

Critical station identification of metro networks based on the integrated topological-functional algorithm: A case study of Chengdu

Zi-Qiang Zeng(曾自强), Sheng-Jie He(何圣洁), and Wang Tian(田旺)

Citation: Chin. Phys. B, 2025, 34 (**2**): 028902. DOI: 10.1088/1674-1056/ad9734
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Node ranking based on graph curvature and PageRank

Hongbo Qu(曲鸿博), Yu-Rong Song(宋玉蓉), Ruqi Li(李汝琦), Min Li(李敏), and Guo-Ping Jiang(蒋国平)
Chin. Phys. B, 2025, 34 (**2**): 028901. DOI: 10.1088/1674-1056/ad9a9b

Identifying influential spreaders in complex networks based on density entropy and community structure

Zhan Su(苏湛), Lei Chen(陈磊), Jun Ai(艾均), Yu-Yu Zheng(郑雨语), and Na Bie(别娜)
Chin. Phys. B, 2024, 33 (**5**): 058901. DOI: 10.1088/1674-1056/ad20d6

Identifying multiple influential spreaders in complex networks based on spectral graph theory

Dong-Xu Cui(崔东旭), Jia-Lin He(何嘉林), Zi-Fei Xiao(肖子飞), and Wei-Ping Ren(任卫平)
Chin. Phys. B, 2023, 32 (**9**): 098904. DOI: 10.1088/1674-1056/acac0b

Important edge identification in complex networks based on local and global features

Jia-Hui Song(宋家辉)
Chin. Phys. B, 2023, 32 (**9**): 098901. DOI: 10.1088/1674-1056/aca6d8

The robustness of sparse network under limited attack capacity

Xiao-Juan Wang(王小娟), Mei Song(宋梅), Lei Jin(金磊), Zhen Wang(王珍)
Chin. Phys. B, 2017, 26 (**8**): 088901. DOI: 10.1088/1674-1056/26/8/088901

Critical station identification of metro networks based on the integrated topological-functional algorithm: A case study of Chengdu

Zi-Qiang Zeng(曾自强)†, Sheng-Jie He(何圣洁), and Wang Tian(田旺)

Business School, Sichuan University, Chengdu 610065, *China*

(Received 3 October 2024; revised manuscript received 20 November 2024; accepted manuscript online 26 November 2024)

As a key mode of transportation, urban metro networks have significantly enhanced urban traffic environments and travel efficiency, making the identification of critical stations within these networks increasingly essential. This study presents a novel integrated topological-functional (ITF) algorithm for identifying critical nodes, combining topological metrics such as K-shell decomposition, node information entropy, and neighbor overlapping interaction with the functional attributes of passenger flow operations, while also considering the coupling effects between metro and bus networks. Using the Chengdu metro network as a case study, the effectiveness of the algorithm under different conditions is validated. The results indicate significant differences in passenger flow patterns between working and non-working days, leading to varying sets of critical nodes across these scenarios. Moreover, the ITF algorithm demonstrates a marked improvement in the accuracy of critical node identification compared to existing methods. This conclusion is supported by the analysis of changes in the overall network structure and relative global operational efficiency following targeted attacks on the identified critical nodes. The findings provide valuable insight into urban transportation planning, offering theoretical and practical guidance for improving metro network safety and resilience.

Keywords: critical node, metro network, topological structure, functional operation

PACS: 89.75.Hc, 89.40.Bb, 05.10.–a, 45.70.Vn

DOI: 10.1088/1674-1056/ad9734 **CSTR:** 32038.14.CPB.ad9734

1. Introduction

Urban rail transit systems are integral to the public transportation framework of large cities, playing a crucial role in driving urban economic growth.[1,2] With the continuous expansion of urban populations and the increasing demand for transportation, surface transportation systems, such as buses, face considerable pressure.[3] Due to their high capacity and operational efficiency, metro systems are indispensable in alleviating urban traffic congestion and enhancing travel efficiency.[4,5] However, the rapid expansion of metro lines and stations has led to a substantial increase in the complexity of metro networks.[6] This growing complexity, coupled with fluctuating passenger flow demands, exacerbates the risk of operational disruptions.[7] For instance, on July 20, 2021, a severe rainstorm in Zhengzhou caused extensive flooding at Wulongkou station on metro line 5, significantly damaging the metro network infrastructure and impairing its functionality.[8] Thus, to enhance the resilience and reliability of metro systems, it is critical to identify and prioritize key stations. These critical nodes are essential in maintaining the network's integrity and functionality during disruptions,[9] while also serving as focal points for future development and strategic planning, thereby ensuring the sustainable and reliable operation of metro systems.[10]

The identification of critical nodes in networked systems has evolved considerably over time.[11] Initial methods were based exclusively on network topological structures, utilizing both local and global centrality measures. By calculating node centrality indices and ranking nodes, the critical nodes could be identified. These approaches include local information-based measures such as degree centrality[12] and eigenvector centrality,[13] as well as global information-based measures such as betweenness centrality,[14] closeness centrality[12] and K-shell decomposition.[15] As research progressed, it became evident that a node's significance is influenced not only by its intrinsic properties and its immediate neighbors but also by its overall position within the network.[16,17] To improve the accuracy of local indices while mitigating the computational complexity of global indices, hybrid methods that integrate both local and global information have been employed in various fields.[18] For example, Engsig *et al.* proposed the DomiRank centrality metric,[19] which integrates local and global topological information through a tunable parameter to identify fragile neighborhoods in networks, emphasizing the dominance of specific nodes. Based on the three dimensions of self-centrality, local centrality and global centrality of network nodes, Ai *et al.* proposed the SLGC method[20] for identifying influential nodes in a network. In addition, Huang *et al.* introduced the adjacency information entropy method, which identifies critical nodes by considering both the node's position and the positions of its neighboring nodes within the network.[21] Duan *et al.* developed a network connectivity model based

†Corresponding author. E-mail: zengziqiang@scu.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

on a random connection matrix and minimal low-dimensional structure, employing information entropy to rank the significance of the network nodes.[22]

However, in transportation networks, while the topological structure is relatively stable, the role of these networks in satisfying urban travel demands is highly complex.[23] Particularly in metro systems, the distribution of passenger flow around stations varies significantly across both spatial and temporal dimensions, with passengers having diverse route choices from origin to destination.[24] From this perspective, identifying critical nodes based solely on topological structures,[25] without considering the network's functional characteristics as a transportation system, may fail to accurately capture real-world conditions. Consequently, hybrid approaches that combine topological features with functional attributes have been developed to address these limitations.[26] In the context of urban rail transit systems, researchers have increasingly focused on methods that incorporate passenger flow and network functionality. Xu *et al.* developed a dynamic travel network model based on the tripartite graph concept and proposed a node centrality ranking method that integrates origin-destination flow.[27] Jing *et al.* introduced a probability metric based on path redundancy to identify critical nodes in urban rail networks, proposing the probability degree method to quantify node importance.[28] Xia *et al.* examined subway passenger flow patterns and developed the SIRank algorithm,[29] an improved PageRank model, to identify critical nodes by incorporating human mobility data. Zhang *et al.* proposed the comprehensive vote-rank algorithm,[30] which integrates mutual information and clustering techniques to analyze passenger flow patterns and identify critical stations under various scenarios. Wang *et al.* further extended this work by proposing the modified weighted K-shell model, which identifies critical nodes in multimodal transportation networks by considering both traffic demand and the interactions between multi-order neighbor nodes.[31]

In summary, the identification of critical nodes in transportation networks has progressed from purely topological approaches to more integrated methods that account for both structural and functional characteristics. This holistic perspective is crucial for ensuring the resilience and reliability of urban rail transit systems, especially given the increasing complexity and rising passenger demands. Aligned with this concept, this paper first evaluates the topological significance of metro stations by examining their position within the overall network and the influence of neighboring stations. In terms of functional importance, the study assesses station significance by analyzing passenger flow intensity and incorporates the coupling effect of the bus transit system on metro passenger flow. By integrating these structural and functional dimensions, a comprehensive framework for identifying key nodes in the urban transportation network is constructed. The identified critical nodes are then subjected to continuous disruption simulations to examine the metro network's reliability in terms of structural integrity and functional operational capacity. Finally, this method is applied to the Chengdu rail transit system, and practical recommendations are proposed for future metro development. The primary contributions and innovations of this study are as follows:

(i) Establishing an integrated topological-functional algorithm for critical node identification. This approach comprehensively incorporates both network structure and functionality, significantly enhancing the accuracy of identifying critical nodes.

