Dynamics of functional failures and recovery in complex road networks

Xianyuan Zhan, Satish V. Ukkusuri, and P. Suresh C. Rao

Phys. Rev. E **96**, 052301 — Published 1 November 2017

DOI: 10.1103/PhysRevE.96.052301

Dynamics of Functional Failures and Recovery in Complex Road Networks

Xianyuan Zhan and Satish V. Ukkusuri*
Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA

P. Suresh C. Rao
Lyles School of Civil Engineering and Agronomy Department, Purdue University, West Lafayette, IN 47907, USA
(Dated: October 18, 2017)

We propose a new framework for modeling the evolution of functional failures and recoveries on complex networks, with traffic congestion on road networks as the case study. Different from conventional approaches, we transform the evolution of functional states into an equivalent dynamic structural process – dual vertex splitting and coalescing embedded within the original network structure. The proposed model successfully explains traffic congestion and recovery patterns at the city-scale based on high-resolution data from two megacities. Numerical analysis shows that certain network structural attributes can amplify or suppress cascading functional failures. Our approach represents a new general framework to model functional failure-recovery on flow-based networks, and allows understanding of the interplay between structure and function for flow-induced failure propagation and recovery.

I. INTRODUCTION

Flow-induced functional failures are common phenomena on many human engineered and natural flow-based networks. The functional performance of such networks is reflected by how efficiently flow is propagated across the network, and their functional failures are mostly due to flow overload (e.g. traffic congestion on transportation network; power surge on electric grids; flooding on drainage network). Traffic congestion on urban road networks is a typical case of functional failure processes on flow-based networks. It can be perceived as a form of temporary partial functional failure resulting from high traffic load. Under congestion, certain segments of roads are temporarily closed or operating with reduced efficiency, causing partial or full functional losses in the network. Compared with structural disruptions, functional failures such as traffic congestions are more frequent, and pose significant operational and monetary loss to urban communities.

It is always desired to design infrastructure networks that suppress the emergence and cascading of functional failures. However, even the first step towards this goal, modeling the functional failure process and the resulting network performance, has been shown to be a difficult and not well-understood problem. The structural characteristics and functional features of flow-based networks interact in complex ways that jointly determine how and where the functional failures emerge, how the functional failures propagate and how recovery occurs [1–8].

Traditional approaches on functional failure analysis seek to obtain the network functional performance by solving for flow patterns on the network using optimization-based methods [9, 10] or traffic simulation [11, 12]. This detaches network functional performances from structural details, leading to an incomplete understanding of the underlying mechanisms of the failure-recovery processes. Incorporating the propagation characteristics of real-world flows and their overloading behaviors on structural analysis has long been recognized as a challenging research question [13–15]. This is because the flow patterns on the network are governed by both the flow propagation principles (e.g., traffic equilibrium in traffic networks [9, 10], routing behavior in information networks [16–18], minimum energy dissipation principle in river networks [19, 20], etc.) as well as the network structure [21]. Overlaying flow propagation principles on the structure of networks is difficult, which has been shown to be analytically tractable only under special cases, such as trees [20], star-like and homogeneous structured networks [14] or ring-and-hub structure [2, 22]. To model the functional failures and recoveries in real-world networks, new analytical tools need to be developed to capture the non-trivial interactions between network structure and functions.

We propose a vertex split-recovery model for examining traffic congestion evolution process on urban road networks. Unlike traditional studies that distribute traffic flow on the network and use road capacities to identify congestion [9, 10], the proposed model transforms congestion as a dynamic structural process on the network. The model is built upon a dual representation of road networks augmented with functional states. We show that the congestion/recovery on a road network under this representation is equivalent to the splitting/coalescing of dual vertices. To construct the model, we collect high temporal resolution network traffic state data from two megacities in China (Beijing and Shanghai). Based on the insights from empirical observations, we model the vertex split-recovery process as a composite of four stochastic processes: 1) *self-splitting*, vertex split due to network-wide loading of traffic; 2) *self-contagion*, conges-

* Corresponding author. E-mail: sukkusur@purdue.edu.

tion propagates along the same roads (same dual vertex); 3) *neighbor-contagion*, congestion cascades to neighboring roads (neighboring dual vertex); and 4) *recovery*, congestion on road segments recovers (coalescing of split dual vertices). The proposed model can explain the congestion evolution phenomenon in the real-world data, and provides new insights on the interplay of structure and function in flow-based networks.

II. FUNCTIONAL DUAL MAPPING AND VERTEX SPLIT-RECOVERY

Conventional representation of road networks perceives intersections as vertices and road segments as edges, referred to as primal networks. Recently, the dual representation of road networks is gaining more attention by researchers, which represents the intersections as dual edges, and merges the individual road segments into meaningful stretches of roads represented as dual vertices based on certain criteria, e.g. axial direction, name of the street, road classes or continuity [23–25]. The concept of dual representation is different from the "dual graph" in graph theory, which relates to the faces of planar graphs. The Hierarchical Intersection Continuity Negotiation (HICN) model [25, 26] is among the best dual representation approaches for road networks. According to HICN, two consecutive road segments belong to the same road if they have same road class and the convex angle they form is close to 180 degree. The benefits of using dual representation are that the dual-mapped networks are no longer constrained by the planar embedding, and uncover the underlying network hierarchy. For example, the important roads in the network tend to be long and connected to many other roads, which leads to large vertex degree in the dual-mapped network. An important finding is that both small-world and scale-free properties are observed in the dual representation of road networks [23–25]. It is found that the dual vertex degree distributions of different road networks have similar power-law exponent γ varying between 2-3 [23, 24].

We consider an extended form of HICN dual mapping that incorporates the functional states of the network, referred as *functional dual mapping*. The key idea is to perform dual mapping on a function state encoded network. This treatment enables converting the congestion evolution, a functional process into a structural process (see Fig.1). Fig.1 provides an illustration of functional dual mapping. Consider the road 2 in network G in Fig.1(a), under dual representation, road 2 will be mapped to a dual vertex in Fig.1(c) that consists of road segments KH, HE, EB. Assume road segment HE gets congested and temporarily lost its functional connectivity. From the functional perspective, the current network is equivalent to a network with road segment HE removed (referred as G'). If dual mapping is performed on G', road 2 will be represented as two disconnected dual vertices $2'$ and $2''$ in Fig.1(d). Under functional dual mapping, the congestion and recovery on road segments are equivalent to dual vertex splits and merging of the previously split dual vertices in the functional dual-mapped networks (referred as *function augmented dual network*). We refer to the original dual-mapped network as the *base dual network* $G_D(V_D, E_D)$, which corresponds to the network without any congestion.

To track the dynamic vertex split-recovery process that represents the congestion evolution, the functional dual mapping is applied and creates a series of function augmented dual networks $G_D^t(V_D^t, E_D^t)$ for each time step t (see Appendix A). The variations of dual degree distributions within a day for the function augmented dual networks can be found in Fig.2. The degree distribution of the base dual network can be fitted into a power law distribution $p(k) \sim k^{-\gamma}$ for $k \geq k_{min}$. In the function augmented dual network, high degree dual vertices are more likely to experience vertex split, causing faster probability decay at the tail. This lead to the deviation of power law distribution, but can be better fitted to power law with exponential cutoff distribution $p(k) \sim k^{-\gamma} e^{-\kappa k}$.

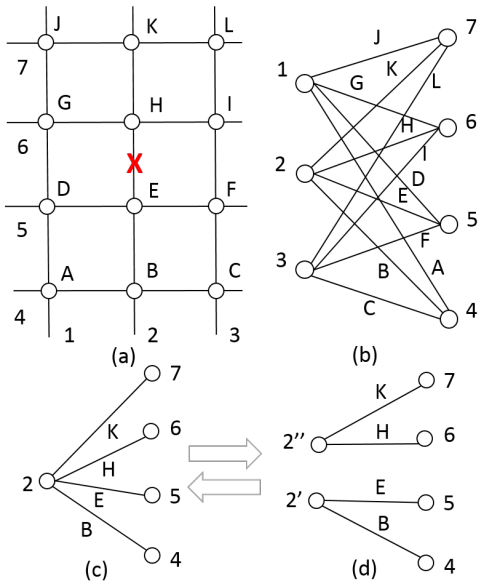
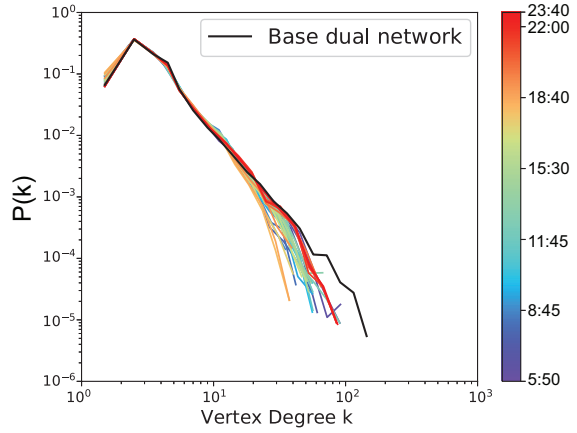


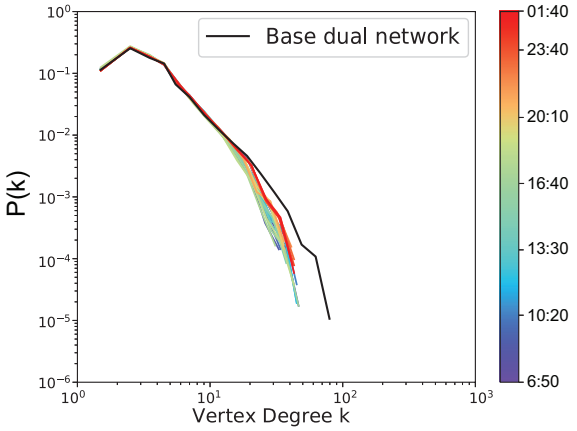
FIG. 1: Illustration of functional dual mapping and vertex split-recovery process. (a) Primal representation of the road network, the roads are labeled as numbers and intersections are labeled as letters. (b) The dual representation of the road network in (a). (c) and (d) are the split and recovery of dual vertex 2 due to congestion on HE segment of road 2.

