Commun Nonlinear Sci Numer Simulat 98 (2021) 105758

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Research paper

Efficient traffic-aware routing strategy on multilayer networks

Yaqin Hu[c], Mingyue Xu[a], Ming Tang[b,d,*], Dingding Han[a,*], Ying Liu[e,f]

[a] School of Information Science and Technology, Fudan University, Shanghai 200433, China
[b] School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China
[c] School of Communication and Electronic Engineering, East China Normal University, Shanghai 200241, China
[d] Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China
[e] School of Computer Science, Southwest Petroleum University, Chengdu 610500, China
[f] Big Data Research center, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:
Received 10 July 2020
Revised 1 February 2021
Accepted 10 February 2021
Available online 18 February 2021

Keywords:
Multilayer networks
Network congestion
Traffic capacity
Routing strategy

ABSTRACT

Most modern infrastructures are coupled with or are dependent on one another to form multilayer networks. Contrary to a single-layer transportation network, multilayer transportation networks not only have micro-interactions between nodes within each layer, they are also characterized by macro-differences, where the transport speeds of layers differ. Based on a multilayer network model, we articulate an efficient traffic-aware routing (ETAR) strategy. An adjustable macro-parameter and an adjustable micro-parameter are introduced to evaluate the relative importance of transport paths within different layers and reflect the traffic awareness mechanism integrating the global static and the local dynamic information, respectively. A packet can adaptively choose its transport layer and routing direction during the transport process. By optimizing the combination of inter-layer macro-parameter and intra-layer micro-parameter, the ETAR strategy can not only reasonably distribute the traffic load from the low-speed layer to the high-speed layer, but also prevent packets from taking a long detour and waiting at hubs for a long time. The simulation results on both synthetic and empirical networks indicate that the ETAR strategy greatly enhances the traffic capacity of multilayer networks, and that the average path length and the average transport time of packets are guaranteed to be reasonable. We also find that increasing the network size of high-speed layer can obviously improve the traffic capacity, but an appropriate mean degree can help to enhance the network traffic throughput.

1. Introduction

Since the introduction of Erdos-Renyi random network [1], the small-world network [2] and the scale-free network [3], many real-world systems with intricate relationships, such as transportation networks [4], communication networks [5] and social networks [6,7], are often described by complex networks. In the modern society, because of the exponentially increasing network size and load transportation, networks need to have a higher traffic capacity. Once the data being transported on the network exceeds the threshold of the traffic capacity, it results in congestion [8,9] or even the entire system with collapse. Thus it is imperative to enhance the traffic capacity of networks and meanwhile the transport speed of packets cannot be significantly reduced.

* Corresponding authors.
E-mail addresses: mtang@ce.ecnu.edu.cn (M. Tang), ddhan@fudan.edu.cn (D. Han).

https://doi.org/10.1016/j.cnsns.2021.105758
1007-5704/

There are a lot of researches on enhancing the traffic capacity of networks which can be summarized as: (1) optimizing network topology [10–14], (2) reasonably allocating network resources [15–17] and (3) improving the routing strategy [18–23]. Refs. [10,11] showed that the traffic capacity of networks is enhanced by adding a few links (or edges). But the cost of adding edges is much higher than deleting edges. Thus Ref. [13] enhanced the traffic capacity of scale-free networks by closing or cutting off some links between the hubs. The optimization of resource allocation concentrates on the node processing capacity [15], the link bandwidth [16], and the node queue length [17]. Compared with optimization of network topology and resource allocation, implementing efficient routing strategies is less costly and thus more attractive. For example, Ref. [21] proposed a global dynamic routing strategy to enhance the traffic capacity of networks based on the queue length information of nodes.

With the development of society and economy, many infrastructures in the real world are no longer isolated, but are coupled with or dependent on one another to form multilayer networks [24–26]. There are many types of interactions between nodes and a single-type edge can not well describe the relationships between nodes. For instance, a geographical location can be reached by different transportations [27] and users can spread information through different online social platforms [28,29]. Recent studies showed that the transport process on multilayer networks is very different from that on single-layer networks [30–34]. Therefore, in the era of big data, the research of transport dynamics on multilayer networks is inevitable and of great significance. Solé-Ribalta et al. [33] demonstrated that the structure of multiplex networks may cause congestion, and the cooperative phenomenon is associated with the Braess paradox, which means the independently selected optimal path to improve local performance may reduce the traffic capacity of the whole networks. Manfredi et al. [34] proposed a multilayer network model with limited node queue length and processing capacity, and revealed the relationship between the optimal buffer length and coupling strength. Moreover, Chen et al. [35] studied the traffic-driven epidemic spreading in multiplex networks and found that the outbreak of an epidemic can be suppressed by the multiplex structure.

One key point of designing the routing strategy on multilayer networks is how to distribute the load onto each layer. Gao et al. [36] proposed a comprehensive traffic routing strategy to redistribute the load of the low-speed layer to the high-speed layer by considering the macroscopic difference in transport speeds between layers and the microscopic difference in topology between nodes within a layer. Here we explore the transportation packets along the paths other than the most efficient path by introducing the concept of traffic awareness. Every node has an independent buffer queue, which changes dynamically with time, and the buffer queue length and processing capacity of a node determine the waiting time of packets on this node. We articulate the efficient traffic-aware routing (ETAR) strategy to maximize the traffic capacity of multilayer networks. Based on the macro information of the difference in transport speeds between layers, we combine the global structure information obtained from efficient routing [19] (i.e., efficient routing path length of packets) with the local dynamic information of the neighbor nodes' queue length (i.e., the waiting time of packets). And compared with the global dynamic routing strategy [21], the ETAR strategy just considers the local queue length, which can greatly reduce computational complexity and is more suitable for large-scale networks. An optimal parameter combination can reasonably distribute the load into nodes with different layers, and thus the traffic capacity of multilayer networks is enhanced greatly.

This paper is organized as follows. In Section 2, we describe the ETAR strategy on multilayer networks in detail. In Section 3, we test the ETAR strategy on both synthetic and empirical networks. In Section 4, we summarize the results.

