

Percolation of heterogeneous flows uncovers the bottlenecks of demand-serving networks

Homayoun Hamedmoghadam^{a,1}, Mahdi Jalili^b, Hai L. Vu^a, and Lewi Stone^{c,d,1}

^aInstitute of Transport Studies, Faculty of Engineering, Monash University, Melbourne, VIC 3800, Australia; ^bElectrical and Biomedical Engineering, School of Engineering, RMIT University, Melbourne, VIC 3000, Australia; ^cBiomathematics Unit, School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv 6997801, Israel;

^dMathematical Sciences, School of Science, RMIT University, Melbourne, VIC 3000, Australia

This manuscript was compiled on March 9, 2020

1 Whether it be the mobility demand of passengers in transportation
2 systems, or the energy demand of consumers in power grids, the
3 primary purpose of many physical infrastructures is to best serve
4 these different "flow demands" as they fluctuate unevenly across
5 complex networks. To assess the reliability of infrastructure net-
6 works, we propose a new theoretical percolation-based framework
7 that not only incorporates the network structure and link-level per-
8 turbations, but also caters for the heterogeneous flow demand dis-
9 tribution. The framework systematically takes into account that the
10 larger the flow demand between two nodes in the network, the more
11 crucial are their connecting paths, making the analysis more realis-
12 tic than current state-of-the-art approaches. The method also identi-
13 fies "critical links" that have guaranteed decisive effect by rupturing
14 network flows, and this allows us to identify network "bottlenecks."
15 Melbourne's bus and tram Public Transportation (PT) network is dig-
16 itally reconstructed based on >100-million smart-card transaction
17 records, and its functionality under actual perturbations caused by
18 road traffic congestion is analyzed by means of the proposed frame-
19 work. Ameliorating a small number of identified bottlenecks is found
20 to significantly improve network reliability and reduce overall travel
21 times, indicating the efficacy of the method. The generality of the pro-
22 posed method is also demonstrated on random geometric graphs
23 as a generic proxy for spatial networks. Our results suggest that
24 embracing flow demand as a new dimension in percolation analysis
25 leads to a better understanding of complex infrastructure networks.

network analysis | percolation theory | bottleneck identification | flow demand

1 Recent theoretical advances in network science have con-
2 siderably contributed to our understanding of complex
3 systems, cutting across many disciplines from the social and
4 technological sciences, to the fields of ecology and biology (1–
5 12). In many modern studies, percolation theory (13, 14) has
6 been frequently employed to characterize structure, function-
7 ality and resilience of network systems. In this approach, link
8 failure is simulated by a percolation model which progressively
9 removes links from the network (15, 16). The impact is usually
10 measured via reduction in the size of the network's largest
11 connected component, or Giant Component (GC), as links
12 are gradually removed (17–19). Different strategies for simu-
13 lating link failures, e.g. random (error) or targeted (attack)
14 (20), make it possible to study a range of different topological
15 characteristics.

16 In real networks, links are consistently subject to various
17 forms of perturbations which often degrade their performance
18 in a continuous manner, rather than only causing complete
19 failure. To consider this, state-of-the-art approaches (21–23)
20 model the link-level dynamics on a network G by associating
21 each link e_{ij} (connecting node i to node j) with its own

"quality" attribute $q_{ij} \in (0, 1]$ at each time, which indicates
22 the temporal link performance relative to an observed or
23 predetermined maximum level of performance (24, 25). For
24 example, in a road traffic network, where the speed on each
25 road is changing temporally, link quality can be defined as
26 the ratio of instantaneous traffic speed to the limited maximal
27 speed on each link (23, 26), or in a communication network
28 link quality can be the relative instantaneous delivery rate
29 of information packets on a link (27). Typically, percolation
30 is simulated on such networks by increasing a threshold ρ
31 from 0 to 1 and simultaneously removing all links e_{ij} having
32 $q_{ij} \leq \rho$ from the initial network G to obtain G_ρ (28–30); see
33 the illustrated process on a small network in Fig. 1.

34 Of special interest is the critical threshold $\rho = \rho_c$ during
35 the percolation process at which the GC disintegrates into
36 components of smaller size. The percolation threshold ρ_c
37 can be an efficacious measure of global quality of network
38 structure, indicating that the network fails to provide global
39 connectivity with paths of links only having quality above
40 ρ_c (27, 31–33). While this generic critical phenomenon is of
41 vital importance for characterizing networks, we will show
42 that limiting attention exclusively to the GC and its sudden
43 disintegration reveals only a part of the full picture when
44 studying real-world problems.

45 The primary goal in many critical infrastructure networks

46

Significance Statement

A major characteristic of many infrastructural systems is the uneven demand for flow movement between different origin-destination points in the network. Unlike current state-of-the-art percolation analyses of perturbed networks, which assume this origin-to-destination flow demand is homogeneous, our proposed theoretical framework for the first time systematically takes into account the heterogeneity of flow demand. This is a significant addition to the modern network science where the heterogeneity of flow demand is often ignored in the analysis. The paper uncovers new network theoretic indices for assessing both network "reliability" and for finding "bottleneck" links, that will have important practical application in future studies of infrastructural network systems. New techniques for identifying network bottlenecks are shown to outperform the existing state-of-the-art methods.

H.H. conceived the original idea, and conducted the experiments and analyses. H.H. and M.J. and H.V. designed the study. H.H. and L.S. developed the methodology and wrote the manuscript. All authors interpreted the results, reviewed the paper and approved the final version.

Authors declare no competing interests.

¹To whom correspondence should be addressed. E-mail: homayoun.hamed@monash.edu, lewistone100@gmail.com

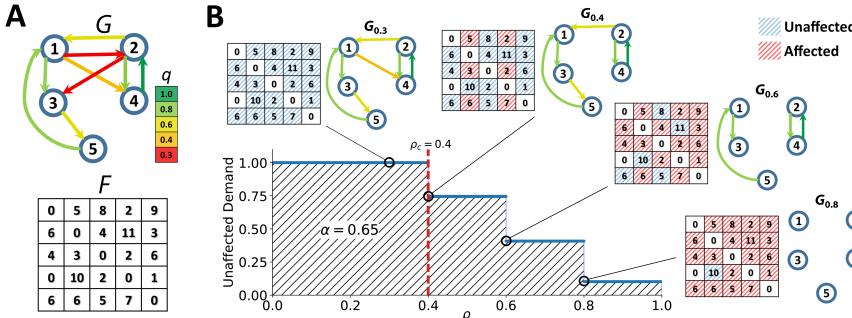


Fig. 1. Demand-serving reliability of an example network. (A) A demand-serving network of size $n = 5$ with color-coded link qualities. Matrix F shows the quantity of flow demand between all pairs of nodes which sums up to 100 units in total. (B) The percolation process is simulated by increasing a threshold ρ while removing links e_{ij} with $q_{ij} \leq \rho$. At multiple points in the process, the network is visualized with its corresponding affected (hatched red) and unaffected (hatched blue) flow demand. In this example, by definition, the system collapses at $\rho_c = 0.4$, when the 5-nodes strongly connected GC fragments into two strongly connected components of sizes 2 and 3, while Unaffected Demand (UD) is still at 75%. The network reliability α (Eq. 2) is calculated as the area under curve of UD between the limits of ρ (hatched in black).

such as communication, power distribution, water supply, and transportation systems, is to serve a certain demand for flow movement; we refer to such systems as *demand-serving networks*. In reality, the flow demand between Origin-Destination (O-D) node pairs is often distributed heterogeneously, and the larger the flow demand between two nodes, the more crucial are their connecting paths. In demand-serving networks, although the global connectivity is lost at percolation criticality, yet there might be a substantial flow demand inside the isolated components in subcritical phase. For example, if the bulk of the flow demand is contained within small and medium-sized clusters (e.g. resulted from the disintegration of the GC), the network can remain highly functional even after the GC collapse (Fig. 1B). In other words, the global dynamics in demand-serving networks is not only controlled by structure and organization of link qualities, but also by the distribution of the flow demand over O-D pairs.

