

# Dynamic Robustness Analysis of a Two-Layer Rail Transit Network Model

Chao Gao<sup>ID</sup>, Member, IEEE, Yi Fan, Shihong Jiang, Yue Deng<sup>ID</sup>, Jiming Liu, Fellow, IEEE, and Xianghua Li<sup>ID</sup>

**Abstract**—Robustness is one of the most important performance criteria for any rail transit network (RTN), because it helps us enhance the efficiency of RTN. Several studies have addressed the issue of RTN robustness primarily from the perspectives of given rail network structures or static distributions of passenger flow. An open problem that remains in fully understanding RTN robustness is how to take the spatio-temporal characteristics of passenger travel into consideration, since the dynamic passenger flow in an RTN can readily trigger unexpected cascading failures. This paper addresses this problem as follows: (1) we propose a two-layer rail transit network (TL-RTN) model that captures the interactions between a rail network and its corresponding dynamic passenger flow network, and then (2) we conduct the cascading failure analysis of the TL-RTN model based on an extended coupled map lattice (CML). Specifically, our proposed model takes the strategy of passenger flow redistribution and the passenger flow capacity of each station into account to simulate the human mobility behaviors and to estimate the maximum passenger flow appeal in each station, respectively. Based on the smart card data of RTN passengers in Shanghai, our experiments show that the TL-RTN robustness is related to both external perturbations and failure modes. Moreover, during the peak hours on weekdays, due to the large passenger flow, a small perturbation will trigger a 20% cascading failure of a network. Having ranked the cascade size caused by the stations, we find that this phenomenon is determined by both the hub nodes and their neighbors.

**Index Terms**—Two-layer rail transit network model, robustness, cascading failure, passenger flow redistribution, dynamic.

Manuscript received 29 January 2020; revised 18 October 2020 and 5 January 2021; accepted 3 February 2021. Date of publication 18 February 2021; date of current version 8 July 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61976181 and Grant 11931015, in part by the Hong Kong Research Grants Council under Grant HKBU12201619, and in part by the Natural Science Foundation of Chongqing under Grant cstc2018jcyjAX0274 and Grant cstc2019jcyj-zdxdmX0025. The Associate Editor for this article was A. Nunez. (*Corresponding authors: Jiming Liu; Xianghua Li.*)

Chao Gao is with the School of Artificial Intelligence, Optics, and Electronics (iOPEN), Northwestern Polytechnical University, Xian 710072, China, and also with the College of Computer and Information Science, Southwest University, Chongqing 400715, China.

Yi Fan, Shihong Jiang, and Yue Deng are with the College of Computer and Information Science, Southwest University, Chongqing 400715, China (e-mail: cgao@swu.edu.cn).

Jiming Liu is with the Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong (e-mail: jiming@comp.hkbu.edu.hk).

Xianghua Li is with the School of Artificial Intelligence, Optics, and Electronics (iOPEN), Northwestern Polytechnical University, Xian 710072, China (e-mail: li\_xianghua@nwpu.edu.cn).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TITS.2021.3058185>, provided by the authors.

Digital Object Identifier 10.1109/TITS.2021.3058185

## I. INTRODUCTION

WITH the acceleration of urbanization, urban rail transit has become the preferred method of transportation due to its advantages of safety, rapidity and environmental protection [1], [2]. However, due to sudden incidents (such as signal failures, extreme weather/hazards, and entertainment activities) [3], [4], some stations in a rail transit network (RTN) may experience different extents of localized failures and disruptions (i.e., the station is out of service) [5], [6]. It is true that the station disruptions caused by such kind of accidents is not that serious like train derailment and collision. However, this type of perturbation will in turn spread across the RTN due to the coupling relationship between stations and the dynamic distribution of passenger flow. Eventually, the entire interconnected system of an RTN may be completely paralyzed due to such cascading effects. Therefore, studies on RTN robustness has attracted widespread concern in the domains of transportation science [7] and computer science [8]. These studies can help transportation managers formulate emergency management schemes reasonably in the case of rail transit failures and provide reliable services for passengers.

To date, several studies have addressed the issue of RTN robustness primarily from a given rail structure [9]–[11], [13], [14] or static distribution of passenger flow [5], [15]–[20]. However, the rail structure and the dynamic distribution of passenger flow are not independent of each other [21]. For example, if any station or line breaks down in the RTN, the passengers who are currently at a disrupted station or planning to go to the disrupted station will spontaneously re-plan their routes to minimize the impact of this disruption, which further leads to a dynamic redistribution of passenger flow in the RTN [22], [23]. As shown in Fig. 1, if the passenger flow of a station (i.e.,  $V_C$ ), including the passengers transferred from other stations (i.e.,  $V_A$ ), exceeds its capacity threshold, congestion or even failure will be triggered in this station [24], [25]. Therefore, the interactions between the rail structure and the dynamic distribution of passenger flow may cause unpredicted cascading failures in the RTN [26].

## A. Motivations

Due to the complexity and uncertainty of dynamic passenger flow in a rail network, it is necessary for us to design an analytical model to characterize this dynamic relationship

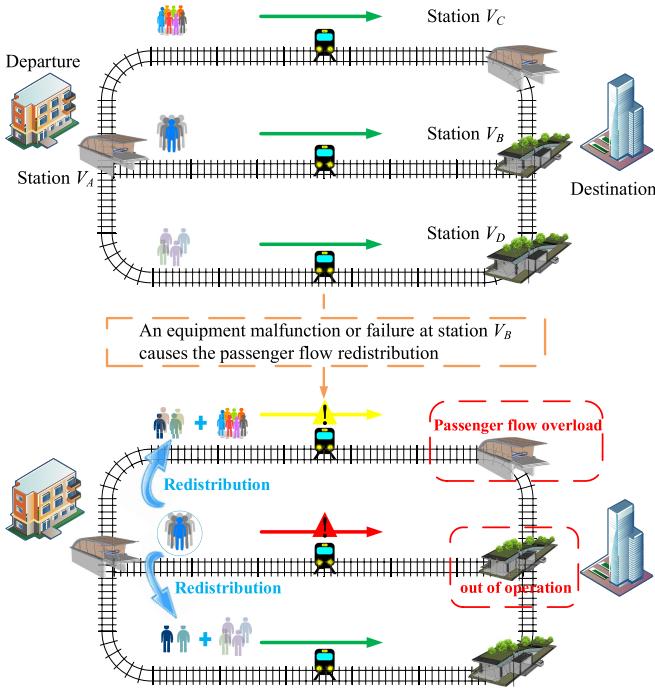


Fig. 1. An example of passenger flow redistribution caused by an equipment malfunction or failure of a station. Due to the sudden incidents (such as signal failures, extreme weather/hazards, and entertainment activities), station  $V_B$  is out of service. The passengers who plan to go to  $V_B$  have to reroute dynamically to other stations that are close to the destination (i.e.,  $V_C$  or  $V_D$ ). Such a rerouting process could cause  $V_C$  to be out of service because of the overload of passenger flow. The red arrow denotes that  $V_B$  is out of service. Passengers at  $V_A$  have to re-plan their travels. The green arrow denotes that such a line is in normal operation even when it encounters a passenger flow redistribution. The yellow arrow denotes that when passengers arrive at the next station, the passenger flow of this station will be overloaded.

between the rail network and passenger flow to reveal and understand the cascading failure process of RTNs. Previous studies have proposed cascading failure models (such as the load capacity model [27], binary impact model [28], sand model [29] and coupled map lattice (CML) model [30]). Among them, the CML model is widely used to analyze the cascading failure process of traffic networks due to its high efficiency and easy implementation [22], [31]. Although this model adopts discrete spatiotemporal representations, the state values remain continuous. For example, a node with a state value denotes a station in CML. At each iteration step, the state value of each station will be updated based on the external perturbations and the state values of its neighbors. If such a state is greater than or equal to a threshold, this station will fail that may be most likely to cause other nodes to fail. When there are no more failed nodes in the network, the cascading failure process of the rail transit network ends. However, some constraints limit their applications in investigating the cascading failure process of RTNs.

First, the spatiotemporal characteristic of passenger flow is one of the main features of an RTN [32]–[34]. In existing CML-based studies, this characteristic is formulated as a static weight of a rail network [22], [23] or is not addressed. For example, researchers only focus on the cascading failure process of small-world or scale-free networks [35], [36].

Second, the interaction between the rail network and dynamic passenger flow could cause an unexpected cascading failure in the RTN [26]. For example, the passenger flow redistribution in a disrupted station will trigger a capacity overload of other stations [37], [38]. Therefore, it is of urgent importance for us to reveal the effects of spatiotemporal passenger flow on the robustness and vulnerability of RTNs [5], [39]. To address the above issues, the following questions are addressed.

- How can the cascading process of an RTN triggered by both the static rail network and the dynamic passenger flow be characterized?
- How can the effect of passenger flow redistribution on the RTN robustness be characterized?

### B. Main Contributions

Based on the above, this paper reveals the dynamic robustness of an RTN by extending the CML model to characterize the cascading failure process caused by the interaction between the static rail network and dynamic passenger flow.

First, we propose a two-layer rail transit network (TL-RTN) model for characterizing the interaction between a static rail network and its corresponding dynamic passenger flow. The lower layer is a rail network, which is constructed based on the structural information of the urban rail transit system. The upper layer is a passenger flow network, which is constructed based on the check-in and check-out data of passengers. We can extract such data from the metro automatic fare collection system. Then, by characterizing the passenger flow travels along the rail network, we extend the CML model to simulate two cascading processes in both the rail network and the passenger flow network simultaneously. Moreover, we incorporate the passenger flow redistribution in the disrupted station and the passenger flow capacity of each station into the extended CML model for describing the spontaneous rerouting behaviors of passengers and heterogeneous capacities (i.e., the maximum passenger flow) of different stations in the RTN, respectively [40]. Finally, we initialize the extended CML model based on the Shanghai rail transit line and passenger flow data, and conduct the cascading failure analysis of the TL-RTN. The main contributions of this paper can be summarized as follows:

- 1) We propose a TL-RTN model for characterizing the interaction between a static rail network and its corresponding dynamic passenger flow network.
- 2) We extend the CML model to characterize the coupling cascading process of TL-RTN given the strategy of passenger flow redistribution and the passenger flow capacity of each station.
- 3) We systematically conduct extensive experiments to reveal the dynamic robustness of Shanghai TL-RTN.

The remainder of this paper is organized as follows. Sec. II reviews the related work. Sec. III first illustrates the process of constructing a TL-RTN and then proposes an extended CML model for characterizing the cascading failure process of a TL-RTN. Sec. IV implements some analyses to reveal the dynamic robustness of the Shanghai TL-RTN. Sec. V further discusses the Shanghai TL-RTN robustness in terms of

small external perturbations and coupling strength coefficients. Finally, Sec. VI concludes this paper.

## II. RELATED WORK

Previous studies have proposed some analytical models to analyze RTN robustness and measure the efficiency of an RTN. This section first surveys existing studies on RTN robustness in Sec. II-A. Then, Sec. II-B focuses on the coupled map lattice (CML) model which is a basis for characterizing the interaction between a static rail network and its corresponding dynamic passenger flow network in this paper.

### A. The RTN Robustness

A rail transit network (RTN) can be formulated as a network. Existing studies have proposed some metrics from the perspective of complex networks to estimate RTN robustness. For example, the largest component [41]–[43], the average shortest path length [44], natural connectivity [45] and global efficiency [46] are used to measure the impact of rail structure [9]–[13] or the distribution of passenger flow [5], [15]–[20] on RTN robustness.

Considering the effect of network structural dynamics on the RTN robustness, Fisk *et al.* analyzed subway networks under different attack strategies [10], through which they found that subway networks with consequential high connectivity and low degree are robust to targeted attacks, and most subway networks are robust to random attack. Derrible *et al.* found the scale-free and small-world characteristics of subway networks by investigating the structure of 33 cities' subway systems [9]. Moreover, they proposed a strategy to enhance RTN robustness by increasing the number of transfer stations. After comparing the random and targeted attacks to the transport network in London and Paris, Ferber *et al.* proved that the network will be paralyzed if 0.5% of the transport stations suffer sustained damage [11]. Yang *et al.* assessed the RTN robustness in the face of different attack modes [12], through which they found the Beijing subway network has high survivability and robustness under random attacks and is relatively low when the hubs are subject to targeted attacks. Zhang *et al.* found that the Shanghai subway network is robust to random attacks but vulnerable to malicious attacks by assessing the changes in network efficiency and functionality loss [13].