(ii) Proposing a method for evaluating the functional operational importance of metro networks. This method defines the functional importance of nodes based on passenger flow through metro stations and the weight proportion of their corresponding paths, effectively capturing the dynamic functional attributes of the metro network.

(iii) Expanding the understanding of the interaction effects in multimodal traffic. In exploring the functional importance of metro systems, this study considers the coupling and interaction effects of the public transportation system, integrating multimodal traffic evaluation into metro network analysis. This broadens the scope for future research in this field.

(iv) Defining a relative global operational efficiency evaluation index. A new index is introduced to assess the operational level of metro systems, offering a more precise and effective representation of the transportation system's capacity.

The rest of this paper is organized as follows. Section 2 outlines the problem statement concerning the identification of critical nodes in urban metro networks. Section 3 proposes the integrated topological-functional algorithm and details the network performance evaluation metrics. In Section 4, a case study of the Chengdu metro network is presented to validate the effectiveness of the proposed algorithm. Finally, Section 5 concludes with a summary of the findings and provides recommendations for future research directions.

2. Problem statement

Current research on identifying critical nodes in metro networks predominantly relies on topological metrics such as degree centrality, betweenness centrality and K-shell decomposition. While these approaches offer valuable insight into the structural importance of nodes, they often fail to capture the functional dynamics of metro systems, particularly passenger flow variations and the interactions between metro stations and other transportation modes. This oversight limits the real-world applicability of these methods, as metro networks are not merely static structures but dynamic systems characterized

by fluctuating demand patterns and intricate intermodal interactions. Therefore, developing a new model that can evaluate both the static network topology and the dynamic functional operations is a crucial step towards accurately identifying key nodes in metro networks.

This study seeks to address these limitations by introducing an integrated topological-functional (ITF) algorithm, which synthesizes both structural and functional dimensions of metro networks. The primary research question guiding this study is: How can critical nodes in a metro network be identified more accurately by integrating topological properties with real-time operational data? By answering this question, the research aims to develop a more comprehensive and realistic framework for critical node identification, ultimately enhancing the resilience and operational efficiency of metro systems. This investigation is crucial for future urban transportation planning and for strengthening the robustness of metro networks in the face of potential disruptions.

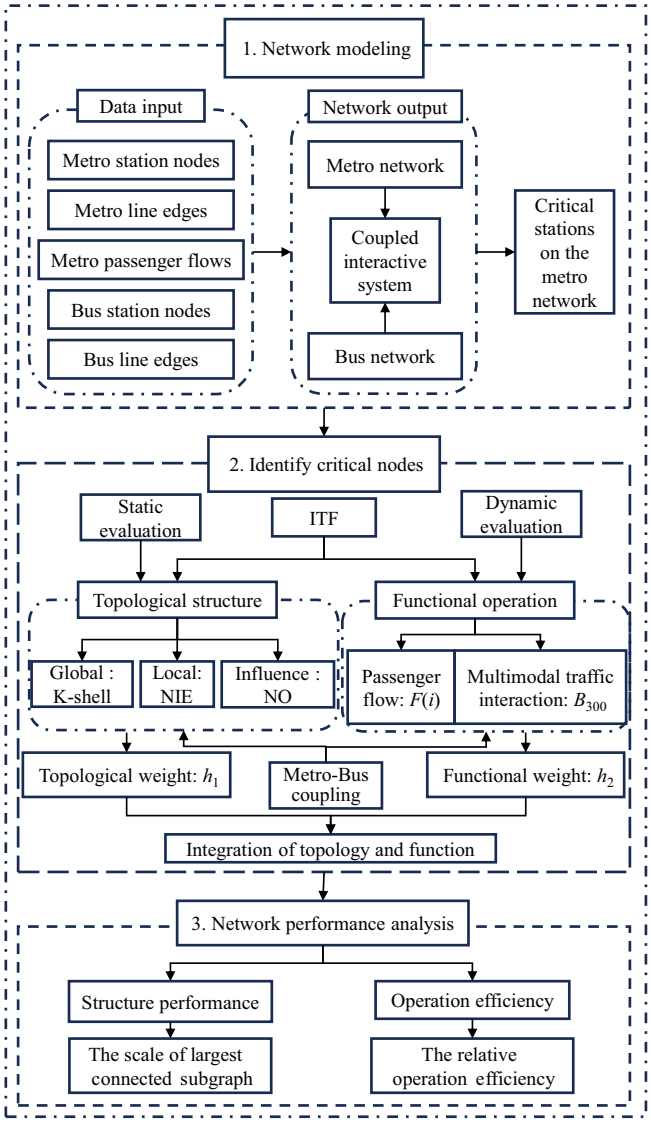


Fig. 1. The framework of the study.

The research framework is divided into three main steps, as illustrated in Fig. 1. First, an urban metro network model is constructed using structural data from metro stations and lines, alongside operational traffic data and passenger flow characteristics. In the second step, a comprehensive algorithm is developed to identify critical nodes by assessing both the topological significance and functional operational capacity of metro network nodes, utilizing corresponding evaluation metrics. Finally, in the third step, the Chengdu metro network is used as a case study to analyze the impact of removing critical nodes on the network's structural integrity and functionality. The performance of the proposed algorithm is then compared with traditional methods for critical node identification to validate the model's effectiveness.

3. Methodology

3.1. Metro network model

Previous approaches for constructing metro networks have typically employed methods such as space L, space P, space R and space C.[32] The topological structure of metro networks in different cities varies depending on the method used.[33] In the space L model, metro stations are represented as nodes, and the intervals between two adjacent stations are depicted as edges in the corresponding topology. Compared to the other three methods, space L more accurately captures the overall structure of metro networks. Therefore, in this paper, we adopt the space L model to construct the metro network since it is better suited for representing the real-world metro system. A metro network can be represented as an undirected network $G = (V, E, F)$, where $V = \{v_i | i \in (1, n)\}$ and $E = \{e_{ij} | i \neq j\}$ are the node and edge sets, respectively, which correspond to the stations and tracks between stations in the metro network. When nodes v_i and v_j are connected, $e_{ij} = 1$, otherwise $e_{ij} = 0$. As a vital form of urban transportation, the passenger flow served by the metro is a key indicator of its operational efficiency. To maintain consistency, we define the passenger flow intensity set of a metro network as $F = \{f_i | i \in (1, n)\}$, where f_i represents the number of passengers served by a given station v_i. Based on the constructed metro network, we introduce several widely used methods for identifying node importance, including degree centrality (DC), betweenness centrality (BC), node information entropy (NIE) and K-shell decomposition, which are described as follows.

