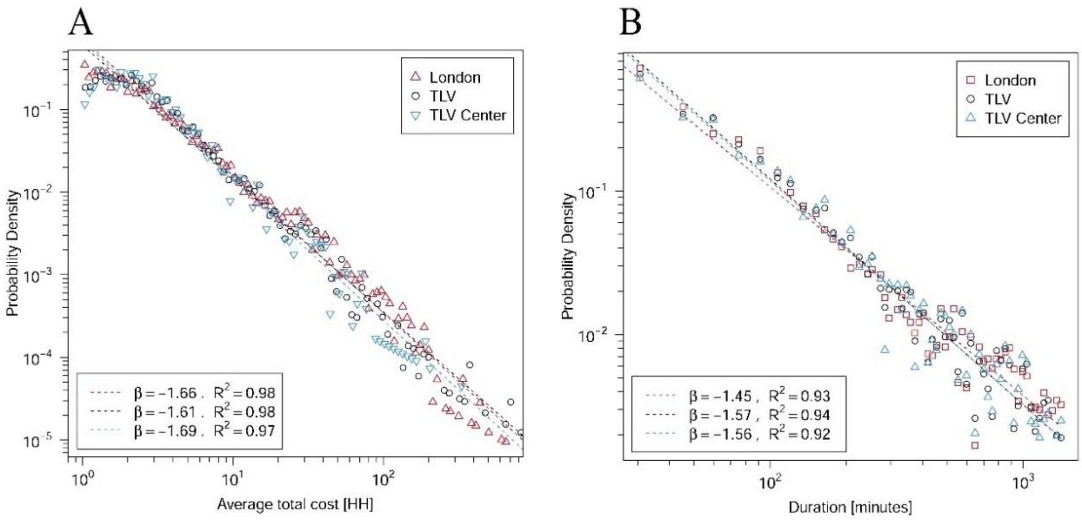


*Figure 3: Analysis of the PDFs of the RJTs in London, Tel Aviv, and Tel Aviv Center (based on the data of all 5 days). (A) PDF of average cost (in VH) (B) PDF of duration (in minutes).*

These findings may be explained by the fact that, while different cities have different physical constraints, historical development, and socio-economic processes, urban road networks were developed based on similar principles, i.e., similar parameters of demand (urban travel) and supply (road infrastructure).

We examined these distributions on data of separated days and hours and found that the results of the analysis of all these three temporal scales (large, intermediate, and short) show that the probability of having traffic congestion of a given cost is scale-free for all cities and in all time spans. Such PDFs can be useful for forecasting the existence of costly traffic congestions (above a certain threshold) and the volume of their costs at different time scales. However, the values of these congestions (in VH units) and the exponents that govern their behaviour decrease with size, depend on spatio-temporal features and represent the different attributes of the different areas. These attributes can relate to the morphology of the street network (45, 47) or other factors such as different types of transportation methods available in each location, working hours norms, etc. Furthermore, while the distributions remain similar on different days in the same city, we also examined how much the roads involved in the jam trees remain the same on different days.

## 3.2. Meso-dynamics

When zooming into the meso-dynamics of the traffic congestions, we unveil local characteristics that reflect shifts in the location of bottlenecks over time.

We analysed the repetition of bottlenecks on different days and found that most of the bottlenecks are irregular and the same bottlenecks usually do not repeat daily (Fig. 4A-C). In all three datasets, close to 60% of all bottlenecks appear only in one day of the week. About 20% appear in two days and less than 10% of the bottlenecks with the same level of cost, appear in three days. Even when ignoring their cost levels, the number of bottlenecks that appear once or twice exceeds 60%. Thus, we see that most heavy traffic congestions do not repeat daily. We also compared the above results to the analysis of the bottlenecks' duration (in terms of hours) and found a similar behavior, i.e., bottlenecks that lasted longer tend to repeat slightly more frequently than the shorter ones.

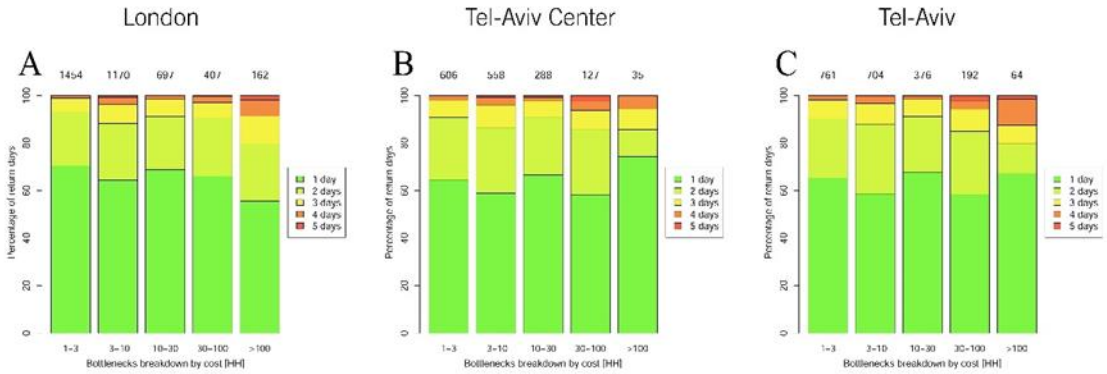



*Figure 4: Average and maximal Traffic Quality (TQ) values and standard deviation (σ) for DSCs in London and Tel Aviv. The maximal values refer to maximal values that appeared in at 85th percentile of the samples and the average of the top 15% values.*

Next, we examined the dynamics of the bottlenecks that affected the jammed trees. For that, we analysed (for each street that was part of any of the traffic congestions during our examination time window) the number of different bottlenecks it was connected to. We found congested streets are connected to a different number of bottlenecks (ranging between 1-22) regardless of their cost. However, the bottlenecks in London and Tel Aviv Center are located relatively in proximity to each other (their median distance is less than 1 KM) while the bottlenecks in Tel Aviv are spread over a wider area (up to 2.4 KM), which can be explained by the length of the Ayalon Highway (see above). This means, that while the traffic congestion can be associated with a specific area in the city, and even with some specific streets (38, 40), the location of the bottlenecks that causes the congestion changes constantly on different days and hours.

## 3.3. Urban Fingerprints

To understand at what time during the day the jam trees, connected to a specific trunk, will have the greatest impact on the traffic system, we analysed the dynamics of the number of jam trees throughout the day. By calculating the probability density of the number of trunks (each represents a jam tree) during the day, we found two distinct peaks: one between 7:00–9:00 and the other between

17:00–19:00 in both Beijing and Shenzhen, these represent the rush-hour windows in both cities (Fig. 5A and Fig. 5B). Additionally, in both cities, we found more jam trees during the afternoon rush hours in comparison to the morning rush hours. The number of jam trees in each time unit (Fig. 5A and 5B) depicts the frequency of existing jam trees but it does not give us enough information about their actual importance (i.e., size or cost). We, therefore, calculate the total momentary cost by summing up the momentary cost of all jam trees in the traffic network at a given time, and analysed how this value evolves over time. As seen in Fig. 5C and Fig. 5D, the evolution of the total momentary cost also presents, for both cities, two peaks during rush hours. These results not only indicate, as 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 found that the total momentary cost during rush hours in Beijing is approximately 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 hour of more than 15000 Vehicle Hours (VH) in Beijing and over 3000 VH in Shenzhen. However, it is worth noting that when it comes to the average momentary cost at each moment, as seen in Fig. 5E and Fig. 5F, the situation can be rather different. In Shenzhen, the average momentary cost during the above 2 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. This implies that during the morning rush hours, the traffic congestion is more intense compared to the evening rush hours, with higher average momentary cost and lower number of trees.