III. EMPIRICAL OBSERVATIONS

To explore the behavior of real-world vertex split-recovery process, we perform empirical analyses using network congestion evolution data from two megacities in China: Beijing (road network within 4th ring road,



(a)



(b)

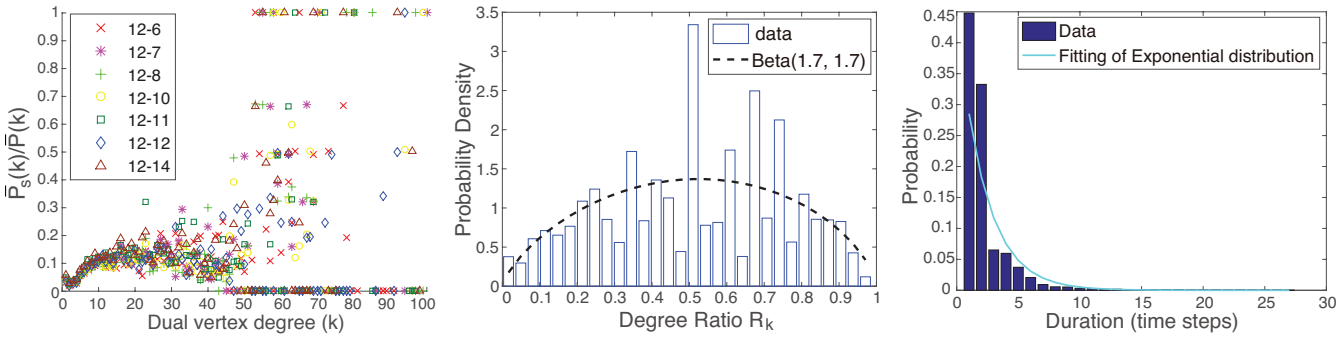
FIG. 2: Plots of the evolution of dual degree distributions for function augmented dual networks: (a) Beijing road network (2015/12/14), (b) Shanghai road network (2016/7/7). The black line and the lines with color represent the degree distribution of the base dual network and the function augmented dual networks for different times in the day. Logarithmic binning is used for better clarity. The degree distribution of the base dual network can be fitted into a power law distribution $p(k) \sim k^{-\gamma}$ for $k \geq k_{min} = 3, \gamma = 2.58$ (Beijing) and $k \geq k_{min} = 4, \gamma = 2.48$ (Shanghai).

contains 17,148 road segments) and Shanghai (road network within the Middle ring road, contains 18,173 road segments). The network-wide link travel times of these two megacities were collected from an on-line digital map service in China (Baidu Map) every 40-60min from 6:00 to 24:00 using a data crawler, which allows for tracking the congestion evolution at an hourly basis (Beijing: around 40min, Shanghai: around 60min). Seven days of network state data from Beijing road network in December 2015 (12/6-12/8, 12/10-12/12 and 12/14), and six days of data from Shanghai in July 2016 (7/7, 7/9-7/13) were collected. All road travel time data were converted to road speeds. The collected link travel times were further compiled into binary functional states (failed state 0: congested; working state 1: not congested) at each time step. We identify a link as congested when its speed is less than 20% of its speed limit. For simplicity, we model the road network as an undirected network. For a road segment that carries bidirectional traffic, we consider the segment in the failed state if traffic of any of the two directions gets congested. Fig.3 presents some observed behaviors for the vertex split-recovery process using the seven days' congestion evolution data from Beijing road network.

The vertex degree and splitting histories of dual vertices are found to be the two major factors governing the vertex splitting process. Important roads in the network usually serve as backbones of the network and tend to have high vertex degrees in the dual-mapped network. These roads are more likely to carry larger amount of traffic, thus prone to congestion. Fig.3a confirms the intuition that the overall normalized vertex split probability shows a positive correlated trend with dual vertex degree, especially low degree dual vertices. The extreme probability values (0 or 1) for high dual degree vertices are caused by their small sample sizes. There are only 1 or 2 such dual vertices in the function augmented dual networks at certain time step, thus more likely to yield probability of 0 or 1.

Numerical analysis also reveals the facilitative impact of splitting history on future vertex splits. Suppose a dual vertex is split at some time step, and then congestion is more likely to propagate along the same road as well as to neighboring roads, causing further vertex splits at later time steps. Some typical examples of this phenomenon are the propagation of traffic kinematic waves [27] and the queue spillover at oversaturated roads [28]. It is observed that the conditional probability of vertex split under the existence of vertex split on the original dual vertex (self-split) in previous time step (P(split|has previous split) = 0.23) is significantly higher than the case when there is no previous self-splits (P(split|no previous split) = 0.03). Similarly, the conditional vertex split probability under the existence of neighbor splits in previous time step (P(split|has previous neighbor split) = 0.05) is about twice as large as the case when none of the neighboring dual vertices split (P(split|no previous neighbor split) = 0.03). This phenomenon is also related to the formation of gridlock in road networks [29], where traffic congestion spreads across neighboring roads and cause severe local network functional failure. Both observations confirm that split history increases the vertex split probability, and the impact of historical self-split is much larger than the splits occurring on neighboring dual vertices.

We also examined the changes of dual degree after each vertex is split. To simplify the analysis, we only considered binary splits. If there are multiple vertex splits at a time step, we decompose them into a series of binary vertex splits. For each binary split, assume the original dual vertex with degree k splits to two sub-vertices with dual degree k' and k'', the ratio $R_k = k'/k$ (referred



(a) (b) (c)

FIG. 3: Behavior of vertex split-recovery process using the seven days' congestion evolution data from Beijing road network. (a) is the plot of normalized vertex split probability. The y-axis in (a) is the normalized vertex split probability $\bar{P}_s(k)/\bar{P}(k)$. $\bar{P}(k)$ and $\bar{P}_s(k)$ are the average proportions of degree k dual vertices and the degree k dual vertices with vertex split in the function augmented dual networks. Both $\bar{P}(k)$ and $\bar{P}_s(k)$ are obtained by averaging the results from function augmented dual networks of all time steps during the day. (b) is the probability distribution of the degree ratio R_k for binary vertex splits, which can be fitted to a Beta distribution ($Beta(1.697, 1.697)$). (c) is the distribution of the duration time t between the time that the vertex gets split and the time it recovers, which can be approximated by exponential distribution.

as degree ratio) is found to be roughly approximated by symmetric Beta distribution $Beta(\beta+1, \beta+1)$ (Fig.3b), where $\beta \sim 0.7$. This indicates that the conditional split probability of a degree k dual vertex can be modeled as: $P(k'|k) \propto [k'(k-k')]^{\beta}$. It suggests that the dual vertices are more likely to split into two dual vertices with similar dual degrees rather than highly unbalanced dual degrees.

Fig.3c presents the distribution of time duration for a split vertex to recover. The time duration is measured as the number of data collection steps (around 40min for Beijing road network). Due to Internet delay during the data collection process, the actual length of each time step can be slightly different (typically 2-5min). It is found that the duration time t can be approximated by an exponential distribution ($P_c(t) = \theta e^{-\theta t}, t > 0$). The recovery process can be captured using a remarkably simple statistical distribution which is common in modeling the survivor times of many physics, biology and economics processes [30, 31]. Moreover, the exponential duration time distribution for the dual vertices that remain in split condition also implies a constant recovery rate in the system.

IV. VERTEX SPLIT-RECOVERY MODEL

We propose a vertex split-recovery model for the functional failure process on urban road networks based on the equivalent representation of road congestion and vertex split/recovery. We introduce following model assumptions based on the insights from the previous empirical observations:

1. Degree 1 dual vertices do not split.

2. *Binary split and no degree loss*: a dual vertex with degree k splits into two sub-vertices with degree i and $k-i$.

3. *Conditional splits*: the resulting degree of the split sub-vertices of a dual vertex with degree k follow some conditional splitting probability distribution $P(i, k-i|k)$.

4. *Self-splitting*: each dual vertex can split with some self-split rate $\rho(k, \eta(t))$ due to the network loading of flow, where $\eta(t)$ is a functional measure of network loading level at time t. We further assume $\rho(k, \eta(t)) = (k-1)\eta(t)$.

5. *Self-contagion*: if a dual vertex with degree k has unrecovered split, it will continue splitting with rate $g(k) = \tau(k-1)$, where τ is a fixed rate.

6. *Neighbor-contagion*: a split dual vertex will cause its neighbor to split with fixed rate λ. Moreover, if the neighboring dual vertex with degree k already splits into sub-vertices, the impact on each sub-vertex with degree k_i is $\lambda\frac{k_i-1}{k-1}$ (ratio of the potential number of splits for the sub-vertex and the original dual vertex).

7. *Recovery*: each vertex split recovers with fixed rate θ.

With above model assumptions, we can simulate the vertex split-recovery process on the function augmented dual network. However, directly solving the detailed network configuration at a particular time step will be analytically intractable. Instead, we are interested in the expected stationary solution of the vertex split-recovery process at the stable state under constant network loading level ($\eta(t) = \eta$). This solution is relatively easy to obtain while providing sufficient insights about the final

impact of vertex split-recovery process. In the following, we present a two-level model to obtain this expected stationary solution. The microscopic-level model characterizes the expected behavior for a degree k dual vertex after S splits, which focuses only on an individual dual vertex. The macroscopic-level model characterizes the evolution of the number of splits for each dual vertex in the entire network.

A. Microscopic-level model

If we remove the recovery history of the dual vertex (if a dual vertex has split but later recovered, we ignore this split) until time t, the splitting history of a dual vertex can be represented as a *ranked planar Markov branching tree* [32]. Markov branching trees were introduced by Aldous [33] as a class of random binary or multifurcating phylogenetic models, which is widely used in phylogenetic studies [33–35]. A *ranked plane tree* is defined as a tree that we distinguish the left and right child vertices of an internal vertex, and every internal vertex is labeled by an integer keeping track of the ordering in which the splits occur during the construction of the tree. The internal ordering are necessary in our case, since each dual vertex is comprised of an ordered set of road segment, and such ordering are preserved during the vertex splits caused by functional failures. Fig.4 presents an illustration of the ranked planar tree representation of the vertex splitting history.