(i) Degree centrality (DC) is the simplest measure of a node's influence within a network. A node with a higher degree is considered more influential because it can extend its impact to a greater number of neighboring nodes. The normalized degree centrality is defined as

$$\mathrm{DC}(i) = \frac{k_i}{N-1}, \quad \forall i, \tag{1}$$

where N is the number of nodes in G, k_i is the degree of a node v_i, which is defined as the number of directly connected neighbors of node v_i.

(ii) Betweenness centrality (BC) shows the hub characteristics of the node in the network, which is described as the number of shortest paths that pass through a node. The larger the number of shortest paths that pass through the node, the larger the betweenness of the node is, the more important the node in the network is. The normalized betweenness centrality is defined as

$$\mathrm{BC}(i) = \sum_{s\neq i\neq t} \frac{n_{st}^{i}}{n_{st}}, \quad \forall i, \tag{2}$$

where n_{st} represents the total number of shortest paths between node v_s and node v_t, and n_{st}^{i} represents the number of those paths that pass through node v_i. However, as a path-based centrality measure, betweenness centrality emphasizes finding the shortest paths within the network. This may not be consistent with the actual operations of metro systems, where the same route can carry varying passenger volumes at different times, corresponding to different transportation functions.[34] Relying solely on the static network topology, while overlooking the dynamic changes in the functional roles of routes, suggests that betweenness centrality may be less suitable for evaluating metro networks.

(iii) Node degree-based information entropy In information theory, the concept of entropy, as proposed by Shannon,[35] is used to quantify the uncertainty of information. The greater the entropy, the more uncertain or unpredictable the information. To evaluate the influence of neighboring nodes on a given node, the concept of entropy is extended to complex networks, referred to as node information entropy (NIE). Node degree-based information entropy captures the local properties of nodes; the higher the entropy, the more influential the node.[36] Assuming that the degree of node v_i is k_i, the information content contained in node v_i is $I_i = k_i/\sum_{j=1}^{n} k_j$, where n is the number of nodes in G, so the node information entropy is defined as

$$\mathrm{NIE}(i) = -\sum_{j\in\Gamma(i)} I_j \cdot \ln I_j, \quad \forall i, \tag{3}$$

where $j \in \Gamma(i)$ is the set of neighbors of node v_i.

(iv) K-shell decomposition The K-shell decomposition method is used to quantify the importance of nodes by categorizing them into different shells.[37] The algorithm begins by removing all nodes with a degree of 1, iteratively repeating the process until no such nodes remain. All the nodes and edges removed in the first iteration form the 1-shell. This procedure is repeated to obtain the 2-shell, 3-shell, and so on. Ultimately, each node is assigned to a specific shell, and all nodes within the same shell have an identical core value.[38] The core value reflects a node's structural importance and spreading capability. Nodes in higher shells are considered to have a greater influence within the network.

3.2. Integration of the topological-functional algorithm

The aforementioned methods for identifying node importance primarily focus on the topological characteristics of the network. Metrics, such as betweenness centrality and the K-shell decomposition assess node importance from a global network perspective. However, relying solely on betweenness centrality, which identifies critical nodes based on the shortest paths between nodes in the network, falls short of reflecting the dynamic changes in passenger flow along routes within the metro system as a transportation network due to the fixed route structure of these systems. The K-shell method, while stratifying nodes based on their position within the overall network, does not effectively distinguish between the importance of nodes within the same layer. In contrast, degree centrality and information entropy focus on node importance from a local network perspective. Degree centrality is useful for identifying highly connected nodes, but in metro networks, only a few nodes exhibit high degrees, and this metric does not accurately capture the importance of most nodes. Information entropy, which takes into account both the degree of a node and the degrees of its neighboring nodes, provides a more nuanced assessment of critical nodes. Nevertheless, a comprehensive evaluation of a node's importance requires considering both global and local properties, as well as the mutual influence between nodes. For urban metro networks, which serve not only a transportation function but also possess significant social attributes, topological structure alone is insufficient to capture the full importance of a node. The functional characteristics, particularly the dynamics of passenger flow, are essential factors in determining the criticality of a node within the network.

In summary, this study proposes a comprehensive algorithm for identifying critical stations in urban metro networks by integrating topological features with passenger flow dynamics. This algorithm combines traditional structural analysis with the functional operational role of the metro system, while also accounting for the coupling and interaction effects of multimodal transportation. The specific details of the algorithm are as follows:

Definition 1: Neighbor overlapping interaction In an urban metro network, the interaction between adjacent nodes plays a significant role in the transmission of information between nodes. For instance, if a metro station experiences a failure, the interaction with neighboring stations allows passengers to reroute through adjacent stations, thereby mitigating the transportation disruption. Therefore, when evaluating the importance of nodes in a metro network, the interaction between nodes is irreplaceable. This paper proposes the neighbor

overlap ratio R_{st} to characterize the interaction between nodes:

$$R_{st} = \frac{C_{st}}{k_s - 1 + k_t - 1 - C_{st}}, \tag{4}$$

where k_s and k_t represent the degrees of nodes v_s and v_t, and C_{st} denotes the number of common neighbors between nodes v_s and v_t. The more common neighbors two nodes share, the more similar the nodes are, and thus, the closer their relationship. Each node typically has multiple neighboring nodes, and these neighbors influence the interaction between nodes. Therefore, the neighbor overlapping interaction of node v_i denoted as $\mathrm{NO}(i)$ is

$$\mathrm{NO}(i) = \sum_{s,t \in \Gamma(i)} R_{st}, \ \ \forall i, \tag{5}$$

where $\Gamma(i)$ represents the set of neighbors of node v_i.

Definition 2: Topological importance In this paper, the topological importance of a node is measured by considering three aspects: the global properties of the node within the network, its local properties, and the interactions between nodes. Specifically, the K-shell value, node information entropy and neighbor overlapping interaction are selected to represent these aspects. The topological importance of a node in the metro network is evaluated as the product of these three metrics. The topological importance of a network node is represented as

$$T(i) = K(i) \cdot \mathrm{NIE}(i) \cdot \mathrm{NO}(i), \ \ \forall i, \tag{6}$$

where $K(i)$ represents the K-shell value of the node, $\mathrm{NIE}(i)$ represents the degree-based information entropy of the node and $\mathrm{NO}(i)$ represents the neighbor overlapping interaction of the node. These three indicators characterize the structural importance of a node from three different perspectives: the global properties of the node within the entire network, the local properties of the node, and the mutual influence between nodes.

Definition 3: Functional operational importance Based on the transportation functionality of the metro network, this paper selects the passenger flow through metro nodes as the key indicator for measuring the functional importance of nodes. In addition, considering the various transportation modes available to urban residents,[39] the number of bus stations within a 300-meter radius of a metro station is taken into account. The more bus stations within this area, the greater the increase in metro passenger flow due to bus transportation, indicating a higher importance of the metro station.[40] The functional importance of a network node in terms of passenger flow is represented as

$$F(i) = \left(1 + B^i_{300}\right) \cdot f_i \cdot \sum_{s \neq i \neq t} \frac{\sum_{j \in L^i_{st}} f_j}{\sum_{k \in L_{st}} f_k}, \ \ \forall i, \tag{7}$$

where L_{st} represents the set of shortest paths between nodes v_s and v_t, and L^i_{st} represents the subset of those shortest paths that pass through node v_i. In addition, f_i represents the passenger flow intensity of node v_i, f_j represents the passenger flow intensity through the shortest path in the set of shortest paths L^i_{st} that pass node v_j, f_k represents the passenger flow intensity through the shortest path in the set of shortest paths L_{st} that pass node v_k, while B^i_{300} denotes the number of bus stations within a 300-meter radius of a metro station.