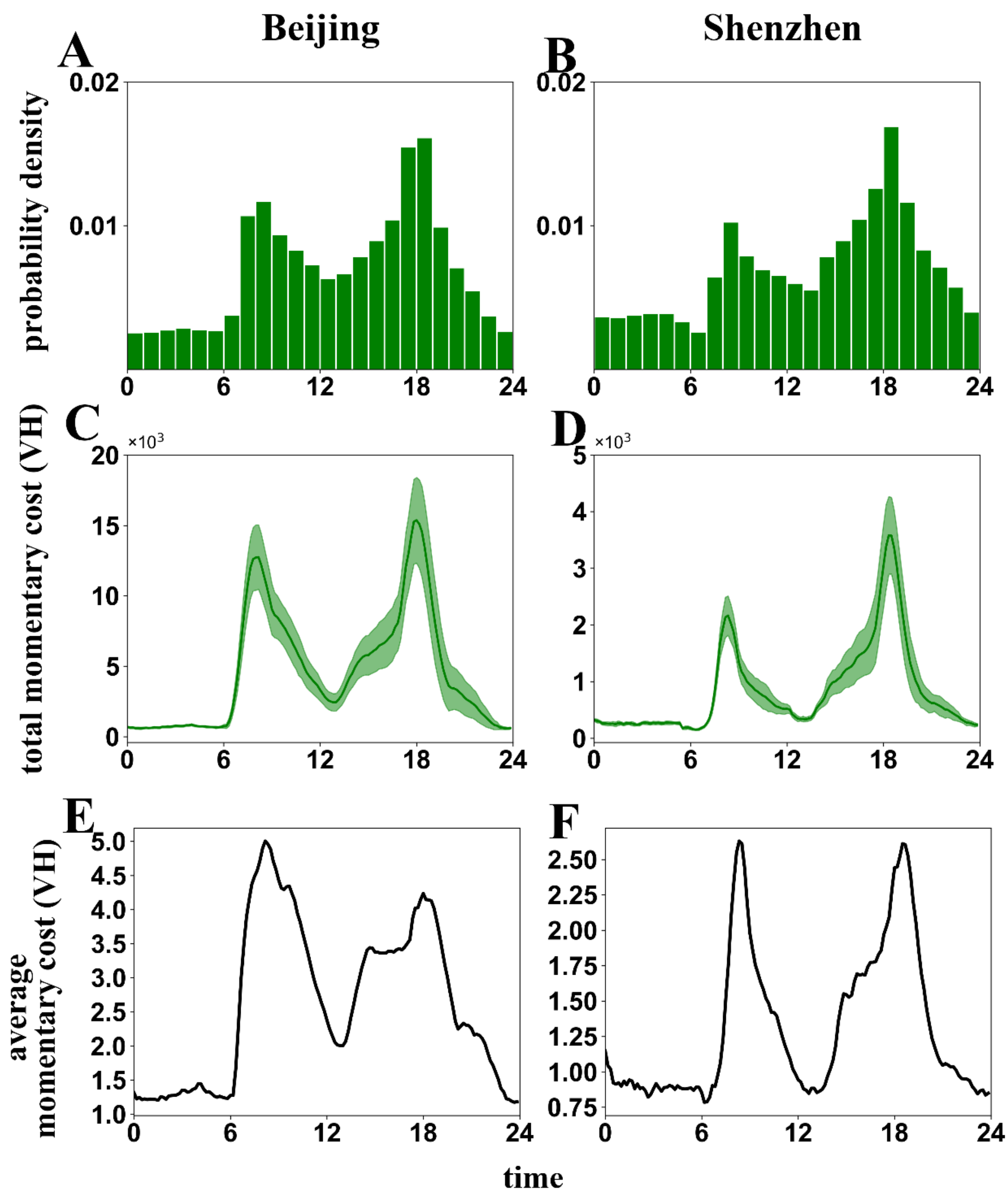



*Figure 5: Distribution of jam tree costs of urban traffic. (A) and (B) The distribution of jam costs associated with a given trunk in a typical working day in (A) Beijing and (B) Shenzhen. (C) and (D) Probability density of the jam cost exponent β for 17 working days for (C) Beijing and (D) Shenzhen. We also apply two-sample Kolmogorov-Smirnov test (KS test) to check if the distributions of daily exponents in two cities are indeed different and obtain a p-value of nearly*

*0 (much lower than the common threshold of 0.05). The distinct values of exponents in these two cities suggests the exponent β is a fingerprint of a city.*

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 whole day and considered it as the daily cost associated with this trunk. The identification of highly cost trunks is important for identifying, managing, and improving the traffic in these locations in order to improve global urban traffic. As can be seen in Fig. 6A and Fig. 6B, for a typical working day (Oct. 26, 2015), we found that, in both cities, the probability density of the daily cost of jam trees follows a power-law distribution, with exponents of 1.81 and 1.97, for Beijing and Shenzhen respectively. Furthermore, when considering the working days only (over a month), we also found that the exponents of all the cost distributions of jam trees for individual days have very similar characteristics with similar exponents. As can be seen in Fig. 6C and Fig. 6D, the cost distribution exponent in Beijing is 1.84±0.05, while in Shenzhen the value is 2.09±0.07. We also calculate the exponents of three other cities and obtained values of 1.80±0.05 for Shanghai, 1.78±0.07 for Hangzhou, and 1.75±0.05 for Jinan, respectively. Therefore, we argue that these exponents could be considered as a "city signature" that characterises the pattern of traffic congestion in a city. This is because the cost distributions are similar on different days for the same city, but rather different in two different cities. Note that smaller distribution exponents represent a more congested city, as they indicate the existence of more large-scale traffic jams in the city. Indeed, it is observed based on the cost distribution that the overall traffic performance in Beijing is worse than that in Shenzhen.