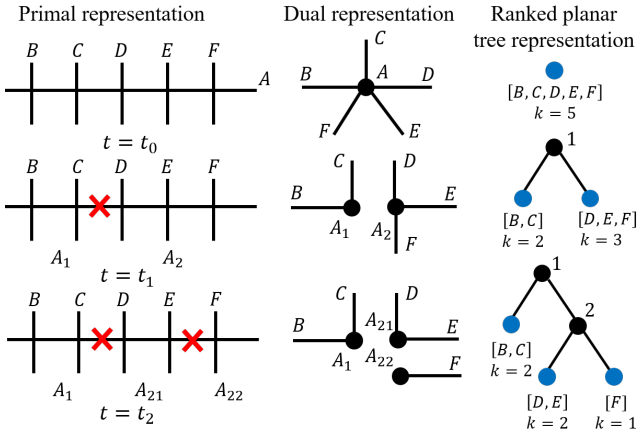


FIG. 4: Illustration of the ranked planar tree representation of vertex splitting history. Road A (dual degree $k = 5$) experiences congestion in time t_1 and t_2, which result in two dual vertex splits. Under the ranked planar tree representation, it can be represented as two consecutive branching of the tree.

Here we consider a special generating process that approximates the ranked planar tree for the splitting history of a dual vertex with degree k having S splits:

1. First split: the first split partition the dual vertex into leaf vertices with degree k_1 and $k-k_1$ according to the conditional splitting distribution $P(k_1, k-k_1|k)$.
2. At ith split: in all leaf vertices, selecting a leaf vertex j of degree k_j with probability proportional to $k_j - 1$. Split vertex j into two new leaf vertices with degree k_{j1} and $k_j - k_{j1}$ according to $P(k_{j1}, k_j - k_{j1}|k_j)$.
3. Stop after all S splits performed.

Based on the empirical observations, we define the conditional splitting distribution has the form of $P(i, k-i|k)$. That is

$$P(i, k-i|k) = \frac{[i(k-i)]^\beta}{\sum_{j=1}^{k-1}[j(k-j)]^\beta} \tag{1}$$

When $k \to \infty$, the normalized splitting location i/k asymptotically follows symmetric Beta distribution ($Beta(\beta+1, \beta+1)$). This generative process is similar to the incremental construction method in the Beta-splitting model for evolutionary trees proposed by Sainudiin and Véber [32], which constructs the evolutionary tree by incrementally partitioning an interval (vertices are represented as intervals). The value of β represent different splitting behaviors:

- $\beta = 0$: uniform split. The split location is chosen uniformly at $k-1$ possible locations.
- $\beta > 0$: split location favors more in the middle section. This is the case reflected in the empirical observation, where we have $\beta \sim 0.7$.
- $\beta < 0$: split location favors more on splitting in the two ends.

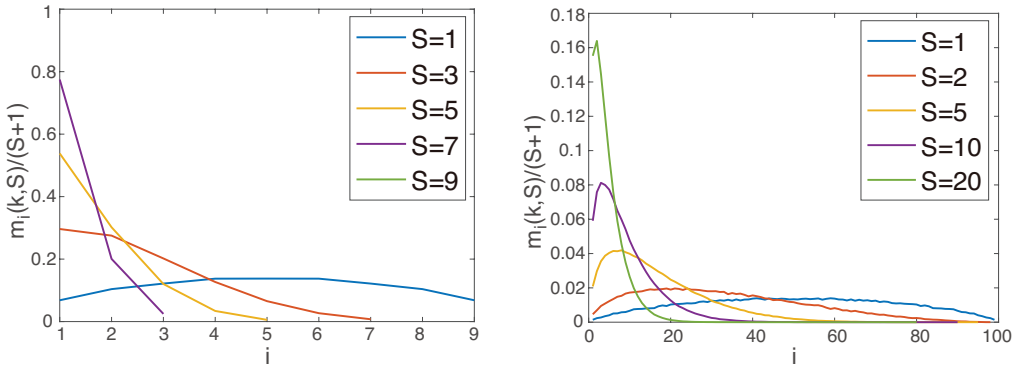
In the microscopic-level model, we are interested in the expected degrees of leaf vertices for a degree k dual vertex after S splits. Define $M(k,S) = [m_1(k,S), m_2(k,S), \cdots]^T$, where $m_i(k,S)$ is the expected number of the split sub-vertices with degree i for a dual vertex with degree k after S splits. For $S = 1$, it can be easily shown that

$$m_i(k,1) = \begin{cases} \frac{2[i(k-i)]^\beta}{\sum_{j=1}^{k-1}[j(k-j)]^\beta}, & i < k \\ 0, & i \geq k \end{cases} \tag{2}$$

For $S > 1$ and $i = 1, 2, \cdots, k-S-1$, it can be proved following recursive formulation holds (see Appendix B):

$$m_i(k,S+1) = \frac{k-S-i}{k-S-1} m_i(k,S) + \frac{1}{k-S-1}\sum_{s=i+1}^{k-S}(s-1)m_s(k,S)m_i(s,1) \tag{3}$$

Although the closed form expression for above recursive formulation is only known for the reduced case $\beta = 0$ (see Appendix C), $m_i(k,S)$ can be computed by recursively solving Eq.3 or using Monte Carlo simulation based on the generating process. Numerical tests show that the expected degree distribution for the split sub-vertices $\frac{m_i(k,S)}{S+1}$ asymptotically follows a Beta distribution $Beta(\tilde{\alpha}(k,S), \tilde{\beta}(S))$, where $\tilde{\alpha}(k,S)$ is monotonically increasing with the increase of both k and S, and $\tilde{\beta}(S)$ decreases with the increase of S (Fig.5).



(a) (b)

FIG. 5: Monte Carlo simulation results of the expected proportion of the degree i split sub-vertices after S splits $(m_i(k,S)/(S+1))$ for degree k dual vertices: (a) $k=10$, (b) $k=100$. Computed using 50,000 simulation runs of the generating process of the microscopic-level model.

B. Macroscopic-level model

The macroscopic-level model focuses on the entire network rather than individual dual vertices. It characterizes the evolution of the number of splits for each dual vertex in the entire network at stable state $(t\to\infty)$ under constant network loading level $(\eta(t)=\eta)$. We model the evolution of the number of splits of a dual vertex v with degree k in the network as a continuous time Markov chain involving k states $(0,1,...,k-1)$. Each state represents the number of split S_v . To simplify the model, we further assume that the states can only increase or decrease S_v by 1 at each transition. This can be perceived as a special generalization of the Susceptible-Infected-Susceptible (SIS) model in epidemiology [36]. The differences are that SIS model only considers two states: susceptible (S) and infected (I) and the transition between these two states; while in the macroscopic-level model, we have transitions among k states.

For a dual vertex v in the based dual network, let k_i be the degree of its ith child vertex after splitting. Define the sign function $sgn(S_v)=1$ if $S_v>0$; 0 otherwise. The rate of transition between states can be modeled as follows:

- Transition rate from $S_v\to S_v+1$ $(S_v<k-1)$:

$$\sum_{i=1}^{S_v+1}\Big[\rho(k_i,\eta)+g(k_i)sgn(S_v)+\lambda\frac{k_i-1}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big] = (k-1-S_v)\Big[\eta+\tau\cdot sgn(S_v)+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big] \quad (4)$$

- Transition rate from $S_v\to S_v-1$ $(S_v>0)$: θS_v

Eq.4 computes the overall rate of increase in S_v contributed by its S_v+1 child vertices, where $\rho(k_i,\eta)$ and $g(k_i)sgn(S_v)$ are the self-splitting and self-contagion rates of the ith child vertex; $\lambda\frac{k_i-1}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)$ is the overall neighbor-contagion rate on the ith child vertex from all the split neighboring dual vertices $(sgn(S_u)>1)$.

Let probability distribution $X^v(t)=(x_0^v(t),x_1^v(t),...,x_{k-1}^v(t))$, where $x_S^v(t)$ is the probability of dual vertex v at time t having S splits. Suppose all dual vertices initially do not have vertex splits $(x_0^v(0)=1,\forall v\in V_D)$, using the previous transition scheme, the state equation for each dual vertex v at time t in the network can be written as:

$$\begin{aligned}
\frac{dx_0^v(t)}{dt}&=-x_0^v(t)(k-1)\Big[\eta+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]+\theta x_1^v(t)\\
\frac{dx_1^v(t)}{dt}&=-x_1^v(t)(k-2)\Big[\eta+\tau+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]\\
&-\theta x_1^v(t)+x_0^v(t)(k-1)\Big[\eta+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]+2\theta x_2^v(t)\\
\frac{dx_S^v(t)}{dt}&=-x_S^v(t)(k-1-S)\Big[\eta+\tau+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]\\
&-S\theta x_S^v(t)+x_{S-1}^v(t)(k-S)\Big[\eta+\tau+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]\\
&+(S+1)\theta x_{S+1}^v(t),\quad 1<S<k-1\\
\frac{dx_{k-1}^v(t)}{dt}&=x_{k-2}^v(t)\Big[\eta+\tau+\frac{\lambda}{k-1}\sum_{(u,v)\in E_D}sgn(S_u)\Big]\\
&-(k-1)\theta x_{k-1}^v(t)
\end{aligned} \quad (5)$$

The state equations model the changes of $x_S^v(t)$ caused by four different types of state transitions: 1) the decrease due to transition from S to $S+1$; 2) the decease due to recovery from S to $S-1$; 3) the increase due to transition from $S-1$ to S; and 4) the increase due to recovery from $S+1$ to S.

Directly solving such a large coupled differential equation system is highly complex. We utilize the *degree-based mean field* (DBMF) approximation [37–40] to gain some insights on the behavior of the stationary solution. The DBMF approximation for dynamical processes on networks assumes that all vertices of degree k are statistically equivalent. Under DBMF approximation, we only need to consider $X^k(t)=(x_0^k(t),x_1^k(t),...,x_{k-1}^k(t))$, where $x_S^k(t)$ is the proportion of degree k dual vertices

that has S splits. Specifically, we consider two cases: (1) no degree correlation: the average probability that a dual vertex has a split neighbor at time t, $\phi(t)$ is same for all dual vertices; (2) general degree correlation: the average probability that a dual vertex of degree k has a split neighbor $\tilde{\phi}(t|k)$ depends on its degree.