The goal of the present paper is to add further realism to percolation-based network analysis by inclusion of the heterogeneous flow demand. Building upon state-of-the-art percolation-based network analyses (22, 23, 29), we introduce a calculus of network flows whereby known volumes of flux between each O-D pair become involved in the percolation process. This allows us to measure network's ability to provide pathways of high quality (well-performing) links to satisfy the flow demand, and to identify the network bottlenecks. Here, we first restrict our attention to a real transportation network as an exemplary instance of demand-serving networks, and then demonstrate the generality of our proposed framework. We show that the incorporation of the flow demand as an extra dimension to percolation-based network analysis can lead to different conclusions compared to those obtained solely from studying structural critical phenomenon.

80 The case of a real demand-serving network

81 We demonstrate the application of the proposed framework, on
82 the bus and tram (on-road) PT system in Melbourne, Australia,
83 modeled using smart-card transaction data collected during
84 September and October 2017. On-road PT systems are in
85 constant conflict with road conditions, such as crowds, traffic,
86 and signals, all negatively affecting the traveling flows by
87 decelerating the PT vehicles. Separation of high demand O-D
88 points by local pockets of congestion is an issue of considerable
89 concern in transportation systems. The concept of O-D travel
90 demand is fundamental to transportation theory (34), but only
91 in this paper has it been made use of in studying passenger
92 flows as a percolation-like process on dynamical transportation

networks.

Network $G(V, E, d, t)$ was generated for each day d and a moving 2-hour window centered at each time t of the day (see Materials and Methods and SI Note 1). Each node $i \in V$ corresponds to a cluster of closely situated bus and tram stops. A directed link $e_{ij} \in E$ connects node i to node j if there is at least one direct service between the two nodes. The flow demand matrix F was generated within the time window associated with each network, with each entry f_{ij} counting the number of passengers traveling from node i to node j via a single trip or multiple chained trips. Melbourne's on-road PT network is comprised of approximately 5,500 (2,800) nodes, 10,500 (4,500) links, and a flow demand derived from 470,000 (210,000) single-leg trips on average, during a normal weekday (weekend day).

In order to take into account the link-level perturbations caused by road conditions, similar to (22, 23) we assign a quality attribute to each link e_{ij} , calculated as:

$$q_{ij}(d, t) = \frac{\min_{t'}(\tau_{ij}(d, t'))}{\tau_{ij}(d, t)}. \quad [1]$$

Here, $\tau_{ij}(d, t)$ is the travel time on the link e_{ij} at time t of day d . The quality attribute $q_{ij}(d, t)$ indicates the link-level temporal performance of the network. At any point in time, a high-quality link has relatively low travel time (high velocity) compared to the rest of that day. In the following, for simplicity we refer to the network and its attributes without the day and time parameters. Figure 2A shows the spatial distribution of q_{ij} over the Melbourne's on-road PT network at 8:00 on 1 September 2017.

The percolation process illustrated in Fig. 2B for the Melbourne's PT network, indicates the loss of global connectivity at the percolation threshold of $\rho_c = 0.39$. However, as our analysis shows, at this point of criticality over 80% of trips can still be carried out between their O-D node pairs. This highlights a problem with interpreting ρ_c as a reliability index when the interest is on passenger flow demand. Note that this is not just a purely academic point. Stated in other words, previous methods (22, 23, 32), argue that if flows cannot reach location B from A without traversing low quality links, the transportation network has a problem. This is of course correct from a network connectivity perspective. On the other hand, a given passenger flow demand informs us about where the passenger flows begin and where they want to get to. Therefore, although location A is separated from location B with low quality links, it might not actually be important as much as the separation of other points, say, C from D where the flow

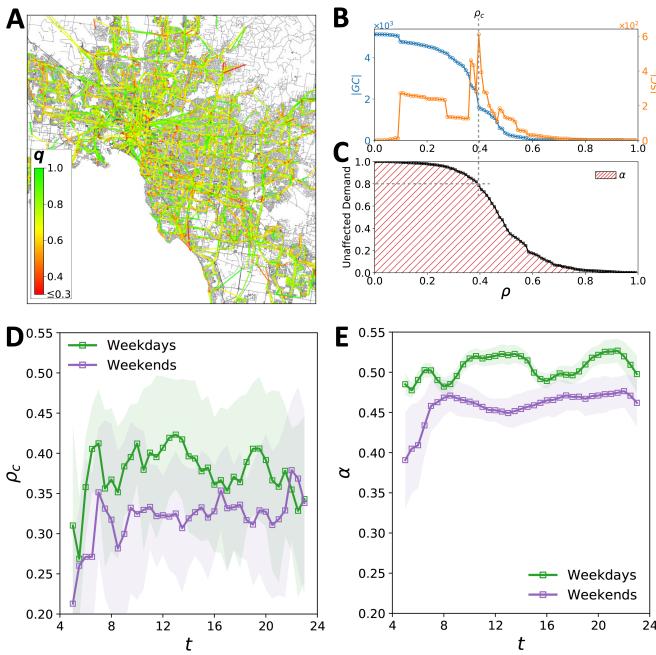


Fig. 2. Reliability of the Melbourne's bus and tram Public Transportation (PT) network. (A) The network representation of the PT system with colors indicating the link qualities q_{ij} for the period 7:00–9:00 (i.e. $t = 8$ and $\delta = 2$ h) on 1 September 2017. (B) Percolation process on the network illustrated in A, demonstrated by $|GC|$ and $|SC|$ as functions of threshold ρ ; the critical threshold $\rho = \rho_c$ is determined as the point of maximal $|SC|$ (vertical dashed gray line). (C) The same process as in B, demonstrated via UD_ρ , which shows the value of $\approx 80\%$ at percolation criticality (indicated by gray dashed lines). The proposed index α corresponds to the hatched area. Temporal evolution of (D) ρ_c and (E) α , during the day separately for weekdays (green) and weekends (purple). In D and E, curves show the mean over September and October 2017 and the shaded areas indicate the standard deviation of values around their mean. Note the much wider standard deviations for ρ_c in C.

links of the lowest quality (colored red) are removed, but this does not affect the flow demand between any pair of nodes, and thus $UD_{0.3} = 1$. When $\rho = 0.4$, however, removal of the link $1 \rightarrow 4$ impacts the UD as it prevents flows from reaching nodes 2 or 4 from either node 1, 3, or 5, by any pathway. The quantities of affected flows sum up to 25, thus the UD drops to $UD_{0.4} = 0.75$.