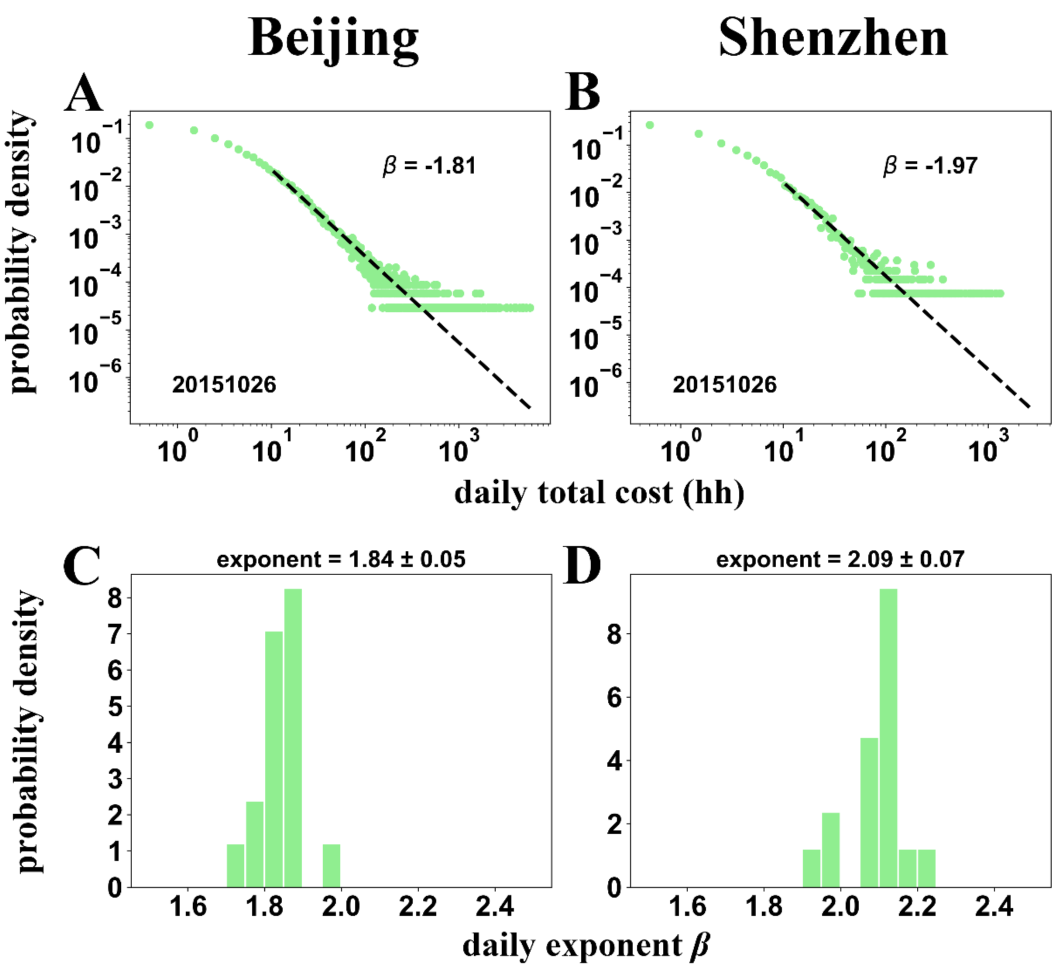



*Fig. 6. Distribution of jam tree costs of urban traffic. (A) and (B) The distribution of jam costs associated with a given trunk in a typical working day in (A) Beijing and (B) Shenzhen (C) and (D) Probability density of the jam cost exponent $\beta$ for 17 working days for (C) Beijing and (D) Shenzhen. We also apply a two-sample Kolmogorov-Smirnov test (KS test) to check if the distributions of daily exponents in two cities are indeed different and obtain a p-value of nearly 0 (much lower than the common threshold of 0.05). The distinct values of exponents in these two cities suggests the exponent $\beta$ is a fingerprint of a city.*

Based on the above results, we can conclude that while specific traffic congestion, at the microscopic scale, may vary from day to day, the overall macroscopic pattern of traffic congestion in a city (i.e., the distribution of costs throughout the day) remains stable and consistent from day to day. This raises the following question: Is there an indicator that can characterise at the macro-scale the daily pattern of traffic congestion in a given city? As the traffic system is expected to present different characteristic during different periods (i.e., rush hours and non-rush hours) the effectiveness of the above indicator, i.e., the exponent of cost distribution $\beta$, needs to be further validated by testing it at different hours during the day. To do so, we analysed the evolution of exponent $\beta$ over time during the entire day. We analysed the exponents of the cost distribution of jam trees in 5 different cities in China, during different times of the day over different days. For each city, we calculate the tree cost every 20 minutes and analysed the corresponding momentary tree cost distribution. Our findings show that for each city the cost exponent of the jam trees exhibits a similar temporal pattern during working days (as shown in Figure 7A-E). It is worth noting that the patterns are similar for the same city but different for different cities. The "W-like" shape curves for the momentary cost exponent in each day are also consistent with the results presented earlier, indicating that the traffic jam is typically more significant during morning or evening rush hours. Focusing on the evolution of the exponent of the tree cost distribution in different cities, it is interesting to learn 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.

## 3.4. Asymmetry in the congestion dynamics

Following the time evolution process of a jam tree, we divide the size S (i.e., the number of congested roads in the jam tree) into two stages: growth and recovery stages. The growth stage of an EJT ends and the recovery stage begins when the size S reaches its maximum ($S_P$). The entire lifespan T of the EJT is defined as the sum of the duration TG of the growth stage and the duration TR of the recovery stage, i.e.,

$$T = T_G + T_R, \quad (7)$$

where the growth duration is the time-interval from the earliest stage when a trunk emerges to the time it reaches its maximal size $S_P$ , and the recovery duration is the time-interval from the time the jam tree reaches $S_P$ to the time it has been completely dissolved and the traffic flow is fluent.

To explore the evolution durations of the growth and the recovery stages, we study the lifecycle of each jam tree in the entire road networks of Beijing and Shenzhen during the entire day over a period of 30 days. To this end, we computed the vectors of the growth duration $T_G = \{T_G^1, \cdots, T_G^N\}$ and the recovery duration $T_R = \{T_R^1, \cdots, T_R^N\}$. Here, $N$ is the total number of jam trees, which is over 670,000 on a workday in Beijing and over 180,000 in Shenzhen. The distributions of the durations $T_G$ and $T_R$ on Friday, October 16, 2015, in both cities are presented in Figs. 7 A and B. The results show that the distribution of $T_R$ is significantly broader than that of $T_G$. i.e., EJTs typically reach their maximum size SP within 100 minutes, but can take them up to 1000 minutes to return.to a free flow state. The distributions of $T_G$ and $T_R$ are both well approximated by power-law distributions

$$p(T_G) \sim (T_G)^{-\beta_G} \tag{8}$$
$$p(T_R) \sim (T_R)^{-\beta_R} \tag{9}$$

with exponents $\beta_G$ close to 5.9 on Oct. 16 2015 for both cities, and $\beta_R$ equal to 2.5 for Beijing and 2.69 for Shenzhen. In addition, the distribution of the ratio $r$ between $T_R$ and $T_G$ for each EJT also follows a power-law distribution $p(r) \sim (r)^{-\beta_r}$ with exponents $\beta_r$ around 2.71±0.05 on workdays in Beijing, and 3.03±0.05 on workdays in Shenzhen. The average value $<r>$ for all the jam components is evaluated by

$$<r> = \frac{1}{N}\sum_{1\le i\le N} r_i = \frac{1}{N}\sum_{1\le i\le N}\frac{T_R^i}{T_G^i}, \tag{10}$$

where the values of $<r>$ are 2.06 for Beijing and 1.66 for Shenzhen on Oct. 16 2015. This shows that on average, the duration of the recovery stage of a congestion is 2.06 times as long as its growth duration in Beijing, and 1.66 times in Shenzhen.