1. DBMF approximation with no degree correlation

If we assume there is no degree correlation in the network, then the overall neighbor-contagion rate for a dual vertex v with degree k can be approximated as [40, 41]:

$$\lambda \sum_{(u,v)\in E_D} sgn(S_u) \simeq \lambda k \phi(t) \tag{6}$$

where $\phi(t)$ is the average probability that a dual vertex has a split neighbor at time t. If there is no degree correlation, $\phi(t)$ is given as [40]:

$$\phi(t) = \sum_{k'>1} \frac{k'P(k')}{\langle k\rangle}(1 - x_0^{k'}(t)) \tag{7}$$

where $P(k)$ is the proportion of degree k dual vertices in the based dual network, and $\langle k\rangle$ is the average degree of the base dual network. The above expression gives the average probability of finding a split dual vertex following a randomly chosen dual edge. For simplicity, let

$$\Lambda_1^k(t) = \frac{\eta + \frac{\lambda k}{k-1}\phi(t)}{\theta}, \quad \Lambda_2^k(t) = \Lambda_1^k(t) + \frac{\tau}{\theta} \tag{8}$$

Under DBMF approximation, the state equations Eq.5 can be simplified as

$$\begin{aligned}
&\frac{dx_0^k(t)}{dt} = -x_0^k(t)(k-1)\theta\Lambda_1^k(t) + \theta x_1^k(t)\\
&\frac{dx_1^k(t)}{dt} = -x_1^k(t)(k-2)\theta\Lambda_2^k(t) - \theta x_1^k(t) +\\
&\quad x_0^k(t)(k-1)\theta\Lambda_1^k(t) + 2\theta x_2^k(t)\\
&\frac{dx_S^k(t)}{dt} = -x_S^k(t)(k-1-S)\theta\Lambda_2^k(t) - S\theta x_S^k(t)\\
&\quad + x_{S-1}^k(t)(k-S)\theta\Lambda_2^k(t) + (S+1)\theta x_{S+1}^k(t),\\
&\quad 1 < S < k-1\\
&\frac{dx_{k-1}^k(t)}{dt} = x_{k-2}^k(t)\theta\Lambda_2^k(t) - (k-1)\theta x_{k-1}^k(t)
\end{aligned} \tag{9}$$

We are interested in the stable proportion of dual vertices of degree k that has i splits ($x_i^k(t)$) when $t \to \infty$, which correspond to the stationary distribution of the continuous time Markov chain. The stationary distribution can be analytically solved by setting $\frac{dx_S^k}{dt} = 0, \forall S = \{0, 1, 2, ..., k-1\}, \forall k$. Solving the system of equations, we obtain for $k > 1$:

$$x_i^{k*} = \binom{k-1}{i}\Lambda_1^k\big(\Lambda_2^k\big)^{i-1}x_0^{k*}, \quad i = 1, 2, \cdots, k-1 \tag{10}$$

where Λ_1^k, Λ_2^k are $\Lambda_1^k(t), \Lambda_2^k(t)$ with $\phi(t)$ value taken at the stationary solution (ϕ^*). As $\sum_{i=0}^{k-1} x_i^{k*} = 1$, thus

$$\begin{aligned}
x_0^{k*} &= \left[1 + \frac{\Lambda_1^k}{\Lambda_2^k}\sum_{i=1}^{k-1}\binom{k-1}{i}\big(\Lambda_2^k\big)^i\right]^{-1}\\
&= \left[1 + \frac{\Lambda_1^k}{\Lambda_2^k}\Big[\big(1+\Lambda_2^k\big)^{k-1} - 1\Big]\right]^{-1}
\end{aligned} \tag{11}$$

Solving for ϕ^* using Eq.7, we have

$$\begin{aligned}
\phi^* =& \frac{1}{\langle k\rangle}\sum_{k'>1} k'P(k')\cdot\\
&\frac{-1 + \left[1 + \frac{\eta+\tau+\frac{\lambda k'}{k'-1}\phi^*}{\theta}\right]^{k'-1}}{\frac{\tau}{\eta+\frac{\lambda k'}{k'-1}\phi^*} + \left[1 + \frac{\eta+\tau+\frac{\lambda k'}{k'-1}\phi^*}{\theta}\right]^{k'-1}}
\end{aligned} \tag{12}$$

Above nonlinear equation can be solved numerically. Using the solution of ϕ^*, we can obtain x_0^{k*} and x_i^{k*}.

2. DBMF approximation with general degree correlation

For the case when considering the general degree correlation, the overall neighbor-contagion rate can be approximated as

$$\lambda \sum_{(u,v)\in E_D} sgn(S_u) \simeq \lambda k \tilde{\phi}(t|k) \tag{13}$$

where $\tilde{\phi}(t|k)$ is the average probability that a degree k dual vertex has a split neighbor. It can be computed using $P(k'|k)$, the conditional probability of a degree k dual vertex that has a degree k' neighbor as follows [40]:

$$\tilde{\phi}(t|k) = \sum_{k'} P(k|k')(1 - x_0^{k'}) \tag{14}$$

$P(k'|k)$ can be evaluated empirically from the actual base dual network, which incorporates more structural details of the network. Under such condition, the $\Lambda_1^k(t)$ and $\Lambda_2^k(t)$ in the no degree correlation case now become

$$\tilde{\Lambda}_1^k(t) = \frac{\eta + \frac{\lambda k}{k-1}\tilde{\phi}(t|k)}{\theta}, \quad \tilde{\Lambda}_2^k(t) = \tilde{\Lambda}_1^k(t) + \frac{\tau}{\theta} \tag{15}$$

It can be shown the state equation Eq.9 still holds, but with $\Lambda_1^k(t)$ and $\Lambda_2^k(t)$ replaced by $\tilde{\Lambda}_1^k(t)$ and $\tilde{\Lambda}_2^k(t)$. The final stationary solutions can be derived as

$$x_i^{k*} = \binom{k-1}{i}\tilde{\Lambda}_1^k\big(\tilde{\Lambda}_2^k\big)^{i-1}x_0^{k*}, \quad i = 1, \cdots, k-1 \tag{16}$$

$$x_0^{k*} = \left[1 + \frac{\tilde{\Lambda}_1^k}{\tilde{\Lambda}_2^k}\Big[\big(1+\tilde{\Lambda}_2^k\big)^{k-1} - 1\Big]\right]^{-1} \tag{17}$$

where $\tilde{\Lambda}_1^k, \tilde{\Lambda}_2^k$ are $\tilde{\Lambda}_1^k(t), \tilde{\Lambda}_2^k(t)$ with $\tilde{\phi}(t|k)$ taken the value of $\tilde{\phi}(\infty|k)$ ($\tilde{\phi}(t|k)$ value at stationary solution ${x_i^k}^*$.

Under the general degree correlation, there is no simple form stationary solution due to the coupling of dual vertices of different degree. The state equations for ${x_0^2}^*, {x_0^3}^*, {x_0^{k_{max}}}^*$ now becomes all coupled, as the computation of $\tilde{\Lambda}_1^k$ and $\tilde{\Lambda}_2^k$ involves $\tilde{\phi}(\infty|k)$ and all ${x_0^k}^*, k = 1, 2, ..., k_{max}$. However, ${x_i^k}^*$ is still numerically solvable for finite networks, which can be obtained by solving a large system of nonlinear equations involving Eq.14-17 for all dual vertex degrees.

C. Dual vertex degree distribution under stationary solution

Combining the analytical results from the microscopic and macroscopic-level model, the expected total number of dual vertices N^* and the dual vertices of degree k, N_k^* in the stable function augmented dual network are:

$$N^* = n \sum_{k=1}^{k_{max}} P(k) \sum_{S=0}^{k-1} (S+1) {x_S^k}^* \tag{18}$$

$$N_k^* = n \sum_{k'=k}^{k_{max}} P(k') \sum_{S=0}^{k'-1} m_k(k', S) {x_S^{k'}}^* \tag{19}$$

where n and k_{max} are the total number of dual vertices and the maximum degree in the base dual network. The expected degree distribution for the function augmented dual network under stable condition can thus be obtained as $P^*(k) = N_k^*/N^*$.

V. NUMERICAL RESULTS

A. Fitting to real world data

The stationary solution of vertex split-recovery model under DBMF approximation is completely determined by three key parameters, namely: 1) normalized network loading level $w_1(t) = \eta(t)/\theta$; 2) normalized self-contagion rate $w_2 = \tau/\theta$ and 3) normalized neighbor contagion rate $w_3 = \lambda/\theta$. Here, $w_1(t)$ is a dynamic variable that measures the relative loading level of the entire network; w_2 and w_3 are fixed parameters governed by network structure.

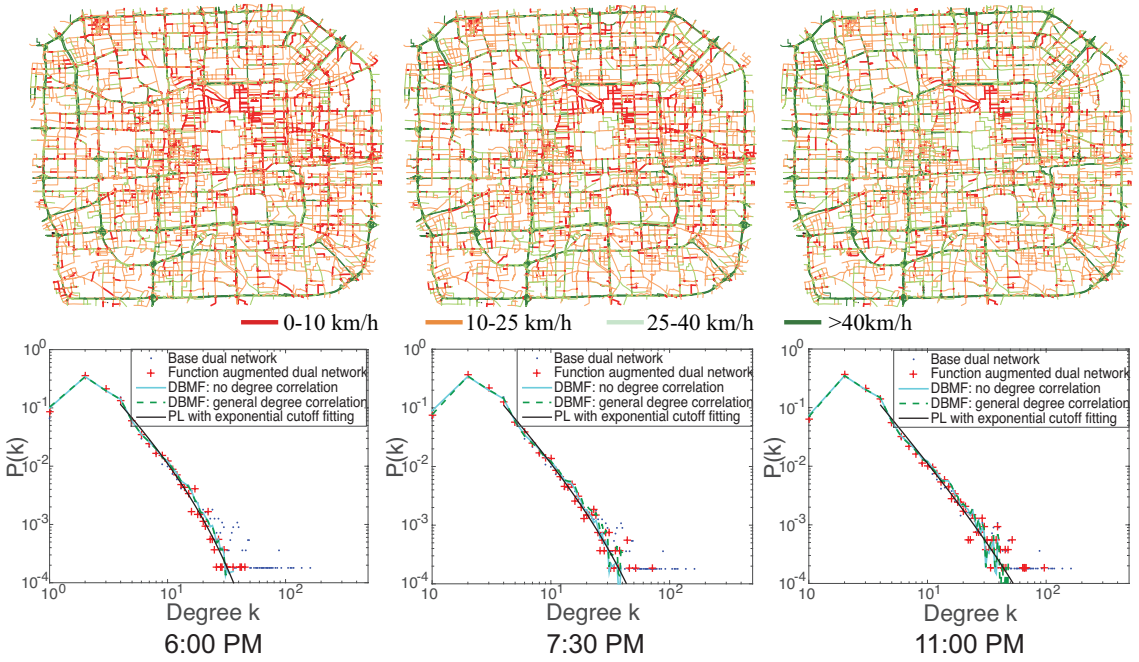
We fit the proposed model to empirical data to uncover the actual $w_1(t), w_2$ and w_3 values in the test networks. The fitting is achieved by minimizing the overall statistical divergence between the model predicted and empirical dual degree distributions of the function augmented dual networks. Let $P_T^*(k)$ and $Q_T(k)$ be the model predicted and empirical dual degree distribution for time period T. We minimize the overall statistical divergence of the two distributions over the entire day, defined as $min_{w_1(t),w_2,w_3} \sum_T J(P_T^*||Q_T)$, where $J(P_T^*||Q_T)$ is the Jensen-Shannon divergence [42] between distribution P_T^* and Q_T. Jensen-Shannon divergence is a popular measure for evaluating the dissimilarity between two probability distributions and widely used in statistics and information theory.