In terms of the effect of passenger flow on RTN robustness, researchers have proposed some simulated-based methods and metrics to characterize the dynamic change of passenger flow, which help us measure and understand the interplay between the passenger flow and RTN. For example, Jenelius *et al.* proposed an individual-based model to explore the value of new public transport links for improving the robustness of a traffic network [15]. In their model, a passenger is denoted as an agent that can characterize the dynamic travel choices of such a passenger. To further improve the robustness of transit systems, Cats *et al.* implemented an agent-based transport model to simulate the evolution of traffic network reliability and on-board crowding [16]. Considering the situation that passengers will continue to take the subway unless there exists no accessible route when a failure happens, Hu *et al.* used

some existing metrics to measure the structure and robustness of the Beijing metro network [17]. Then, a new hybrid weighted metric was proposed to estimate the importance of a station by considering the passenger flow. Xiao *et al.* proposed new dynamic metrics that represent the local and global features of node's degree and betweenness, through which they estimated the vulnerability of the Beijing metro network [18]. The results show that the heterogeneity and vulnerability of such networks vary over time when we take passenger flow into consideration.

Furthermore, researchers have explored the effect of dynamic passenger flow on RTN vulnerability or resilience. For example, Cai *et al.* analyzed the effects of passenger flow and capacity constraints on the topological vulnerability of the Beijing metro network [19]. They apply the logit-based path choice model and the travel time model to characterize the passenger flow and the congestion level, respectively. The results show that the metro network will be vulnerable if the passenger flow volume becomes larger and the travel choices of passengers are fewer. Lu *et al.* presented an accessibility-based measurement for analyzing RTN vulnerability and illustrating the characteristics of passenger flow in the Shenzhen rail network [5]. The results show that the consequences of disruptions in terms of network accessibility clearly vary with the characteristics of passenger flow at different stations. The undisrupted station is vulnerable to failures from surrounding stations. Chopra *et al.* presented a multi-pronged framework to reveal the resilience of the London metro system by analyzing the topological and passenger flow features [20]. They found that the London metro system is robust to random failures and that a few critical stations are vulnerable to disruptions. Meanwhile, they proposed a metric called the fracture coefficient that determines the functional vulnerability of an edge in terms of the extent of fracture caused by its removal. The higher the fracture coefficient is, the lower the resilience of the London metro system.

Overall, the above studies only consider the influence of the inherent rail structure or the static distribution of passenger flow on the network robustness. The dynamic interaction between the rail network and passenger flow network is not addressed in these studies. To investigate such dynamic interactions, a two-layer network model is defined in this paper. In this type of network, the lower layer is a given transport network, which is regarded as a physical layer. The upper layer is a traffic flow network, which is regarded as a logical layer. The layered network model can further reveal the underlying law of traffic network robustness [24] by characterizing the interaction between the static rail network and dynamic passenger flow network.

### B. CML-Based Analysis

Due to its advantages of low computational cost and high numerical simulation efficiency, researchers widely apply the CML model to study the spatiotemporal behavior of complex systems [30] and cascading failures in complex networks [22], [23], [35], [36], [47]. For example, Wang *et al.* found that the cascading failure processes occur much more easily in

the small-world and scale-free networks than in globally coupled networks [35]. Moreover, we can observe the synergetic principle of the CML-based cascading failure process in scale-free networks [36]. By investigating the effectiveness of edge tolerance in scale-free and random networks based on the CML model, Cui *et al.* found that a larger tolerance parameter can more efficiently delay the cascading failure process in scale-free networks than in random networks [47].

To further explore the dynamic characteristics of passenger flow, Huang *et al.* proposed a CML-based model to analyze the bus passenger flow network [22]. They take the topology of the bus transport network and the dynamic redistribution of passenger flow in a station into account. This redistribution is based on the local edge weight, where a passenger is more likely to reroute to stations with more passengers. In fact, the path choices of passengers most likely depend on the destination of a trip and the real-time traffic information by related APP or the subway websites, leaving out the concerns of the comfort degree and weather conditions [48]. Therefore, such redistribution may make the analysis results biased. In view of this, Shen *et al.* proposed a new strategy of  $\eta$ -based passenger flow redistribution in their bidirectional CML model [23]. This strategy is based on the edge betweenness, where a passenger has a higher probability of rerouting to the station linking the shortest paths. Based on this, they analyzed the cascading failure process of the Nanjing RTN during the morning peak hour (i.e., 7:30-7:45 am on weekday), through which they found that such a network is more vulnerable to targeted failure than to random failure. However, we know that there are different periodicities of passenger flow during a day (such as during morning and evening peak hours [49]). If a station is disrupted during off-peak hours, such a disruption will affect fewer people than one that occurs during peak hours. Therefore, even if the strategy of passenger flow redistribution remains unchanged during different time periods, its effect on RTN robustness can be quite different. It is important and valuable for us to investigate the effects of passenger flow with periodical features on the RTN robustness.

In this paper, we analyze the effects of both a static rail network and dynamic passenger flow on RTN robustness. First, a two-layer rail transit network (TL-RTN) model is constructed to characterize the dynamic passenger flow traveling along a static rail network. Then, an extended CML model is used to characterize the coupling cascading failure processes in the rail network and the passenger flow network. In particular, some unpredicted failures, caused by the interaction between the passenger flow redistribution and the allowed maximum passenger flow of each station, can be observed, which play an important role in the dynamic robustness of TL-RTN.

### III. PRELIMINARIES

The formulation of a two-layer rail transit network (TL-RTN) model and the improved coupled map lattice (CML) model are introduced in Sec. III-A and Sec. III-B, respectively.

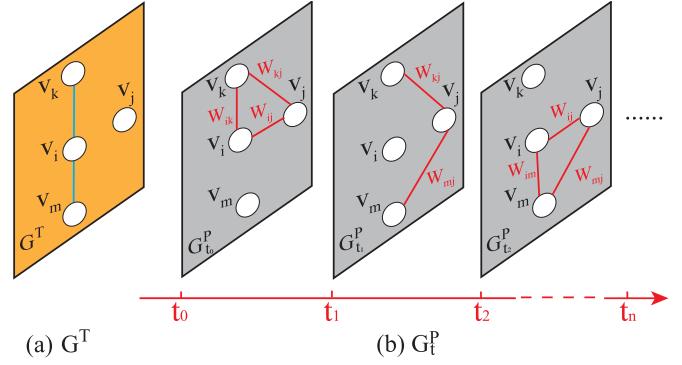


Fig. 2. An example of TL-RTNs at different times. (a) The lower layer  $G^T$  is a rail network. (b) The upper layer  $G_t^P$  is a passenger flow network at time  $t$ .  $w_{ij}(t)$  denotes the passenger flow between the two stations  $v_i$  and  $v_j$ . The blue and red solid lines denote edges in the rail network and the passenger flow network, respectively.

#### A. The Layered Model of RTN

To incorporate the dynamic passenger flow into the analysis of RTN robustness, a TL-RTN model is constructed in this paper, and the following definitions are presented.

**Definition 1:** A TL-RTN model (i.e.,  $G_t$ ) consists of two layers. The lower layer (i.e.,  $G^T$  in Fig. 2(a)) denotes a rail network based on the physical rail lines, and the upper layer (i.e.,  $G_t^P$  in Fig. 2(b)) is a passenger flow network at time  $t$  based on the check-in and check-out data of passengers taking the subway.

In  $G_t$ , a set of nodes in  $G^T$  and  $G_t^P$  are identical. Once a node fails in  $G^T$ , the corresponding node will simultaneously fail in  $G_t^P$ . Meanwhile, the failed nodes in  $G_t^P$  are removed and the weights of failed edges are reassigned based on the redistribution strategy of passenger flow. In this way, a situation is implemented where trains will not stop at closed stations influenced by some emergencies like weather conditions or congestion (i.e., the jump of passenger flows [19], [20]). Such congestion may be caused by an unbalanced redistribution of passenger flows, or entertainment activities near a station.

A rail network  $G^T = \langle V, E^T \rangle$  is constructed based on the L-space method [50]. A station is formulated as a node, and an edge denotes the track between two adjacent stations. In general, subway trains can travel and return along the same route from one metro station to another [51]. Thus, the rail network can be denoted as an undirected and unweighted graph, which is shown in Def. 2.

**Definition 2:** In  $G^T$ ,  $V = \{v_i | i \in [1, N]\}$  and  $E^T = \{e_{ij}^T = (v_i, v_j) | i, j \in [1, N], i \neq j\}$  represent a set of nodes and a set of edges, respectively.  $N$  denotes the total number of stations.

A passenger flow network  $G_t^P = \langle V, E^P, W \rangle$  is constructed based on the origin-destination (*OD*) data of passengers taking the subway. Generally, passengers are inclined to continue taking subway after entering the subway system unless there is no direct route passing their destinations [17], [18]. The *OD* data consist of *OD* pairs. Each pair  $(v_i, v_j, w_{ij}(t))$  denotes the passenger flow volume  $w_{ij}(t)$  between stations  $v_i$  and  $v_j$  at time  $t$ . The  $(v_i, v_j)$  is defined as an alternating sequence of nodes and links between  $v_i$

and  $v_j$  [7]. Therefore, the passenger flow network can be formulated as an undirected and weighted graph, which is shown in Def. 3.

**Definition 3:** In  $G_t^P$ ,  $V = \{v_i | i \in [1, N]\}$  is same as that in  $G^T$ , denoting a set of stations.  $E^P = \{e_{ij}^P = (v_i, v_j) | i, j \in [1, N], i \neq j\}$  denotes a set of passenger routes. Specifically,  $e_{ij}^P$  represents that a route exists between two stations  $v_i$  and  $v_j$ .  $W = \{w_{ij}(t) | i, j \in [1, N], i \neq j\}$  represents the weight between two stations, which denotes the total number of passengers who choose this path at time  $t$ .

Moreover, some metrics are applied for measuring the centrality of a node in  $G_t$  as follows.

**Definition 4:** The **degree** of  $v_i$  ( $D_i$ ) denotes the number of nodes linked with it in a network, shown in Eq. (1).

$$D_i = \sum_{j=1, j \neq i}^N e_{ij} \quad (1)$$

where  $N$  denotes the total number of stations. If an undirected link exists between  $v_i$  and  $v_j$ ,  $e_{ij} = 1$ ; otherwise,  $e_{ij} = 0$ .

**Definition 5:** The **betweenness** of  $v_i$  ( $B_i$ ) is defined in Eq. (2).

$$B_i = \sum_{s, t \in [1, N]} \frac{\sigma_{st}(v_i)}{\sigma_{st}} \quad (2)$$

where  $\sigma_{st}(v_i)$  denotes the number of the shortest paths between  $v_s$  and  $v_t$  past  $v_i$ .  $\sigma_{st}$  denotes the number of the shortest paths between  $v_s$  and  $v_t$ . The larger  $B_i$  is, the more important the connectivity role  $v_i$  plays in a network.

**Definition 6:** The **strength** of  $v_i$  ( $S_i(t)$ ) is denoted as Eq. (3).

$$S_i(t) = \sum_{j=1, j \neq i}^N w_{ij}(t) \quad (3)$$

where  $w_{ij}(t)$  denotes the weight of an edge  $e_{ij}$  at time  $t$ . The larger  $S_i(t)$  is, the greater the passenger flows between  $v_i$  and other nodes are.

### B. Cascading Failure Based on the CML Model

In this section, we first propose an extended CML model that is suitable for describing the cascading failure process of a TL-RTN in Sec. III-B.1. Then, Sec. III-B.2 introduces the cascading failure process of a TL-RTN.

*1) The Extended CML Model for TL-RTN:* In a TL-RTN ( $G_t$ ), a cascading failure may be caused by the interaction between a rail network ( $G^T$ ) and a passenger flow network ( $G_t^P$ ). For example, as shown in Fig. 2, if  $v_i$  fails in  $G_t$ , the passenger flow of this node will be redistributed to other nodes at  $G^T$ , which will further cause changes in  $G_t^P$ . At the same time, if the passenger flow in a station  $v_k$  (including the passengers from the disrupted station  $v_i$ ) exceeds the capacity threshold of  $v_k$ , it will fail in the next time step.

Therefore, we have to incorporate the passenger flow redistribution and the passenger flow capacity of each station ( $Q_i$ ) into our model to simulate dynamic behaviors of passengers and estimate the maximum passenger flow appeal in each

station, respectively.  $Q_i$  is calculated based on the nonlinear load-capacity model proposed [52], as shown in Eq. (4).

$$Q_i = C_i + \beta (C_i)^\alpha \quad (4)$$

where  $\alpha$  and  $\beta$  are two capacity control parameters ranging from 0 to 1. When one parameter equals 1, the model can degenerate into the linear load-capacity one.  $C_i$  is calculated by Eq. (5).

$$C_i = \max_{T \in \{t_1, t_2, t_3, \dots\}} \{S_i(T)\} \quad (5)$$

where  $\max_{T \in \{t_1, t_2, t_3, \dots\}} \{S_i(T)\}$  denotes the maximum value of the strength at  $T$ . In particular, the networks at the time set  $\{t_1, t_2, t_3, \dots\}$  reflect the different passenger flow volumes.