To eliminate the dimensional differences between the topological importance and passenger flow importance of a node, both measures are standardized as follows:

$$T'(i) = \frac{T(i)}{\max\left(T(i)\right)}, \ \ \forall i, \tag{8}$$

$$F'(i) = \frac{F(i)}{\max\left(F(i)\right)}, \ \ \forall i. \tag{9}$$

The final node importance is obtained by assigning weights to both aspects of importance

$$\mathrm{ITF}(i) = h_1 \cdot T'(i) + h_2 \cdot F'(i), \ \ \forall i, \tag{10}$$

where h_1 and h_2 represent the degrees of focus on structural stability and functional operational ability of the metro network, and they satisfy $h_1 + h_2 = 1$. Typically, when $h_1 > h_2$, this indicates a focus on the topological structure of the metro network, which is often a consideration in future urban planning and development. On the other hand, when $h_1 < h_2$, the emphasis is on the metro system's role in alleviating current urban traffic congestion, serving as a measure of meeting the present needs of city residents. When $h_1 = h_2$, this reflects equal importance given to both the structural and functional aspects of the metro network.

3.3. Network evaluation index

When certain edges or nodes are removed from a network, the structure and functionality of the system can experience significant alterations, potentially leading to network failure. These components represent the critical nodes and edges within the network. In the context of an urban metro network, structural stability reflects the system's capacity to maintain normal operations, while functional reliability indicates the network's ability to meet passenger traffic demands.[41] To assess the Chengdu metro network, we propose two key evaluation metrics, one focused on the network's topological structure and the other on passenger service levels. By deliberately targeting critical nodes and observing the resulting changes in these two metrics, we can validate the effectiveness of the critical node identification method proposed in this study. The two metrics are as follows:

(i) Scale of largest connected subgraph The robustness of urban metro networks refers to the network's ability to

maintain its function when its components are subjected to attacks. The largest connected subgraph[42] is used to describe the local connectivity of urban metro networks and is adopted as a metric to assess the network's robustness under attack. It is defined as follows:

$$S = \frac{n'}{n}, \tag{11}$$

where n' represents the number of nodes in the largest connected subgraph after the network has been attacked, and n is the total number of nodes in the initial network.

(ii) Relative global operational efficiency Global network efficiency describes the overall connectivity of an urban metro network.[43] Inspired by this, we propose the global operational efficiency to represent the operational level of passenger flow within the metro network:

$$O = \frac{1}{n(n-1)} \cdot \sum_{s \neq t} \frac{f_s + f_t}{d_{st}}, \tag{12}$$

where d_{st} represents the shortest distance between nodes v_s and v_t, and f_s and f_t represent the passenger flow intensities of nodes v_s and v_t, respectively. To eliminate the data dimensional differences caused by variations in node passenger flow, we define the relative global operational efficiency of the network as the ratio of the network's global operational efficiency after attacking key nodes to its initial global operational efficiency, which is shown as

$$\mathrm{OE} = \frac{O'}{O}, \tag{13}$$

where O represents the initial global operational efficiency and O' represents the global operational efficiency after the network has been attacked. This results in the relative global operational efficiency, which effectively captures the operational level of the metro network.

4. Results

In this study, the Chengdu metro system is selected as the case study. Chengdu, the largest city in southwestern China, has seen rapid urban transit development. As of July 2024, the Chengdu metro system operates 293 stations across 13 lines, with a cumulative passenger volume surpassing 10 billion trips. Concurrently, the Chengdu bus system, which interacts with and complements the metro network, has established 683 bus routes and 13132 bus stations at the same time. Figure 2 presents a topological visualization of the Chengdu metro-bus network in 2024, illustrating the combined efforts of these two extensive and complex public transportation systems in supporting the daily travel needs of Chengdu's residents.

To better capture the interaction between the metro and bus systems in meeting passenger transportation demands, and to account for the contribution of bus stop passenger flow to metro station ridership, we define a 300-meter radius around each metro station as a geographic buffer zone (as depicted in Fig. 3). Within this zone, a higher number of bus stations correlates with a greater positive effect on metro station passenger flow since the distance aligns with most passengers' preferred bus-to-metro transfer distance.[44] Conversely, beyond this buffer zone, bus operations tend to reduce metro ridership, as passengers are more likely to choose buses for their entire journey.

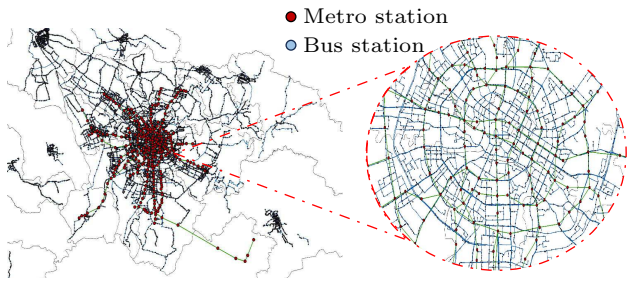


Fig. 2. The Chengdu metro–bus network.

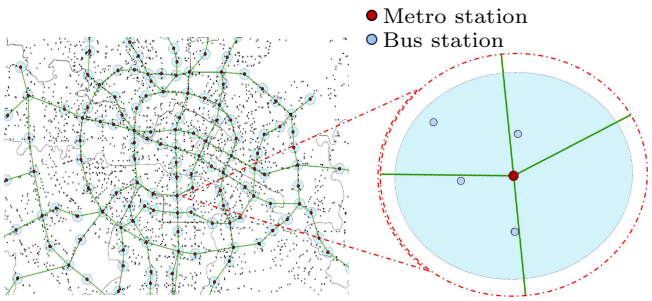


Fig. 3. Metro station buffer.

To obtain a more accurate and comprehensive understanding of the Chengdu metro network, a network model consisting of 293 nodes and 336 edges was constructed using Python, as depicted in Fig. 4. The topological characteristics of the network, including the node degree distribution and the shortest path length distribution, were subsequently analyzed (see Fig. 5). The basic features are as follows. (i) Transfer and non-transfer stations: Of the 293 metro stations, 68 are transfer stations, while the remaining 225 are non-transfer stations. Transfer stations are primarily concentrated along lines 1, 7, 6, 3, 18 and 19, indicating that these lines serve as major transit routes traversing significant parts of the city. This highlights the crucial role these lines play in facilitating daily commutes and activities for Chengdu's residents. (ii) Node degree distribution: As shown in Fig. 5(a), most metro stations exhibit a node degree of 2, accounting for approximately 76.8% of the total nodes in the network. The maximum node degree is 6, observed at the "Sichuan provincial people's hospital" station, which serves as a transfer hub for lines 2, 4 and 5. Meanwhile, a fitting analysis of the node degree distribution was performed using Python, and the results are shown in Fig. 5(b), revealing that the node degree distribution of the

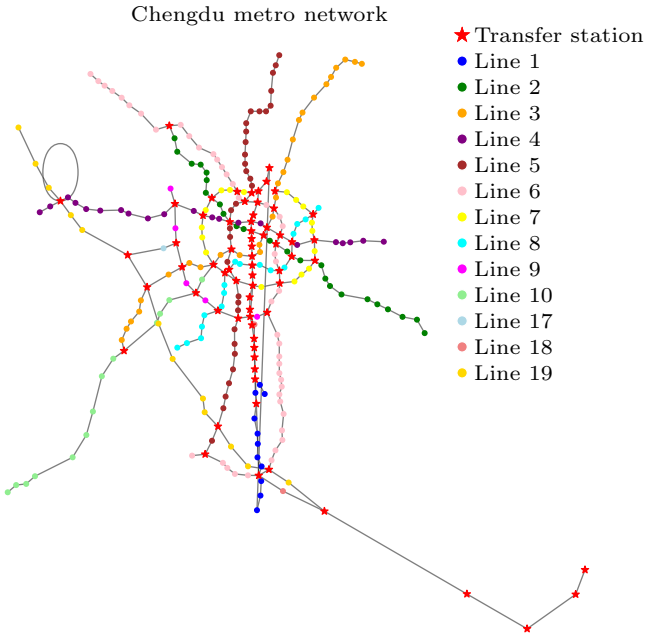


Fig. 4. Topological structure of Chengdu metro network.