2. Efficient traffic-aware routing strategy on multilayer networks

2.1. Double-layer coupled network model

To simplify the experiment, we use a double-layer network which has a low-speed layer and a high-speed layer. The low-speed layer is usually sparse, such as the bus network on the bus-subway coupled network or the railway network on the railway-air coupled network, which is labled as layer A. In contrast, the high-speed layer is more dense, such as the subway network or the air network, which is labled as layer B. In general, the construction cost of high-speed layer is much higher than that of the low-speed layer. Therefore, the size of high-speed network is much smaller and we set the nodes in layer B as the subset of layer A.

The double-layer network is constructed as follows. (1) Layer A is an uncorrelated configuration model (UCM) [37] with a power-law degree distribution as $p(k) \sim k^{-\gamma}$, where γ is the degree exponent, and the maximum degree $k_{max} \sim \sqrt{N_A}$ (N_A is the size of layer A). (2) Randomly select N_B ($N_B = pN_A$, p is the connection probability) nodes in layer A to reform an ER network as layer B; (3) Connect the counterpart nodes in layers A and B one by one. If a node of layer A without counterpart in layer B wants to transport data to layer B, it must be through the node that has a counterpart in layer B.

2.2. Traffic transportation dynamical model

The nodes in layer A are considered to be both the hosts that can generate packets and the routers that can transport packets, while the nodes in layer B can only transport packets. The processing capacity of a node i is C_i^F which means node i in layer F can process C_i^F packets in each time step, and the remaining packets of this node should wait for the next time step to be processed. To simplify the simulation, we set $C_i^A = C_i^B = 1$ in this paper. The packets can be directly

transported between counterpart nodes without cost, which means the processing capacity of these nodes is not consumed. The forwarding process is as follows:

1) At each time step, R packets are generated in layer A, and the source and destination are randomly selected for each packet.
2) The queue length of each node is unlimited, and the packets are transported under first-in-first-out (FIFO) rule. In each time step, each node processes the first C_i packets of the current queue.
3) Calculate the selection probability of each neighbor node according to the preset routing strategy, then select the nearest neighbor node as the next hop to transport a packet.
4) Once a packet reaches its destination, it will be deleted. Otherwise it will continue to wait to be transported on the networks.

2.3. Efficient traffic-aware routing strategy

In order to synthesize the macro-condition of different transport speeds of layers, and the micro-condition of the interactions of nodes within the same layer, we propose an efficient traffic-aware routing (ETAR) strategy to maximize the traffic capacity of multilayer networks. For a packet, the path between its source and destination on the networks is

$$p(s \to t) := s \equiv x_F^0, x_F^1, \ldots, x_F^l, \ldots, x_F^{n-1}, x_F^n \equiv t, \tag{1}$$

where x_F^l represents the lth hop node and $F \in \{A, B\}$. The data starting from the source node s will run n hops to arrive at the destination node t. We define the efficient traffic-aware distance between the node s and its neighboring node i toward the destination of a packet in the ETAR strategy as

$$D_{si} = \alpha_F[h_F d_{it}^E + (1 - h_F)q_i], \tag{2}$$

which is used to choose the nearest neighboring node with the smallest value of efficient traffic-aware distance as the next hop for this packet, as described in Section 2.2. In Eq. (2), α_F is a macro-parameter evaluating the speeds of different layers. The smaller the value of α_F, the faster the transport speed. If the neighboring node is in layer A, then it is α_A, otherwise it is α_B. h_F is an adjustable micro-parameter used to indicate the level of traffic awareness. Similarly, if the neighboring node is in layer A, then it is h_A, otherwise it is h_B. q_i represents the buffer queue length of the neighboring node in the current time step [38,39], and d_{it}^E is the minimum efficient distance from the neighbor node to the destination t calculated by efficient routing strategy [19]. The efficient distance L of all possible paths from the neighboring node i to the destination t is calculated in the efficient routing strategy, and the minimum value of L among all the paths is set as the value of d_{it}^E. The efficient distance of a path L from node i to node t is calculated as follows:

$$L(p(i \to t), \beta_F) = \sum_{l=0}^{n-1} k(x_F^l)^{\beta_F}, \tag{3}$$

where $k(x_F^l)$ represents the degree of the node x_F^l, and β_F is an adjustable exponent. For simplicity, the difference between layer A and B is not considered here and we set $\beta_F = 1$ [19].

The ETAR strategy integrates multiple information, including the transport speeds of different layers on the macro-level and the role of different nodes in the same layer on the micro-level, and also takes into account the global static information (i.e., the degrees of all nodes) and the local dynamic information (i.e., the buffer queue length in different time steps). $h_F = 1$ indicates that only the shortest efficient distance calculated from node degrees on the path is considered, while $h_F = 0$ indicates that only the buffer queue length of the buffer is considered.

2.4. Definition of statistical parameters

In order to describe the network congestion, we use the order parameter introduced in Ref. [40]. This parameter is calculated as follows:

$$H = \lim_{t \to \infty} \frac{C}{R} \frac{\langle \triangle W \rangle}{\tau}, \tag{4}$$

where $\triangle W = W(t + \tau) - W(t)$. $W(t)$ is the total number of packets on the network at time t, $\triangle W$ is the increment of the packets on the networks from time t to time $t + \tau$ and $\langle \cdots \rangle$ represents the average of them. Further, C is the node processing capacity, R is the number of packets generated at each time step, and τ is the time interval. H actually represents the ratio of the packets generated to those processed in each time step on the networks. At $H = 0$, the packets generated and processed are in a balanced state, i.e., the network is in the free-flow state. With the increase of R, as the processing capacity of nodes is limited, the packets generated in each time step could not be processed in time. Consequently, the packets accumulate and the network experiences congestion where $H > 0$. The phase transition point of H corresponds to the traffic capacity of the network R_c. When $R < R_c$, the network is in free-flow state, and there is no packet accumulation. When $R = R_c$, the network is in a critical state. And at $R > R_c$, the network becomes congested, when packets cannot reach the destination nodes in time.