Moreover, some variables are defined to record the dynamic changes in  $G^T$  and  $G_t^P$  when a cascading failure occurs in  $G_t$ .  $A = (a_{ij})_{N \times N}$  denotes the connection information of nodes in  $G^T$ , which follows the iterative evolution of Eq. (7). If a node  $v_i$  is connected to  $v_j$  with an edge, then  $a_{ij} = 1$ ; otherwise,  $a_{ij} = 0$ .  $w_{ij}(t)$  is the weight of an edge  $e_{ij}^P$  in  $G_t^P$ , which follows the iterative evolution of Eq. (7).  $T_i(s)$  and  $P_i(s)$  denote the state of  $v_i$  at time step  $s$  in a rail network ( $G^T$ ) and in a passenger flow network ( $G_t^P$ ), respectively. The process for updating the states in  $v_i$  is formulated in Eqs. (7) and (7).

$$T_i(s+1) = |(1 - \varepsilon_1)f(T_i(s)) + \varepsilon_1 \sum_{j=1, j \neq i}^N \frac{a_{ij} f(T_j(s))}{K_i}| \quad (6)$$

$$\begin{aligned} P_i(s+1) = & |(1 - \varepsilon_2)f(P_i(s)) \\ & + \varepsilon_2 \sum_{j=1, j \neq i}^N \frac{w_{ij}(s) f(P_j(s))}{Q_i}| \\ & i \in [1, N] \end{aligned} \quad (7)$$

where  $K_i$  is the degree of  $v_i$ .  $\varepsilon_1$  and  $\varepsilon_2$  represent the coupling strength coefficients of  $G^T$  and  $G_t^P$ ,  $0 < \varepsilon_1, \varepsilon_2 < 1$ . In real life, the coupling strength coefficients indicate the impact of the states of a station (i.e., normal or disrupted) on its neighbors. For example, if  $\varepsilon_1 = 0.4$ , 40% of the states of a station come from the impact of their neighbors' states. The nonlinear function  $f$  represents the dynamic behavior of the node itself. Here  $f(x) = 4x(1-x)$  is a chaos mapping, where  $0 \leq x \leq 1$  and  $0 \leq f(x) \leq 1$  based on [35]. The absolute value in the above equation guarantees the non-negative state of each node.

If  $v_i$  is in a normal state at time  $s_0$  (i.e.,  $0 < T_i(s_0) < 1, 0 < P_i(s_0) < 1$ ),  $v_i$  is in a normal state at time step  $s$  ( $s \leq s_0$ ). Conversely, if  $T_i(s_0) \geq 1$  or  $P_i(s_0) \geq 1$ ,  $v_i$  will fail at time  $s_0$ , which means that there will be an external perturbation in  $v_i$ . For example, the signal failures and extreme weather/hazards will trigger a failure at the rail network (i.e.,  $G^T$ ), and the jump of passenger flows that may be caused by entertainment activities will trigger a failure at the passenger network (i.e.,  $G_t^P$ ). At the early stage of failure, the disruption stations are in the minority and only one station disrupts most of the time. Thus, there are many different choices for passengers to reroute. However, concerning the convenience and reliability, the neighbors of disruption stations should be the preferred plan. Therefore, this paper designs the strategy of

passenger flow redistribution based on  $\eta$ -based passenger flow redistribution [23]. Moreover, a cascading failure in  $v_i$  will be caused by the passenger redistribution of failed neighbors of  $v_i$  in which the traffic capacity of  $v_i$  exceeds its threshold. In this case, the two states of  $v_i$  are always equal to zero at any later time, i.e.,  $T_i(s) \equiv P_i(s) \equiv 0, s > s_0$ . If the initial states of all nodes are within the range and there is no external perturbation, the entire system will always stay in a normal state and the subway network can operate normally.

2) *The Cascading Failure Process of TL-RTN:* Given an extended CML model, a cascading process is triggered by a failed node  $v_i$ , which may be generated from an external perturbation at time step  $s$  in a rail network. Eq. (8) adds such external perturbation  $R$  into Eq. (7). Generally, the external perturbation follows two types of failure modes i.e., random and targeted failure modes, for a node [53], [54]. Based on the extended CML model, we can characterize the cascading failure process and further estimate the effects of passenger flow redistribution and failure modes on the RTN robustness. The whole process is illustrated in Alg. 1.

$$T_i(s) = |(1 - \varepsilon_1)f(T_i(s-1)) + \varepsilon_1 \sum_{j=1, j \neq i}^N \frac{a_{ij}f(T_j(s-1))}{K_i}| + R \quad (8)$$

As illustrated in Alg. 1, an external perturbation first occurs in station  $v_c$  at time step  $s_0$ . This station is out of service, and the subway line cannot carry passengers at this station. The state of  $v_c$  is updated by Eq. (8). Meanwhile, the corresponding  $v_c$  in  $G_t^P$  will fail at the same time, which can further cause a weight redistribution in this network. Therefore, for all  $s > s_0$ ,  $T_c(s) \equiv 0$  and  $P_c(s) \equiv 0$ . All neighboring nodes in both the rail network and passenger flow network directly connecting to  $v_c$  will be affected by the states of  $v_c$  at time step  $s_0+1$ . The states of other nodes in  $G^T$  and  $G_t^P$  are updated by Eq. (7) and Eq. (7), respectively. If there is a node  $v_i$  whose state value surpasses 1 in  $G^T$  or  $G_t^P$ , then a new round of failures and capacity overload can be triggered simultaneously. If there is a node  $v_i$  whose state value is less than 1 both in  $G^T$  and  $G_t^P$ , but the strength of  $v_i$  ( $S_i(s)$ ) is larger than the capacity of  $v_i$  ( $Q_i$ ), then a new round of passenger flow redistribution will be triggered at this time. And  $v_i$  should be added to the overloaded node set  $O(s)$ . Here, the total number of failed nodes in a network at time step  $s$  is denoted by Eq. (9).

$$I(s) = I_T(s) + I_P(s) - I_{T \cap P}(s) \quad (9)$$

where  $I(s)$  is the total number of failed nodes at time step  $s$ .  $I_T(s)$  and  $I_P(s)$  are the numbers of failed nodes at time step  $s$  in  $G^T$  and  $G_t^P$ , respectively.  $I_{T \cap P}(s)$  is the number of common failed nodes at time step  $s$  in both  $G^T$  and  $G_t^P$ . When there are no more failed nodes in  $G^T$  and  $G_t^P$ , the cascade size of  $G_t$  is represented by  $I$ , that is,  $I \equiv \lim_{s \rightarrow \infty} \frac{I(s)}{N}$ .

Specifically, taking Fig. 3 as an example, if node  $v_i$  is overloaded or failed, the steps of passenger flow redistribution are as follows.

- (1) Let  $N_T(v_i)$  be a normal neighbor set of  $v_i$  in  $G^T$ .  $N_T(v_i) = \{v_k, v_m\}$ .

---

**Algorithm 1:** Cascading Failure Based on the Extended CML Model

---

```

Input: The TL-RTN and failure modes for  $v_c$  at time step  $s_0$ 
Output: The cascade size  $I$  of TL-RTN
1 Initializing:  $\varepsilon_1, \varepsilon_2, Q_i, R$ , the maximal iteration step  $T$  and the initial state of each node in  $G^T$  and  $G_t^P$ 
2 for  $s$  from 1 to  $T$  do
3   for  $i$  from 1 to  $N$  do
4     /* calculating the state value of each node */
5     if  $s == s_0$  and  $i == c$  then
6       | Calculating  $T_i(s)$  based on Eq. (8);
7     end
8     else
9       | Calculating  $T_i(s)$  based on Eq. (7);
10      | Calculating  $P_i(s)$  based on Eq. (7);
11      | Calculating  $S_i(s)$  based on Eq. (3);
12      if  $S_i(s) > Q_i$  and  $T_i(s) < 1$  and  $P_i(s) < 1$  then
13        | Adding  $v_i$  to the overloaded nodes set  $O(s)$ ;
14      end
15    end
16  /* passenger flow redistribution */
17  Updating  $I_T(s), I_P(s), I_{T \cap P}(s)$ ;
18  Calculating  $I(s)$  based on Eq. (9);
19  for  $v_i$  in  $O(s)$  or  $I(s)$  do
20    if  $v_i$  is overloaded then
21      | Getting  $N_T(v_i), N_P(v_i)$ ;
22      | Updating  $E_P$ ;
23      | for  $e$  in  $E_P$  do
24        |   Establishing  $e$  in  $G_t^P$ ;
25        |   Reassigning weight to  $e$  based on Eq. (10),
26          |   Eq. (11) and Eq. (12);
27      end
28    end
29    else if  $v_i$  is failed then
30      | Getting  $N_T(v_i), N_P(v_i)$ ;
31      | Removing  $v_i$  from  $G_t^P$ ;
32      | Updating  $E_P$ ;
33      | for  $e$  in  $E_P$  do
34        |   Establishing  $e$  in  $G_t^P$ ;
35        |   Reassigning weight to  $e$  based on Eq. (10),
36          |   Eq. (11) and Eq. (12);
37      end
38    end
39  end
40 end
41 end
42 end
43 end
44 end
45 end
46 end
47 end
48 end
49 end
50 end
51 end
52 end
53 end
54 end
55 end
56 end
57 end
58 end
59 end
60 end
61 end
62 end
63 end
64 end
65 end
66 end
67 end
68 end
69 end
70 end
71 end
72 end
73 end
74 end
75 end
76 end
77 end
78 end
79 end
80 end
81 end
82 end
83 end
84 end
85 end
86 end
87 end
88 end
89 end
90 end
91 end
92 end
93 end
94 end
95 end
96 end
97 end
98 end
99 end
100 end
101 end
102 end
103 end
104 end
105 end
106 end
107 end
108 end
109 end
110 end
111 end
112 end
113 end
114 end
115 end
116 end
117 end
118 end
119 end
120 end
121 end
122 end
123 end
124 end
125 end
126 end
127 end
128 end
129 end
130 end
131 end
132 end
133 end
134 end
135 end
136 end
137 end

```

---

- (2) Let  $N_P(v_i)$  be a normal neighbor set of  $v_i$  in  $G_t^P$ .  $N_P(v_i) = \{v_j\}$ .
- (3)  $E_P = \{e_{kj}^P, e_{mj}^P\}$  denotes a set of new edges between each node in  $N_T(v_i)$  and  $N_P(v_i)$ . Each edge in  $E_P$  is established in  $G_t^P$ . Furthermore, if  $v_i$  is failed,  $v_i$  should be removed from  $G_t^P$ .

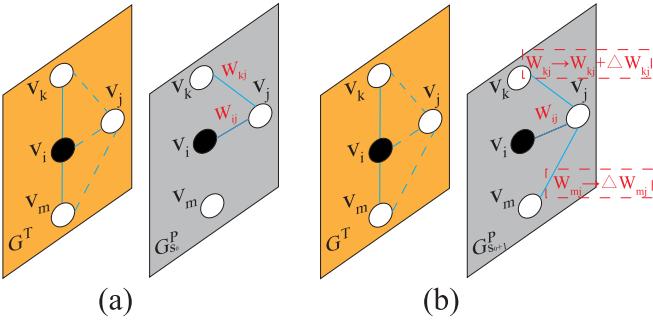


Fig. 3. Illustration of the method of passenger flow redistribution.  $v_i$ ,  $v_j$ ,  $v_k$  and  $v_m$  are nodes in a TL-RTN.  $v_i$  and  $v_m$  are the neighbors of  $v_i$  in the lower layer network (i.e., the rail network  $G^T$ ).  $v_j$  is the neighbor of  $v_i$  in the upper layer network (i.e., the passenger flow network  $G_{s0}^P$ ). The blue links denote edges in  $G^T$  and  $G_{s0}^P$ . (a) A failed or overloaded node  $v_i$  means that passengers cannot enter or leave this station. (b) The passenger flow originally between  $v_i$  and  $v_j$  is allocated to  $v_k$  and  $v_m$ .  $e_{kj}^P$  and  $e_{mj}^P$  are established in  $G_{s0}^P$ . If  $v_i$  is failed, then  $e_{ij}^P$  is deleted from  $G_{s0}^P$ . The weight of  $e_{kj}^P$  is the original weight of an edge (i.e.,  $w_{kj}(s_0)$ ) plus the passenger flow volume assigned from  $v_i$  to  $v_k$  (i.e.,  $\Delta w_{kj-i}$ ). Because the original weight of  $e_{mj}^P$  is zero, the weight of  $e_{mj}^P$  is the passenger flow volume assigned from  $v_i$  to  $v_m$  (i.e.,  $\Delta w_{mj-i}$ ).