We now present our key index for assessing reliability of the demand-serving networks under link-level perturbations. We define the *demand-serving reliability* of the network, denoted by α , as the area under the curve of the UD_ρ over the domain of ρ (hatched area under curve in Fig. 1B), which can be formulated as:

$$\alpha = \int_0^1 UD_\rho d\rho = \int_0^1 \frac{\text{tr}(R_\rho F)}{\mathbf{1}_n^\top F \mathbf{1}_n} d\rho. \quad [2]$$

where $\text{tr}(\cdot)$ is the trace of the $n \times n$ square matrix, and $\mathbf{1}_n$ is a column vector of all n elements equal to one. In simple terms α gives an indication of average proportion of demand in the network that is unaffected as links are removed over the whole percolation process. As seen in Eq. 2 above, it is also possible to formulate UD_ρ , and as a result α , in simple mathematical terms making use of the network's so-called reachability matrix R (R_ρ corresponds to the network under percolation G_ρ) and the flow demand matrix F (see Materials and Methods).

Let $|GC|$ be the size of the GC. In Materials and Methods, we show that when flow demand is distributed uniformly (i.e. f_{ij} equals a constant for any reachable pair of nodes i and j) then on any large-enough undirected network $|GC| \approx n \cdot \sqrt{UD}$ at any threshold during the percolation. So first, only by assuming a uniform flow demand over the network, UD is able to replicate the percolation analysis based on monitoring the GC; this is also confirmed experimentally later in the paper. Second, with heterogeneous flow demand, the above relation is no longer true and UD provides its unique description of the system dynamics with respect to the flow demand distribution. As a result, α can be a very useful indicator of network reliability as it captures the heterogeneity of flow demand through UD's behavior during the percolation.

Demand-serving network reliability α

Our approach is based on monitoring what we refer to as *Unaffected Demand* (UD), and requires keeping track of flows between all node pairs on the network during the percolation process. Network flow demand can be described by a square matrix $F = [f_{ij}]$ of order n equal to the network size, where each entry f_{ij} quantifies the flow from node i to node j (see Fig. 1A). At any threshold ρ , the flow demand f_{ij} is deemed to be "unaffected" if there is at least one directed path from i to j remaining on G_ρ . To assist in interpreting this, consider a link that is part of a path that begins from origin node i and reaches destination node j . If the link is removed, and there is still at least one other directed pathway from i to j , the flow demand remains unaffected by the link removal. We define UD_ρ as the fraction of total flow demand that is not affected by removal of links performing with quality below the threshold ρ ; see Materials and Methods for detailed formulation.

It is instructive to examine how UD_ρ varies with increasing ρ on the example network shown in Fig. 1A, where the total flow demand is 100 by some arbitrary unit of measurement and $UD_0 = 100/100$ initially. When $\rho = 0.3$ (Fig. 1B), two

demand is far greater. The above issues motivate us to develop a new approach to capture the reliability of demand-serving networks, as an alternative to the critical percolation threshold, in order to enhance the understanding of such network systems.

Demand-serving network reliability α

Our approach is based on monitoring what we refer to as *Unaffected Demand* (UD), and requires keeping track of flows between all node pairs on the network during the percolation process. Network flow demand can be described by a square matrix $F = [f_{ij}]$ of order n equal to the network size, where each entry f_{ij} quantifies the flow from node i to node j (see Fig. 1A). At any threshold ρ , the flow demand f_{ij} is deemed to be "unaffected" if there is at least one directed path from i to j remaining on G_ρ . To assist in interpreting this, consider a link that is part of a path that begins from origin node i and reaches destination node j . If the link is removed, and there is still at least one other directed pathway from i to j , the flow demand remains unaffected by the link removal. We define UD_ρ as the fraction of total flow demand that is not affected by removal of links performing with quality below the threshold ρ ; see Materials and Methods for detailed formulation.

It is instructive to examine how UD_ρ varies with increasing ρ on the example network shown in Fig. 1A, where the total flow demand is 100 by some arbitrary unit of measurement and $UD_0 = 100/100$ initially. When $\rho = 0.3$ (Fig. 1B), two

80% of trips can nevertheless be carried out between their O-D points ($UD_{0.39} = 0.8$), which empirically demonstrates how characterizing a network based on ρ_c alone can be misleading when flow demand distribution is heterogeneous. The reliability indicator α leverages the concept of UD to account for the fact that pathways between different node pairs are not equally crucial to the reliability of the network. In effect, during the link removal process UD does not necessarily decline with the same rate as pairwise connectivity (see SI Note 2 and Fig. S3). In Melbourne's PT network, the number of connected node pairs decrease faster than UD, meaning that demand is higher within clusters of high-quality links in the network.

We examined both ρ_c and α on Melbourne's PT network over the main functioning hours of the system during two months of September and October 2017, separately for weekdays and weekends. Temporally, ρ_c had unusually large fluctuations over the day, and there appeared to be no repeating pattern on a day to day comparison (Fig. 2D). For further details regarding these intense fluctuations compare Figs. S4A and S4C. In contrast, the proposed reliability measure α followed a clear daily pattern (Fig. 2E) with variations that have relatively small standard deviation. The approximately 10% drops in α at 8:00 and 16:00-18:00 on weekdays are associated with morning and evening peak commuting periods, when high rates of congestion and large numbers of commuters predictably increase the conflict between PT system and road conditions. Consistent daily evolution of α with circadian rhythm of urban human mobility, and its low variability over different days suggest its success in unravelling the repeating daily pattern in complex interactions between major constituents of the system, namely, supply network structure, link-level perturbations, and passenger flow demand (also see SI Note 3). The results also suggest that Melbourne's PT network is relatively stable over a day, despite multiple periods of intense traffic, which is partially due to more available PT services during the rush hours which increase the number of links and thus network density.

Figure 2E shows that α varied temporally in two different ranges during weekdays and weekends, implying that the network functions in two separate modes. Despite larger flow demand and more extensive congestions, α was larger for weekdays compared to weekends. This is explained by the observed phenomenon in Fig. S3, which shows that a higher volume of flow has to traverse lower quality links during weekends compared to weekdays, which expectedly lowered the demand-serving reliability of the network.