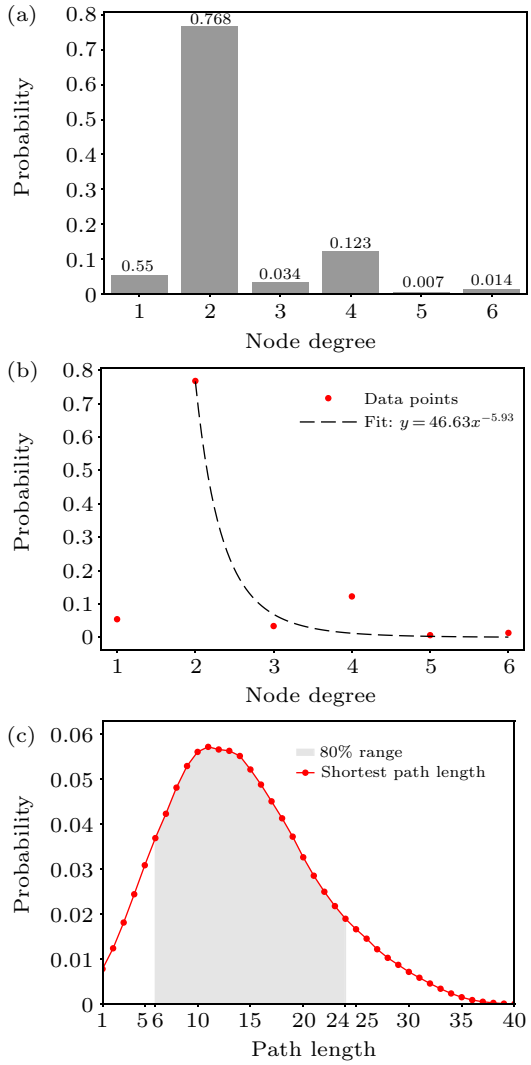


Fig. 5. The network characteristic analysis of the Chengdu metro network. (a) Node degree distribution; (b) node degree distribution curve; (c) shortest path length distribution.

Chengdu metro network can be approximated by the function $y = 46.63x^{-5.93}$. (iii) Shortest path length distribution: From the perspective of the entire metro network topology, when any two nodes are selected as the origin and destination for metro travel, the journey typically passes through multiple nodes and edges in the network graph. Therefore, the path length for this trip is represented by the number of edges traversed along the entire route. Based on the probability distribution of the shortest path lengths within the Chengdu metro network as depicted in Fig. 5(c), and excluding extremely short or long travel demands, 80% of the shortest path lengths fall within the range of 6–24. The average shortest path length is 14.16, meaning that, on average, a passenger needs to pass through approximately 14 stations to travel between any two points within the Chengdu metro network.

4.1. Characteristic analysis of passenger flow

In general, urban metro networks possess a fixed topological structure, while passenger flow is dynamic and varies over time, influencing the operational characteristics of individual stations. As a result, it is crucial to analyze the fluctuations in passenger flow and evaluate their impact on station performance to more accurately identify critical stations within the Chengdu metro network.

For this study, we collected operational data from the Chengdu metro between May 1 and May 31, 2024. The data set includes the daily total number of passengers entering and exiting each station. To analyze the relationships in passenger flow across different dates, we calculated the Pearson correlation coefficients based on the daily total passenger flows of all stations combined. The Pearson correlation coefficients represent the degree of similarity in passenger flow patterns between different days. A heatmap was generated to visualize these correlations, where each cell indicates the correlation coefficient between two specific dates, with diagonal items representing the self-correlation of a given day. Figure 6 illustrates this heatmap, depicting the Pearson correlation coefficients of passenger flows between any two days in May 2024 for the Chengdu metro network.

This figure clearly illustrates the significant variations in passenger flow between weekdays and non-working days in the Chengdu metro system. Specifically, the daily passenger flow remained relatively stable on weekdays throughout May, whereas it decreased on non-working days such as weekends and holidays. Moreover, due to the rescheduling of the Labor Day holiday, May 11, which was originally a Saturday, was designated as a working day. A comparison of May 11 with May 12, a non-working day, reveals notable differences in passenger flow between the two days. These observations highlight the importance of accounting for the spatiotemporal variations in passenger flow when identifying critical stations in urban metro networks.