- (4) The weight of the established edges are assigned based on Eq. (10). Here,  $e_{kj}^P$  is taken as an example.

$$w_{kj}(s+1) \rightarrow w_{kj}(s) + \Delta w_{kj-i} \quad (10)$$

where  $w_{kj}(s+1)$  represents the weight of edge  $e_{kj}^P$  at time step  $s+1$ . Because there is an edge between  $v_k$  and  $v_j$  at time step  $s$  in  $G_t^P$ ,  $w_{kj}(s)$  is the weight of the original edge  $e_{kj}^P$ . However, if there is no edge between  $v_k$  and  $v_j$  at time step  $s$  in  $G_t^P$ ,  $w_{kj}(s) = 0$ .  $\Delta w_{kj-i}$  denotes the passengers assigned by  $v_i$ . The calculation of  $\Delta w_{kj-i}$  is shown in Eq. (11), where  $EB(e_{ki}^T)$  represents the edge betweenness of  $e_{ki}^T$ , which is calculated based on Eq. (12).  $v_n$  is any node in  $N-T(v_i)$ ,  $\sum_{s \in N-T(v_i)} EB(e_{si}^T)$  is the sum of the edge betweenness.

$$EB(e_{kj}^T) = \sum_{s,t \in [1,N]} \frac{\sigma_{st}(e_{kj}^T)}{\sigma_{st}} \quad (12)$$

where  $\sigma_{st}(e_{kj}^T)$  denotes the number of the shortest paths between  $v_s$  and  $v_t$  past  $e_{kj}^T$ .  $\sigma_{st}$  denotes the number of the shortest paths between  $v_s$  and  $v_t$ .

In order to estimate of robustness of a rail transit network, some metrics are defined as follows.

**Definition 7:** The **The relative size of the largest component** of  $G_t$  is denoted as  $LC$  and formulated as Eq. (13).

$$LC = \frac{N'}{N} \quad (13)$$

where  $N'$  denotes how many nodes are left in the largest component when partial nodes are failed.  $N$  means the total number of nodes.  $LC$  characterizes the robustness of  $G_t$  in the network scale. A higher value of  $LC$  usually denotes higher robustness of the network under attack.

**Definition 8:** The **efficiency** of  $G_t$  is denoted as  $E$  and formulated as Eq. (14).

$$E = \frac{1}{N(N-1)} \sum_{i,j=1(i \neq j)}^N \frac{1}{d_{ij}} \quad (14)$$

where  $d_{ij}$  is the number of edges along the shortest path from  $v_i$  to  $v_j$  in  $G_t$ . If there is no reachable path between  $v_i$  and  $v_j$ , define  $1/d_{ij} = \infty$ .  $E$  is used to estimate the network connectivity. The higher the value of  $E$  is, the better the network connectivity and the higher the efficiency are.

#### IV. SIMULATION AND ANALYSIS

Sec. IV-A introduces the datasets and statistical analysis of the Shanghai RTN. Having constructed the Shanghai TL-RTN in Sec. IV-B, we design experiments that exploring the Shanghai TL-RTN robustness in Sec. IV-C.

##### A. Datasets and Statistical Analysis

In this section, we first introduce the characteristics of the datasets in Sec. IV-A.1. Then, we analyze passenger travel patterns based on the previous datasets in Sec. IV-A.2.

**1) Dataset Description:** This paper uses two datasets as follows: one is the Shanghai rail transit line and station data. The other is the Shanghai rail transit smart card data.

**(1) The Shanghai rail transit line and station data.** We collected this dataset from the Shanghai Metro.<sup>1</sup> This dataset includes the information of 14 subway lines and 289 subway stations from April 2015. In addition, stations are labeled to facilitate the exploration of passenger travel patterns and construction of the Shanghai TL-RTN in Table A of Supplementary File. Stations with the same name but different routes are represented by the same *ID*, which means that they are the same node in a network. Stations with different names are represented by a different *ID*, conveying that they are different nodes. One special example is the two Pudian Road which are located on different areas; thus, they are treated as two different nodes. Therefore, subway lines can be denoted as  $L_i$ ,  $i = 1, 2, \dots, 13, 16$ , and subway stations can be denoted as  $v_i$ ,  $i \in [1, 289]$ . Fig. 4 visualizes the Shanghai rail transit lines based on the latitude and longitude information of all stations.

**(2) The Shanghai rail transit smart card data.** This dataset was collected by Shanghai Public Transportation Card Co. Ltd and released by the organized committee of Shanghai Open Data Apps.<sup>2</sup> This dataset covers 11 million individuals and 123 million trips during April 2015. Given that the raw data may miss something or contain some errors, we cleaned and processed erroneous data. Examples of such processed data is shown in Table I, through which we can grasp the detailed passenger smart card transaction information.

<sup>1</sup><http://service.shmetro.com/>

<sup>2</sup><http://data.sh.gov.cn/>

TABLE I

EXAMPLES OF THE SMART CARD DATA FOR THE SHANGHAI RAIL TRANSIT. USERID DENOTES THE SMART CARD NUMBER OF A PASSENGER. DATE AND TIME DENOTE THE CHECK-IN TIME OR CHECK-OUT TIME OF A PASSENGER. AMOUNT IS THE PAYMENT FOR A TRIP. IN THE CASE WHERE THE USERID IS THE SAME, THE RECORD IS A CHECK-IN DATA POINT IF THE AMOUNT IS 0; THE RECORD IS A CHECK-OUT DATA POINT IF THE AMOUNT IS GREATER THAN 0. THEREFORE,  $(v_1, v_{235})$ ,  $(v_{11}, v_{143})$ ,  $(v_{48}, v_{224})$  ARE ALTERNATING SEQUENCES OF NODES AND LINKS. IF THE TOTAL PASSENGER FLOW VOLUME BETWEEN  $v_1$  AND  $v_{235}$  IS  $w_{1,235}$ ,  $(v_1, v_{235}, w_{1,235})$  IS AN *OD* PAIR

Record	UserID	Date	Time	Station	Line	Amount
1	602141128	2015-04-01	07:51:08	$v_1$ (Xinzhuang)	Line 1	0.00
2	602141128	2015-04-01	09:07:57	$v_{235}$ (East Changji Road)	Line 11	6.00
3	2201252167	2015-04-01	18:43:14	$v_{11}$ (South Shaanxi Road)	Line 1	0.00
4	2201252167	2015-04-01	19:20:33	$v_{143}$ (Changzhong Road)	Line 7	4.00
5	2001530605	2015-04-01	08:41:14	$v_{48}$ (Guanglan Road)	Line 2	0.00
6	2001530605	2015-04-01	09:28:17	$v_{224}$ (Guoquan Road)	Line 10	4.00

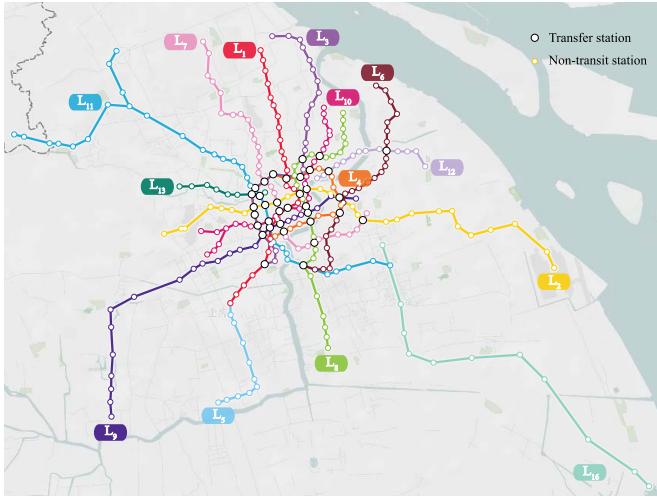


Fig. 4. Illustration of Shanghai rail transit lines and stations, in which each line is highlighted with a different color. There are 14 lines and 289 stations, including 35 transfer stations.

2) *Statistics of Passenger Travel Patterns*: The temporal and spatial characteristics of passenger travel greatly influence the dynamic distribution of passenger flow, and understanding and revealing the laws of passenger travel patterns can help us better construct the Shanghai TL-RTN. With the help of the Shanghai rail transit smart card data, researchers can explore passenger travel patterns [55]–[57]. We first select three weeks of data (i.e., April 6, 2015 to April 26, 2015) for passenger flow statistics. The results are shown in Fig. 5, from which we can observe the residents' lifestyle switching back and forth between weekday mode (from Monday to Friday) and weekend mode (Saturday, Sunday and Tomb Sweeping Festival on April 6) [58]. Moreover, we use a box-whisker plot to further analyze the distribution laws of passenger flow during the three-week study. Fig. 6 illustrates that passenger

flow varies greatly on weekdays but that the passenger flows among each period on weekends fluctuate slightly. Thus, we classify each weekday into several periods according to the passenger flow laws as follows.

- (1) Morning off-peak hours are from 6:00 to 7:00.
- (2) Morning peak hours are from 7:00 to 9:00.
- (3) Afternoon off-peak hours are from 9:00 to 17:00.
- (4) Evening peak hours are from 17:00 to 19:00.
- (5) Evening off-peak hours are from 19:00 to 00:00.

Furthermore, we quantitatively analyze the passenger travel patterns of the Shanghai RTN for three time periods, i.e., morning (8:00–9:00), afternoon (12:00–13:00) and evening (18:00–19:00) on weekdays and weekends. First, we use chord diagrams to describe the passenger flow characteristics between the two lines. As shown in Fig. 7, we can observe that the passenger flow volume in each line during peak hours on weekdays is larger than that of other periods. For example, the passenger flow volumes in  $L_1$  and  $L_2$  are larger than those of other lines.  $L_{16}$  has the smallest passenger flow volume. As shown in Fig. 7(a) and Fig. 7(b), there is an inverse relationship between the passenger travel patterns during morning and evening on weekdays. For example, most passengers travel during morning on weekdays from  $L_5$  to  $L_1$ , while most passengers travel during the evening on weekdays from  $L_1$  to  $L_5$ . This phenomenon indicates that many stations in  $L_1$  are near the work area, and many stations in  $L_5$  are near the residential area. As shown in Fig. 7(b) and Figs. 7(d)–7(f), the passenger flow volume in each line greatly reduces when people take a break, leading to a variance and disorder in passenger travel. Another phenomenon is that passengers on  $L_1$  and  $L_8$  travel in opposite directions on weekdays and weekends, which indicates that many stations of  $L_8$  are near areas that attract passengers for entertainment.

Then, the passenger flow characteristics between the two stations are as shown in Fig. 8. The passenger flow volumes

$$\Delta w_{kj-i} = \begin{cases} w_{ij}(s) * \frac{EB(e_{ki}^T)}{\sum_{v_n \in (N_T(v_i))} EB(e_{ni}^T)}, & P_i(s) \geq 1 \\ (S_i - Q_i) * \frac{EB(e_{ki}^T)}{\sum_{v_n \in (N_T(v_i))} EB(e_{ni}^T)} * \frac{w_{ij}(s)}{S_i}, & S_i > Q_i \text{ and } P_i(s) < 1 \end{cases} \quad (11)$$

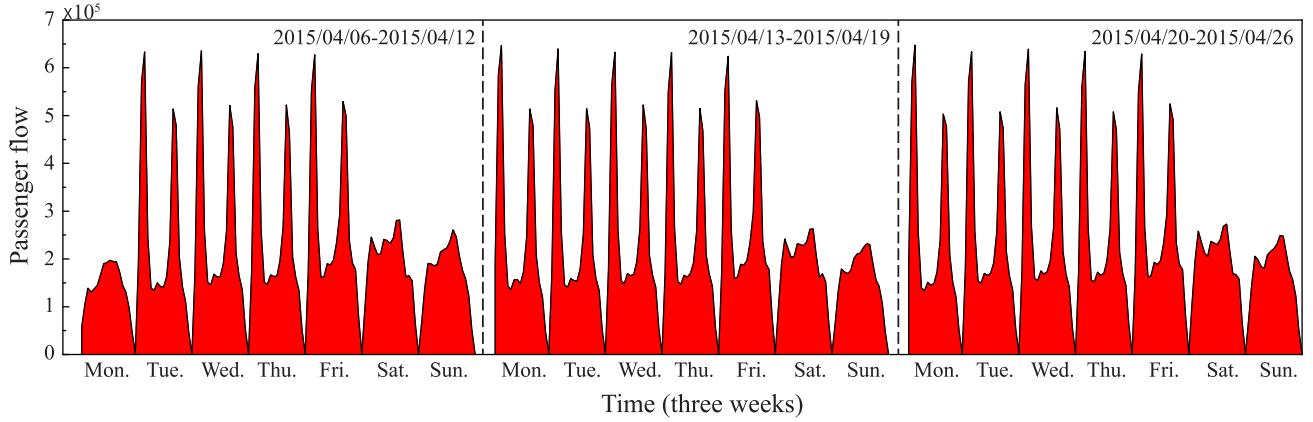


Fig. 5. The dynamic changes in passenger flow in three weeks. Each day is divided by hour from 6:00 to 24:00. Thus, there are 18 time periods every day. The statistical results show that the travel patterns of residents are basically identical in each week.