Fig.6 presents sample fitting results of the vertex split-recovery model (see the Supplemental Material (SM) for complete fitting results [43]). The larger extent of deviation for dual vertex degree distribution under severe congestion against no congestion condition can be clearly observed. It is observed that the vertex split-recovery model captures the expected behavior of dual degree distribution of the function augmented dual networks for both cities under different congestion levels. The fitted results for w_2 and w_3 are found to be different in Beijing and Shanghai road networks, but within same range. The mean and standard deviation of w_2 and w_3 in Beijing road network under DBMF with no degree correlation are $\bar{w_2} = 0.0744, \sigma(w_2) = 8.3 \times 10^{-3}, \bar{w_3} = 4.3 \times 10^{-4}, \sigma(w_3) = 3.3 \times 10^{-4}$ ($\bar{w_2} = 0.0687, \sigma(w_2) = 0.0109, \bar{w_3} = 3.6 \times 10^{-4}, \sigma(w_3) = 2.0 \times 10^{-4}$ for DBMF with general degree correlation). The values for Shanghai road network are $\bar{w_2} = 0.125, \sigma(w_2) = 0.015, \bar{w_3} = 5.0 \times 10^{-4}, \sigma(w_3) = 7.0 \times 10^{-6}$ ($\bar{w_2} = 0.0915, \sigma(w_2) = 0.0099, \bar{w_3} = 8.6 \times 10^{-4}, \sigma(w_3) = 4.0 \times 10^{-5}$ for DBMF with general degree correlation). Shanghai road network exhibits larger w_2 and w_3 values than Beijing road network, likely because it has a more decentralized and homogeneous network structure. Beijing road network has a typical ring-and-radial structure, where the major ring roads and radial express ways have high traffic capacity; whereas Shanghai road network is more grid-like. This makes congestion overall easier to propagate along same or neighboring roads in Shanghai road network (higher w_2 and w_3 values).

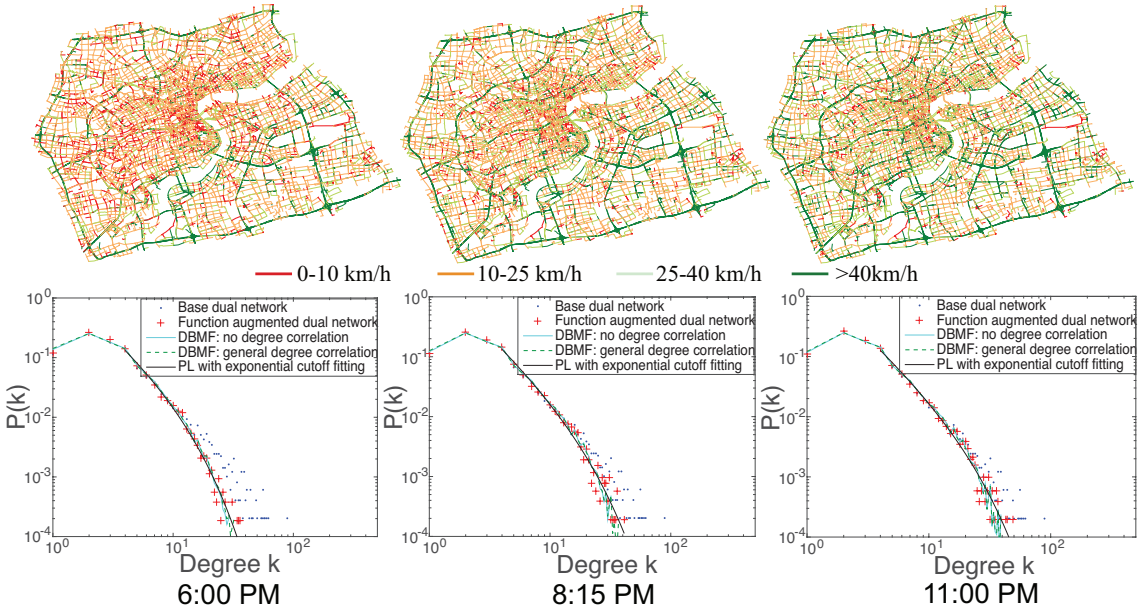
Fig.7 presents the evolution of fitted $w_1(t)$ for a typical weekday and a weekend for both Beijing and Shanghai road networks. It agrees well with the real-world network traffic loading level, where high traffic loading in morning and evening peaks as well as low traffic loading in off-peak hours are well captured. Beijing and Shanghai have shown slightly different traffic loading patterns, where higher traffic loading is observed from noon to early afternoon period (12:00 - 17:00) in Shanghai network. These results show that $w_1(t)$ can be a good measure for the actual traffic loading level at network-scale, which reflects the explanatory power of the vertex split-recovery model.

To further validate the vertex split-recovery model, we compare the total number of dual vertices and the vertex splits between the empirical data and the expected values obtained from the vertex split-recovery model. The expected number of vertex splits N^* is computed using Eq.19, and the expected total number of vertex splits N^{ES} is computed as follows:

$$N^{ES} = n \sum_{k=2}^{k_{max}} P(k) \sum_{s=1}^{k-1} s \cdot {x_s^k}^* \tag{20}$$



(a)



(b)

FIG. 6: Evolution of dual degree distributions for: (a) Beijing road network (2015/12/10), and (b) Shanghai road network (2016/7/11). The top figures in (a) and (b) are the road speed plots at three sample time steps. The bottom figures in (a) and (b) present the corresponding dual vertex degree distributions for: 1) base dual network (no congestion); 2) function augmented dual network (actual congestion scenario); 3) network configurations predicted by vertex split-recovery model under different DBMF approximation schemes (no degree correlation and general degree correlation); and 4) power-law with exponential cutoff fittings of the dual degree distributions.

where $x_s^{k^*}$ is the stationary solution of the proportion of degree k dual vertices that has s splits. The results of N^* and N^{ES} under both DBMF approximation with no degree correlation and general degree correlation are computed. Two metrics, namely mean absolute error (MAE) and mean relative error (MRE), are used for evaluation, which are computed as follows:

$$\mathcal{E}_{MAE} = \frac{\sum_{i=1}^{n} |d_i - \hat{d}_i|}{n}, \quad \mathcal{E}_{MRE} = \frac{\sum_{i=1}^{n} |d_i - \hat{d}_i|}{\sum_{i=1}^{n} d_i}$$

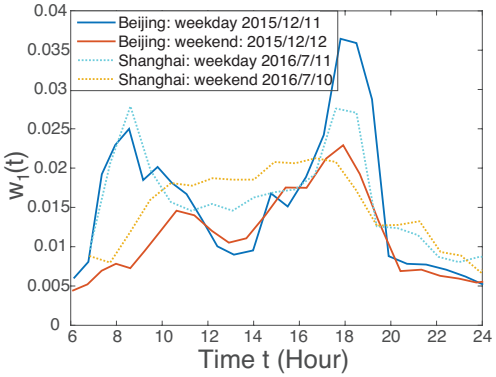


FIG. 7: Fitting result for $w_1(t)$ using DBMF approximation with general degree correlation.

where d_i is the empirical observation i and $\hat{d}_i$ is the value computed using the vertex split-recovery model. Fig.8-9 present the sample results for a typical weekday and a weekend for both Beijing and Shanghai road networks (complete results see the Supplemental Material [43]).

The comparison results show that N^* and N^{ES} computed from the vertex split-recovery model generally agree with the patterns of empirical data. It is found that for the total number of dual vertices N^*, the MRE and MAE for Beijing road network are less than 2.5% and 150 respectively. For Shanghai road network, the MRE and MAE are less than 1.5% and 80 respectively. For the total expected number of vertex splits N^{ES}, the MRE and MAE for Beijing road network are less than 16% and 105 respectively. The results of MRE and MAE for Shanghai road network are less than 11% and 72 respectively. On weekends, when the traffic loading level is lower, even smaller MAE and MRE results are achieved for both cities.

It should be noted that N^* and N^{ES} are obtained using the stationary solution of the vertex split-recovery model, which corresponds to the stable network configuration under constant traffic loading level after sufficiently long observation period. As the real world traffic loading is constantly changing, such stable network configuration is not easily achievable. It is reasonable to expect certain level of discrepancy between the empirical data and the predictions from the model. Considering the relatively low level of MAE and MRE values, the vertex split-recovery model provides a reasonably well approximation of the real world congestion evolution process.

B. Expected functional loss for dual vertices

There is a minor difference in the dual vertex degree distribution results between DBMF approximation with no degree and general degree correlation (Fig.6). This is because considering general degree correlation only improves the approximation of the neighbor-contagion process. Given that the normalized neighbor-contagion rate w_3 is very small compared with other parameters, the impact of considering degree correlation is limited.

As w_3 is very small, if set $w_3 = 0$, we can approximate the expected number of vertex splits for dual vertices with degree k ($k > 1$) under stationary solution in close form:

$$N_k^{ES} = \sum_{s=1}^{k-1} s \cdot x_s^{k*} \approx \frac{w_1(k-1)(1+w_1+w_2)^{k-2}}{1+\frac{w_1}{w_1+w_2}[(1+w_1+w_2)^{k-1}-1]} \quad (21)$$

When normalized by the maximum possible number of vertex split $k-1$, it can be shown $\frac{N_k^{ES}}{k-1} \leq \frac{w_1+w_2}{1+w_1+w_2}$ and the equal sign is attainable only when $k \to \infty$. The quantity $N_k^{ES}/(k-1)$ can be perceived as a measure of the expected level of functional loss for dual vertices of degree k , for which $N_k^{ES}/(k-1) = 0$ indicates no congestion present and 1 indicates all the segments of the road are congested.