Bottleneck identification

The framework suggests a new method for identifying network bottlenecks. Here, we introduce the link *criticality score* which quantifies the importance of each link quality for the demand-serving reliability of the network. The idea here is to examine all O-D pairs and seek out the links between them which are potentially the most responsible for controlling the quality of flow. The criticality score proposed here takes advantage of the work on the maximum capacity paths problem (37) to theoretically tie up the reachability of destinations from origins to percolation on the network. Suppose there is a set of different directed pathways Ψ_{OD} that connect node O to node D (see Fig. 3A). On each pathway $\psi \in \Psi_{OD}$, we search for the link with the minimum quality (Fig. 3B). Among those

particular links, we choose the link with the maximum quality (Fig. 3C), refer to it as the *limiting link* associated with the ordered pair (O, D) and denote it by e_{OD}^* . For simplicity, let us assume that each link quality value on the network is unique. Then, there will be only a single limiting link between any reachable pair of nodes. For a link e_{ij} , if $e_{OD}^* = e_{ij}$ for only a single pair (O, D) , then the link criticality score s_{ij} will be the fraction of the total demand that flows from O to D , i.e.

$$s_{ij} = \frac{f_{OD}}{\mathbf{1}_n^T \mathbf{F} \mathbf{1}_n}. \quad [3]$$

The index relies on the feature that during the percolation, as soon as the threshold ρ reaches the value q_{ij} , removal of the limiting link e_{ij} results in complete rupture of all pathways from O to D . If the link e_{ij} is the limiting link between several node pairs (see Fig. S4), Eq. 3 extends to:

$$s_{ij} = \sum_{O,D \in V, e_{OD}^* = e_{ij}} \frac{f_{OD}}{\mathbf{1}_n^T \mathbf{F} \mathbf{1}_n}. \quad [4]$$

We have identified an important relationship that connects the link quality (q_{ij}), the link criticality score (s_{ij}), and the network reliability (α), namely:

$$\sum_{e_{ij} \in E} s_{ij} \cdot q_{ij} = \alpha, \quad [5]$$

as proven in *Materials and Methods*. It can be rigorously shown that for any link e_{ij} , increasing q_{ij} within a non-empty range will increase the network reliability α , with the magnitude of increase being proportional to s_{ij} (see SI Note 4). (This is a nontrivial problem since alteration of the quality of any link in the network can change the criticality score of multiple links.) As such, links with larger criticality scores have greater influence on network reliability α . Therefore, after ranking the links according to their criticality scores, a desired number of the top critical links can be identified as network bottlenecks. For our experiments on Melbourne's PT network, we chose the top 2% in the ranking of critical links as the network bottlenecks.

Bottlenecks of Melbourne's PT network

In Melbourne's on-road PT network, link criticality scores vary over time. Therefore, we calculate the mean criticality score of each link over the course of the available data, and identify the network bottlenecks as links with the largest mean criticality scores, separately for weekdays and weekends. The identified bottlenecks are shown to be robust and occur as highly critical links on most days (Fig. S8).

The spatial distribution of the link criticality scores is portrayed in Fig. 4A. In the actual network, flow demand is significantly high to or from urban hotspot locations (38). Pockets of traffic congestions and crowds, which decrease the quality of PT network links, usually form around these hotspots. As a result, during both weekdays and weekends (Fig. S9A), links with large criticality scores were found to be generally in alignment with the urban morphology, as they were situated in urban hotspots and the areas surrounding them. Specifically, the biggest urban shopping center is surrounded by links with high criticality score, and top bottlenecks were mostly distributed around the single most significant hotspot of Melbourne which is the Central Business District (CBD).

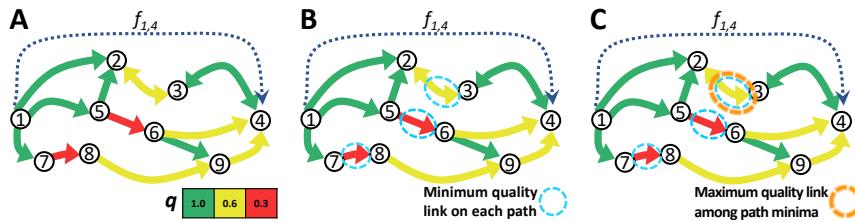


Fig. 3. Finding the limiting link between an origin-destination node pair. (A) As an example, we look for the limiting link associated with O-D node pair (1, 4) with flow demand $f_{1,4}$. (B) The available paths from node 1 to node 4 (and their associated minimum-quality links) are 1 → 2 → 3 → 4 ($e_{2,3}$), 1 → 5 → 2 → 3 → 4 ($e_{2,3}$), 1 → 5 → 6 → 4 ($e_{5,6}$), 1 → 5 → 6 → 9 → 4 ($e_{5,6}$), 1 → 7 → 8 → 9 → 4 ($e_{7,8}$). (C) Among the minimum-quality links of these paths, $e_{2,3}$ has the maximum quality. Just below the threshold $\rho = 0.6$, still two paths connect nodes 1 to 4, but then with $e_{2,3}$ removed, node 4 becomes unreachable from node 1. The limiting link for node pair (1, 4) is $e_{2,3}$ as increase in $q_{2,3}$ will keep the pair connected for longer during the percolation. The ratio of $f_{1,4}$ to the total demand is added to $s_{2,3}$ to reflect the importance of $e_{2,3}$ for demanded flow from node 1 to 4.

Also, universities are good examples of urban hotspots which are only fully active on weekdays. Among the top network bottlenecks, we observed links to and from major universities (Fig. 4B), emerging only on weekdays (see Fig. S9B). Given that the proposed method does not incorporate any geo-spatial information from the network, the surprising alignment between the locations pinned by identified bottlenecks and the hotspots of Melbourne's urban area, suggests that the method is capturing actuality.

The impact of bottleneck amelioration

It is interesting to compare the effectiveness of our proposed Criticality Score-based (CS) bottleneck identification scheme, to other well-established bottleneck identification schemes (see SI Note 5). In particular, we compare against the widely-used edge betweenness centrality measure (39), here referred to as EB bottlenecks. The bottlenecks found by this scheme are the links bridging between the locally well-connected communities of the network. Alternatively, bottlenecks can be identified among the links removed at percolation criticality as used in (22), which we refer to as PC bottlenecks. These bottlenecks, termed “red bonds” in percolation theory (40), glue the GC together by connecting the communities of higher quality links.

To make comparisons between these approaches, we ameliorated specific bottlenecks and monitored the response of the network in terms of changes to the demand-serving reliability α . In practice, the most obvious proposal for enhancing the reliability of an on-road PT network is to reduce the conflict of PT vehicles with road conditions at network bottlenecks, which can be done for example by giving signal priority to PT vehicles or allocating segregated (exclusive) PT lanes. Here, the bottlenecks are taken to be the top 2% most critical links in the network over time, according to each approach. Let B denote the set of bottlenecks identified by one of the schemes. We ameliorated the bottlenecks by synthetically increasing the qualities of bottleneck links $e_{ij} \in B$, to unity ($q_{ij} = 1$). Figures 4C compares the impact of amelioration on bottlenecks for the three different approaches, as a function of time during weekdays (see Fig. S11 for the results on weekends). Complete amelioration of the CS bottlenecks resulted in more than 22% (25%) improvement in α , on average during weekdays (weekends). Amelioration of EB and PC bottlenecks increased α over any day by approximately 8% and 6% on average, respectively.