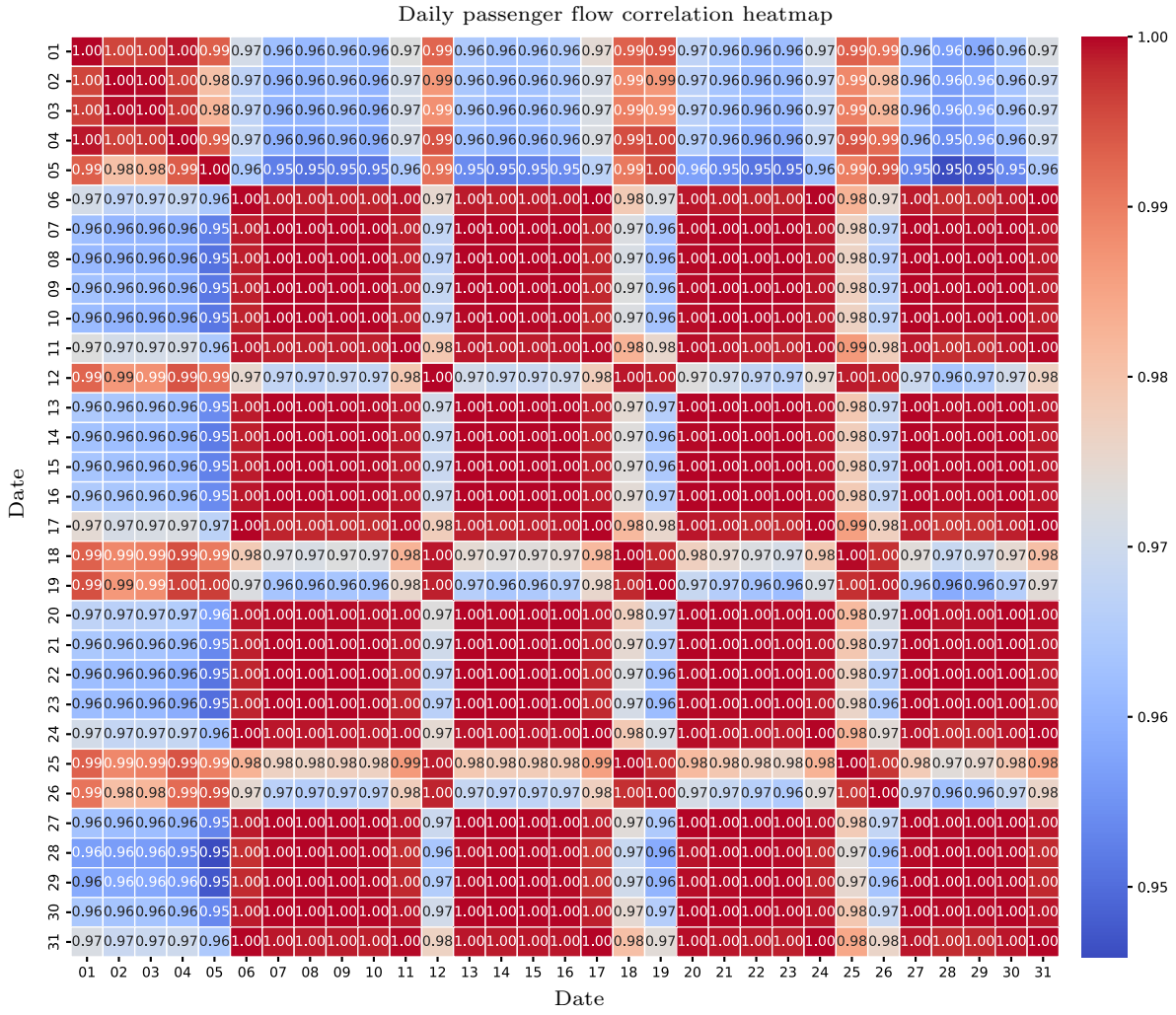


Fig. 6. Daily passenger flow correlation heatmap in Chengdu metro.

4.2. Identification results based on the ITF algorithm

Based on the topology and passenger flow data of the Chengdu metro network, the proposed integrated topological-functional (ITF) algorithm was applied to identify and rank the critical stations. To emphasize the importance of the metro network structure for future urban development while addressing the necessity of meeting current travel demands, the weight parameters for the network structure and function were set at $h_1 = 0.6$ and $h_2 = 0.4$, respectively. On this basis, Table 1 presents the top 10 critical stations for both working and non-working days. From the data in Table 1, it can be observed that 60% of the top 10 critical stations remain the same between working days and non-working days, although their rankings differ. This finding underscores the importance of considering passenger flow variations when identifying critical stations. In addition, when examining the number of bus stations within a 300-meter radius of the top six common critical metro stations, the "Chunxi road" station experiences the greatest impact on metro passenger flow due to the travel demand from passengers at six nearby bus stations. In contrast,

Table 1. Critical station identification results.

Rank	Working days	B_{300}	Non-working days	B_{300}
1	Chengdu east railway station	3	Chengdu east railway station	3
2	Chunxi road	6	Chunxi road	6
3	Chengdu west railway station	4	Tianfu square	3
4	Xipu railway station	2	Xipu railway station	2
5	South railway station	3	South railway station	3
6	Tianfu square	3	Chengdu west railway station	4
7	Century city	4	Western China int'I expo city	2
8	Third Tianfu street	6	Kuanzhaixiangzi alleys	2
9	Sichuan provincial people's hospital	2	Gaoshengqiao	4
10	Haichang road	4	Chengdu zoo	3

the remaining critical stations have around 2–4 bus stations in their vicinity. This highlights the significant coupling effect between the metro network system and the bus transportation system.

Similarly, under this condition, figure 7 illustrates the geographical distribution of the critical stations in the Chengdu metro network. Orange circles denote critical stations on working days, green squares represent critical stations on non-working days, and red stars indicate stations that are critical on both working and non-working days. Notably, 90% of the critical stations identified on working days and 80% of those on non-working days are transfer stations. These results highlight the importance for the Chengdu metro to prioritize the protection and efficient management of these key stations during daily operation. In addition, the analysis reveals that critical stations on working days are predominantly concentrated in the city center, while those on non-working days are distributed across both the city center and the outskirts. This pattern further reflects Chengdu's dual role as a city well-suited for both living and working.

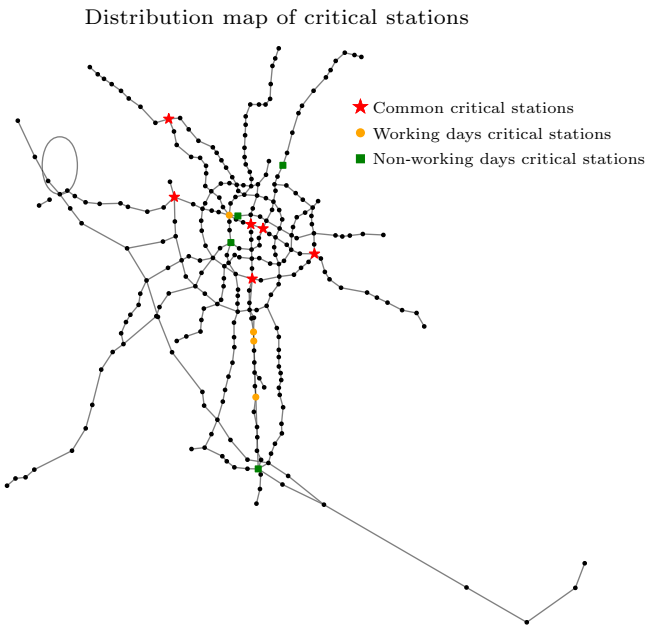


Fig. 7. Distribution map of critical stations.

4.3. Metro network performance analysis

In this paper, we identify the critical stations within the Chengdu metro network by considering both its topological structure and operational functionality. To validate the superiority of the proposed ITF method over other approaches, deliberate attacks were simulated to observe changes in the network's structure and functionality under attack. Specifically, the nodes in the metro network were ranked based on their importance as identified by various critical node identification methods, including ITF, K-shell, degree centrality, betweenness centrality, information entropy and passenger flow. Different attack intensities were then applied by selectively disrupting nodes in order of their importance, removing varying proportions of the most critical nodes. The impact of these attacks was evaluated using two performance indicators: the scale of the largest connected subgraph, which reflects the network's structural integrity and the relative global operational efficiency, which quantifies the network's functionality. Under the same attack intensity, the greater the structural disruption and functional damage, the more important the attacked nodes are, thereby demonstrating the effectiveness of the corresponding critical node identification method. Through these experiments, the ITF method was shown to outperform other methods in identifying nodes critical to the Chengdu metro network.

4.3.1. Structure performance analysis

The robustness of urban metro networks refers to their ability to maintain functionality when key components are compromised or subjected to attacks. The largest connected subgraph is utilized to describe the local connectivity of urban rail networks and serves as a metric to assess the network's robustness under these conditions.

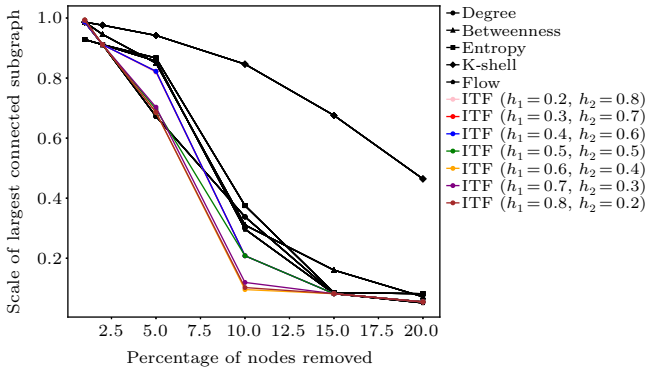


Fig. 8. The largest connected subgraph of Chengdu metro network on working days.