In order to better characterize the transport of packets and facilitate the understanding of the ETAR strategy on multilayer networks, the following indicators of coupling strength between layers are introduced. The first is the coupling coefficient of paths Λ between layers, which describes the way in which packets are transported between two layers, and is given by

$$\Lambda = \lim_{t\to\infty} \frac{\sum_{s\neq t} \sigma_B^{st}(\alpha_F, h_F)}{\sum_{s\neq t} \sigma^{st}(\alpha_F, h_F)}, \tag{5}$$

where $\sum_{s\neq t} \sigma^{st}(\alpha_F, h_F)$ is the number of paths between the source s and the destination t under given α_F, h_F, and $\sum_{s\neq t} \sigma_B^{st}(\alpha_F, h_F)$ indicates the number of paths that contain at least one edge in layer B. $\sum_{s\neq t} \sigma_B^{st}(\alpha_F, h_F) = 0$ means that the path only passes through layer A. When $\Lambda \approx 0$, the paths of all packets are mainly in layer A, and almost no packet passes through layer B. When $\Lambda \approx 1$, the transport paths of almost all packets pass through layer B.

Next, we define the proportion coefficient of the edges passing through layer B on all paths as

$$\Delta = \lim_{t\to\infty} \frac{\sum_{s\neq t} e_B^{st}(\alpha_F, h_F)}{\sum_{s\neq t} e^{st}(\alpha_F, h_F)}, \tag{6}$$

where $e_B^{st}(\alpha_F, h_F)$ is the number of edges that pass through layer B in the path from the source to the destination under given α_F, h_F. $\sum_{s\neq t} e^{st}(\alpha_F, h_F)$ is the number of edges that pass through the entire network. When $\Delta = 0$, it means that the transportation does not occur in layer B. But with the increase of Δ, more and more packets are transported through layer B.

We define the average path length of packets as

$$\langle L\rangle = \lim_{t\to\infty} \frac{1}{n} \sum_{s\neq t} L^{st}(\alpha_F, h_F), \tag{7}$$

where $L^{st}(\alpha_F, h_F)$ is the path length of packets from the source to the destination under given α_F, h_F in the ETAR routing strategy, as described in Sections 2.2 and 2.3, and n is the number of all packets that can reach the destination. The shorter the average path length is, the lower the probability of packet detours, and the higher possibility of reaching the destination.

The average transport time of packets on the networks is defined as

$$\langle T\rangle = \lim_{t\to\infty} \frac{1}{n} \sum_{s\neq t} T^{st}(\alpha_F, h_F), \tag{8}$$

where $T^{st}(\alpha_F, h_F)$ is the travel time of packets to reach the destination from the source under given α_F, h_F. The smaller the average transport time is, the faster the packets can reach the destination.

We define the ratio of the accumulated data packets in layer B to those on the entire network as

$$\Phi = \lim_{t\to\infty} \frac{\sum_{N_B} q_i}{\sum_N q_i}, \tag{9}$$

where q_i is the queue length of node i, $\sum_{N_B} q_i$ is the number of packets in layer B, and $\sum_N q_i$ is the number of packets on the entire network. When $\Phi = 0$, there is no packet accumulated in layer B. But when $\Phi = 1$, all the packets are accumulated in layer B.

3. Simulation results

3.1. The efficiency of the ETAR strategy

The simulations are carried out on double-layer networks. The size of layer A is $N_A = 1000$ and the degree distribution exponent is $\gamma = 3.0$. The size of layer B is $N_B = 500$ and the average degree is $\langle k_B\rangle = 6$. To eliminate randomness, the results are obtained by averaging over 100 independent simulations on 10 different double-layer networks.

The main objective of our work is to study the influence of the macro-parameter α_B and the micro-parameter h_B on the traffic capacity of double-layer networks. First, the transportation process in the isolated layer A is simulated by setting the value of α_A to 1. Fig. 1 shows that the isolated layer A reaches its maximum traffic capacity $R_c^A = 35$ at about $h_A = 0.8$. Therefore, in the subsequent simulation experiments, h_A is fixed as 0.8. After adding the layer B as high-speed network, we study how to combine α_B and h_B to make the traffic more reasonably distributed among layer A and layer B. Making full use of network resources can not only make the double-layer networks reach the maximum traffic capacity, but also effectively restrict the transport time of packets.

Packets are transported on the double-layer networks under the ETAR strategy. From Fig. 2(a), we can see that the order parameter H shows a monotonous increase with the packet generation rate R, and there is a phase transition point which corresponds to the critical traffic capacity R_c of networks. Before reaching the threshold, the order parameter H tends to 0

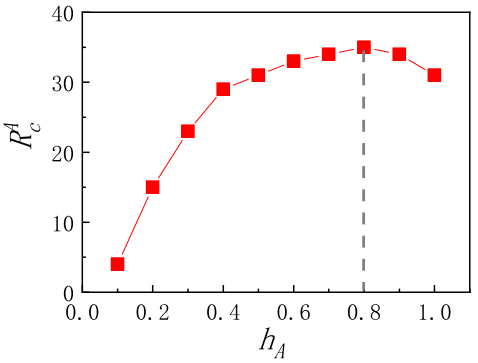

Fig. 1. The maximum traffic capacity for the ETAR strategy on isolated layer A when $\alpha_A = 1$. Other parameters are set as $N_A = 1000$, $k_{min} = 2$, $k_{max} \sim \sqrt{N_A}$, $\gamma = 3$.

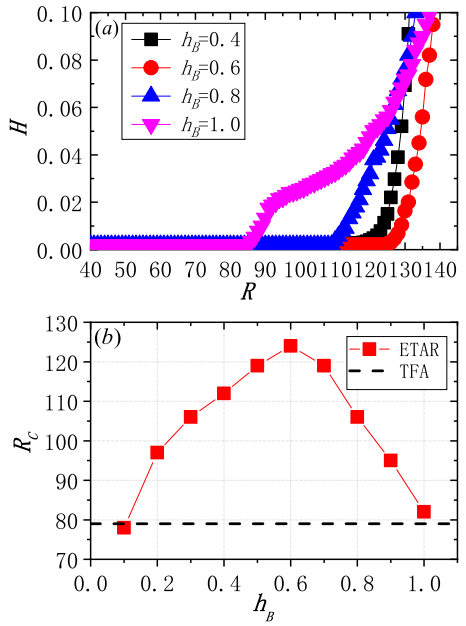

Fig. 2. The traffic capacity for the ETAR strategy on synthetic networks when $\alpha_B = 0.7$. (a) The order parameter H versus the packet generation rate R for different h_B values. (b) The traffic capacity R_c versus micro parameter h_B, where the black dotted line represents the maximum traffic capacity for the TFA strategy. The other parameters are set as $N_B = 500$, $\langle k_B \rangle = 6$, $\alpha_A = 1$, $h_A = 0.8$.