The above results show the strong asymmetry between jam agglomeration and dissipation. The agglomeration or dissipation of a jam tree associated with a bottleneck depends on whether the flow demand $D$ of the bottleneck could be dissolved by its capacity $C$. The ratio $\lambda = (D-C)/C$ can represent the ability of the congestion to accumulate ($\lambda > 0$) or dissipate ($\lambda < 0$). If the demand for the bottleneck is equal to its capacity per time unit, the branches in the EJT may remain unchanged. If the demand becomes larger than its capacity ($\lambda > 0$), the size of an EJT increases as new branches are added to it. Otherwise, its size decreases by losing branches that are no longer congested. The lowest bound of $\lambda$ is $-1$, whereas the upper bound of $\lambda$ can be larger than 1, revealing that the dissipation ability of traffic congestions is potentially less than their agglomeration. For example, if the congestion accumulates with $\lambda = 3$, the jam queues accumulated in one time unit require at least three times longer to dissolve. This can explain why the recovery stage takes usually longer than the growth stage.

Examining the growth and the recovery durations on other days, we found that the power law exponents $\beta_R$ of the distribution of the recovery duration are more consistent and similar across different days, compared to the $\beta_G$ values of growth duration. This may indicate the stable self-adaptability of the traffic system (Figs. 7 C and D). Note also that both exponents ($\beta_R$, and $\beta_G$) are slightly but consistently larger on holidays than those on workdays and weekends. In Figs. 7 E and F, we show the average ratio $<r>$ between the recovery duration and growth duration for 30 days during October 2015 in Beijing and Shenzhen. The results show that $<r>$ has a stable value for a specific type of days. In both cities, $<r>$ on workdays is larger than that on non-working days. This means that on workdays, the recovery duration of the jam trees is longer than their growth duration, compared to non-working days. These findings may be explained by the higher travel demand on workdays compared to holidays, leading to larger traffic congestion and making it harder for the congestion to dissipate.

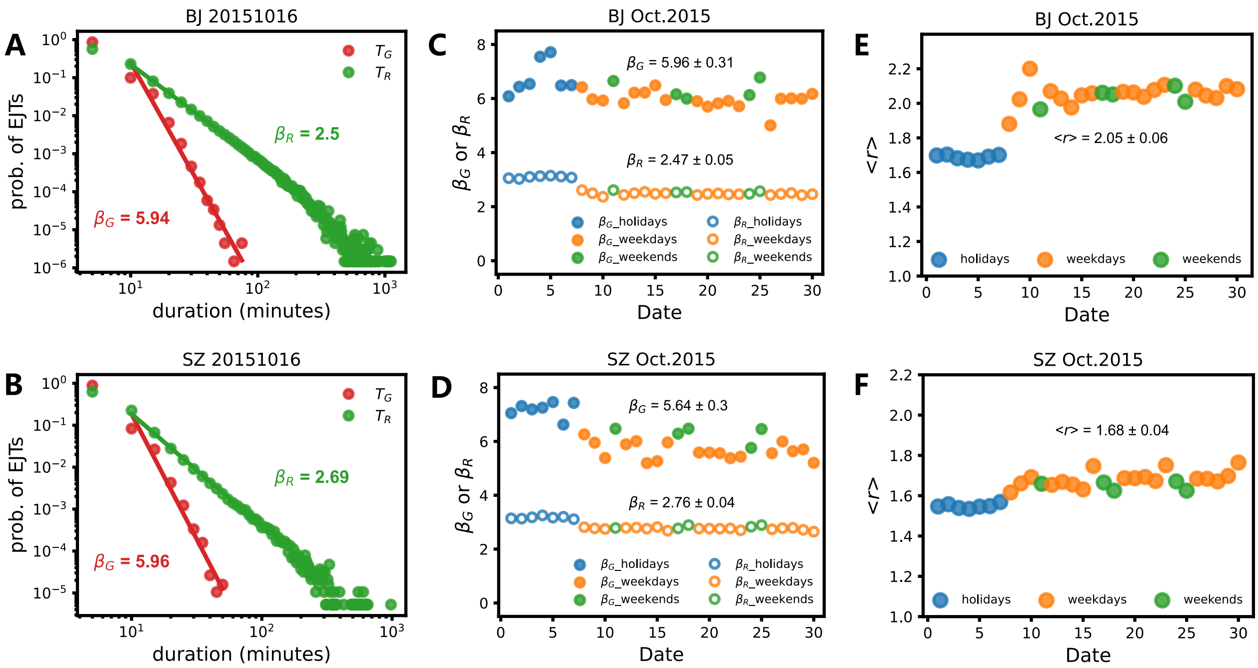



*Figure 7.* Distributions of growth duration and recovery duration. A and B. Growth duration $T_G$ and recovery duration $T_R$ follow power-law distributions. The data is taken for Friday, October 16, 2015, for (A) Beijing and (B) Shenzhen. C and D show the exponents $\beta_G$ and $\beta_R$ for each day in the entire month for (C) Beijing and (D) Shenzhen. E and F show the values of $< r >$ for each day in the entire month for (E) Beijing and (F) Shenzhen. The days between October 1 and October 7, are the seven days of the Chinese National holidays. October 11, 17, 18, 24, and 25 are the weekends, and the other 18 days are workdays. The mean values and standard errors on subfigures C to F are based on weekdays.

## 3.5. Correlation between spatial and temporal evolution

To further explore the relationship between temporal and spatial evolutions, we examined the correlation between the spatial maximal size $S_P$ and the temporal growth speed of the congestion. For that, we defined and calculated the growth speed of each EJT, where the average growth speed $V_A$ is defined as the maximal size $S_P$ divided by its growth duration $T_G$,

$$V_A = \frac{S_P}{T_G}. \tag{11}$$

A larger $V_A$ means that a bottleneck induces more congested branches per time unit, which is set to be 5 minutes here. Figs. 8 A and B show the relationship between $S_P$ and $V_A$ on October 16, 2015, in Beijing and Shenzhen. In the box plots, $S_P$ is classified into different groups by the values of $V_A$. It is seen that EJTs with a larger average speed are more likely to reach a larger size. The linear correlation between the maximal size $S_P = \{S_P^1, \cdots, S_P^N\}$ and the average growth speed $V_A = \{V_A^1, \ldots, V_A^N\}$ of these EJTs is also characterized by high values of Pearson correlation coefficients $\rho_{S_P,V_A}$, ( Fig. 8). When omitting small jam trees with $S_P = 1$ the value of $\rho_{S_P,V_A}$ on Oct. 16 2015 is 0.75 for Beijing and 0.79 for Shenzhen, indicating that $S_P$ and $V_A$ have a strong positive linear correlation. We also found that in both cities, the values of the correlations $\rho_{S_P,V_A}$ in the studied month are stable for a specific type of days (Fig. 8 C and D). The correlation is larger on holidays than on workdays, and also larger in Shenzhen than in Beijing.