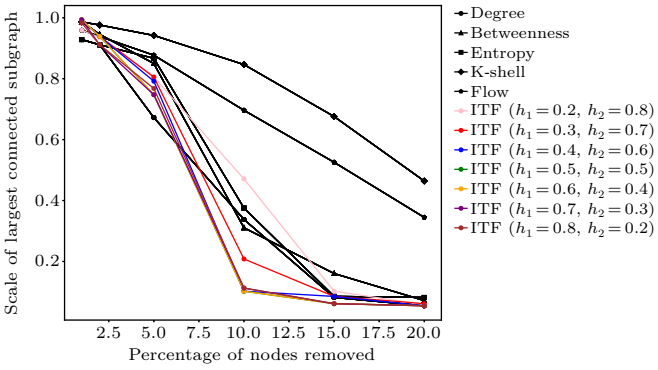


Fig. 9. The largest connected subgraph of Chengdu metro network on non-working days.

Figures 8 and 9 depict the changes in the size of the largest connected subgraph in the Chengdu metro network under targeted attacks on working days and non-working days, respectively. It is observed that as the proportion of attacked

critical nodes increases, the scale of the largest connected subgraph progressively decreases.

On working days, when the proportion of removed nodes is less than 5.1%, deliberate attacks on the critical nodes identified by degree centrality cause the greatest damage to the Chengdu metro network, regardless of the parameter settings h_1 and h_2. However, when the proportion of removed nodes exceeds 7.5%, the ITF method proposed in this paper consistently outperforms degree centrality in identifying critical nodes, regardless of the parameter settings. Attacks on these ITF-identified nodes lead to a faster reduction in the size of the largest connected subgraph. Once the proportion of removed nodes exceeds 10%, the scale of the largest connected subgraph drops significantly, indicating severe structural damage to the entire network. Interestingly, when the proportion of removed nodes lies between 5.1% and 7.5%, the accuracy of these two methods in identifying critical nodes within the metro network is influenced by the parameter settings. Since h_1 gradually increases from 0.2 to 0.8, the critical attack intensity at which the ITF method surpasses degree centrality gradually decreases. This highlights that as the importance of the network's topological structure grows, the ITF method, which integrates both network topology and passenger flow operational functionality, becomes increasingly accurate in identifying critical nodes. In contrast, traditional methods based solely on topological metrics, such as K-shell, betweenness centrality and information entropy, perform less effectively.

On non-working days, the size of the largest connected subgraph in the Chengdu metro network displays similar characteristics under increasing proportions of attacked critical nodes. The ITF method again demonstrates higher accuracy in identifying critical nodes. However, methods that rely solely on passenger flow data are less effective in this context. This is because, in a megacity like Chengdu, passenger flow demand on working days is strongly correlated with the metro network structure. On non-working days, however, the variability in metro travel demand due to residents' and tourists' diverse leisure activities introduces a degree of randomness. As a result, passenger flow-based methods fail to accurately identify the true critical nodes in the metro network on non-working days.

4.3.2. Operational efficiency analysis

The global efficiency of a network is a key metric that characterizes the overall connectivity of an urban metro system, indicating whether the metro network can maintain normal operation. Building on this concept, this paper introduces a new metric, the relative global operational efficiency, designed to evaluate the overall operational performance of passenger flow operation within urban metro networks.

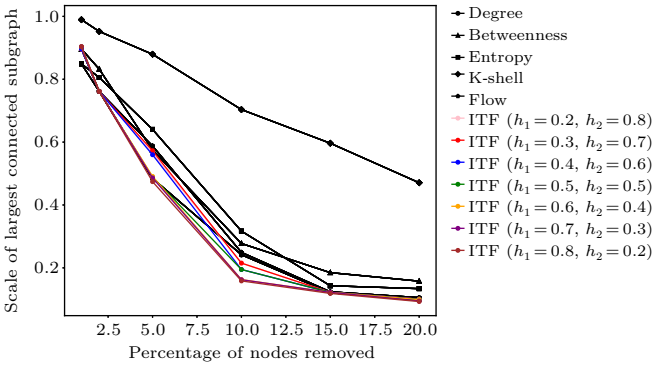


Fig. 10. The passenger flow operational efficiency of Chengdu metro network on working days.

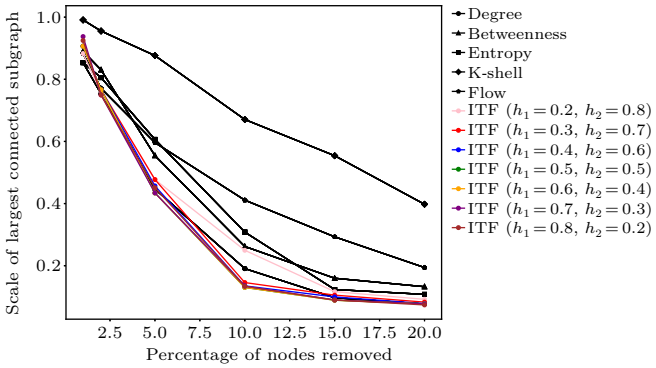


Fig. 11. The passenger flow operational efficiency of Chengdu metro network on non-working days.

Figures 10 and 11 depict the changes in the relative global operational efficiency of the Chengdu metro network when subjected to attacks on critical nodes during both working days and non-working days. As expected, the relative global operational efficiency decreases progressively as the proportion of attacked critical nodes increases.

On working days, when the proportion of removed nodes is less than 5%, the method based on degree centrality identifies critical nodes more effectively, regardless of the parameter settings h_1 and h_2. However, when the proportion of removed nodes exceeds 8.5%, attacks targeting the critical nodes identified by the ITF method cause the greatest damage to the Chengdu metro network, regardless of the parameter settings. In this scenario, the passenger operational efficiency of the metro declines sharply shortly after the node failures. When the proportion of removed nodes exceeds 10%, the relative global operational efficiency of the network has already dropped significantly, leading to severe losses in the metro network's operational functionality. Similarly, when the proportion of removed nodes lies between 5% and 8.5%, the accuracy of these two methods in identifying critical nodes is influenced by the parameter settings. As h_1 gradually increases from 0.2 to 0.8, the critical attack intensity at which the ITF method outperforms degree centrality gradually decreases. This indicates that as the topological importance of the Chengdu metro network increases, the ITF method becomes increasingly accurate in identifying critical nodes. By integrating both net-

work topology and passenger operational functionality, the ITF method effectively balances these two aspects. In contrast, traditional methods based purely on network topology, such as K-shell, betweenness centrality and information entropy, demonstrate lower effectiveness in accurately identifying critical nodes within the Chengdu metro network.

On non-working days, the relative global operational efficiency of the Chengdu metro network shows a similar pattern, with a gradual decrease as the proportion of attacked critical nodes increases. Once again, the ITF method exhibits greater accuracy in identifying critical nodes. The primary difference lies in the critical threshold of removed nodes, which decreases from 8.5% on working days to approximately 6.7% on non-working days. Furthermore, the method based solely on passenger flow performs poorly, lagging significantly behind methods such as betweenness centrality and information entropy in identifying critical nodes.

5. Conclusions and future research

With the rapid development of urban metro networks and the swift expansion of transportation infrastructure, developing a highly accurate, reliable and cost-effective method for identifying critical network nodes is of both theoretical and practical significance. Building on traditional methods that identify critical nodes based on network structural characteristics, this paper proposes the integrated topological-functional algorithm, which combines network topology and operational functionality to assess the importance of nodes in urban metro networks. The ITF method is compared with traditional node importance identification methods, including degree centrality, betweenness centrality, K-shell and node degree-based information entropy. Subsequently, deliberate attacks are conducted on the identified critical nodes, and the network's structure and functionality are evaluated using two key metrics, the size of the largest connected subgraph and the relative global operational efficiency. The Chengdu metro network is selected as a case study to validate the accuracy and reliability of the ITF method.