and the network is in free-flow state. Once R exceeds the threshold R_c, the order parameter H starts to increase gradually. It means that the packet generation rate exceeds the traffic capacity. In other words, the number of packets generated in each time step is greater than the number of the eliminated, then packets begin to accumulate on the coupled networks, which causes a congestion. In addition, it can be seen from Fig. 2(b) that the maximum traffic capacity R_c of the multilayer networks exhibits a non-monotonic relationship with the micro-tunable parameter h_B. According to the results of the numerical simulations, the traffic capacity reaches an optimal value of $R_c = 124$ at $h_B \approx 0.6$, which is 2.5 times higher than that of an isolated layer A. It illustrates the efficiency of the ETAR strategy, which makes reasonably use of resource in the high-speed layer to offload the low-speed layer and thus enhances the traffic capacity of multilayer networks. With the degree information of nodes, the ETAR strategy enables the traffic to be distributed to the other nodes from hubs. At the same time, the dynamic change of the queue length is considered to reduce the load on the busy nodes. In addition, we compare the maximum transmission capacity of the traffic-flow assignment (TFA) strategy [36] with that of the ETAR strategy in Fig. 2(b), and it can be seen that the capacity of ETAR is larger than that of the TFA strategy to a large extent.

To further verify effectiveness of the ETAR strategy, we investigate the average path length and the average transport time under different packet generation rates. According to Fig. 2(b), there is a maximum $R_c = 124$ at about $h_B = 0.6$. We set the packet generation rates $R = 50, 110, 140$, corresponding to the free-flow state, slight congestion state and heavy congestion state of transportation systems with different h_B values, respectively. It can be seen from Fig. 3(a) that the average path length $\langle L \rangle$ in Eq. (7) does not decrease monotonically with the micro-parameter h_B, but tends to a small value when h_B

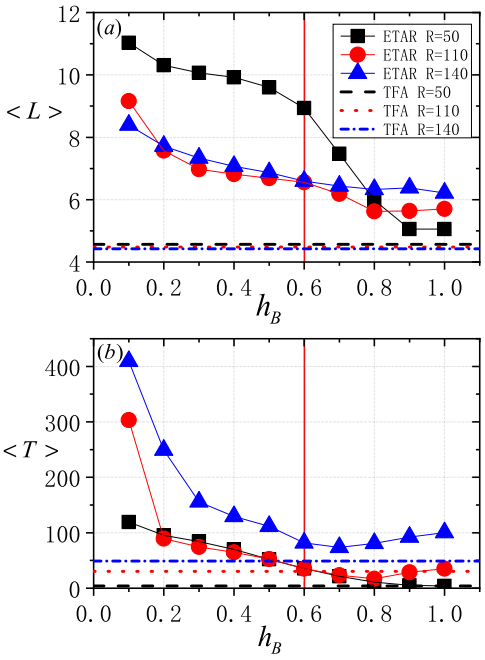

Fig. 3. The performance of the ETAR strategy and the TFA strategy in different phases on synthetic networks. (a) the average path length $\langle L\rangle$ and (b) the average transport time $\langle T\rangle$ versus h_B. Other parameters are set as $\alpha_A = 1$, $h_A = 0.8$, and $\alpha_B = 0.7$.

Table 1
The rate of packet arrival of the ETAR strategy and the TFA strategy.

R	50	110	140
ETAR	98.576%	**98.547**%	**96.704**%
TFA	98.665%	78.521%	63.397%

is greater than a critical value, e.g., $h_B > 0.8$ for $R = 110$. As h_B increases, the packets tend to pass through the nodes in efficient paths between their sources and destinations, resulting in the decrease of $\langle L\rangle$. Fig. 2(b) shows that a small (or a large) value of h_B causes a traffic jam due to a long detour (or the accumulation) of more packets from (on) hubs. Once h_B is greater than a critical value at a fixed R, e.g., $h_B > 0.8$ for $R = 110$, a congestion will occur in the transportation systems. There are a great number of packets piled up on most of the nodes of the double-layer networks. According to Eq. (2), those paths with small efficient distance are chosen to be next hop of packets more easily, which leads to a saturation effect of $\langle L\rangle$.

As the packet generation rate increases, the average path length $\langle L\rangle$ will increase in most cases. A greater packet generation rate R enhances the traffic load on the networks, and many packets accumulate on some nodes and can not arrive at their destinations in time. In this case, the other packets who choose to detour may reach the destinations early, so the path length of these packets increases. But when $h_B \epsilon [0, 0.7]$ for $R = 50$, there is a large value of $\langle L\rangle$, which means that most packets take a long detour from efficient paths to arrive at the destination early in the free-flow phase, see the case of $R = 50$ in Fig. 3(b). Different from almost the same value of $\langle T\rangle$ for a large h_B, Fig. 3(b) shows that $\langle T\rangle$ has a minimum value at a large h_B. An appropriate value of h_B can prevent packets from not only taking a long detour but also waiting at hubs for a long time, leading to a minimum value of $\langle T\rangle$. When $h_B = 0.6$, although $\langle L\rangle$ and $\langle T\rangle$ of the ETAR strategy are slightly higher than that of the TFA strategy in most cases, it can be seen from Table 1 that the TFA strategy results in much lower arrival rates of packets, especially for the cases of $R = 110$ and 140. This is because a smaller $R_C = 79$ of the TFA strategy makes the traffic system be in the congested phase more easily.