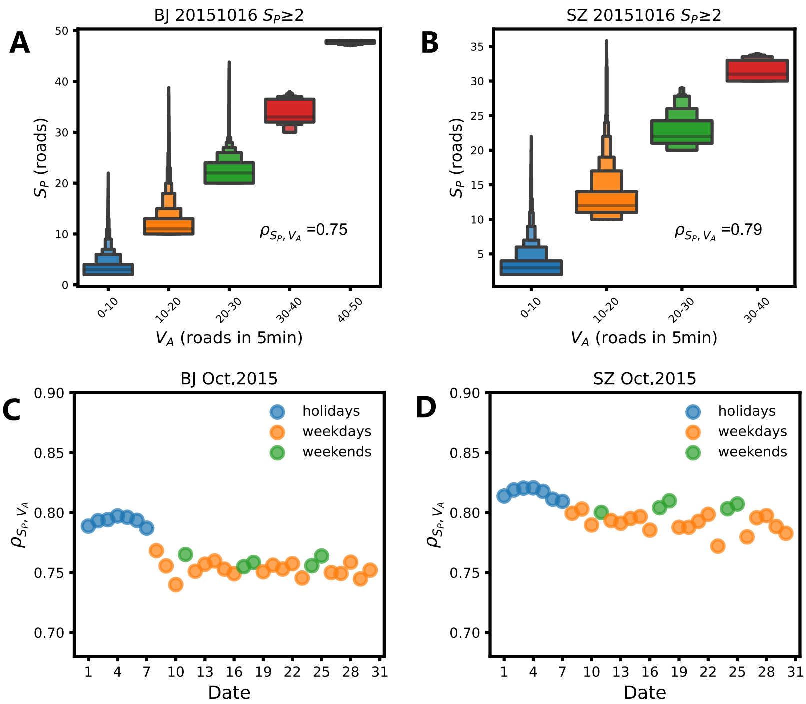



*Fig. 8. Correlation of the maximal size SP and the temporal growth speed of EJTs. A and B. Box plots of size SP grouped by average growth speed VA on Friday, October 16, 2015, in (A) Beijing*

*and (B) Shenzhen. The black lines in the bottom, middle and upper of the largest box are respectively 25%, 50%, and 75% percentiles of SP. Due to the long upper tail of the distributions, the 75%-100% percentile is equally divided into smaller boxes. C and D. Pearson correlation* $\rho_{S_P,V_A}$ *for 30 days during October 2015 in (C) Beijing and (D) Shenzhen. Holidays represent the 7 National-days, and the workdays include 18 days.*

# 4. Summary of the methodology

We showed that although some universal power-laws distributions that appear daily govern the macroscopic spatio-temporal behaviour of sizes of traffic congestion incidents, there are also strong indications that local attributes affect traffic dynamics, as the same traffic bottlenecks usually do not reappear on different days. In other words, the macro-stability, presented by the scaling characteristics of the traffic bottlenecks that represent the seeming regularity of traffic load both in time and space, overshadows the existence of rich meso-dynamics, where the bottlenecks that create these JT loads change significantly their location in time and space. This means that in order to manage traffic congestion in different locations and at different times and determine priorities regarding which ones should be addressed first, there is a need to implement unique solutions that track traffic and evaluate the relative effect of each bottleneck in real-time on the entire road network. Such solutioned are yet to be developed.

In the next stage, we extended the jam tree model in order to detect more potential configurations of jam trees. Particularly, we addressed the case where several jam trees share the same trunk or branches, a situation that frequently appears in complex traffic systems such as traffic networks in megacities. In this sense, the extended model is found useful for megacities or urban agglomerations. We found that both the number of existing jam trees and their average cost increase every day during rush hours. We also calculated the distribution of the cost of all the jam trees during a day and found that the cost distribution of the jam trees follows a power-law distribution with a similar daily exponent for all the analysed cities. Furthermore, we analysed the evolution of the cost distribution exponent over time during the entire day and found that the patterns of the cost distribution on the jam trees during the day, can be considered as the fingerprint of the urban traffic in a city. This is because the patterns indicated by the evolution of the above exponent are consistent each day for a specific city, but different for different cities. The unique patterns of traffic in urban areas (i.e., the suggested fingerprints of the urban traffic) may provide valuable insights for establishing new traffic management goals and assessing the effectiveness of various traffic improvement strategies.

Lastly, we analysed the spatiotemporal features of the Evolution of Jam Trees (EJTs) throughout the entire day over a month time and found that the growth and recovery stages of EJTs develop asymmetrically and that the duration of each stage follows a power-law distribution. Additionally, the recovery time of the EJTs is approximately twice as long, as the time it takes them to form. This trend is similar to patterns observed in other natural phenomena, such as the slower recovery of COVID-19 cases in terms of both their spatial distribution and number of infections compared to the rate of its outbreak.

# 5. An example - Comparing different control scenarios in Utrecht

We used our proposed methodology to examine different control scenarios for the Utrecht model. In this model we analysed a network of 878 nodes and 1570 directed links with two volumes of demand: 1. medium demand for non-rush hours and 2. high demand for rush hours.

In the following, we address them as:

- NC: Fixed time control settings / no adaptive control
- MP_0: Max Pressure in all nodes
- MP_25: Max Pressure (single) in 25% of selected nodes
- PC: Perimeter Control (single)
- PC_MP_20: Max Pressure in 20% of the nodes+ Perimeter Control
- MP_15: Max Pressure in 15% of selected nodes
- PC_MP_25: Max Pressure in 25% of the nodes + Perimeter Control

We used the mean link speed in every control cycle (in km/h, ranging from 0 to 25), and 240 out of 320 (med/high) control cycles (where every control cycle lasts for 90 seconds). Based on the bottleneck methodology, we calculated the momentary cost for every JT in the system for the different scenarios (other than NC and MP_0).

When we tested the distributions of the cost of JTs under different strategies, we found that the PC and PC_MP_25 strategies present better results in comparison to other strategies (figure 9).

When examining the different control strategies under medium demand, we found different results; in this case, most of the scenarios yields similar and good result in terms of traffic. The only exception is the MP_0 (Max Pressure in all nodes) strategy which yielded higher JTs cost in comparison to other strategies.