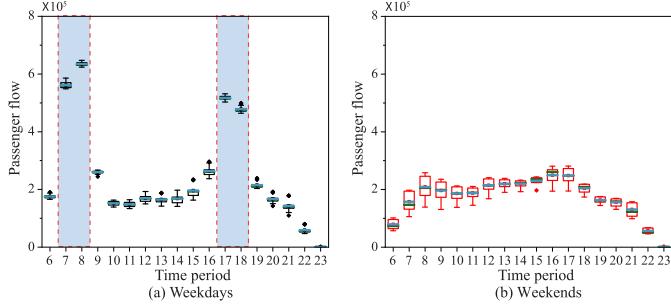


Fig. 6. The dynamic distribution of passenger flow during different time periods in three weeks. The x-axis denotes 18 h in a day starting at 6 am. The y-axis denotes the total passenger flow volume. The statistical results show that there are two peak hours in a day, i.e., morning peak hours are from 7:00 to 9:00 and evening peak hours are from 17:00 to 19:00.

during morning and evening on weekdays are greater than those of other time periods. Based on Figs. 8(a) and 8(c), we find that the routes with the large passenger flow volumes are from  $v_{194}$  (Jiuting) to  $v_{199}$  (Caohejiang Hi-Tech park), from  $v_{23}$  (Tonghe Xincun) to  $v_{16}$  (Shanghai Railway Station), from  $v_{192}$  (Shehan) to  $v_{199}$  (Caohejiang Hi-Tech Park) and from  $v_1$  (Xinzhuang) to  $v_{13}$  (People's Square). Generally, during peak hours, many passenger origin stations are located in residential area and their destination stations are located in the working area [49]. Therefore, there are many commuters near  $v_{194}$ ,  $v_{23}$ ,  $v_{192}$  and so on, and there are many working areas near  $v_{199}$ ,  $v_{16}$ ,  $v_{13}$  and so on. In addition, the passenger flow volumes during the three periods on weekends are relatively steady, and the routes with the large passenger flows are from  $v_1$  to  $v_8$  and from  $v_1$  to  $v_{13}$ . We know that  $v_8$  and  $v_{13}$  are near famous shopping districts, so people are inclined to visit some places for leisure or just remain at home on weekends.

### B. Initialization of Shanghai TL-RTN

Based on the analyses in Sec. IV-A.2, we find that the dynamic distribution of passenger flow among weekdays is similar, but the passenger travel patterns are diverse on weekdays. In addition, several studies have shown that subway stations are often crowded on weekdays [58]. Therefore,

TABLE II  
THE NODE PROPERTY METRICS OF THE RAIL NETWORK (I.E.,  $G^T$ ) AND THREE PASSENGER FLOW NETWORKS (I.E.,  $G_{MP}^P$ ,  $G_{OP}^P$  AND  $G_{EP}^P$ ).  $N$  AND  $E$  REPRESENT THE TOTAL NUMBERS OF NODES AND EDGES, RESPECTIVELY.  $\text{Max}_D_n$ ,  $\text{Max}_B_n$  AND  $\text{Max}_S_n$  REPRESENT THE IDs OF THE NODES WITH THE LARGEST DEGREE, THE LARGEST BETWEENNESS AND THE LARGEST STRENGTH, RESPECTIVELY.  $\text{Max}_D$ ,  $\text{Max}_B$  AND  $\text{Max}_S$  ARE THEIR CORRESPONDING VALUES. ‘-’ MEANS THERE IS NO SUCH METRICS IN THIS NETWORK

Indicators	$G^T$	$G_{MP}^P$	$G_{OP}^P$	$G_{EP}^P$
$N$	289	289	289	289
$E$	332	34512	25834	32555
$\text{Max}_D_n$	$v_{88}$	$v_{13}$	$v_{16}$	$v_{13}$
$\text{Max}_B_n$	$v_{16}$	$v_{16}$	$v_{16}$	$v_{16}$
$\text{Max}_S_n$	-	$v_{13}$	$v_{13}$	$v_{13}$
$\text{Max}_D$	8	288	288	288
$\text{Max}_B$	0.23	0.001	0.005	0.002
$\text{Max}_S$	-	25390	7668	22482

studies on the RTN robustness when a node is disrupted at different times on weekdays can help us explore the following question: what effects do the dynamic distribution of passenger flow have on the RTN robustness? In this paper, we select three time periods on April 13 (i.e., morning (8:00–9:00), evening (18:00–19:00) and afternoon (12:00–13:00), as discussed in Sec. IV-A.2) to construct Shanghai TL-RTNs.

First, the rail network ( $G^T$ ) is constructed based on Def. 2. Then, three passenger flow networks corresponding to three time periods are constructed based on Def. 3, which are denoted as  $G_{MP}^P$ ,  $G_{OP}^P$  and  $G_{EP}^P$ , respectively. As shown in Fig. 9, the above three passenger flow networks are combined with the rail network (i.e.,  $G^T$ ) to form the Shanghai TL-RTN based on Def. 1. These three Shanghai TL-RTNs are denoted as  $G_{MP}$ ,  $G_{OP}$  and  $G_{EP}$ . Furthermore, Table II counts three node property metrics (i.e., degree, betweenness, strength) of the rail network and three passenger flow networks.

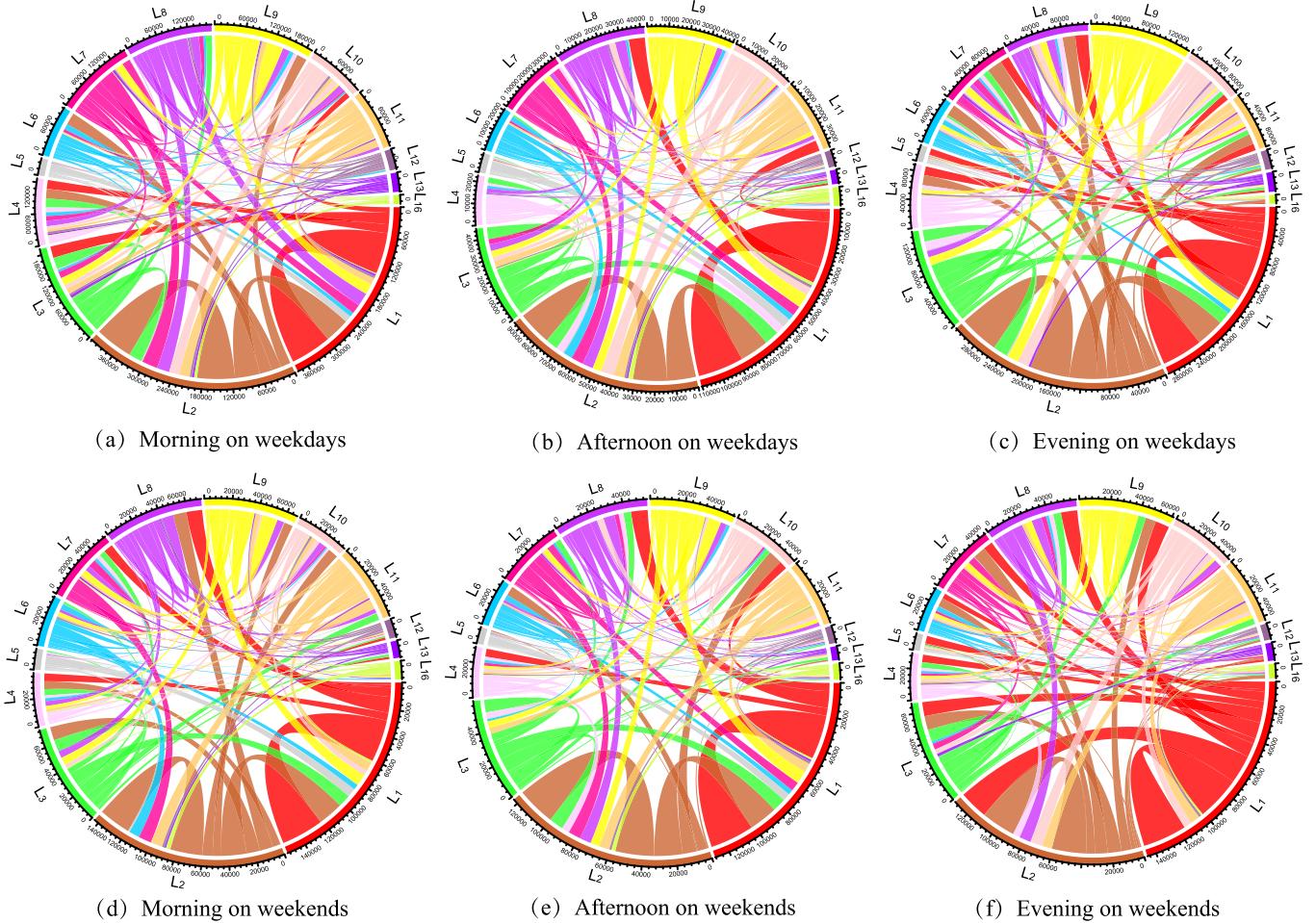


Fig. 7. Subway-lines-based passenger travel patterns.  $L_i$  ( $i = 1, 2, \dots, 13, 16$ ) is the ID of each line. Each circle consists of arcs of various colors. Each color in the outer ring represents a subway line. Each link in the inner ring is the passenger flow volume between a pair of lines. The larger the passenger flow volume is, the thicker the link becomes. The link color represents the travel direction of passengers. Links with the same / different color as the outer ring represent the number of passengers leaving / entering this line. The results show that during the peak hours on weekdays, passenger travel patterns are diverse but regular, and during other time periods, passengers are relatively less willing to travel.

From the statistical results in Table II, we find that in  $G_{MP}^P$  and  $G_{EP}^P$ ,  $v_{13}$  (People's Square) has the largest degree and strength, which indicates that there are many passengers from different areas passing through  $v_{13}$ . Insofar as the topology of the passenger flow network reflects the relationship between the origin and destination stations, the shortest path of the passenger flow network is rare. Therefore, it can be seen that the largest betweenness of passenger flow networks is much smaller than that of the rail network.

### C. Dynamic Robustness Analysis of the Shanghai TL-RTN

According to the cascading failure process of TL-RTN in Sec. III-B.2, we conclude that external perturbations and failure modes have different effects on the network robustness. We will explore the effects of the above factors on the Shanghai TL-RTN robustness in Sec. IV-C.1 and Sec. IV-C.2.

Each experiment is simulated 100 times, and the final experimental results are the averaged value of these 100 simulation runs. Before we start the simulation analysis, some parameters are set as follows:

- Because the initial state of each node is normal, the initial state of each node in  $G^T$  and  $G_t^P$  can be a random number between 0 and 1 [35].
- In Sec. IV-C.1 and Sec. IV-C.2, when relevant experiments are based on the Shanghai TL-RTN, the coupling strength coefficients ( $\varepsilon_1, \varepsilon_2$ ) of  $G^T$  and  $G_t^P$  are set to  $\varepsilon_1 = 0.4$  and  $\varepsilon_2 = 0.6$ ; when relevant experiments are only based on the Shanghai rail network, the coupling strength coefficients of  $G^T$  and  $G_t^P$  are set to  $\varepsilon_1 = 0.4$  and  $\varepsilon_2 = 0$ , which means that the impact of passenger flow is not considered.
- The hourly capacity of a station is calculated based on Eq. (4). Specifically, we set two capacity control parameters  $\alpha = \beta = 0.1$  [52]. The time set is [6, 24], meaning April 13 is divided by hour from 6:00 to 24:00.

*1) The Effect of External Perturbations:* In this section, under the random failure mode, we analyze the effects of external perturbations on the Shanghai TL-RTN robustness with our extended CML model. The random failure mode ( $F_r$ ) is defined as a node that fails randomly in a network.