3.2. The synergistic interaction between layers

In order to understand the mechanism of the synergy on the multilayer networks, we introduce three indexes of the coupling coefficients of paths Λ, edges Δ and packets Φ between layers, see the definition in Section 2.4. When $R = 50$, i.e., the network is in free-flow state, the values of all three parameters in Fig. 4 are monotonically decreasing. The reason may be that when the network is in free-flow state, there are fewer packets stacked on multilayer networks, and the queue length of nodes is very small. That is, the efficient routing strategy dominates the selection of next hop. As h_B increases to a

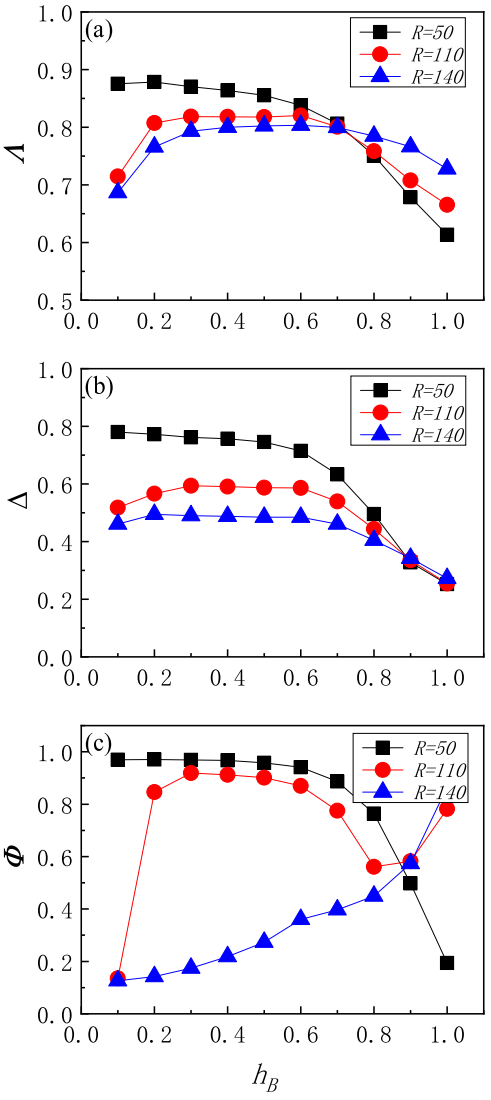

Fig. 4. The coupling strength between layers as a function of h_B for different R values, where (a) the coupling coefficient of paths Λ, (b) the coupling coefficient of edges Δ and (c) the coupling coefficient of packets Φ. Other parameters are set as $\alpha_A = 1$, $h_A = 0.8$, and $\alpha_B = 0.7$.

critical value(i.e., $h_B \approx 0.6$), the term $h_B d_{it}^E$ on layer B increases, but the term $h_A d_{it}^E$ is unchanged. Therefore, more and more packets avoid making a detour and rarely selecting the nodes in layer B as their next hop.

As the packet generation rate increases, a congestion on the networks occurs, and the queue length of accumulated nodes gradually increases. When $R = 140$ in Fig. 4(c), the system is in congestion state with many accumulated packets, especially for layer B. With the gradual increase of h_B, the coefficient $(1 - h_B)$ of queue length in the ETAR strategy decreases, but remains unchanged for the nodes of layer A. Therefore, the packets will choose the neighbors in layer B preferably, which makes the packets piled up in layer B increase. In this case, only a few packets can reach their destinations mainly by using layer A with a small number of stacked packets, resulting in the decrease of both Λ and Δ when $h_B > 0.6$ [see the cases of $R = 110$ and $R = 140$ in Fig. 4(a) and (b)]. When h_B is small, we see that both Λ and Δ increase slightly with h_B. The ETAR strategy with a small h_B mainly relies on local dynamic information(i.e., the queue length of nodes), but ignores global static information(i.e., the dfficient path), leading to the congestion on the networks easily. A large number of packets are piled up in layer B, that is an increase of Φ for the cases of $R = 110$ and $R = 140$ in Fig. 4(c). A few packets choosing layer B can arrive at their destinations, which leads to the slight increase of Λ and Δ for those arrived packets. Moreover, we find that a smaller R results in greater Λ, Δ and Φ when h_B is not large (e.g., $h_B < 0.6$). Compared with layer A, layer B have a denser connection and a more homogeneous structure. If packets are preferentially distributed to layer B and mainly take the local dynamic information of layer B into account at a small h_B, they will arrive at the destinations more quickly [see the case of $R = 50$ in Fig. 3(b)], but have a large $\langle L \rangle$ [see the case of $R = 50$ in Fig. 3(a)].

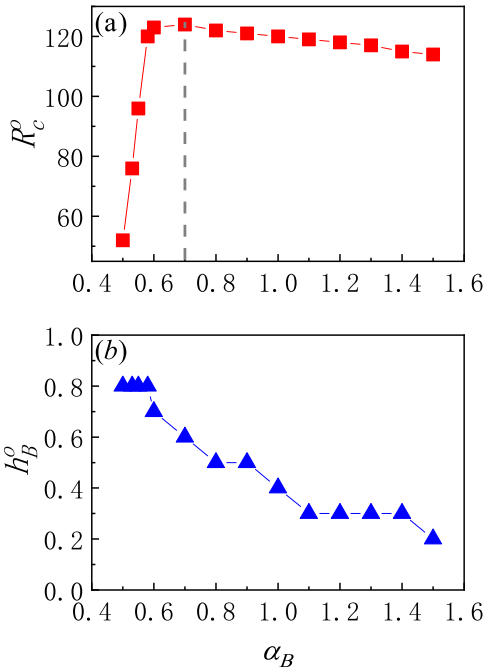

Fig. 5. The optimal performance of the ETAR strategy for different values of the macro parameter α_B. (a) the maximum traffic capacity R_c^o and (b) the optimal micro parameter h_B^o as a function of α_B. The other parameters are set as $\alpha_A = 1$, and $h_A = 0.8$.

3.3. The optimal performation of the ETAR strategy

We test various combinations of the macro-parameter α_B and the micro-parameter h_B in simulations to make the load of the networks distributed more uniformly through the ETAR strategy under different intra-layer speeds. In Fig. 5, each macro parameter α_B has the corresponding optimal micro parameter h_B^o, and the traffic capacity of the networks has different optimal values R_c^o under different combinations. As α_B increases, a smaller h_B^o can make sure that the traffic flow be reasonably distributed to layer B, and thus can enhance the network throughput. This indicates that the ETAR strategy can adjust the routing adaptively to the global static information and the local dynamic information under fixed network resources. Simulation results show that the maximum traffic capacity of multilayer networks is 124, which is obtained under the best combination of $\alpha_B = 0.7$ and $h_B = 0.6$. By using the ETAR strategy, the traffic capacity of multi-layer networks is more than 3.5 times that of single-layer networks.