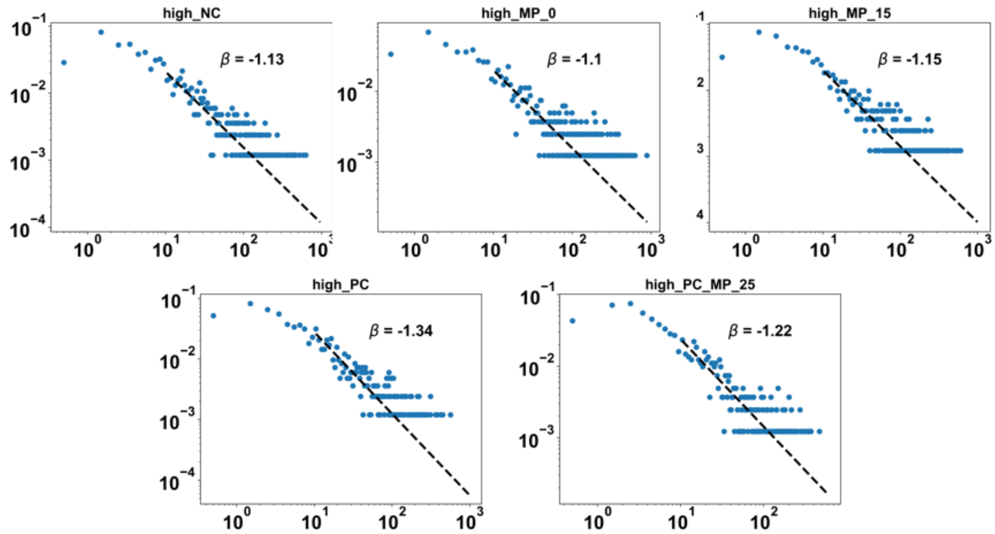



*Fig. 9. Distribution of the cost of JTs under different control strategies under high demand. The X-axis represents the Jam cost of a tree while the Y-axis represents the fraction of the jam trees within a certain cost. Higher absolute values of the exponents, mean the lower frequency of jam trees with large costs, which indicates better traffic.*

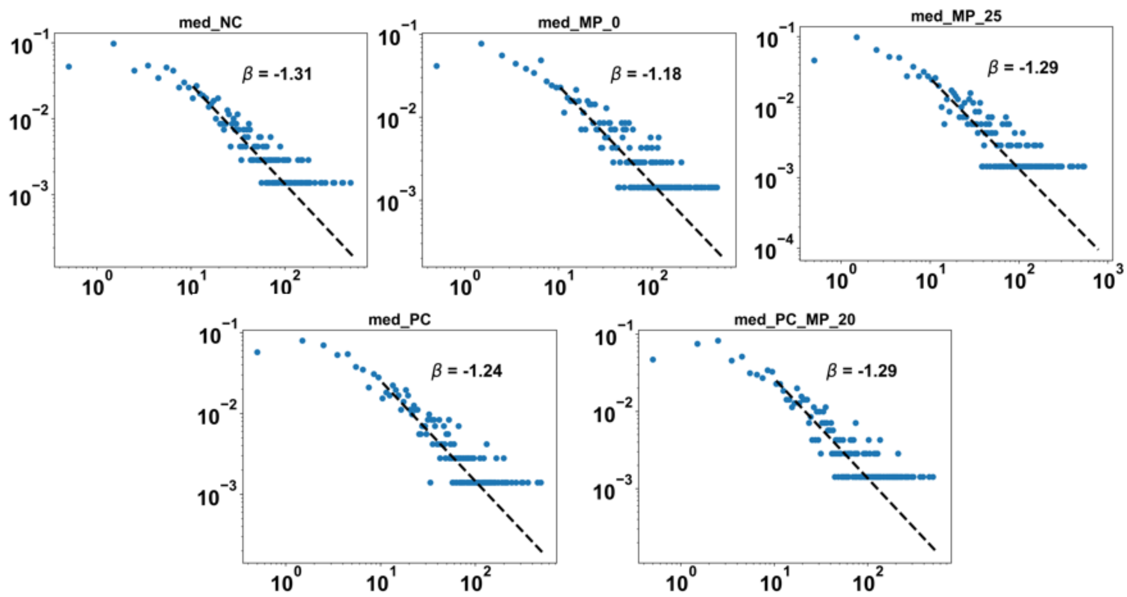



*Fig. 10. Distribution of the cost of JTs under different control strategies under medium demand. The X-axis represents the Jam cost of a tree while the Y-axis represents the fraction of the jam trees within a certain cost. Higher absolute values of the exponents, mean the lower frequency of jam trees with large costs, which indicates better traffic.*

# 6. An algorithmic approach for the adaptation / self-calibration of computer agents

To use an algorithmic approach for the adaptation/self-calibration of computer agents we have developed two different simulation models using two different platforms – Sumo and CityFlow. In the following, we introduce these models and their potential to contribute to this purpose.

## 6.1. Sumo platform (BIU)

This model was developed based on SUMO platform which allows simulating the traffic flow of cars and traffic lights operation systems at the micro level (for elaboration on this model see D3.2). It is based on the following traffic signal control schemes:

1. Fixed-time signal controllers – as explained above.
2. Actuated - Signal controllers based on SUMO optimization algorithm – *actuated* mode. In this mode, the system tries to alternate between phases based on demand at the phase of the intersection, with a high-demand phase receiving priority over a low-demand phase in both duration and timing. A minimum time for each phase is defined.
3. Demand calculation based on the number of vehicles on each road of the intersection. The algorithm determines the phase durations based on demand in each of the roads leading to the intersection (where we set a minimum time for each phase).
4. Using the bottleneck methodology to calculate the cycle time allocation. The methodology (presented above) assesses demand in the entire network and balances each direction at each intersection based on its accumulated impact (cost) on the network.

This model allows to examine the adaptation of computer agents (e.g. traffic lights) according to the JT prioritization methodology (see sections 3,4 in the traffic signal control schemes above). In future tasks we will calibrate the model based on the data gained from other simulations and the pilots to study how this methodology allows for the adaptation of traffic lights in order to improve urban traffic flow.

This model will be used (in future tasks) to evaluate how the proposed bottleneck prioritization method may influence traffic control systems and improve traffic flow.

## 6.2. CityFlow platform (ETHZ)

We have also designed other traffic simulation scenarios in another microscopic traffic simulator called CityFlow [50]. While SUMO is a versatile traffic software, one of the other considerations is fast, computationally efficient simulations that can be extended for the purposes of deep reinforcement learning (RL) simulations. CityFlow was specifically designed to address this shortcoming of using SUMO for RL simulations, achieving up to 20x speedup compared to SUMO when using the TRACI python API, however, it does sacrifice some other aspects of traffic modeling such as not being able to simulate buses and bicycle/pedestrian flows.

CityFlow was developed primarily for RL studies of traffic signal control. CityFlow simulations require the following:

1. **Road network:** We used synthetic road networks for initial studies of our algorithms (figure 11). However, in future tasks (Task 1.3) we will extend our simulations to account for realistic road networks as well.
2. **Vehicle Demands:** We use random vehicle OD demands for all possible pairs of links on the edges of the grid road networks.
3. **Traffic signal control schemes:** Traffic controllers choose from one of the 8 phases to be active at any given time. A traffic phase consists of compatible traffic movements (go straight, turn right, or turn left). The chosen phase will have all of its member movements given a green light. We then implemented three different traffic signal control schemes: three as baselines, and one as our test algorithm.
    a. *Random:* This traffic controller chooses the active phase at random.

b. *Actuated:* This controller follows a predetermined cyclic order of phase activations. The controller chooses to allocate green time to a phase according to the vehicle demand experienced by that phase.
c. *Demand:* The controller chooses the phases that serve the highest vehicle demand.
d. *Analytic+:* Analytic+ is an adaptive, self-organizing method that relies on optimization and stabilization rules [51].