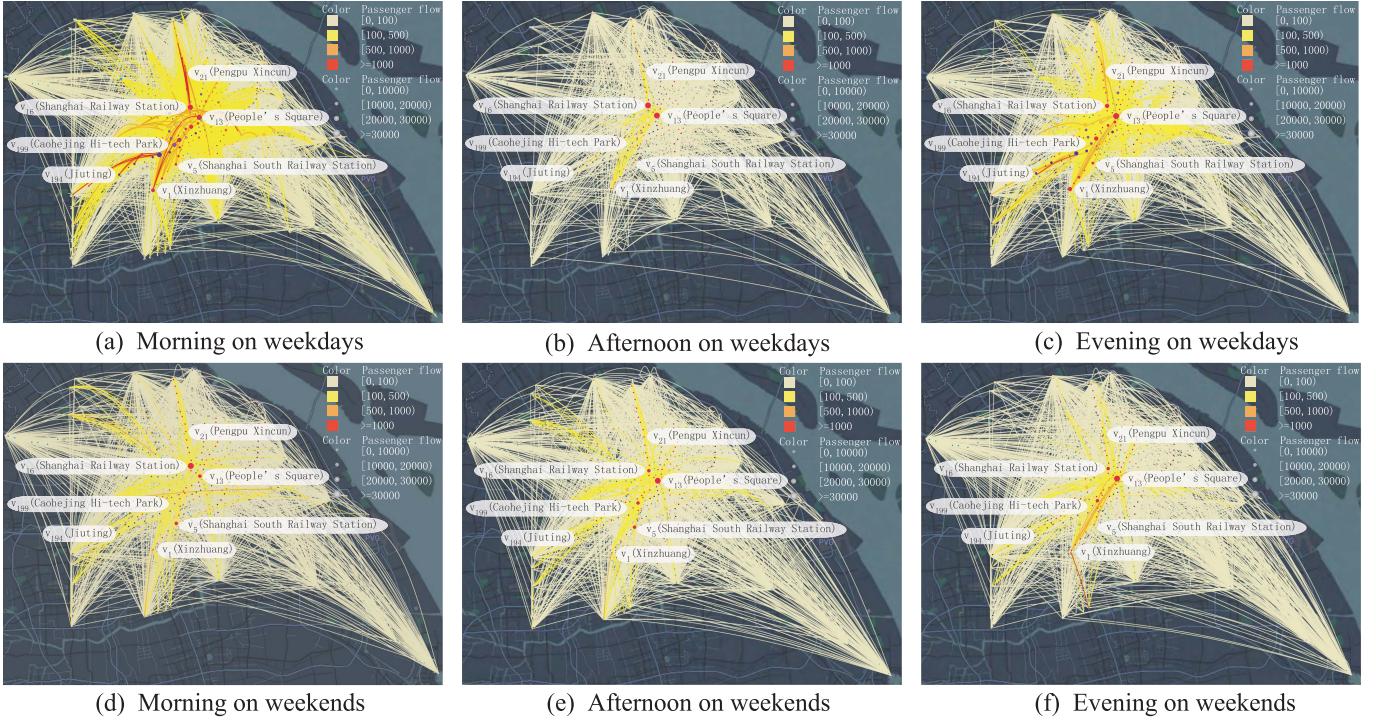


Fig. 8. Subway-stations-based passenger travel patterns. The results show that travel patterns include weekday and holiday modes. In the weekday mode, original stations with large passenger flows are located in the residential area, and destination stations with large passenger flows are located in working area. In the holiday mode, people generally choose to remain at home or relax in the famous business districts (such as  $v_8$  (Xujiahui) and  $v_{13}$  (People's Square)).

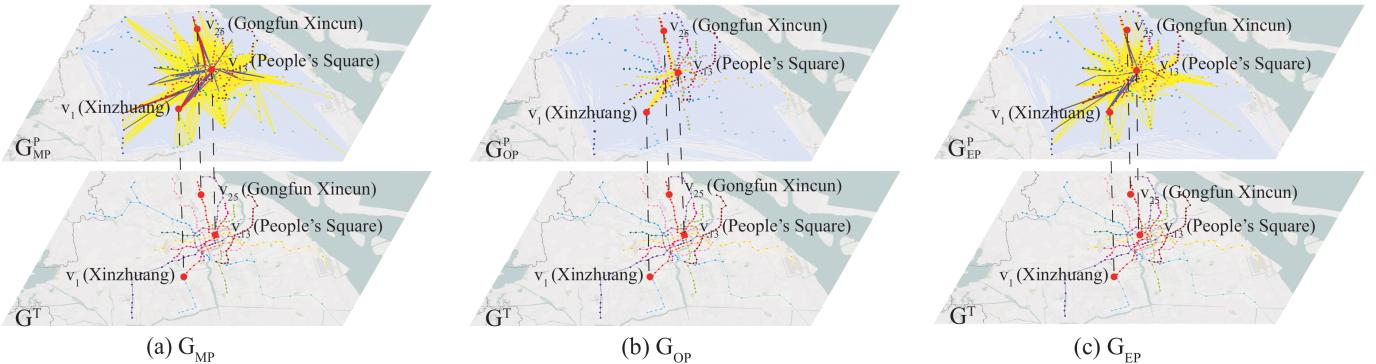


Fig. 9. Shanghai TL-RTNs during three time periods on weekdays. (a)  $G_{MP}$  is the Shanghai TL-RTN during morning (8:00–9:00). (b)  $G_{OP}$  is the Shanghai TL-RTN during afternoon (12:00–13:00). (c)  $G_{EP}$  is the Shanghai TL-RTN during the evening (18:00–19:00). The results show that the Shanghai RTN during the different time period has different structures due to the dynamic distribution of passenger flow.

Meanwhile, to clearly reveal the impact of dynamic passenger flow on the network robustness, we conduct a comparative analysis, that is, considering the RTN robustness only from the perspective of rail structure.

As shown in Fig. 10, the comparison experiments on the relationship of the external perturbation  $R$  with three robustness metrics, namely, the cascade size  $I$ , the network efficiency  $E$  and the relative size of the largest component  $LC$ . Judging from the changing trends of such three metrics, we find that the network robustness becomes weaker as  $R$  increases. Moreover, the network falls into the global cascading failures when  $R = 2.5$ . Thus, all these robustness metrics can reflect the trend that the network robustness changes with the external perturbation. More specifically, an inserted logarithmic chart

in Fig. 10(a) is shown a clear gap under a small external perturbation (i.e.,  $R < 2$ ). Taking this figure as an example, we further analyze the relationship between the cascade size  $I$  and the external perturbation  $R$  as follows.

When  $1 \leq R \leq 1.5$ , there are 8.46% failed nodes in  $G_{MP}$ , and there is almost no cascading failure in  $G_{OP}$ ,  $G_{EP}$  and  $G^T$ . When  $1.5 < R < 2$ , cascading failures occur in  $G_{MP}$ ,  $G_{OP}$ ,  $G_{EP}$  and  $G^T$ . In particular, the cascade size of  $G_{MP}$  reached 18.6% – 54.9%, which is the largest among the above four networks. These findings are understandable in that although the external perturbation is small, the passengers at the disrupted stations will reroute spontaneously to other normal stations, and this phenomenon will cause such stations to become crowded or even be disrupted. In other

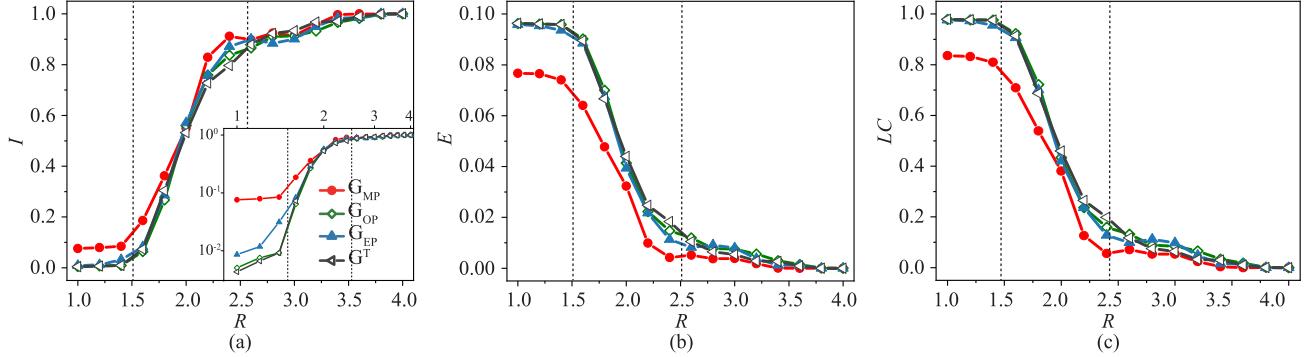


Fig. 10. Cascading failures of  $G_{MP}$ ,  $G_{OP}$ ,  $G_{EP}$  and  $G^T$  under  $F_r$ .  $R$  represents different external perturbations.  $I$  is the cascade size.  $E$  is the network efficiency.  $LC$  is the relative size of the largest component.  $G_{MP}$ ,  $G_{OP}$ ,  $G_{EP}$  and  $G^T$  represent the Shanghai TL-RTN during three time periods (morning (8:00–9:00), evening (18:00–19:00) and afternoon (12:00–13:00)) and rail network (defined in Eq. 2), respectively. The results show that when the perturbation is small, the cascading failure still occurs in peak hours. As the perturbation increases,  $I$ ,  $E$ ,  $LC$  are more related to the rail network.

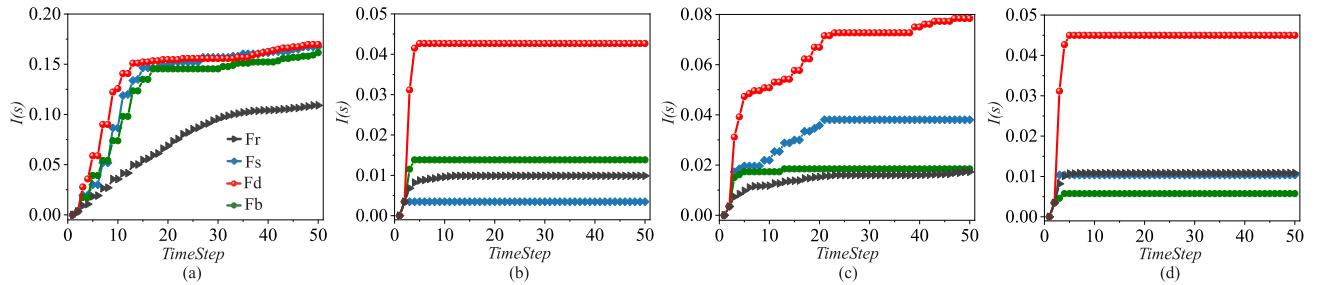


Fig. 11. The robustness of (a)  $G_{MP}$ , (b)  $G_{OP}$ , (c)  $G_{EP}$  and (d)  $G^T$  under different failure modes when  $R = 1.5$ . Based on the relationship between the cascade size  $I(s)$  and the time steps of failure propagation  $TimeStep$ , we can conclude that different failure modes have different effects on network robustness, and if the passenger flow volume is larger, the effects of failure modes on the network robustness will become more obvious.

words, the passenger flow volume will amplify the effect of external perturbations on the network robustness. The larger the passenger flow volume is, the worse is the tolerance of the RTN to station failure. In Sec. V-A, we will further explore the amplification of external perturbations by the dynamic passenger flow.

When  $2 \leq R < 2.5$ , a large-scale cascading failure will occur in all of the networks, and the gap of the cascade size among these networks becomes very small. When  $R \geq 2.5$ , global cascading failures occur in all of the networks. After  $R \geq 2.5$ , the cascading failure processes of the above networks are basically similar. These findings suggest that the failure propagation speed in the rail network is much faster than that in the passenger flow network. Therefore, we can conclude that when the external perturbation is sufficiently large, the network robustness can be analyzed and estimated only by the failure propagation in the rail network [59].

2) *The Effect of Different Failure Modes:* In this section, we analyze the cascading failure processes of  $G_{MP}$ ,  $G_{OP}$ ,  $G_{EP}$  and  $G^T$  under different failure modes in the case of a small external perturbation ( $R = 1.5$ ). In addition to the random failure mode ( $F_r$ ), we define a targeted failure mode. Based on Table II, in  $G_{MP}^P$ ,  $G_{OP}^P$  and  $G_{EP}^P$ ,  $Max\_S$  is  $v_{13}$  and  $Max\_B$  is  $v_{16}$ ; in  $G^T$ ,  $Max\_D$  is  $v_{88}$  and  $Max\_B$  is  $v_{16}$ . Therefore, in the targeted failure mode, there are the following three failure strategies:

- (1)  $F_s$ : This failure mode is defined as the node with the largest strength fails in  $G_{MP}^P$ ,  $G_{OP}^P$  and  $G_{EP}^P$  (i.e.,  $v_{13}$ ).

- (2)  $F_b$ : This failure mode is defined as the node with the largest betweenness in  $G_{MP}^P$ ,  $G_{OP}^P$ ,  $G_{EP}^P$  and the node with the largest betweenness in  $G^T$  fails (i.e.,  $v_{16}$ ).
- (3)  $F_d$ : This failure mode is defined as the node with the largest degree fails in  $G^T$  (i.e.,  $v_{88}$ ).

As shown in Fig. 11(a), different failure modes cause different cascading failures in  $G_{MP}$ . Especially,  $G_{MP}$  under  $F_r$  has the smallest cascade size which are similar in other failure modes. Results show that during peak hours, the RTN is more robust under random failure mode than targeted failure modes [11], [41], [42]. However, as shown in Figs. 11(b)–(d),  $G_{OP}$ ,  $G_{EP}$  and  $G^T$  under different failure modes only cause a 1% – 8% cascade size. The cascade size caused by  $G_{OP}$  and  $G_{EP}$  in  $F_r$  is larger than  $G^T$  in  $F_r$ . Results suggest that if the passenger flow volume is large, the RTN robustness in the same failure mode would decrease.

## V. DISCUSSION

In this section, we discuss the Shanghai TL-RTN robustness in small external perturbations and coupling strength coefficients. Sec. V-A identifies some important stations under small external perturbations. Sec. V-B explores the effect of coupling strength coefficients on the Shanghai TL-RTN robustness.