The investigation was extended by verifying the impact of bottleneck amelioration on reducing the delay in passenger

travel times, as a more intuitive efficiency measure for transportation networks. In order to calculate the delay caused by perturbations, we first generated an unperturbed copy of the network at each time of a day by changing the actual travel time on each link to the minimum travel time observed on that link during the day. We assumed that each trip took place on the directed path with minimum sum of the link travel times, between origin and destination nodes. Then, for any particular network, delay was calculated as the absolute difference between the total travel time on the actual and the unperturbed copy of the network. Delay indicates the extent of the impeding effect of perturbations on passenger trips.

Separately for weekdays and weekends, we simulated the amelioration of the top CS, EB, and PC bottlenecks (the top 2% most critical links based on each scheme). Ameliorating the CS bottlenecks decreased the delay per passenger trip from 5.3 min (5.7 min) to 3.8 min (4.2 min) during weekdays (weekends). Figure 4D shows the temporal delay per passenger on the actual and ameliorated temporal network during the first five weekdays in September 2017; Fig. S12 extends Fig. 4D over the months of available data. Amelioration of EB (PC) bottlenecks had almost half (less than half) the reduction effect of CS bottlenecks on delay per trip. The time saved by amelioration of CS bottlenecks compared to that of EB (PC) bottlenecks is almost twofold (more than twofold). The result of ameliorating the top CS bottlenecks was close to 2,000 hours of passenger travel time saved only during morning rush hours (7:00-9:00) on the network, and approximately 11,000 hours of passenger travel time over a normal weekday.

Generality of the proposed framework

In order to emphasize the generality of the proposed framework, we used a Random Geometric Graph model (RGG) as a generic proxy of spatial networks and showed the ability of the framework to reflect the true global flow properties of the network. Here, undirected RGG networks were generated by first distributing 10^4 nodes uniformly at random on the plane $[0, 100]^2$, and then connecting any pair of nodes with Euclidean distance below $r = 1.5$. Each link quality was drawn uniformly at random from $(0, 1]$, making percolation a random link removal process which depends only on the network topology. RGGs are built of clusters with high intra-connectivity glued together by bridging links (Fig. 5A). This structure demonstrates a clear phase transition during the percolation process, when enough inter-cluster links are removed to make a significant part of the network isolated from the rest (Fig.

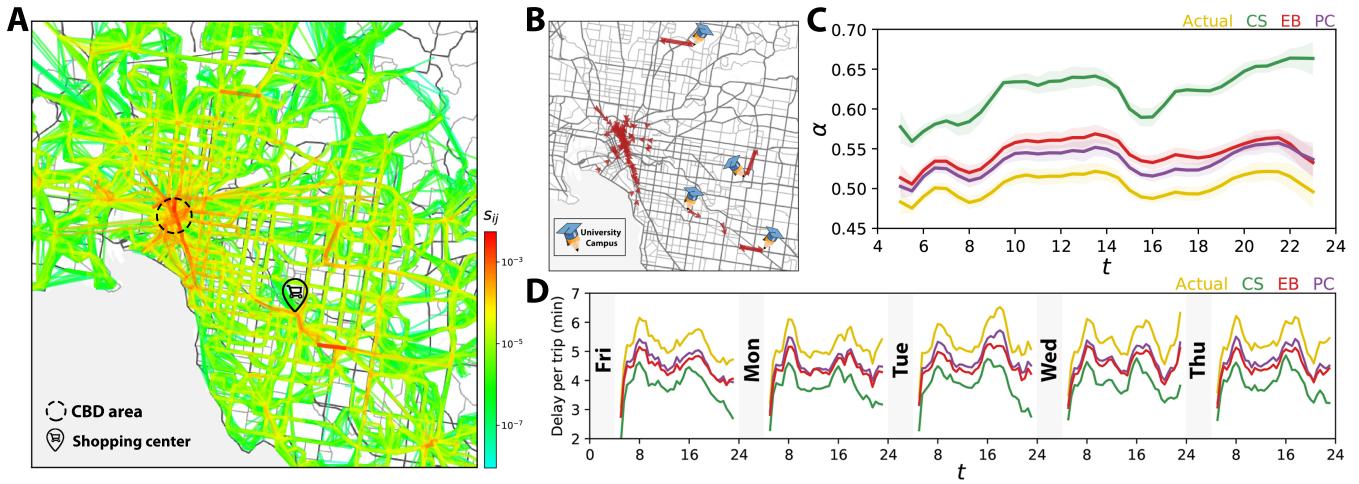


Fig. 4. Bottleneck identification and amelioration on a real-world network. (A) Spatial distribution of the link criticality scores s_{ij} over Melbourne's Public Transportation (PT) network during weekdays. The Central Business District (CBD) and the biggest shopping center in Melbourne are pinned on the map. (B) Top 100 weekday bottlenecks of Melbourne's PT network, identified based on link criticality scores. Major university campuses outside Melbourne's CBD area are pinned on the map. C and D compare the impact of ameliorating perturbations on bottlenecks identified based on Criticality Score (CS), Edge Betweenness centrality (EB), and Percolation Criticality (PC); the number of identified bottlenecks by each approach is equal to 2% of the number of links that appear on the network over time. (C) Daily evolution of α calculated for the actual (yellow) and ameliorated networks associated with different bottleneck identification approaches. Results show the average (solid line) and standard deviation (shaded area) over the weekdays of September and October 2017. (D) Delay per trip in minutes caused by road perturbations, on the actual and networks in which bottlenecks are ameliorated.

426 5B).

427 Over each instance of RGG, we distributed a fixed total
 428 flow demand volume, according to three different scenarios,
 429 namely, uniform, short-range, and long-range. In the uniform
 430 demand scenario, the total flow demand volume was divided
 431 equally among all reachable (O, D) node pairs; i.e. all entries
 432 of F , corresponding to a reachable node pair, are equal to a
 433 constant. Let d_{ij} be the Euclidean distance between nodes i
 434 and j , and d_{max} the distance between the most distant node
 435 pair in the network. Then, to generate the short-range (long-
 436 range) flow demand scenarios, we repeatedly picked a node
 437 pair (O, D) uniformly at random and then with probability
 438 $0.2e^{-0.2d_{OD}}$ ($0.2e^{-0.2(d_{max}-d_{OD})}$) added one unit to the flow
 439 demand between that O-D pair f_{OD} , until the fixed total flow
 440 demand volume is completely allocated over the network.