Fig.10 presents the impact of w_1 on $N_k^{ES}/(k-1)$ for dual vertices of different degrees. $N_k^{ES}/(k-1)$ converges to $\frac{w_1+w_2}{1+w_1+w_2}$ as k becomes large. This leads to an interesting observation, that although high dual degree vertices are more likely to experience vertex split, they seem to converge to the same level of expected functional loss. In the vertex split-recovery model, dual vertices with more splits have higher recovery rate (more congested segments are likely to recover), which suppresses dual vertices from further splitting. This is similar to the existence of the equilibrium state between the susceptible and infected population in SIS model. One possible real world explanation is that, a severely congested road will lose considerable fraction of functional connectivity, which might forbid further entry and propagation of traffic flow; this in turn could impede the further worsening of congestion. Another possible explanation is related to the moving jam effect [44, 45], that the road segments get congested and recovered as the jam moves along the road. In this situation, although congestion is present, the road still maintains certain functional level.

C. Network Performance

The vertex split-recovery model can serve as a simple but powerful tool in analyzing network performance under real or hypothetical traffic loads. Fig.11a plots the impact of w_1 on the average probability that a dual vertex has a split neighbor ϕ^*. A high value of ϕ^* corresponds to a heavily congested network. We find that Shanghai network performs better than Beijing road network under light traffic loading levels, but tends to perform slightly worse under high loading situations. As mentioned previously, Shanghai network is less hierarchical and grid-like, making it more resistive to the emergence of congestion, but at the same time, opens more channels for the further propagation of congestion, causing worse performance at high loading levels. On the contrary, Beijing road network depends more on major ring and radial expressways,

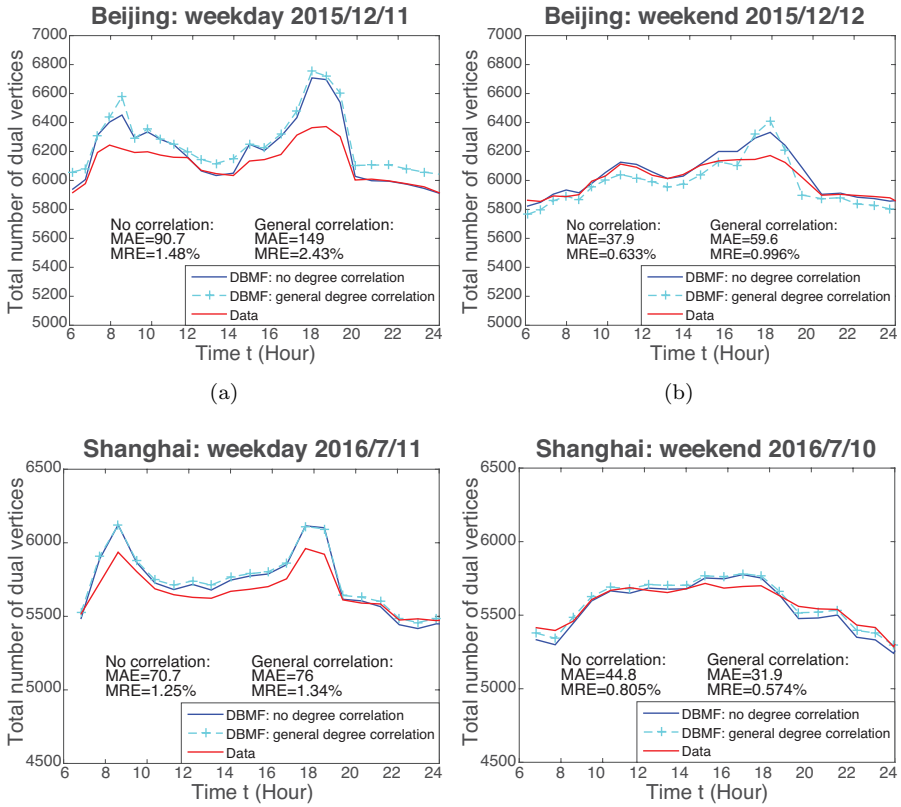


FIG. 8: Comparison of the total number of dual vertices from emprical data (red line) and the expected total number of dual vertices (N^*) obtained from the vertex split recovery model for Beijing and Shanghai road network.

which increases the chances of the emergence of congestion. Existence of high capacity major roads also allows for handling higher traffic volume and the hierarchical structure reduces potential congestion propagation pathways under high traffic loading. The differences of the network performances is further reflected on the dual degree distributions of the functional augmented dual networks. Fig.11b and 11c present the comparison of the two networks under the same set of network loading levels. Larger extent of deviation of the dual degree distribution can be observed in Shanghai network under high network loading levels compared to Beijing road network.

Another interesting finding is the close relationship of vertex split-recovery model with the emergence of exponential cutoff behavior in the dual degree distribution. As shown in Fig.6, power law with exponential cutoff distribution ($P(k) \sim k^{-\gamma}e^{-\kappa k}$) can serve as an excellent approximation of the stationary solution of vertex split-recovery model for $k \geq k_{min}$ ($k_{min} = 3$ for Beijing network and $k_{min} = 4$ for Shanghai network). Unfortunately, the complex mathematical form of the stationary solution forbids the analytical derivation of this corresponding relationship. However, when compared with the fitted normalized network load level w_1 and the parameters of γ and κ, a remarkably simple linear relationship can be shown. Fig.12 presents the sample results of the $\gamma - w_1$ and $\kappa - w_1$ relationships for Beijing and Shanghai road networks. We observe that $\gamma \approx \alpha_0 - \alpha_1 w_1$ and $\kappa \approx \alpha_2 w_1$ (the constant term in $\kappa - w_1$ linear fitting is relatively small and can be ignored), where α_0, α_1 and α_2 are positive constants. This analysis shows that the normalized network loading level w_1 plays a central role that causes the dual vertex degree distribution of function augmented dual network deviates from power law distribution, and the emergence of the exponential cutoff behavior.

VI. CONCLUSION

Our theoretical analysis, combined with observational data from two megacities in China, enables examination and prediction of how functional failures (traffic congestion) and recoveries evolve on structurally intact road networks. We successfully described the deviation of the dual vertex degree distribution from power law distribution, and the emergence of exponential cutoff behavior under congestion using a vertex split-recovery model. The model links the network-level functional loss with the traffic loading level, and provides a statistical char-

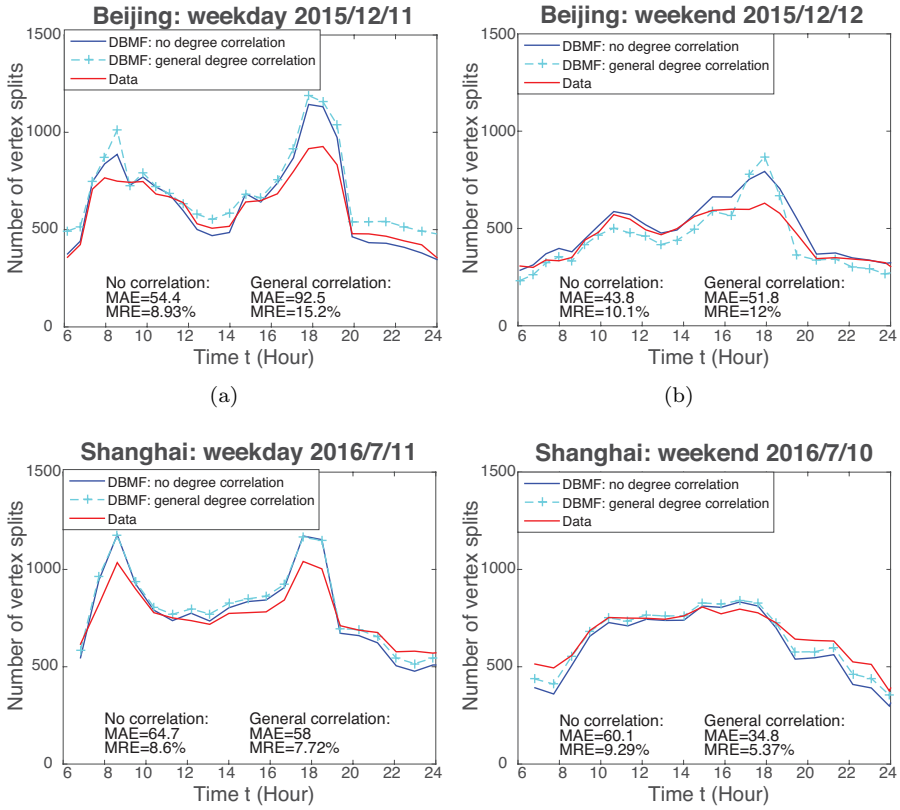


FIG. 9: Comparison of the total number of unrecovered vertex splits at each time step in empirical data (red line) and the expected number of vertex splits (N^{ES}) obtained from the vertex split recovery model for Beijing and Shanghai road network.

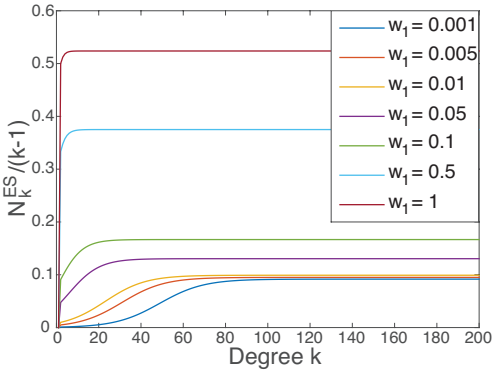
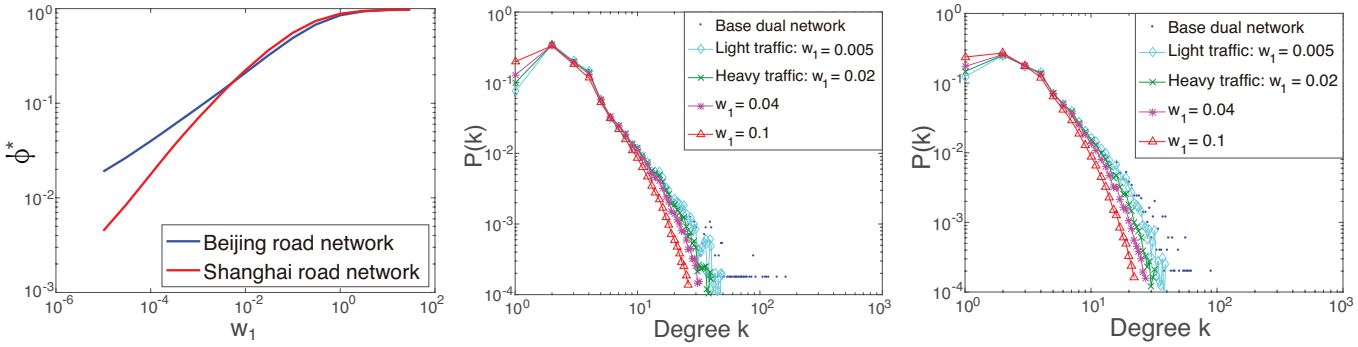


FIG. 10: Impact of w_1 on the expected level of functional loss ($N_k^{ES}/(k-1)$) for dual vertices with different degrees ($w_2 = 0.1, w_3 = 0$).