### A. Important Stations Under Small External Perturbations

Due to the temporal and spatial characteristics of passenger travel, the cascading failures caused by stations under different external perturbations at different times vary. Based on the

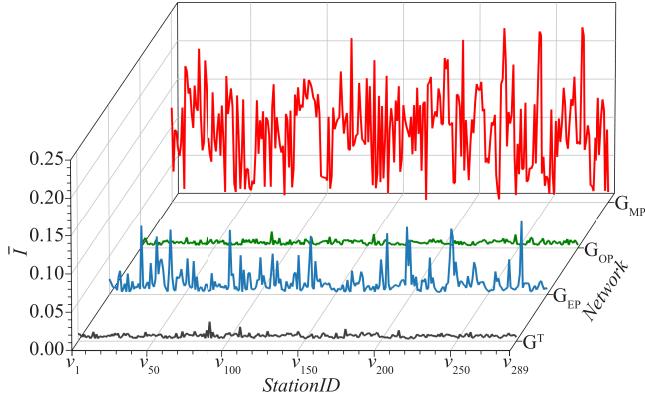


Fig. 12. When  $R \in [1, 1.5]$ , the average cascade size  $\bar{I}$  caused by each node failure in  $G^T$ ,  $G_{MP}$ ,  $G_{OP}$  and  $G_{EP}$  (i.e., the Shanghai rail network and TL-RTNs during three periods: morning (8:00–9:00), evening (18:00–19:00) and afternoon (12:00–13:00), respectively). The results show that the cascade size caused by disrupted stations during peak hours is larger than that during other time periods. Thus, RTN robustness is weak during peak hours.

analysis in Sec. IV-C.1, we know that a station under a small external perturbation can also cause a large-scale cascading failure, even though this phenomenon is rarely considered in relevant studies. Therefore, in this section, we further explore the average cascade size  $\bar{I}$  in  $G_{MP}$ ,  $G_{OP}$ ,  $G_{EP}$  and  $G^T$ , which is caused by each node in the above networks under a perturbation  $R$  from 1.0 to 1.5.  $\bar{I}$  is calculated by Eq. (15).

$$\bar{I} = \frac{\sum_{r \in [1, 1.5]} I_r}{6} \quad (15)$$

where  $I_r$  denotes the cascade size  $I$  when the external perturbation is  $r$ .

Figure 12 shows that in  $G_{MP}$ , all nodes under a small perturbation will cause a cascading failure. Moreover, the averaged cascade size  $\bar{I}$  caused by each node in  $G_{MP}$  is larger than that in other networks. Then, we rank the stations in  $G_{MP}$  based on their averaged cascade size which is calculated by averaging the cascade size after the perturbation  $R \in [1, 1.5]$ . We list the top 10 stations in Table III. Surprisingly, the top two of the lists is  $v_{221}$  (North Sichuan Road), which is neither the largest degree node nor the largest betweenness node of  $G_{MP}$ . However, this result is still understandable because the neighbor nodes of  $v_{221}$  are  $v_{83}$  (Hailun Road) and  $v_{220}$  (Tiantong Road). Both  $v_{83}$  and  $v_{220}$  are transfer stations, indicating that large passenger flows of  $v_{83}$  and  $v_{220}$  have a great tendency to pass through  $v_{221}$ . Therefore, there is a large passenger flow in  $v_{221}$  that cannot be counted by  $OD$  data. Moreover,  $v_{221}$  is not a transfer station, so its capacity for the passenger flow volume is relatively small. Therefore,  $v_{221}$  is prone to passenger flow overload, which can lead to cascading failure. There are many similar stations in Table III. For example, the neighbor of  $v_{202}$  (Dapuqiao) is  $v_{201}$  (Jiashan Road), and the neighbor of  $v_{205}$  (Shangcheng Road) is  $v_{88}$  (Century Avenue). In real life, we are supposed to prevent such stations from being disrupted.

In addition, hub nodes such as  $v_{10}$  (Changshu Road),  $v_{11}$  (South Shaanxi Road) and  $v_{16}$  (Shanghai Railway Station) in the Shanghai TL-RTN all have great importance for the

TABLE III  
RANKING RESULT OF STATIONS IN  $G_{MP}$  BASED ON THE AVERAGE CASCADE SIZE  $\bar{I}$ . WHEN  $R \in [1, 1.5]$ , THE AVERAGE CASCADE SIZE OF THE TOP TEN STATIONS IS GREATER THAN 15%, AND SOME OF THEM ARE NOT ONLY HUB NODES THAT WE NORMALLY FOCUS ON BUT ALSO THE NEIGHBORS OF HUB NODES

Ranking	$G_{MP}$	Average cascade size
1	$v_{10}$ (Changshu Road)	19.19%
2	$v_{221}$ (North Sichuan Road)	18.59%
3	$v_{11}$ (South Shaanxi Road)	17.53%
4	$v_{202}$ (Dapuqiao)	17.22%
5	$v_{16}$ (Shanghai Railway Station)	16.69%
6	$v_{94}$ (Luban Road)	16.40%
7	$v_{205}$ (Shangcheng Road)	16.32%
8	$v_{171}$ (Laoximen)	16.19%
9	$v_{90}$ (Lancun Road)	15.84%
10	$v_{170}$ (Lujiabang Road)	15.84%

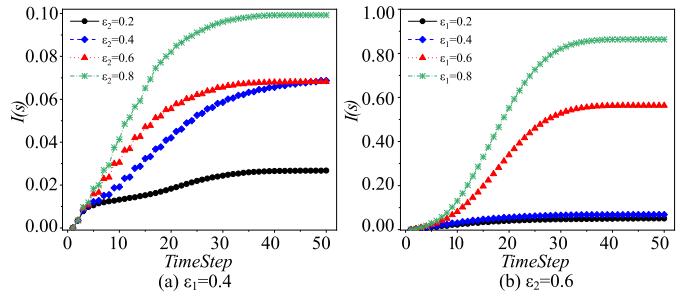


Fig. 13.  $G_{MP}$  under random failure ( $F_r$ ) when  $R = 1.5$ ; (a)  $\epsilon_1 = 0.4$ ,  $\epsilon_2$  taking the values of 0.2, 0.4, 0.6, and 0.8; and (b)  $\epsilon_2 = 0.6$ ,  $\epsilon_1$  taking the values of 0.2, 0.4, 0.6 and 0.8. Based on the relationship between the cascade size  $I(s)$  and the time steps of failure propagation  $TimeStep$ , we can conclude that as the coupling strength coefficients increase, the cascade size caused by disrupted stations becomes increasingly larger.

network topology and large passenger flows. There is no doubt that  $v_{10}$ ,  $v_{11}$  or  $v_{16}$  with small perturbations can probably lead to large-scale cascading failures. Therefore, when a perturbation is small, large-scale cascading failures are caused by both hub nodes and neighbors of these hub nodes.

### B. The Effect of Coupling Strength Coefficients

In real life, the state of a station (i.e., normal or disrupted) has some impact on its neighbors. For example, when any station breaks down in an RTN, the passengers who are currently at or planning to go to disrupted stations will reroute to other stations where normal service is available, especially, the neighboring stations of a disrupted station. However, the above process might cause the passenger flow volume to exceed the capacity threshold of these normal stations; thus, such stations will become congested or even disrupted. The coupling strength is used to evaluate the impact of the state of a station on its neighbors. Therefore, this section discusses the effect of different coupling strength coefficients on the Shanghai TL-RTN robustness.

In the case of the different values of  $\epsilon_1$  and  $\epsilon_2$ , random failure ( $F_r$ ) is carried out on  $G_{MP}$  with the external

perturbation  $R = 1.5$ . As shown in Fig. 13, either  $\varepsilon_1$  or  $\varepsilon_2$  increases, the cascade size becomes larger, and  $\varepsilon_1$  has a greater impact on the network robustness than  $\varepsilon_2$ . Fig. 13(b) shows that when  $\varepsilon_1$  increases to 0.8, although the external perturbation is small,  $G_{MP}$  is almost globally cascading failure under  $F_r$ . These findings indicate that when constructing subway systems, it is necessary to reduce the coupling strength between stations in the rail structure. In addition, Fig. 13(a) shows that when  $\varepsilon_2$  increases to 0.8, the cascade size is close to 10%. Therefore, we still cannot ignore the influence of the coupling strength in the passenger flow network. In particular, during peak hours, a large number of passengers reroute to other normal stations, which is equivalent to increasing the coupling strength coefficient between two stations and then promoting the spread of failure [22]. In real life, it is necessary to optimize the distribution of passenger flow. For example, by setting up entrances or exits and with intentional guidance, passengers can easily be evacuated from the station and distributed homogeneously among stations.

## VI. CONCLUSION

To reflect the influence of the dynamic passenger flow on the rail transit network (RTN) robustness, we first propose a two-layer network model. Based on this, we construct the Shanghai TL-RTN. Then, we propose an extended CML model to describe the cascading failure process of the Shanghai TL-RTN, through which we can explore the effects of external perturbations and failure modes on the Shanghai TL-RTN robustness and identify some important stations. To conclude, the main findings of this paper are as follows.

First, a 20% cascade size still occurs during the peak hours on weekdays when the Shanghai TL-RTN is under a small external perturbation, from which we conclude that a large passenger flow can amplify the effect of external perturbations on RTN robustness. The larger the passenger flow volume is, the worse is the tolerance of the RTN to station failures. Meanwhile, through sorting stations of the Shanghai TL-RTN during the morning peak hour ( $G_{MP}$ ) based on the averaged cascade size, we find that this phenomenon is not only determined by the hub nodes in the network but also related to the neighbors of hub nodes. Moreover, different failure modes have different effects on network robustness. When the RTN is under the same failure mode at different times, the network robustness will decrease if the passenger flow in disrupted stations increases. In addition, when external perturbations are sufficiently large, the network robustness is affected by the inherent rail structure to a great extent. Finally, as the coupling strength coefficients increase, the cascade size becomes larger, and the RTN robustness becomes weaker.

The above findings provide valuable prior knowledge for the traffic management to accurately understand and estimate the degree and tendency of a fault propagation in the case of emergencies at different stations. Based on these knowledge, the traffic managers will prepare certain pre-arranged planning for dealing with different types of emergencies in the future. The future work focuses on estimating the passenger flow capacity of each station by a more accurate method [60] to

make the extended CML model better simulate the cascading failure process of TL-RTN. Furthermore, we will study the redistribution strategy of passenger flows to improve the RTN robustness [61].

## REFERENCES

- [1] Y. Huang, L. Yang, T. Tang, F. Cao, and Z. Gao, "Saving energy and improving service quality: Bicriteria train scheduling in urban rail transit systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3364–3379, Dec. 2016.
- [2] Y. Zhou, J.-B. Sheu, and J. Wang, "Robustness assessment of urban road network with consideration of multiple hazard events," *Risk Anal.*, vol. 37, no. 8, pp. 1477–1494, Apr. 2017.
- [3] D. Duan *et al.*, "Universal behavior of cascading failures in interdependent networks," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 45, pp. 22452–22457, Oct. 2019.
- [4] K. L. Clark, U. Bhatia, E. A. Kodra, and A. R. Ganguly, "Resilience of the U.S. National airspace system airport network," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 12, pp. 3785–3794, Dec. 2018.
- [5] Q.-C. Lu and S. Lin, "Vulnerability analysis of urban rail transit network within multi-modal public transport networks," *Sustainability*, vol. 11, no. 7, p. 2109, Apr. 2019.
- [6] B. P. Y. Loo and K. Y. K. Leung, "Transport resilience: The occupy central movement in Hong Kong from another perspective," *Transp. Res. A, Policy Pract.*, vol. 106, pp. 100–115, Dec. 2017.
- [7] X. Yang, A. Chen, B. Ning, and T. Tang, "Measuring route diversity for urban rail transit networks: A case study of the Beijing metro network," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 2, pp. 259–268, Feb. 2017.
- [8] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, Jan. 2019.
- [9] S. Derrible and C. Kennedy, "The complexity and robustness of metro networks," *Phys. A, Stat. Mech. Appl.*, vol. 389, no. 17, pp. 3678–3691, Sep. 2010.
- [10] P. Angeloudis and D. Fisk, "Large subway systems as complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 367, pp. 553–558, Jul. 2006.
- [11] C. V. Ferber, B. Berche, T. Holovatch, and Y. Holovatch, "A tale of two cities," *J. Transp. Secur.*, vol. 5, no. 3, pp. 199–216, Jul. 2012.
- [12] Y. Yang, Y. Liu, M. Zhou, F. Li, and C. Sun, "Robustness assessment of urban rail transit based on complex network theory: A case study of the Beijing subway," *Saf. Sci.*, vol. 79, pp. 149–162, Nov. 2015.
- [13] J. Zhang, X. Xu, L. Hong, S. Wang, and Q. Fei, "Networked analysis of the Shanghai subway network, in China," *Phys. A, Stat. Mech. Appl.*, vol. 390, nos. 23–24, pp. 4562–4570, Nov. 2011.
- [14] S. Wang, W. Lv, L. Zhao, S. Nie, and H. E. Stanley, "Structural and functional robustness of networked critical infrastructure systems under different failure scenarios," *Phys. A, Stat. Mech. Appl.*, vol. 523, pp. 476–487, Jun. 2019.
- [15] O. Cats and E. Jenelius, "Planning for the unexpected: The value of reserve capacity for public transport network robustness," *Transp. Res. A, Policy Pract.*, vol. 81, pp. 47–61, Nov. 2015.
- [16] O. Cats, J. West, and J. Eliasson, "A dynamic stochastic model for evaluating congestion and crowding effects in transit systems," *Transp. Res. B, Methodol.*, vol. 89, pp. 43–57, Jul. 2016.
- [17] Y. Hu, F. Chen, P. Chen, and Y. Tan, "The influence of passenger flow on the topology characteristics of urban rail transit networks," *Int. J. Mod. Phys. B*, vol. 31, no. 12, May 2017, Art. no. 1750181.
- [18] X.-M. Xiao, L.-M. Jia, and Y.-H. Wang, "Correlation between heterogeneity and vulnerability of subway networks based on passenger flow," *J. Rail Transp. Planning Manage.*, vol. 8, no. 2, pp. 145–157, Sep. 2018.
- [19] H. Cai, J. Zhu, C. Yang, W. Fan, and T. Xu, "Vulnerability analysis of metro network incorporating flow impact and capacity constraint after a disaster," *J. Urban Planning Develop.*, vol. 143, no. 2, Jun. 2017, Art. no. 04016031.
- [20] S. S. Chopra, T. Dillon, M. M. Bilec, and V. Khanna, "A network-based framework for assessing infrastructure resilience: A case study of the London metro system," *J. Roy. Soc. Interface*, vol. 13, no. 118, May 2016, Art. no. 20160113.
- [21] Z. Xu, Q. Zhang, D. Chen, and Y. He, "Characterizing the connectivity of railway networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1491–1502, Apr. 2020, doi: 10.1109/TITS.2019.2909120.
- [22] A. Huang, H. M. Zhang, W. Guan, Y. Yang, and G. Zong, "Cascading failures in weighted complex networks of transit systems based on coupled map lattices," *Math. Problems Eng.*, vol. 2015, pp. 1–16, May 2015.