We investigate the impacts of the network size N_B, the average degree $\langle k_B \rangle$ and the clustering coefficient C_B of layer B on the optimal performance of the ETAR strategy. As shown in Fig. 6, we can see that the maximum traffic capacity R_c^o and the optimal micro parameter h_B^o increase almost linearly with N_B. It may be due to the increase of network size of layer B and the corresponding increase in the number of efficient paths in the layer B. While increasing $\langle k_B \rangle$, there is a trend of R_c^o increasing first and then decreasing. When the average degree of the networks is larger, the packets have more chances to reach the destinations through the shortcuts. However, the packets under the ETAR strategy with a small h_B^o may waste more time and traffic resource due to detours than waiting at nodes if the average degree is too large. But in order to save costs, the average degree of systems in the real world is usually not too large. The ETAR strategy can work well on the networks with the right average degree. As the clustering coefficient increases [41], the network tends to be decomposed into smaller tightly connected clusters (i.e., modular or community structures), which makes the average shortest distance between nodes increase [42] and leads to the limitation of the traffic capacity of the network. From Fig. 6(f), the optimal value of h_B generally changes from 0.6 to about 1. It is probably because the packet transmission within community structures is easily done by using the shortest path strategy, which the other packets only are delivered through bridge nodes (i.e., the shortest distance) between communities to arrive at their destinations in other communities.

3.4. Empirical networks

To further verify the efficiency of the ETAR strategy, we implement simulation experiments on a real-world multiplex network. The data constitute a multimodal transportation networks of le-de-France [43], whose layers correspond to different transportation modes and are interconnected into a multilayer network. This multilayer network consists of three layers: the road network, the subway network and the train network which is a hybrid commuter rapid transit system in France serving Paris and its suburbs. We select the subway network and the train network to construct a double-layer network. The

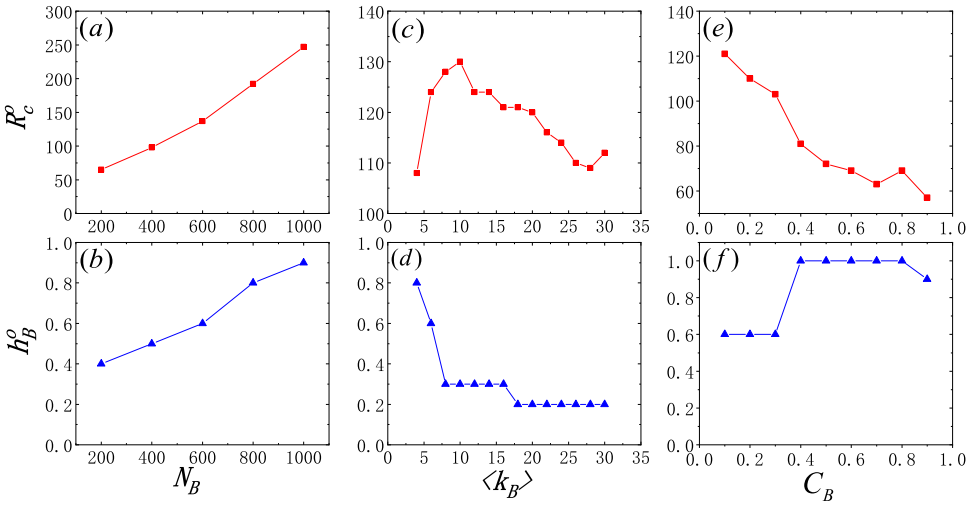

Fig. 6. The effects of network size N_B, mean degree $\langle k_B \rangle$ and the clustering coefficient C_B of layer B on the optimal micro parameter h_B^o and the maximum traffic capacity R_c^o of the ETAR strategy. The maximum traffic capacity R_c^o (top) and the optimal micro parameter h_B^o (bottom) versus N_B[(a) and (b)], $\langle k_B \rangle$ [(c) and (d)] and C_B [(e) and (f)]. The other parameters are set as $\langle k_B \rangle = 6$ [(a) and (b)], $N_B = 500$ [(c) and (d)], $\alpha_A = 1$, $h_A = 0.8$, and $\alpha_B = 0.7$.

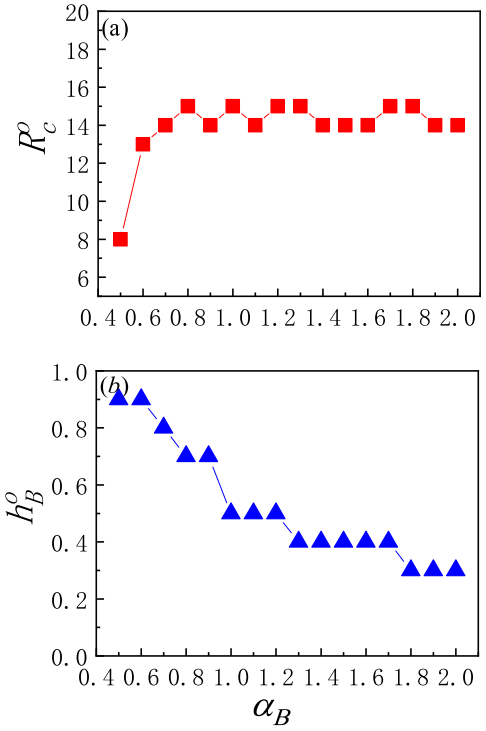

Fig. 7. The performances of the ETAR strategy on the empirical networks for different values of the macro parameter α_B. (a) The maximum traffic capacity R_c^o and (b) The optimal micro parameter h_B^o as a function of α_B. The other parameters are set as $\alpha_A = 1$ and $h_A = 0.6$.

nodes correspond to the subway or train stations, and the edges represent the subway or train lines. On the network, layer A is the subway network with 303 nodes and 356 edges, and layer B is the train network with 241 nodes and 244 edges. There are 64 inter-layer edges.