441 We simulated the percolation for each one of the above flow
 442 demand distributions, on 100 realizations of RGG structure.
 443 During the percolation, we monitored the GC and SC, which
 444 are independent from the demand distribution, and also UD
 445 for the three demand distribution scenarios (Fig. 5B&C). Remarkably, in Fig. 5C for the case of uniform flow, the
 446 percolation diagram as a function of ρ is the same for UD
 447 as it is for the square of $|GC|$ (normalized with the network
 448 size). Thus, simulation results confirm the previously discussed
 449 theoretical relationship $UD_\rho \approx (|GC_\rho|/n)^2$ between evolution
 450 of the GC and UD when demand is uniformly distributed
 451 over the network (Fig. 5C). This shows that by assuming
 452 a uniform flow demand over the network, our method can
 453 provide an analogous analysis to that of monitoring the GC.
 454 Furthermore, UD shows logical sensitivity to non-uniformity of
 455 flow demand distributions on network reliability. Long-range
 456 flows are more likely to break in the event of GC collapse,
 457 which UD reflected with a faster decrease as a function of
 458 increasing ρ (Fig. 5C), resulting a lower demand-serving
 459 reliability ($\alpha = 0.43$) compared to the case of uniform demand
 460 ($\alpha = 0.50$). In contrast, highly-connected local communities
 461 in RGG preserve the short-range flows during the percolation,

462 hence the network becomes more reliable which was reflected
 463 by a higher α ($= 0.58$).

464 Here, we verify the sensibility of the bottlenecks identified
 465 based on link criticality scores for RGG networks. We use the
 466 link overlap $\eta \in [0, 1]$ to determine whether a link belongs to
 467 a community (high overlap) or acts as an inter-community
 468 bridge (low overlap); overlap of a link e_{ij} is defined as $\eta_{ij} =$

$$\frac{|\Gamma(i) \cap \Gamma(j)|}{|\Gamma(i) \cup \Gamma(j)| - 2}$$
 where $\Gamma(i)$ is the neighborhood set of node i .
 469 In Fig. 5A links are color-coded according to their overlap
 470 index. The criticality score of intra-community (high overlap)
 471 links were found to be higher for short-range flow demand
 472 scenario compared to the long-range scenario (Fig. 5D). This
 473 is consistent with the fact that short-range flows are more likely
 474 to have their origin and destination within a community, which
 475 makes the flow-carrying role of intra-community links more
 476 critical. Inter-community (low overlap) links have a bigger
 477 role in bridging between the remote points of the network,
 478 thus, with higher flow exchange between the distant nodes
 479 these links become more critical for reliability of the network.
 480 As expected, criticality score of inter-community links were
 481 higher in long-range flow scenario compared to short-range
 482 flow scenario.

483 Conclusion

484 In this paper, we proposed the demand-serving network reliabil-
 485 ity index α , which quantifies the ability of networks to provide
 486 pathways of high-quality links to satisfy the flow demand be-
 487 tween O-D nodes. By constructing the theoretical relationship
 488 between the link qualities and percolation dynamics on the
 489 network, we introduced a bottleneck identification approach.
 490 Our proposed framework is generally applicable to demand-
 491 serving networks including most physical infrastructures such
 492 as transportation, communication, power distribution, and
 493 water supply networks, where the inherent demand for flow
 494 movement between different nodes makes the connectivity
 495 between some nodes more important than others.

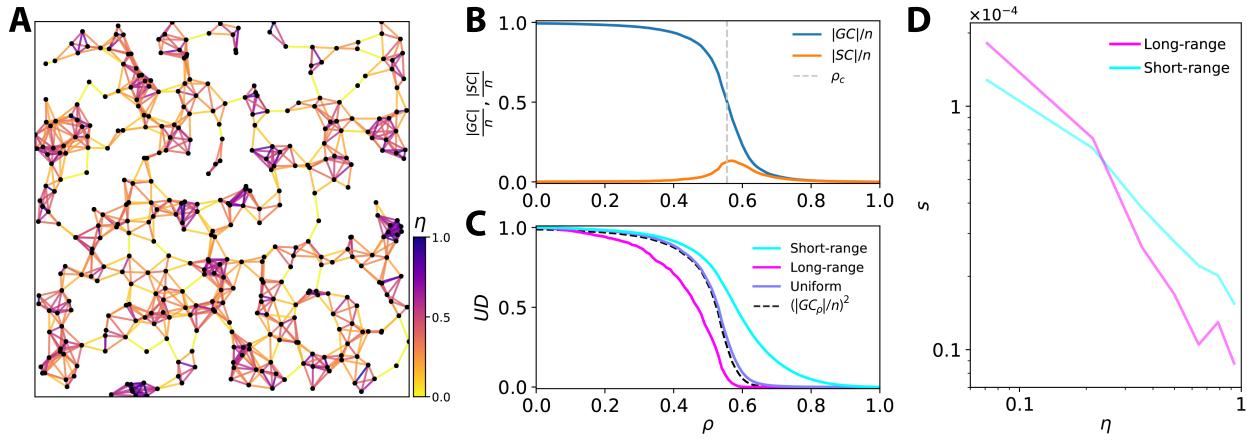


Fig. 5. Capturing true properties of the demand-serving networks. (A) A sample RGG with $n = 400$ spread over the plane $[0, 20]^2$, with each link colored based on its overlap index η . (B) Normalized $|GC|$ and $|SC|$ during the percolation process averaged over 100 realizations of RGG structure with $n = 10^4$ nodes and random link qualities. (C) Unaffected Demand (UD) versus ρ for different flow demand scenarios on the RGG structure, averaged over 100 realizations. As predicted, evolution of size of the GC during the percolation process is approximately equal to $n \sqrt{UD_\rho}$ when the flow demand is uniform; compare the blue and the dashed black curves. (D) Criticality score s versus overlap η of links, compared for short-range and long-range flow demand scenarios.

With link quality attribute describing the link-level dynamics in a network, percolation analysis (21–23, 29, 32) can be the lead for understanding the global quality of network structure. However, when dealing with demand-serving networks, the proposed index α leverages the introduced "unaffected demand" concept to account for the heterogeneity of node-to-node flow volumes, and thus unveils the otherwise obscure global flow quality on the network. Similarly, the bottleneck identification method introduced here, is a more comprehensive approach for such networks and proved more effective than other schemes reported in the literature in terms of both improving the network reliability and reducing the delay caused by link-level perturbations in the presented real-world application.

With ever-increasing availability of data, this study can be a starting point for diverse experiments on real-world networks with known flow demand. Our aim is to also open new avenues for more sophisticated theoretical work on the implications of flow demand on global dynamics of demand-serving networks, which should aid us to obtain a more profound understanding of these complex systems.

Materials and Methods

519

520 **Smart-card data.** The data used in the real-world case study, are the
521 smart-card transaction records, collected by the Automated Fare
522 Collection (AFC) system for PT in Melbourne, Australia. Passen-
523 gers are supposed to perform a scan-on transaction at the start
524 and a scan-off at the end of their trip. Every smart-card transac-
525 tion record contains multiple attributes, namely, anonymized card
526 identifier, PT mode (bus, tram, or train), vehicle identifier (unique
527 number for each bus or tram vehicle), stop identifier, timestamp,
528 and transaction type (scan-on/off). We used an average of over
529 2,120,000 and 912,000 daily transactions associated with all PT
530 modes on weekdays and weekends, respectively, collected during
531 61 days of September and October 2017. After applying a cleaning
532 process, we used the data to generate the temporal network of on-
533 road PT supply and its corresponding passenger travel flow demand
534 (see SI Note 1 for details).