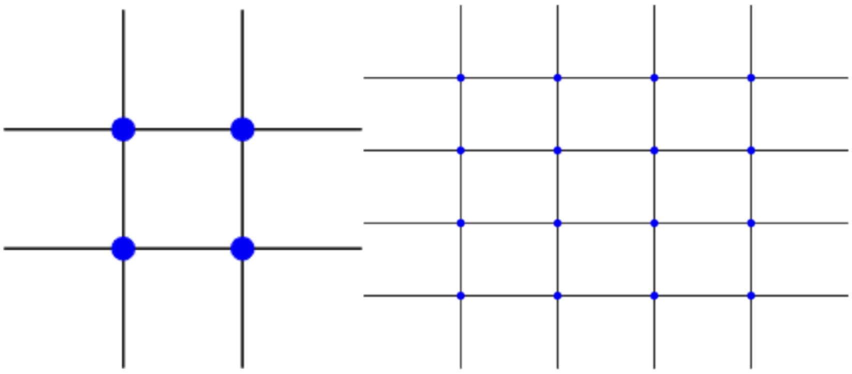


*Figure 11. Synthetic networks for the CityFlow traffic simulations. Left: 2x2 grid network. Right: 4x4 grid network. Each link is bi-directional, and all intersections have its own traffic signal.*

The results of these simulations show that the analytic+ algorithm performs better than the three benchmark algorithms (Figure 12). This shows that the self-organized nature of the Analytic+ algorithm reduces travel times in the synthetic network, without any explicit coordination with adjacent intersections. We will expand on these results further in Task 1.3, where we can integrate these findings into developing reinforcement learning algorithms that are inspired by effective self-calibrating models, as well as to further study the resilience and robustness of these systems.

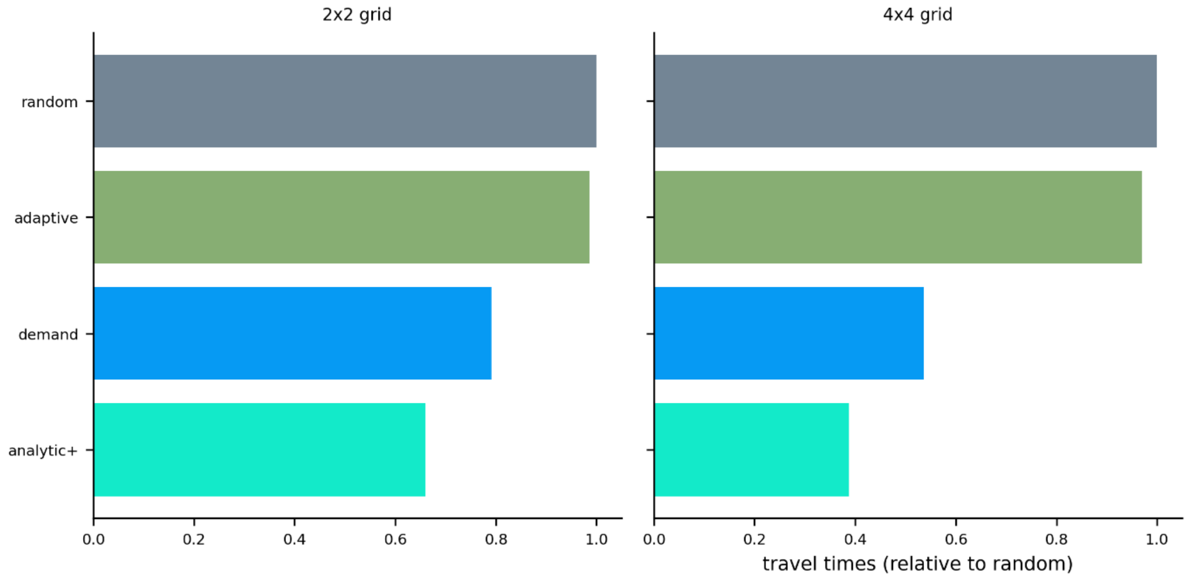



*Figure 12. Average travel times of the two synthetic networks running different control algorithms. Lower is better.*

In future tasks, we will use the JT methodology to evaluate the outcome of different traffic signal control schemes resulting from this model. By doing so, we will be able to estimate what scheme performs better under different constraints, types, and volumes of traffic flow.

# 7. Publications for D1.2


- Serok, N., Havlin, S., & Blumenfeld Lieberthal, E. (2022). Identification, cost evaluation, and prioritization of urban traffic congestions and their origin. *Scientific Reports*, *12*(1), 13026.
- Zeng, G., Serok, N., Blumenfeld Lieberthal, E., Duan, J., Li, D., Havlin, S. (preprint 2023) Unveiling Fingerprints of Urban Traffic based on Jam Patterns.
- Duan, J., Zeng, G., Blumenfeld Lieberthal, E., Li, D., Serok, N., Huang, H-J., Havlin, S., (preprint 2023) Spatio-temporal evolution of traffic congestions: an early signal of heavy bottlenecks

# 8. Code repository

The code of our analysis is available at GitHub:

https://github.com/nimrodSerokTAU/bottlenecks-prioritization

# 9. Bibliography


1. D. Helbing, A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks. *Journal of Physics A: Mathematical and General* **36** (2003).
2. M. Batty, The size, scale, and shape of cities. *Science* **319** (2008).
3. D. Schrank, T. Lomax, "Urban Mobility Report 2009" (2009).
4. M. Barthélemy, Spatial networks. *Physics Reports* **499** (2011).
5. B. Pishue, US Traffic Hot Spots: Measuring the Impact of Congestion in the United States (2017).
6. M. Treiber, A. Kesting, C. Thiemann, How Much does Traffic Congestion Increase Fuel Consumption and Emissions? Applying a Fuel Consumption Model to the NGSIM Trajectory Data. *Transportation Research Board* (2008).
7. J. I. Levy, J. J. Buonocore, K. von Stackelberg, Evaluation of the public health impacts of traffic congestion: A health risk assessment. *Environmental Health: A Global Access Science Source* **9** (2010).
8. J. Currie, R. Walker, Traffic congestion and infant health: Evidence from E-Zpass. *American Economic Journal: Applied Economics* **3** (2011).
9. Henao, Impacts of Ridesourcing - Lyft and Uber - on Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior. *ProQuest Dissertations and Theses* (2017).
10. Schaller, Unsustainable? The growth of app-based ride services and traffic, travel and the future of New York City. *Schaller Consulting* (2017).
11. S. Gehrke, A. Felix, T. Reardon, Fare choices: a survey of ride-hailing passengers in metro Boston. *Metropolitan Area Planning Council* (2018).
12. G. D. Erhardt, *et al.*, Do transportation network companies decrease or increase congestion? *Science Advances* **5** (2019).
13. J. C. Herrera, *et al.*, *Dynamic estimation of OD matrices for freeways and arterials* (Institute of Transportation Studies, UC Berkeley, 2007).
14. J. C. Herrera, *et al.*, Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies* **18** (2010).
15. L. Wynter, W. Shen, Real-time traffic prediction using GPS data with low sampling rates: A hybrid approach. *Transportation Research Board 91st Annual Meeting* (2012).