The case-based numerical results reveal the following key findings: (i) Impact of passenger flow distribution: the spatial and temporal distribution of passenger flow in the Chengdu metro network significantly affects the identification of critical nodes. Certain stations, such as Chengdu east station, consistently rank as important across both working and non-working days, while other stations, such as Tianfu square station, show considerable differences in importance rankings between these scenarios. (ii) Effectiveness of ITF in network resilience: as the proportion of critical node attacks increases, the ITF method demonstrates a more significant decrease in both the size of the largest connected subgraph and the relative global operational efficiency compared to other node importance identification methods. This indicates that the ITF method more accurately identifies critical nodes, resulting in greater damage to the network's structure and functionality when these nodes are compromised. (iii) Weighted integration of topology and functionality: the ITF method incorporates weighted considerations of both network topology and operational functionality. The results show that as the weight assigned to topological importance increases, the accuracy of critical node identification improves. This highlights the superiority of the ITF method over approaches that consider only topological structure or operational metrics in isolation, underscoring the necessity of evaluating both topological and operational attributes when assessing node importance in urban metro networks.

This study demonstrates that the ITF algorithm offers clear advantages in accurately identifying critical nodes, providing valuable insight for future planning and development of the Chengdu metro system. For larger-scale applications of the method, future research could explore the following areas: (i) Impact of multimodal transportation interactions: this study focuses on the interaction between the metro and bus systems, using a 300-meter buffer zone around metro stations to assess the influence on metro passenger flow. However, the optimal range for coupling interactions between transportation modes remains unexplored. In addition, other transportation modes, such as taxis and bicycles, may also impact passenger flow. Future research should investigate the influence of multimodal transportation systems on node importance within metro networks. (ii) Optimization of network performance metrics: while the current study's network scale is moderate, allowing for manageable time complexity in numerical analysis, larger-scale networks with tens of thousands of nodes present challenges in calculating global network-related metrics. Addressing the computational complexity of these evaluations will be crucial in advancing critical node identification research. Developing more efficient metrics and methods will be a key focus for future work.

Acknowledgments

Project supported by the National Natural Science Foundation of China (Grant No. 71971150), the Project of Research Center for System Sciences and Enterprise Development (Grant No. Xq16B05), and the Fundamental Research Funds for the Central Universities of China (Grant No. SXYPY202313).

References

[1] Leite D and De Bacco C 2024 *Nat. Commun.* **15** 7981
[2] De Jong M, Mu R, Stead D, Ma Y and Xi B 2010 *J. Transp. Geogr.* **18** 301
[3] Xu Z and Chopra S S 2023 *Nat. Commun.* **14** 4291

[4] Anupriya, Graham D J, Bansal P, Hörcher D and Anderson R 2023 *Physica A* **609** 128390
[5] Huang W, Shuai B, Sun Y, Wang Y and Antwi E 2018 *Transport. Res. A-Pol.* **111** 292
[6] Derrible S and Kennedy C 2010 *Physica A* **389** 3678
[7] Sun D and Guan S 2016 *Transport. Res. A-Pol.* **94** 348
[8] Yang Z, Dong X and Guo L 2023 *Reliab. Eng. Syst. Safe* **229** 108888
[9] Wang S, Lv W, Zhang J, Luan S, Chen C and Gu X 2021 *Reliab. Eng. Syst. Safe* **207** 107313
[10] Ma F, Shi W, Yuen K F, Sun Q, Xu X, Wang Y and Wang Z 2020 *J. Clean Prod.* **265** 121747
[11] Li M, Liu R R, Lü L, Hu M B, Xu S and Zhang Y C 2021 *Phys. Rep.* **907** 1
[12] Freeman L C 1978 *Soc. Networks* **1** 215
[13] Wang F and Hu H 2021 *Measurement* **179** 109449
[14] Sabidussi G 1966 *Psychometrika* **31** 581
[15] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 *Nat. Phys.* **6** 888
[16] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S and Arenas A 2015 *Nat. Commun.* **6** 6868
[17] Cui D X, He J L, Xiao Z F and Ren W P 2023 *Chin. Phys. B* **32** 098904
[18] Gao C, Fan Y, Jiang S, Deng Y, Liu J and Li X 2022 *IEEE T. Intell. Transp.* **23** 6509
[19] Engsig M, Tejedor A, Moreno Y, Foufoula-Georgiou E and Kasmi C 2024 *Nat. Commun.* **15** 56
[20] Ai D, Liu X L, Kang W Z, Li L N, Lü S Q and Liu Y 2023 *Chin. Phys. B* **32** 118902
[21] Huang W, Li H, Yin Y, Zhang Z, Xie A, Zhang Y and Cheng G 2024 *Reliab. Eng. Syst. Safe* **242** 109766
[22] Duan D L, Ji S Y and Yuan Z W 2024 *Chin. Phys. B* **33** 040502
[23] Lin J and Ban Y 2013 *Transp. Rev.* **33** 658
[24] Kopsidas A, Douvaras A and Kepaptsoglou K 2023 *Appl. Netw. Sci.* **8** 69
[25] Lai Q and Zhang H H 2022 *Chin. Phys. B* **31** 068905
[26] Chen C, Wang S, Zhang J and Gu X 2023 *J. Infrastruct. Syst.* **29** 04022043
[27] Xu M, Wu J, Liu M, Xiao Y, Wang H and Hu D 2019 *IEEE T. Intell. Transp.* **20** 583
[28] Jing W, Xu X and Pu Y 2020 *Reliab. Eng. Syst. Safe* **204** 107204
[29] Xia F, Wang J, Kong X, Zhang D and Wang Z 2020 *IEEE T. Intell. Transp.* **21** 2840
[30] Zhang J, Zhou Y, Wang S and Min Q 2024 *Chaos Soliton. Fract.* **178** 114379
[31] Wang L, Zhang S, Szűcs G and Wang Y 2024 *Reliab. Eng. Syst. Safe* **244** 109956
[32] Shanmukhappa T, Ho I W H, Tse C K and Leung K K 2019 *IEEE Circ. Syst. Mag.* **19** 39
[33] Jin Z, Qinghuai L and Xiaotong H E 2013 *J. Beijing Jiaotong Univ.* **37** 78
[34] Fang X, Lu L, Li Y and Hong Y 2023 *Transport. Res. D-Tr. E* **121** 103819
[35] Shannon C E 1948 *The Bell System Technical Journal* **27** 379
[36] Zhang H, Zhong S, Deng Y and Cheong K H 2022 *IEEE T. Fuzzy Syst.* **30** 3284
[37] Yang X and Xiao F 2021 *Knowl-Based Syst.* **227** 107198
[38] Liu J G, Ren Z M and Guo Q 2013 *Physica A* **392** 4154
[39] Yue M, Kang C, Andris C, Qin K, Liu Y, Meng Q 2018 *T. Gis.* **22** 855
[40] Xu P C, Lu Q C, Xie C and Cheong T 2024 *Transport Res. A-Pol.* **179** 103907
[41] Zhang J, Wang S and Wang X 2018 *Physica A* **496** 72
[42] Nguyen Q, Pham H D, Cassi D and Bellingeri M 2019 *Physica A* **530** 121561
[43] Sheng J, Dai J, Wang B, Duan G, Long J, Zhang J, Guan K, Hu S, Chen L and Guan W 2020 *Physica A* **541** 123262
[44] Hong L, Yan Y, Ouyang M, Tian H and He X 2017 *Reliab. Eng. Syst. Safe* **158** 58