535 **Network and flow demand matrix construction.** To generate the net-
536 work (details provided in SI Note 1) representation of the on-road

PT system in Melbourne at time t on a particular day, the smart-
537 card data was limited to the time window $[t - \delta/2, t + \delta/2]$, where
538 the time window length δ , was set to 2 hours for experiments pre-
539 sented in the main article. First, we clustered the closely located
540 PT stops and mapped each cluster to a node. Using information of
541 smart-card transactions we derived the trajectory of every vehicle on
542 the network, and if there was at least one vehicle traveling from one
543 of the stops associated with node i to a stop associated with node j
544 without stopping, we added a direct link e_{ij} starting at node i and
545 pointing at node j . For each link e_{ij} the average travel time τ_{ij} over
546 the time window was also calculated based on the information from
547 the tracked vehicles. For network of time t , flow demand matrix
548 F measures the demand volumes by the number journeys between
549 nodes, within the time window used for construction of the network.
550 A journey is a chain of one or more trip legs with transfers (but no
551 activities) in between them; see SI Note 1 on how single trip legs
552 are chained to obtain journeys. Throughout this manuscript, unless
553 otherwise stated, a "trip" refers to a journey.

554
Unaffected Demand (UD). To formulate the UD calculation, we use
555 the so-called reachability matrix $R = [r_{ij}]$ (the transitive closure of
556 the network adjacency matrix) which is a square matrix of order n .
557 Each entry r_{ij} is equal to 1 if there is at least one directed path from
558 node i to node j on the network, and $r_{ij} = 0$ otherwise. Let R_ρ
559 be the reachability matrix of network G_ρ . At any threshold ρ , the
560 flow demand f_{ij} is deemed to be unaffected if there is at least one
561 directed path from i to j remaining on G_ρ , i.e. $r_{ij}^\rho = 1$. So, UD_ρ
562 will be the sum of $r_{ij}^\rho f_{ij}$ for all (i, j) pairs of nodes, normalized by
563 the total flow demand:
564

$$UD_\rho = \frac{\mathbf{1}_n^T (R_\rho \circ F) \mathbf{1}_n}{\mathbf{1}_n^T F \mathbf{1}_n} = \frac{\text{tr}(R_\rho F)}{\mathbf{1}_n^T F \mathbf{1}_n}, \quad [6]$$

565 where " \circ " is the entry-wise product of matrices, $\text{tr}(\cdot)$ is the trace of
566 the $n \times n$ square matrix, and $\mathbf{1}_n$ is a column vector of all n elements
567 equal to one.

568
Relation between the evolution of UD and GC during the percolation.
569 Let $|GC_\rho|$ be the size of the GC as a function of ρ , then $|GC_\rho|/n$
570 is called the incipient order parameter which is sometimes used to
571 describe the connectivity of a fragmented network. If we assume a
572 uniform flow demand distribution then on any undirected network,
573 UD_ρ equals the proportion of connected node pairs at ρ , which
574 approaches $(|GC_\rho|/n)^2$ as $n \rightarrow \infty$ (41). So, for large enough
575 networks, monitoring the GC during the percolation is a spacial
576 case of monitoring UD when flow demand is uniform. Therefore, we
577 can accurately predict the evolution of $|GC|$ during the percolation
578 by assuming a uniform flow demand over the network and using
579

580 $|GC_\rho| \approx n\sqrt{UD_\rho}$; this is confirmed experimentally, later in the
581 paper.

582 **Link criticality score.** Suppose that there exists a non-empty set of
583 different directed pathways Ψ_{OD} that route between an origin node
584 O and a reachable destination node D . During the percolation
585 process on the network (whereby ρ is increased from zero to unity),
586 each pathway $\psi \in \Psi_{OD}$ breaks up when the threshold ρ reaches to
587 the minimum link-quality on that path. The limiting link associated
588 with flow from O to D (e_{OD}^*), when removed during the percolation
589 process at $\rho = q_{OD}^*$, breaks the last path(s) connecting O to D and
590 affects the flow between them (f_{OD}). Using the definition of link
591 criticality score in Eq. 4, we can expand the left-hand-side of Eq. 5
592 as:

$$\begin{aligned} \sum_{e_{ij} \in E} s_{ij} \cdot q_{ij} &= \sum_{e_{ij} \in E} \sum_{\substack{O, D \in V, \\ e_{OD}^* = e_{ij}}} \frac{f_{OD}}{\mathbf{1}_n^T F \mathbf{1}_n} \cdot q_{ij} \\ &= \frac{1}{\mathbf{1}_n^T F \mathbf{1}_n} \sum_{\substack{O, D \in V, \\ e_{OD}^* = e_{ij}}} \sum_{e_{ij} \in E} f_{OD} \cdot q_{ij}, \end{aligned} \quad [7]$$

593 and for any pair $O, D \in V$ with non-zero f_{OD} there exist a single
594 limiting link $e_{OD}^* \in E$ with quality q_{OD}^* , so

$$= \frac{1}{\mathbf{1}_n^T F \mathbf{1}_n} \sum_{O, D \in V} f_{OD} \cdot q_{OD}^*. \quad [8]$$

597 By definition, during the percolation process, each entry in the
598 reachability matrix R_ρ switches from 1 to 0 as soon as the last
599 path(s) between its corresponding O-D nodes break. So, we can
600 write:

$$r_{OD}^\rho = \begin{cases} 1, & \rho < q_{OD}^*, \\ 0, & \rho \geq q_{OD}^* \end{cases} \quad [9]$$

602 where r_{OD}^ρ is the (O, D) entry of the reachability matrix R_ρ
603 associated with the network G_ρ . Note that the integral of r_{OD}^ρ with
604 respect to ρ between the limits $\rho = 0$ and $\rho = 1$ is equal to q_{OD}^* .
605 So, from Eq. 8 and Eq. 9 we can write:

$$\sum_{e_{ij} \in E} s_{ij} \cdot q_{ij} = \frac{1}{\mathbf{1}_n^T F \mathbf{1}_n} \sum_{O, D \in V} f_{OD} \cdot \int_0^1 r_{OD}^\rho d\rho, \quad [10]$$

606 where the right-hand-side can be simplified with matrix operations
607 to obtain Eq. 2 which is the definition of the reliability index α , so
608 we can conclude that Eq. 5 holds. In SI Note 4, the definition of
609 the criticality score and the proof of Eq. 5 are generalized further,
610 requiring no assumption on the link quality values.

612 **ACKNOWLEDGMENTS.** We thank Prof. Shlomo Havlin for his
613 extremely helpful discussions of the ideas presented in this paper and
614 for his suggestions. Also, we thank Dr. Meead Saberi for the fruitful
615 discussions, and Nora Estgfaeller for providing valuable feedback
616 throughout this study. M.J. and L.S. are supported by Australian
617 Research Council (ARC) through project No. DP170102303. M.J.
618 is also supported by ARC through project No. DP200101119. H.V.
619 is supported by ARC through project Nos. DP180102551 and
620 DP190102134.