- [23] Y. Shen, G. Ren, and B. Ran, "Cascading failure analysis and robustness optimization of metro networks based on coupled map lattices: A case study of Nanjing, China," *Transportation*, Nov. 2019. [Online]. Available: <https://link.springer.com/article/10.1007/s11116-019-10066-y>, doi: [10.1007/s11116-019-10066-y](https://doi.org/10.1007/s11116-019-10066-y).
- [24] Y. Dong, X. Yang, and G. Chen, "Robustness analysis of layered public transport networks due to edge overload breakdown," *Int. J. Inf. Technol. Comput. Sci.*, vol. 6, no. 3, pp. 30–37, Feb. 2014.
- [25] L. Zhang, G. Zeng, D. Li, H.-J. Huang, H. E. Stanley, and S. Havlin, "Scale-free resilience of real traffic jams," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 18, pp. 8673–8678, Apr. 2019.
- [26] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, "Author correction: The physics of spreading processes in multilayer networks," *Nature Phys.*, vol. 14, p. 523, Feb. 2018.
- [27] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 66, no. 6, Dec. 2002, Art. no. 065102.
- [28] D. J. Watts, "A simple model of global cascades on random networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 9, pp. 5766–5771, Apr. 2002.
- [29] D. S. Lee, K. I. Goh, B. Kahng, and D. Kim, "Sandpile avalanche dynamics on scale-free networks," *Phys. A, Stat. Mech. Appl.*, vol. 338, nos. 1–2, pp. 84–91, Jul. 2004.
- [30] K. Kaneko, "Overview of coupled map lattices," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 2, no. 3, pp. 279–282, Jul. 1992.
- [31] Y. Zhang, Y. Lu, G. Lu, P. Chen, and C. Ding, "Analysis of road traffic network cascade failures with coupled map lattice method," *Math. Problems Eng.*, vol. 2015, pp. 1–8, Jul. 2015.
- [32] X. Han, D. Wang, Y. Liu, and T. Guo, "Passenger flow distribution model under the interruption of urban rail transit network," in *Green Intelligent Transportation Systems* (Lecture Notes in Electrical Engineering). Singapore: Springer, Jul. 2017, pp. 211–221. [Online]. Available: <https://www.springerprofessional.de/en/passenger-flow-distribution-model-under-the-interruption-of-urba/13297016>
- [33] Y. Fan et al., "Dynamic robustness analysis for subway network with spatiotemporal characteristic of passenger flow," *IEEE Access*, vol. 8, pp. 45544–45555, 2020.
- [34] Q. Xu, B. H. Mao, and Y. Bai, "Network structure of subway passenger flows," *J. Stat. Mech., Theory Exp.*, vol. 2016, no. 3, Mar. 2016, Art. no. 033404.
- [35] X. F. Wang and J. Xu, "Cascading failures in coupled map lattices," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 70, no. 5, Nov. 2004, Art. no. 056113.
- [36] Z. J. Bao, Y. J. Cao, L. J. Ding, G. Z. Wang, and Z. X. Han, "Synergetic behavior in the cascading failure propagation of scale-free coupled map lattices," *Phys. A, Stat. Mech. Appl.*, vol. 387, no. 23, pp. 5922–5929, Oct. 2008.
- [37] C. Hong, J. Zhang, W.-B. Du, J. M. Sallan, and O. Lordan, "Cascading failures with local load redistribution in interdependent Watts–Strogatz networks," *Int. J. Mod. Phys. C*, vol. 27, no. 11, Nov. 2016, Art. no. 1650131.
- [38] A. Abdalrahman and W. Zhuang, "PEV charging infrastructure siting based on spatial-temporal traffic flow distribution," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6115–6125, Nov. 2019.
- [39] D. J. Sun and S. Guan, "Measuring vulnerability of urban metro network from line operation perspective," *Transp. Res. A, Policy Pract.*, vol. 94, pp. 348–359, Dec. 2016.
- [40] F. Chen, Q. Wu, H. Zhang, S. Li, and L. Zhao, "Relationship analysis on station capacity and passenger flow: A case of Beijing subway line 1," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 9, no. 2, pp. 93–98, Apr. 2009.
- [41] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.
- [42] Y.-K. Kwon and K.-H. Cho, "Analysis of feedback loops and robustness in network evolution based on Boolean models," *BMC Bioinf.*, vol. 8, no. 1, p. 430, Nov. 2007.
- [43] Z. Wang, C. Wang, X. Li, C. Gao, X. Li, and J. Zhu, "Evolutionary Markov dynamics for network community detection," *IEEE Trans. Knowl. Data Eng.*, early access, May 25, 2020, doi: [10.1109/TKDE.2020.2997043](https://doi.org/10.1109/TKDE.2020.2997043).
- [44] V. G. Cerf, D. D. Cowan, R. C. Mullin, and R. G. Stanton, "A lower bound on the average shortest path length in regular graphs," *Networks*, vol. 4, no. 4, pp. 335–342, 1974.
- [45] Y.-L. Shang, "Local natural connectivity in complex networks," *Chin. Phys. Lett.*, vol. 28, no. 6, Jun. 2011, Art. no. 068903.
- [46] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Phys. Rev. Lett.*, vol. 87, no. 19, Oct. 2001, Art. no. 198701.
- [47] C. Di, G. Zi-You, and Z. Jian-Feng, "Tolerance of edge cascades with coupled map lattices methods," *Chin. Phys. B*, vol. 18, no. 3, pp. 992–996, Mar. 2009.
- [48] O. Cats, H. N. Koutsopoulos, W. Burghout, and T. Toledo, "Effect of real-time transit information on dynamic path choice of passengers," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2217, no. 1, pp. 46–54, Jan. 2011.
- [49] F. Xia, J. Wang, X. Kong, D. Zhang, and Z. Wang, "Ranking station importance with human mobility patterns using subway network datasets," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2840–2852, Jul. 2020, doi: [10.1109/TITS.2019.2920962](https://doi.org/10.1109/TITS.2019.2920962).
- [50] J. Sienkiewicz and J. A. Holyst, "Statistical analysis of 22 public transport networks in poland," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 72, no. 4, Oct. 2005, Art. no. 046127.
- [51] F. Ma, F. Liu, K. Yuen, P. Lai, Q. Sun, and X. Li, "Cascading failures and vulnerability evolution in bus-metro complex bilayer networks under rainstorm weather conditions," *Int. J. Environ. Res. Public Health*, vol. 16, no. 3, p. 329, Jan. 2019.
- [52] F. Ma et al., "Exploring the robustness of public transportation for sustainable cities: A double-layered network perspective," *J. Cleaner Prod.*, vol. 265, Aug. 2020, Art. no. 121747.
- [53] D. Zhao, Z. Wang, G. Xiao, B. Gao, and L. Wang, "The robustness of interdependent networks under the interplay between cascading failures and virus propagation," *EPL Europhys. Lett.*, vol. 115, no. 5, p. 58004, Sep. 2016.
- [54] M. D. Domenico, A. Sole-Ribalta, S. Gomez, and A. Arenas, "Navigability of interconnected networks under random failures," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 23, pp. 8351–8356, May 2014.
- [55] Y. Zhang, E. Yao, J. Zhang, and K. Zheng, "Estimating metro passengers' path choices by combining self-reported revealed preference and smart card data," *Transp. Res. C, Emerg. Technol.*, vol. 92, pp. 76–89, Jul. 2018.
- [56] J. Zhao, Q. Qu, F. Zhang, C. Xu, and S. Liu, "Spatio-temporal analysis of passenger travel patterns in massive smart card data," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3135–3146, Nov. 2017.
- [57] G. Andrienko, N. Andrienko, G. Fuchs, and J. Wood, "Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 9, pp. 2120–2136, Sep. 2017.
- [58] X. Li, J. Guo, C. Gao, Z. Su, D. Bao, and Z. Zhang, "Network-based transportation system analysis: A case study in a mountain city," *Chaos, Solitons Fractals*, vol. 107, pp. 256–265, Feb. 2018.
- [59] J. Zhang, S. Wang, and X. Wang, "Comparison analysis on vulnerability of metro networks based on complex network," *Phys. A, Stat. Mech. Appl.*, vol. 496, pp. 72–78, Apr. 2018.
- [60] X.-Y. Xu, J. Liu, H.-Y. Li, and J.-Q. Hu, "Analysis of subway station capacity with the use of queueing theory," *Transp. Res. C, Emerg. Technol.*, vol. 38, pp. 28–43, Jan. 2014.
- [61] W. Guo, H. Wang, and Z. Wu, "Robustness analysis of complex networks with power decentralization strategy via flow-sensitive centrality against cascading failures," *Phys. A, Stat. Mech. Appl.*, vol. 494, pp. 186–199, Mar. 2018.



**Chao Gao** (Member, IEEE) received the Ph.D. degree in computer science from the Beijing University of Technology, in 2010. He is currently a Full Professor with the School of Artificial Intelligence, Optics, and Electronics (iOPEN), Northwestern Polytechnical University. His main research interests include data-driven complex systems modeling, complex social networks analysis, and nature-inspired computing.



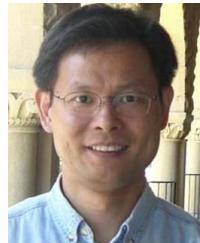
**Yi Fan** is currently pursuing the master's degree with the College of Computer and Information Science, Southwest University, Chongqing, China. His research interests include statistical analysis and complex computational problem solving.



**Shihong Jiang** is currently pursuing the master's degree with the College of Computer and Information Science, Southwest University, Chongqing, China. His research interests include statistical analysis and data-driven modeling.



**Yue Deng** is currently pursuing the bachelor's degree with the College of Computer and Information Science, Southwest University, Chongqing, China. Her research interests include data-driven modeling and statistical analysis.



**Jiming Liu** (Fellow, IEEE) received the M.Eng. and Ph.D. degrees from McGill University. He is currently the Chair Professor of Computer Science with Hong Kong Baptist University. His research interests include data analytics, data mining and machine learning, complex network analytics, data-driven complex systems modeling, and health informatics. He has served as the Editor-in-Chief for *Web Intelligence Journal (IOS)*, and the Associate Editor of *Big Data and Information Analytics (AIMS)*, *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, *IEEE TRANSACTIONS ON CYBERNETICS*, *Neuroscience and Biomedical Engineering* (Bentham), *Computational Intelligence* (Wiley), and among others.



**Xianghua Li** is currently an Associate Professor with the School of Artificial Intelligence, Optics, and Electronics (iOPEN), Northwestern Polytechnical University. Her research interests include nature-inspired computing and statistical analysis.