We first get the optimal value $h_A = 0.6$ on the independent subway layer, and then apply the ETAR strategy to the real double-layer transportation network. In Fig. 7, we can see the relationships between the traffic capacity R_c^o of the network, the optimal micro parameters h_B^o and the macro parameters α_B. Compared with the synthetic networks, the results of the real-world networks are more volatile. This is because the topology of the network in the real world is much more complex.

Table 2
The average path length $\langle L\rangle$ and the average transport time $\langle T\rangle$ under SP, ER and ETAR strategies.

	$\langle L\rangle$			$\langle T\rangle$		
R	8	13	15	8	13	15
SP	10.507	10.263	10.115	6.580	142.208	208.923
ER	10.931	10.943	10.898	5.832	7.361	34.052
ETAR	10.994	11.057	11.611	**5.890**	**7.632**	**10.322**

The maximum traffic capacity is achieved under different combinations of optimal micro parameter h_B and macro parameter α_B, which displays a similier phenomenon with that of the synthetic networks.

To highlight the superiority of the ETAR strategy further, we compare its performance with that of two other strategies: the shortest path (SP) strategy and the efficient routing (ER) strategy with $\beta = 1$ on the real-world multilayer networks. Here we set the parameters of the ETAR strategy as $\alpha_B = 0.8$ and $h_B = 0.7$. The traffic capacities of the network are 8, 13, and 15 in the SP, ER and ETAR strategies, respectively. Table 2 is a comparison of the average path length $\langle L\rangle$ and the average transport time $\langle T\rangle$ under the three strategies. We can see that there are very small differences on $\langle L\rangle$ but huge differences on $\langle T\rangle$ among the three strategies. Once the packet generation rate increases, the average transport time $\langle T\rangle$ under the SP strategy or ER strategy grows significantly. The network traffic turns to congestion state, and many packets cannot reach the destinations. Contrarily, as the ETAR strategy can adaptively adjust the direction of transportation against congestion on the networks, there is no significant increase of $\langle T\rangle$.

4. Conclusion and discussion

Owing to the rapid increase in the amount of data transported across networks and the demand for higher network quality, enhancing the traffic capacity of networks is very significant and urgent. Based on the multilayer networks, we proposed an efficient traffic-aware routing (ETAR) strategy. We first considered the difference in transport speeds of layers and introduced a macro-parameter α_F to evaluate the relative importance of transport paths within different layers. We also introduced a micro-adjustable parameter h_F to integrate the global static information and the local dynamic information in each layer, which reflects the traffic-aware mechanism of the ETAR strategy. We analyzed the extent to which layers A and B are able to engage in the forwarding process synergistically by determining the best combination of macro-factors and micro-adjustable coefficients. In an optimal case, the ETAR strategy can improve greatly the traffic capacity of synthetic double-layer networks. Compared with single-layer networks, the traffic capacity of synthetic double-layer networks is enhanced by more than 2.5 times as a result of adding the high-speed layer. An optimal parameter combination can reasonably distribute the traffic from the low-speed layer to the high-speed layer, which can enhance the traffic capacity of multilayer networks. An appropriate value of micro-parameter can prevent packets from not only taking a long detour but also waiting at hubs for a long time, leading to a minimum value of the average transporting time of packets.

Moreover, we found that increasing the network size of high-speed layer can obviously improve the traffic capacity of multi-mode transportation systems, but an appropriate mean degree can help to enhance the network traffic throughput. We further verified the efficiency of the proposed ETAR strategy on the real-world multi-mode transport network of Ile-de-France, and compared its performance with that of the SP and ER strategies. The ETAR strategy improves the traffic capacity of the multilayer network and reduces the average transport time of packets. Our work not only provides ideas for communication routing, but also for load balancing of transportation networks.

In this paper, an ideal infinite queue length is adopted in the simulations, which is usually assumed in many studies of routing strategies [19,38,44]. In reality, the buffers (i.e., queue lengths) that nodes use to store packets are finite. For example, in a communication network, the capacity of host and router storing packets is limited by cost [17]. In previous studies, the use of limited queue length [45], and the impact of buffer size on packet loss threshold [46] were studied, and it was also applied to two-layer networks [47]. The impact of finite queue length on different routing strategies such as the ETAR strategy on multilayer networks deserves further consideration in the next work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yaqin Hu: Data curation, Software, Writing - original draft. **Mingyue Xu:** Data curation, Validation, Visualization. **Ming Tang:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Dingding Han:** Supervision, Writing - review & editing. **Ying Liu:** Software, Validation, Writing - review & editing.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11875133, 11075057, 11975099 and 61802321), the National Key R&D Program of China (Grant No.2018YFB2101302), the Natural Science Foundation of Shanghai under Grant No. 18ZR1412200, and the Science and Technology Commission of Shanghai Municipality under Grant No. 14DZ2260800.