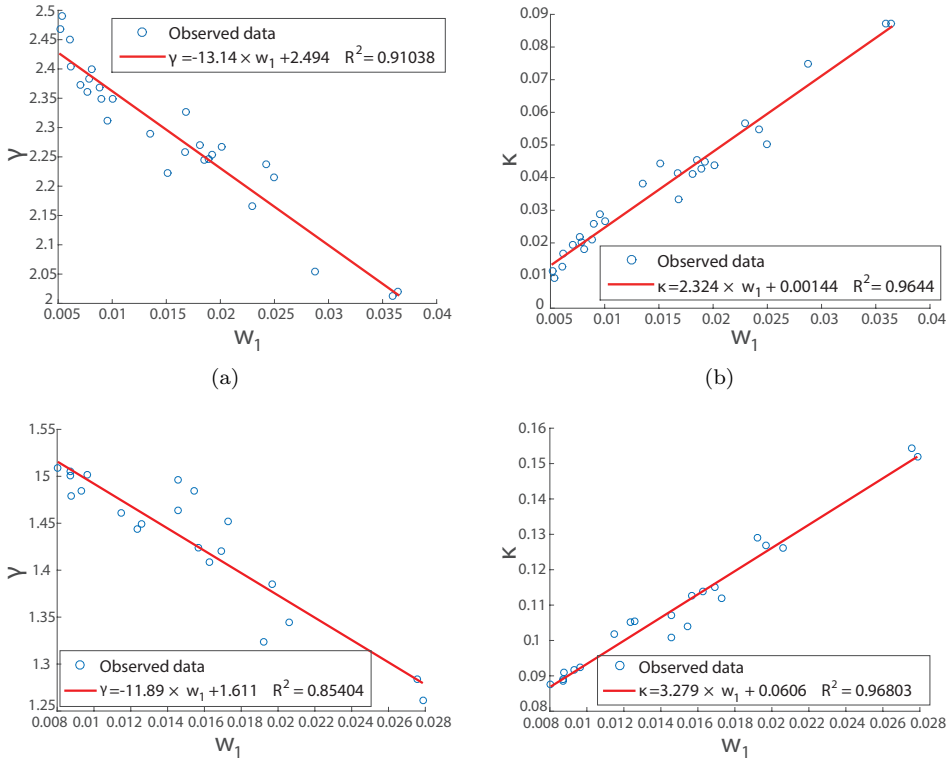
acterization of the likelihood of experiencing functional failures for dual vertices. We also show that certain network topological features can amplify network functional failures through negative feedbacks, while other topological features can suppress congestion cascading. Evolution of functional failure depends on the interaction of network topology and loading. Grid-like road network in Shanghai performs better at low traffic loadings, but propagates congestion at higher loadings. In contrast, ring-and-radial structure of Beijing road network allows higher traffic volumes, but is vulnerable to emergence of congestion.

Performing functional analysis by overlaying real-world flow propagation principles on the structural details of flow-based networks is a challenging problem due to the need of modeling flows. Our work provides a new scientific approach to tackle this complex problem. Instead of analyzing the flow pattern on the network, we can model the flow induced functional process as an equivalent structural process. The vertex split-recovery model is a perfect example showing how traffic congestion on road networks can be modeled as a structural process on a transformed graph, which combines both the flow propagation principle of traffic and network structure. By finding appropriate equivalent structural processes for different types of flow and networks, a similar analysis approach can be applied to other flow-based networks. Future research can be done to further explore such equivalent structural processes, and develop a generalized theory for flow-based complex networks.



(a) (b) (c)

FIG. 11: Comparison of network performance under different normalized network loading levels. (a) Impact of w_1 on the average probability that a dual vertex has a split neighbor at the stationary solution (ϕ^*). Obtained by solving for the stationary solution under the DBMF approximation with no degree correlation assumption. (b) and (c) are plots of the dual degree distributions of Beijing and Shanghai road networks under the same set of normalized network loading levels. The light traffic condition is taken as the typical w_1 value in late night hours; heavy traffic condition is selected as the typical peak hour w_1 value.



(a) (b) (c) (d)

FIG. 12: Relationship of the parameters of power law with exponential cutoff distribution and the fitted w_1 value. (a), (b): linear fitting of $\gamma - w_1$ and $\kappa - w_1$ relationships for Beijing road network (2015/12/11). (c), (d): linear fitting of $\gamma - w_1$ and $\kappa - w_1$ relationships for Shanghai road network (2016/7/11).

ACKNOWLEDGMENTS

This work was partially funded by "NSF Collaborative Research RIPS Type 2: Resilient Simulation for Water, Power, and Road Networks" (lead-PI: P. Suresh C. Rao). The work of Satish Ukkusuri is partly funded by NSF grant 1638311 CRISP Type 2/Collaborative Research: Critical Transitions in the Resilience and Recovery of Interdependent Social and Physical Networks.

APPENDIX

A. Construction of function augmented dual network

The vertex split-recovery process is defined on the function augmented dual networks. We use two procedures to construct the function augmented dual network for each time step. We first construct the base dual network using the hierarchical intersection continuity negotiation (HICN) dual mapping technique [25]. The HICN dual mapping merges consecutive road segments into the same dual vertex if they belong to the same road class and the convex angle they form is close to 180 degree. The details of HICN dual mapping procedure can be found in Algorithm 1. Once the base dual road network is constructed, we overlay the functional states on the network and perform functional dual mapping to track the split-recovery trajectory for each dual vertex in each time step. Details of the functional dual mapping procedure can be found in Algorithm 2.

Algorithm 1: HICN dual mapping procedure

1. Scan the primal network, put all primal edges into the unused edge set E_N. Let the used edge set $E_U = \emptyset$, dual vertex set $V_D = \emptyset$, and dual edge set $E_D = \emptyset$.
2. If $E_N \neq \emptyset$, pick a primal edge e_P from E_N, create a candidate edge $e_{CD} = e_P$; otherwise go to step 4.
3. Grow e_{CD} by recursively executing follows until e_{CD} cannot be extend further:
   (a) Inspect the two end points of e_{CD}. For each of the end point, if it connects to any edge $e \in E_N$, compute the convex angle θ_i for e_{CD} and e_i.
   (b) Merge e_{CD} and e_i if following conditions satisfied:
      i. e_{CD} and e_i are of the same road class.
      ii. $\theta_i = \max\{\theta_1, \theta_2, ...\}$ and $\theta_i < \theta_{max}$, where θ_{max} is the predefined maximum threshold angle ($\pi/3$ used in actual implementation).
   (c) If e_{CD} and e_i can be merged, then $e_{CD} \leftarrow e_{CD} \bigcup \{e_i\}$, $E_N \leftarrow E_N \setminus e_i$, $E_U \leftarrow E_U \bigcup \{e_i\}$
4. Create dual vertex $v_D = e_{CD}$ and let $V_D \leftarrow V_D \bigcup \{v_D\}$. Go back to step 2.
5. Construct the dual-mapped network. For every two dual vertices v_{D_i} and v_{D_j} represented as a set of primal edges, if they contain primal edges that intersecting with each other, construct a dual edge $e_{D_{ij}}$ between v_{D_i} and v_{D_j}.

Algorithm 2: Functional dual mapping procedure

1. Create the base dual road network $G_D^0(V_D^0, E_D^0)$ using Algorithm 1.
2. At time step t. For each dual vertex $v_{D_i}^{t-1} \in V_D^{t-1}$:
   (a) Let $V_D^t = \emptyset, E_D^t = \emptyset$.
   (b) Scan the functional state of each primal edges contained in $v_{D_i}^{t-1}$:
      i. If any primal road segments are at failed state (congested), split the dual vertex $v_{D_i}^{t-1}$ into a set of new dual nodes $V_S = \{V_{D_{i1}}^t, ..., V_{D_{is}}^t\}$ at the location of the failed road segments. Set $V_D^t \leftarrow V_D^t \bigcup V_S$.
      ii. Otherwise, set $v_{D_i}^t = v_{D_i}^{t-1}$. Set $V_D^t \leftarrow V_D^t \bigcup \{v_{D_i}^t\}$.
3. For all dual vertices $\tilde{V}_{D_i} = \{v_{D_{i1}}^t, v_{D_{i2}}^t, .., V_{D_{is}}^t\}$ that originally evolved from $v_{D_i}^0$, check the primal edges in working state (not congested) that contained in $v_{D_i}^0$ but not in dual vertices in $\tilde{V}_{D_i}$. Merge the dual vertices at the location of such primal edges and update the dual vertices in $\tilde{V}_{D_i}$ as well as V_D^t accordingly.
4. Build the dual-mapped network $G_D^t(V_D^t, E_D^t)$. For each pair of dual nodes $v_{D_i}^t$ and $v_{D_j}^t$ in V_D^t, add a dual edge $e_{D_{ij}}$ to E_D^t if v_{D_i}, v_{D_j} share a common primal vertex.