- 621 1. SV Buldyrev, R Parshani, G Paul, HE Stanley, S Havlin, Catastrophic cascade of failures in
622 interdependent networks. *Nature* **464**, 1025–1028 (2010).
- 623 2. M Jalili, M Perc, Information cascades in complex networks. *J. Complex Networks* **5**, 665–693
624 (2017).
- 625 3. D Duan, et al., Universal behavior of cascading failures in interdependent networks. *Proc.
626 Natl. Acad. Sci.* **116**, 22452–22457 (2019).
- 627 4. O Askarishan, et al., Structural balance emerges and explains performance in risky
628 decision-making. *Nat. Commun.* **10**, 1–10 (2019).
- 629 5. A Barja, et al., Assessing the risk of default propagation in interconnected sectoral financial
630 networks. *EPJ Data Sci.* **8**, 32 (2019).
- 631 6. M De Domenico, A Baronchelli, The fragility of decentralised trustless socio-technical sys-
632 tems. *EPJ Data Sci.* **8**, 2 (2019).
- 633 7. J Gao, B Barzel, AL Barabási, Universal resilience patterns in complex networks. *Nature* **530**,
634 307–312 (2016).
- 635 8. L Stone, The google matrix controls the stability of structured ecological and biological net-
636 works. *Nat. Commun.* **7**, 12857 (2016).

- 637 9. SM Hill, et al., Inferring causal molecular networks: empirical assessment through a
638 community-based effort. *Nat. Methods* **13**, 310–318 (2016).
- 639 10. L Stone, The feasibility and stability of large complex biological networks: a random matrix
640 approach. *Sci. Reports* **8**, 1–12 (2018).
- 641 11. M Santolini, AL Barabási, Predicting perturbation patterns from the topology of biological
642 networks. *Proc. Natl. Acad. Sci.* **115**, E6375–E6383 (2018).
- 643 12. L Stone, D Simberloff, Y Artzy-Randrup, Network motifs and their origins. *PLoS Comput. Biol.*
644 **15** (2019).
- 645 13. HE Stanley, *Introduction to phase transitions and critical phenomena*. (Oxford: Clarendon
646 Press), (1971).
- 647 14. A Bunde, S Havlin, *Fractals and disordered systems*. (Springer Science & Business Media),
648 (2012).
- 649 15. AA Ganin, et al., Resilience and efficiency in transportation networks. *Sci. advances* **3**,
650 e1701079 (2017).
- 651 16. V Latora, M Marchiori, Vulnerability and protection of infrastructure networks. *Phys. Rev. E*
652 **71**, 015103 (2005).
- 653 17. M De Domenico, A Solé-Ribalta, S Gómez, A Arenas, Navigability of interconnected networks
654 under random failures. *Proc. Natl. Acad. Sci.* **111**, 8351–8356 (2014).
- 655 18. B Mirzasoleiman, M Babaei, M Jalili, M Safari, Cascaded failures in weighted networks. *Phys.
656 Rev. E* **84**, 046114 (2011).
- 657 19. M Jalili, Error and attack tolerance of small-worldness in complex networks. *J. Informetrics* **5**,
658 422–430 (2011).
- 659 20. R Albert, H Jeong, AL Barabási, Error and attack tolerance of complex networks. *Nature* **406**,
660 378–382 (2000).
- 661 21. A Halu, A Scala, A Khiyami, MC González, Data-driven modeling of solar-powered urban
662 microgrids. *Sci. Adv.* **2**, e1500700 (2016).
- 663 22. D Li, et al., Percolation transition in dynamical traffic network with evolving critical bottlenecks.
664 *Proc. Natl. Acad. Sci.* **112**, 669–672 (2015).
- 665 23. G Zeng, et al., Switch between critical percolation modes in city traffic dynamics. *Proc. Natl.
666 Acad. Sci.* **116**, 23–28 (2019).
- 667 24. L Zhang, G Zeng, S Guo, D Li, Z Gao, Comparison of traffic reliability index with real traffic
668 data. *EPJ Data Sci.* **6**, 19 (2017).
- 669 25. F Wang, D Li, X Xu, R Wu, S Havlin, Percolation properties in a traffic model. *Europhys. Lett.*
670 **112**, 38001 (2015).
- 671 26. M Saberi, et al., A simple contagion process describes spreading of traffic jams in urban
672 networks. *Nat. Commun.* (in press).
- 673 27. P Echenique, J Gómez-Gardenes, Y Moreno, Dynamics of jamming transitions in complex
674 networks. *Europhys. Lett.* **71**, 325 (2005).
- 675 28. LK Gallos, HA Makse, M Sigman, A small world of weak ties provides optimal global in-
676 tegration of self-similar modules in functional brain networks. *Proc. Natl. Acad. Sci.* **109**,
677 2825–2830 (2012).
- 678 29. YN Kenett, et al., Flexibility of thought in high creative individuals represented by percolation
679 analysis. *Proc. Natl. Acad. Sci.* **115**, 867–872 (2018).
- 680 30. J Borge-Holthoefer, Y Moreno, A Arenas, Topological versus dynamical robustness in a lexical
681 network. *Int. J. Bifurc. Chaos* **22**, 1250157 (2012).
- 682 31. SD Reis, et al., Avoiding catastrophic failure in correlated networks of networks. *Nat. Phys.*
683 **10**, 762–767 (2014).
- 684 32. D Li, Q Zhang, E Zio, S Havlin, R Kang, Network reliability analysis based on percolation
685 theory. *Reliab. Eng. & Syst. Saf.* **142**, 556–562 (2015).
- 686 33. R Cohen, S Havlin, *Complex networks: structure, robustness and function*. (Cambridge
687 University Press), (2010).
- 688 34. N Oppenheim, *Urban travel demand modeling: from individual choices to general equilibrium*.
689 (John Wiley and Sons), (1995).
- 690 35. DS Callaway, ME Newman, SH Strogatz, DJ Watts, Network robustness and fragility: Perco-
691 lation on random graphs. *Phys. Rev. Lett.* **85**, 5468 (2000).
- 692 36. LE Olmos, S Çolak, S Shafiee, M Saberi, MC González, Macroscopic dynamics and the
693 collapse of urban traffic. *Proc. Natl. Acad. Sci.* **115**, 12654–12661 (2018).
- 694 37. M Pollack, Letter to the editor—the maximum capacity through a network. *Oper. Res.* **8**,
695 733–736 (1960).
- 696 38. H Hamedmoghadam, M Ramezani, M Saberi, Revealing latent characteristics of mobility
697 networks with coarse-graining. *Sci. Reports* **9**, 1–10 (2019).
- 698 39. M Girvan, ME Newman, Community structure in social and biological networks. *Proc. Natl.
699 Acad. Sci.* **99**, 7821–7826 (2002).
- 700 40. D Ben-Avraham, S Havlin, *Diffusion and reactions in fractals and disordered systems*. (Cam-
701 bridge University Press), (2000).
- 702 41. Y Chen, et al., Percolation theory applied to measures of fragmentation in social networks.
703 *Phys. Rev. E* **75**, 046107 (2007).