References

[1] Erdős P, Rényi A. On the evolution of random graphs. Publ Math InstHung Acad Sci 1960;5(1):17–60.
[2] Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature 1998;393(6684):440.
[3] Barabási A-L, Albert R. Emergence of scaling in random networks. Science 1999;286(5439):509–12.
[4] Han D, Yao Q, Chen Q, Qian J. An assessment method for aviation network optimization based on time-varying small world model. Acta Phys Sin 2017;66(24):248901.
[5] Zhang H, Zhang M. Complex network-based growth and evolution model for internet of things. In: 2014 IEEE 5th International conference on software engineering and service science. IEEE; 2014. p. 120–3.
[6] Miguel MC, Pastorsatorras R. Scalar model of flocking dynamics on complex social networks. Phys Rev E 2019;100(4):1–8.
[7] Zhang Z, Pu P, Han D, Tang M. Self-adaptive Louvain algorithm: fast and stable community detection algorithm based on the principle of small probability event. Physica A 2018;506:975–86.
[8] Yang H, Wang W, Lai Y, Xie Y, Wang B. Control of epidemic spreading on complex networks by local traffic dynamics. Phys Rev E 2011;84(4):045101.
[9] Chen S, Huang W, Cattani C, Altieri G. Traffic dynamics on complex networks: a survey. Math Probl Eng 2012;2012(1024–123X):256–67.
[10] Jiang Z, Liang M, Guo D. Enhancing network performance by edge addition. Int J Mod Phys C 2011;22(11):1211–26.
[11] Huang W, Chow TW. Effective strategy of adding nodes and links for maximizing the traffic capacity of scale-free network. Chaos 2010;20(3):033123.
[12] Nishikawa T, Motter AE. Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions. Proc Natl Acad Sci 2010;107(23):10342–7.
[13] Liu Z, Hu M, Jiang R, Wang W, Wu Q. Method to enhance traffic capacity for scale-free networks. Phys Rev E 2007;76(3):037101.
[14] Huang W, Chow TW. An efficient strategy for enhancing traffic capacity by removing links in scale-free networks. J Stat Mech 2010;2010(01):P01016.
[15] Yang H, Wang W, Wu Z, Wang B. Traffic dynamics in scale-free networks with limited packet-delivering capacity. Physica A 2008;387(27):6857–62.
[16] Tang M, Zhou T. Efficient routing strategies in scale-free networks with limited bandwidth. Phys Rev E 2011;84(2):026116.
[17] Wang W, Wu Z, Jiang R, Chen G, Lai Y. Abrupt transition to complete congestion on complex networks and control. Chaos 2009;19(3):033106.
[18] Bai Y, Han D, Tang M, et al. Multi-priority routing algorithm based on source node importance in complex networks. Int J Mod Phys C 2019;30(07):1–13.
[19] Yan G, Zhou T, Hu B, Fu Z, Wang B. Efficient routing on complex networks. Phys Rev E 2006;73(4):046108.
[20] Ma J, Han W, Guo Q, Zhang S, Wang J, Wang Z. Improved efficient routing strategy on two-layer complex networks. Int J Mod Phys C 2016;27(04):1650044.
[21] Ling X, Hu M, Jiang R, Wu Q. Global dynamic routing for scale-free networks. Phys Rev E 2010;81(1):016113.
[22] Gao X, Guo H, Chen Y, Tang Y, Wang C, Xu S, et al. Global hybrid routing for scale-free networks. IEEE Access 2019;7:19782–91.
[23] Jiang Z, Liang M. Improved efficient routing strategy on scale-free networks. Int J Mod Phys C 2012;23(02):1250016.
[24] Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. J Complex Netw 2014;2(3):203–71.
[25] Tejedor A, Longjas A, Foufoulageorgiou E, Georgiou TT, Moreno Y. Diffusion dynamics and optimal coupling in multiplex networks with directed layers. Phys Rev X 2018;8(3):031071.
[26] Cozzo E, De Arruda GF, Rodrigues FA, Moreno Y. Layer degradation triggers an abrupt structural transition in multiplex networks. Phys Rev E 2019;100(1):012313.
[27] De Domenico M, Solé-Ribalta A, Gómez S, Arenas A. Navigability of interconnected networks under random failures. Proc Natl Acad Sci 2014;111(23):8351–6.
[28] Mucha PJ, Richardson T, Macon K, Porter MA, Onnela J-P. Community structure in time-dependent, multiscale, and multiplex networks. Science 2010;328(5980):876–8.
[29] Zhang L, Su C, Jin Y, Goh M, Wu Z. Cross-network dissemination model of public opinion in coupled networks. Inf Sci 2018:240–52.
[30] Aleta A, Moreno Y. Multilayer networks in a nutshell. Annu Rev Condens Matter Phys 2019;10(1):45–62.
[31] Wu D, Tang M, Liu Z, Lai Y. Impact of inter-layer hopping on epidemic spreading in a multilayer network. Commun Nonlinear Sci NumerSimul 2020;90:105403.
[32] Ma J, Han W, Guo Q, Wang Z. Traffic dynamics on two-layer complex networks with limited delivering capacity. Physica A 2016;456:281–7.
[33] Solé-Ribalta A, Gómez S, Arenas A. Congestion induced by the structure of multiplex networks. Phys Rev Lett 2016;116(10):108701.
[34] Manfredi S, Di Tucci E, Latora V. Mobility and congestion in dynamical multilayer networks with finite storage capacity. Phys Rev Lett 2018;120(6):068301.
[35] Chen J, Hu M-B, Li M. Traffic-driven epidemic spreading in multiplex networks. Phys Rev E 2020;101(1):012301.
[36] Gao L, Shu P, Tang M, Wang W, Gao H. Effective traffic-flow assignment strategy on multilayer networks. Phys Rev E 2019;100(1):012310.
[37] Catanzaro M, Boguná M, Pastor-Satorras R. Generation of uncorrelated random scale-free networks. Phys Rev E 2005;71(2):027103.
[38] Echenique P, Gómez-Gardenes J, Moreno Y. Dynamics of jamming transitions in complex networks. Europhys Lett 2005;71(2):325.
[39] Echenique P, Gómez-Gardeñes J, Moreno Y. Improved routing strategies for internet traffic delivery. Phys Rev E 2004;70(5):056105.
[40] Arenas A, Díaz-Guilera A, Guimera R. Communication in networks with hierarchical branching. Phys Rev Lett 2001;86(14):3196.
[41] Volz E. Random networks with tunable degree distribution and clustering. Phys Rev E 2004;70(5):056115.
[42] Holme P, Kim BJ. Growing scale-free networks with tunable clustering. Phys Rev E 2002;65(2):026107.
[43] Aaron Clauset E.T., Sainz M.. The colorado index of complex networks. https://icon.colorado.edu/; 2016.
[44] Wang W-X, Wang B-H, Yin C-Y, Xie Y-B, Zhou T. Traffic dynamics based on local routing protocol on a scale-free network. Phys Rev E 2006;73(2):026111.
[45] Wu Z-X, Wang W-X, Yeung K-H. Traffic dynamics in scale-free networks with limited buffers and decongestion strategy. N J Phys 2008;10(2):023025.
[46] Chen J, Chen J-Y, Li M, Hu M-B. Traffic dynamics considering packet loss in finite buffer networks. Chin Phys B 2019;28(4):048901.
[47] Ma J, Wang H, Ren H, Xu X, Zhang D, Liu G. Traffic dynamics on two-layer multiplex networks with limited queue resource. Int J Mod Phys B 2019;33(26):1950312.