B. Derivation of the recursive formulation for the microscopic-level model

For $S > 1$ and $i = 1, 2, \cdots, k - S - 1$, it can be proved following recursive formulation holds for the expected number of the split sub-vertices with degree i for a dual vertex with degree k after S splits:

$$m_i(k, S+1) = \frac{k-S-i}{k-S-1} m_i(k, S) + \frac{1}{k-S-1} \sum_{s=i+1}^{k-S} (s-1) m_s(k, S) m_i(s, 1) \tag{22}$$

Proof. Consider a dual vertex with degree k after S splits (remove the effect of historical vertex recoveries). If perform one more split (suppose $S < k-1$), the probability of selecting a leaf vertex with degree s is:

$$m_s(k, S) \cdot \frac{s-1}{k-S-1}$$

as the probability of select the leaf vertex with degree s is proportional to $s-1$. Split this vertex will decrease the expected number of leaf vertices with degree s by one, but increase the expected number of leaf vertices

with degree other than s by $M(s,1)$. Hence the expected changes on M after splitting this vertex is

$$\Delta_s^M = m_s(k,S)\cdot\frac{s-1}{k-S-1}[M(s,1)-E_s]$$

where E_s is a vector that only the sth element is 1, others are zeros. The expected number of sub-vertices with different degrees for the dual vertex with degree k after $S+1$ splits can be obtained by summing Δ_s^M for all possible s ($s=1,2,...,k-S$):

$$\begin{aligned} m_i(k,S+1) &= m_i(k,S)+\frac{1}{k-S-1}\sum_{s=1}^{k-S}\Delta_s^M \\ &= \frac{k-S-i}{k-S-1}m_i(k,S)+ \\ &\quad \frac{1}{k-S-1}\sum_{s=i+1}^{k-S}(s-1)m_s(k,S)m_i(s,1) \end{aligned}$$

We hence obtained the recursive formulation for computing $m_i(k,S)$. Note in above derivation, we used the fact that $m_i(s,1)=0$ for $i\geq 0$. □

C. Close form solution of Eq.3 under $\beta = 0$

The close form expression of the recursive formulation Eq.3 is not known. However, the reduced uniform split case $\beta=0$ can be analytically derived as follows:

$$m_i(k,S)=\begin{cases}\frac{S(S+1)}{k-S}\prod_{j=1}^{S-1}(1-\frac{i}{k-j}), & i\leq k-S \\ 0, & i>k-S\end{cases} \quad (23)$$

Proof. If $\beta=0$, then we have uniform split, that is $P(i,k-i|k)=1/(k-1)$, and

$$m_i(k,1)=\begin{cases}\frac{2}{k-1}, & i<k \\ 0, & i\geq k\end{cases}$$

We prove the validity of the close form expression using induction. When $S=1$, the expression obviously holds. Assume the expression holds for $S\leq c$, then

$$\begin{aligned} m_i(k,c+1) &= \frac{k-c-i}{k-c-1}m_i(k,c)+ \\ &\quad \frac{1}{k-c-1}\sum_{s=i+1}^{k-c}(s-1)m_s(k,c)m_i(s,1) \\ &= \frac{k-c-i}{k-c-1}m_i(k,c)+\frac{2}{k-c-1}\sum_{s=i+1}^{k-c}m_s(k,c) \end{aligned}$$

As

$$m_s(k,c)=\frac{c(c+1)}{k-c}\prod_{j=1}^{c-1}(1-\frac{s}{k-j})=\frac{c(c+1)\prod_{r=1}^{c-1}(k-r-s)}{\prod_{j=1}^{c}(k-j)}$$

$$\sum_{s=i+1}^{k-c}m_s(k,c)=\frac{c(c+1)}{\prod_{j=1}^{c}(k-j)}\sum_{s=i+1}^{k-c}\prod_{r=1}^{c-1}(k-r-s)$$

Let

$$F_s=\prod_{r=1}^{c-1}(k-r-s),\quad G_s=s\cdot F_s$$

Thus

$$F_{s+1}=\frac{k-(c-1)-(s+1)}{k-1-s}F_s=\frac{k-c-s}{k-s-1}F_s$$

Then

$$kF_{s+1}-(s+1)F_{s+1}=(k-c)F_s-s\cdot F_s$$

We have

$$kF_{s+1}-G_{s+1}=(k-c)F_s-G_s$$

Summing the cases for $s=i+1$ to $k-c$,

$$k\sum_{s=i+1}^{k-c}F_{s+1}-\sum_{s=i+1}^{k-c}G_{s+1}=(k-c)\sum_{s=i+1}^{k-c}F_s-\sum_{s=i+1}^{k-c}G_s$$

Note that $F_{k-c+1}=0$, $G_{k-c+1}=0$, thus

$$\sum_{s=i+1}^{k-c}F_{s+1}=\sum_{s=i+1}^{k-c}F_s-F_{i+1},\quad \sum_{s=i+1}^{k-c}G_{s+1}=\sum_{s=i+1}^{k-c}G_s-G_{i+1}$$

Thus

$$c\sum_{s=i+1}^{k-c}F_s-kF_{i+1}+G_{i+1}=0$$

And

$$\sum_{s=i+1}^{k-c}F_s=\frac{k-i-1}{c}F_{i+1}$$

Consequently,

$$\begin{aligned} m_i(k,c+1) &= \frac{k-c-i}{k-c-1}m_i(k,c)+\frac{2c(c+1)}{\prod_{j=1}^{c+1}(k-j)}\cdot \\ &\quad \frac{k-i-1}{c}\prod_{r=1}^{c-1}(k-r-i-1) \\ &= \frac{c(c+1)\prod_{r=1}^{c}(k-r-i)}{\prod_{j=1}^{c+1}(k-j)}+\frac{2(c+1)\prod_{r=1}^{c}(k-r-i)}{\prod_{j=1}^{c+1}(k-j)} \\ &= \frac{(c+1)(c+2)}{k-c-1}\prod_{j=1}^{c}(1-\frac{i}{k-j}) \end{aligned}$$

We proved that the proposed expression also holds for $m_i(k,c+1)$, by induction, the claim holds. □

[1] R. Albert, I. Albert, and G. L. Nakarado, Phys. Rev. E **69**, 1 (2004).
[2] D. J. Ashton, T. C. Jarrett, and N. F. Johnson, Phys. Rev. Lett. **94**, 1 (2005).
[3] B. Danila, Y. Sun, and K. E. Bassler, Phys. Rev. E **80**, 1 (2009).
[4] M. Barthélemy, Phys. Rep. **499**, 1 (2011), 1010.0302.
[5] C. D. Brummitt, R. M. D'Souza, and E. a. Leicht, Proc. Natl. Acad. Sci. U.S.A. **109**, E680 (2012).
[6] D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H. E. Stanley, and S. Havlin, Proc. Natl. Acad. Sci. U.S.A. **112**, 669 (2015).
[7] J. Zhao, D. Li, H. Sanhedrai, R. Cohen, and S. Havlin, Nat. Commun. **7** (2016).
[8] T. J. Brodribb, D. Bienaimé, and P. Marmottant, Proc. Natl. Acad. Sci. U.S.A. **113**, 4865 (2016).
[9] M. Garavello and B. Piccoli, *Traffic flow on networks*, Vol. 1 (American institute of mathematical sciences Springfield, 2006).
[10] M. Patriksson, *The traffic assignment problem: models and methods* (Courier Dover Publications, 2015).
[11] S. Scellato, L. Fortuna, M. Frasca, J. Gómez-Gardeñes, and V. Latora, Eur. Phys. J. B **73**, 303 (2010).
[12] A. Solé-Ribalta, S. Gómez, and A. Arenas, R. Soc. open sci. **3**, 160098 (2016).
[13] D. S. Johnson, J. K. Lenstra, and A. Kan, Networks **8**, 279 (1978).
[14] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002).
[15] A. Tomaszewski, M. Pióro, and M. Żotkiewicz, Networks **55**, 108 (2010).
[16] R. Guimera, A. Arenas, A. Díaz-Guilera, and F. Giralt, Phys. Rev. E **66**, 026704 (2002).
[17] J. Duch and A. Arenas, Phys. Rev. Lett. **96**, 218702 (2006).
[18] D. De Martino, L. DallAsta, G. Bianconi, and M. Marsili, Phys. Rev. E **79**, 015101 (2009).
[19] L. a. Briggs and M. Krishnamoorthy, Proc. Natl. Acad. Sci. U.S.A. **110**, 19295 (2013).
[20] A. Rinaldo, R. Rigon, J. R. Banavar, A. Maritan, and I. Rodriguez-Iturbe, Proc. Natl. Acad. Sci. U.S.A. **111**, 2417 (2014).
[21] A. Solé-Ribalta, S. Gómez, and A. Arenas, Phys. Rev. Lett. **116**, 108701 (2016).
[22] T. C. Jarrett, D. J. Ashton, M. Fricker, and N. F. Johnson, Phys. Rev. E **74**, 1 (2006).
[23] S. Porta, P. Crucitti, and V. Latora, Physica A **369**, 853 (2006).
[24] B. Jiang, Physica A **384**, 647 (2007).
[25] A. P. Masucci, K. Stanilov, and M. Batty, Phys. Rev. E **89**, 1 (2014).
[26] E. Krueger, C. Klinkhamer, C. Urich, X. Zhan, and P. S. C. Rao, Phys. Rev. E **95**, 032312 (2017).
[27] M. J. Lighthill and G. B. Whitham, in *Proc. R. Soc. Lond. A*, Vol. 229 (1955) pp. 317–345.
[28] C. F. Daganzo, Transport. Sci. **32**, 3 (1998).
[29] B. S. Kerner, Phys. Rev. E **84**, 045102 (2011).
[30] D. R. Cox and D. Oakes, *Analysis of survival data*, Vol. 21 (CRC Press, 1984).
[31] J. F. Lawless, *Statistical models and methods for lifetime data*, Vol. 362 (John Wiley & Sons, 2011).
[32] R. Sainudiin and A. Véber, Roy. Soc. Open Sci. **In press.**, Eprint arxiv 1511.08828 (2016).
[33] D. Aldous, in *Random discrete structures* (Springer, 1996) pp. 1–18.
[34] B. Chen, D. Ford, and M. Winkel, Electronic Journal of Probability **14**, 400 (2009).
[35] M. G. B. Blum and O. François, Syst. Biol. **55**, 685 (2006).
[36] R. M. Anderson, R. M. May, and B. Anderson, *Infectious diseases of humans: dynamics and control*, Vol. 28 (Wiley Online Library, 1992).
[37] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[38] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks* (Cambridge University Press, 2008).
[39] S. N. Dorogovtsev, a. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
[40] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015).
[41] M. Newman, *Networks: an introduction* (Oxford university press, 2010).
[42] J. Lin, IEEE Trans. Inf. Theory **37**, 145 (1991).
[43] See Supplemental Material at [SM URL] for additional information and results.
[44] B. S. Kerner, J. Phys. A **33**, L221 (2000).
[45] B. S. Kerner, S. L. Klenov, A. Hiller, and H. Rehborn, Phys. Rev. E **73**, 046107 (2006).