16. M. Rosvall, A. Trusina, P. Minnhagen, K. Sneppen, Networks and cities: An information perspective. *Physical Review Letters* **94** (2005).
17. Z. Wu, L. A. Braunstein, S. Havlin, H. E. Stanley, Transport in weighted networks: Partition into superhighways and roads. *Physical Review Letters* **96** (2006).
18. M. Barthélemy, A. Flammini, Modeling Urban street patterns. *Physical Review Letters* **100** (2008).
19. H. Youn, M. T. Gastner, H. Jeong, Price of anarchy in transportation networks: Efficiency and optimality control. *Physical Review Letters* **101** (2008).
20. B. Jiang, Street hierarchies: A minority of streets account for a majority of traffic flow. *International Journal of Geographical Information Science* **23** (2009).
21. G. Li, *et al.*, Towards design principles for optimal transport networks. *Physical Review Letters* **104** (2010).
22. M. Saberi, *et al.*, A simple contagion process describes spreading of traffic jams in urban networks. *Nature Communications* **11** (2020).
23. Y. Bai, Z. Wu, S. Sun, C. Wang, Automatic identification algorithm for freeway bottleneck in *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering, TMEE 2011*, (2011) https:/doi.org/10.1109/TMEE.2011.6199576.
24. L. Gong, W. (David) Fan, Developing a Systematic Method for Identifying and Ranking Freeway Bottlenecks Using Vehicle Probe Data. *Journal of Transportation Engineering, Part A: Systems* **144** (2018).
25. Z. Zheng, *et al.*, A Fused Method of Machine Learning and Dynamic Time Warping for Road Anomalies Detection. *IEEE Transactions on Intelligent Transportation Systems* (2020) https:/doi.org/10.1109/tits.2020.3016288.
26. Y. Chen, P. Yan, Z. Zheng, D. Chen, Identifying Traffic Bottleneck in Urban Road Networks via Causal Inference in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (2021).
27. W. H. Lee, S. S. Tseng, J. L. Shieh, H. H. Chen, Discovering traffic bottlenecks in an urban network by spatiotemporal data mining on location-based services. *IEEE Transactions on Intelligent Transportation Systems* **12** (2011).
28. S. Dong, Y. Zhang, Research on Method of Traffic Network Bottleneck Identification based on Max-flow Min-cut Theorem in *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering*, *TMEE 2011*, (2011) https:/doi.org/10.1109/TMEE.2011.6199586.

29. R. Tao, Y. Xi, D. Li, Simulation analysis on urban traffic congestion propagation based on complex network in *Proceedings - 2016 IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI 2016*, (2016) https:/doi.org/10.1109/SOLI.2016.7551690.
30. C. Li, W. Yue, G. Mao, Z. Xu, Congestion Propagation Based Bottleneck Identification in Urban Road Networks. *IEEE Transactions on Vehicular Technology* **69** (2020).
31. H. Nguyen, W. Liu, F. Chen, Discovering Congestion Propagation Patterns in Spatio-Temporal Traffic Data. *IEEE Transactions on Big Data* **3** (2016).
32. W. Yue, C. Li, G. Mao, Urban traffic bottleneck identification based on congestion propagation in *IEEE International Conference on Communications*, (2018).
33. J. Ma, *et al.*, On traffic bottleneck in green ITS navigation: An identification method in *IEEE Vehicular Technology Conference*, (2016).
34. H. Qi, M. Chen, D. Wang, Recurrent and non-recurrent bottleneck analysis based on traffic state rank distribution. *Transportmetrica B* **7** (2019).
35. Bui, K. H. N., Lee, O. J., Jung, J. J., & Camacho, D. (2016). Dynamic traffic light control system based on process synchronization among connected vehicles. In Ambient Intelligence-Software and Applications–7th International Symposium on Ambient Intelligence (ISAmI 2016) (pp. 77-85). Springer International Publishing.
36. Hamilton, B. Waterson, T. Cherrett, A. Robinson, I. Snell, The evolution of urban traffic control: changing policy and technology. *Transportation Planning and Technology* **36** (2013).
37. Backfrieder, G. Ostermayer, C. F. Mecklenbrauker, Increased traffic flow through node-based bottleneck prediction and V2X communication. *IEEE Transactions on Intelligent Transportation Systems* **18** (2017).
38. Zhao, C. Z. Xu, S. Liu, J. Zhao, L. Li, Dynamic traffic bottlenecks identification based on congestion diffusion model by influence maximization in metro-city scales. *Concurrency and Computation: Practice and Experience* **33** (2021).
39. Li, *et al.*, Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences of the United States of America* **112** (2015).
40. H. Hamedmoghadam, M. Jalili, H. L. Vu, L. Stone, Percolation of heterogeneous flows uncovers the bottlenecks of infrastructure networks. *Nature Communications* **12** (2021).

41. X. Ban, L. Chu, H. Benouar, Bottleneck identification and calibration for corridor management planning. *Transportation Research Record* (2007) https:/doi.org/10.3141/1999-05.
42. C. Lopez, L. Leclercq, P. Krishnakumari, N. Chiabaut, H. van Lint, Revealing the day-to-day regularity of urban congestion patterns with 3D speed maps. *Scientific Reports* **7** (2017).
43. L. Olmos, S. Colak, M. González, Non-equilibrium dynamics in urban traffic networks. *Nature Communications* (2016).
44. R. Louf, M. Barthelemy, How congestion shapes cities: From mobility patterns to scaling. *Scientific Reports* **4** (2014).
45. S. Zhao, P. Zhao, Y. Cui, A network centrality measure framework for analyzing urban traffic flow: A case study of Wuhan, China. *Physica A: Statistical Mechanics and its Applications* **478** (2017).
46. L. Zhang, *et al.*, Scale-free resilience of real traffic jams. *Proceedings of the National Academy of Sciences of the United States of America* **116** (2019).
47. G. Zeng, *et al.*, Switch between critical percolation modes in city traffic dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **116** (2019).
48. N. Serok, O. Levy, S. Havlin, E. Blumenfeld-Lieberthal, Unveiling the inter-relations between the urban streets network and its dynamic traffic flows: Planning implication. *Environment and Planning B: Urban Analytics and City Science* **46**, 1362–1376 (2019).
49. W. Lai, Fitting power law distributions to data (2016).
50. Zhang, H., Feng, S., Liu, C., Ding, Y., Zhu, Y., Zhou, Z., ... & Li, Z. (2019, May). Cityflow: A multi-agent reinforcement learning environment for large scale city traffic scenario. In The world wide web conference (pp. 3620-3624).
51. Lämmer, S., & Helbing, D. (2008). Self-control of traffic lights and vehicle flows in urban road networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(04), P04019.