International Journal of Modern Physics C
Vol. 30, No. 7 (2019) 1940010 (13 pages)

DOI: 10.1142/S0129183119400102

World Scientific
www.worldscientific.com

Multi-priority routing algorithm based on source node importance in complex networks

Yu Bai[*,†,‡], Ding-Ding Han[*,†,§] and Ming Tang[†]

[*]*School of Information Science and Technology*
Fudan University, Shanghai 200433, P. R. China

[†]*School of Information Science and Technology*
East China Normal University
Shanghai 200241, P. R. China
[‡]*934643344@qq.com*
[§]*ddhan@fudan.edu.cn*

Received 15 October 2018
Accepted 14 May 2019
Published 4 July 2019

Network transmission capacity is an important guarantee for the normal operation of the network. The effective routing strategy avoids the use of nodes with large degree value, which leads to low utilization of nodes and failure to consider the priorities of different packets. On this basis, a routing algorithm based on packet source node classification is proposed. This strategy introduces an adjustable parameter. By adjusting this parameter, the data packets generated at the important nodes are transferred to the nodes with higher degree, which is to say they can reach the destination faster. The data packets generated at the sub-important nodes are transmitted by nodes with smaller degrees, thus reaching the destination relatively slowly. The routing strategy is evaluated in terms of order parameters, average routing time and node utilization. Compared with nonclassified routing and randomly classified routing strategy, the network transmission capacity was increased by 19% and 38%, respectively. Each node in the network was used more evenly. At the same time, the network transmission capacity under different parameters is analyzed theoretically through a series of derivations. In order to explore the performance of routing strategy in actual networks, this paper selects the actual network of web-EPA for simulation. The experimental results show that the proposed routing strategy is 7% and 17% higher than the nonclassified routing and random classified routing, respectively.

Keywords: Classified routing; congestion; node utilization; transmission capability.

PACS Nos.: 11.25.Hf, 123.1K.

1. Introduction

In recent years, complex networks have received extensive attention. Many networks closely related to our life can all be abstracted into complex networks for research, such as Internet, communication network, WWW network and interpersonal

[§]Corresponding author.

communication network. With the advent of the era of Big Data, congestion caused by data overload occurs increasingly, which will push the network at the risk of collapse. Therefore, it is a significant approach to improve the quality of network services through improving network transmission capacity, achieving efficient load transfer. Generally, network transmission capacity can be improved by optimizing network topology[1,2] and designing efficient routing strategies.[3–9] Since it is difficult and costly to change the network topology, it is a common practice to adopt appropriate routing strategies.

The shortest path routing strategy is currently the most widely used routing strategy,[3] with which Dijkstra algorithm or Floyd algorithm is adopted to find the path with the minimum number of hops between any two nodes for a given network topology. However, due to the heterogeneity of scale-free networks, path selection with shortest routing (SR) usually biases the node with high degree of selection, whose processing capability is limited, leading to frequent congestion at the hub node of the network. Yan *et al.*[4] proposed an efficient path strategy through minimizing the sum of degrees along the selected path, effectively avoiding the occurrence of hub nodes. However, this strategy was excessively dependent on the nodes with small degree, resulting in the increase of average path length. Chen *et al.*[5] presented a new algorithm on the basis of generalized betweenness centrality which redistributed traffic load from nodes with large betweenness centrality to nodes with small betweenness centrality, thus enhancing the traffic capacity of the network. This strategy reduced the possibility of congestion at the expense of transmission time, which is essentially similar to the one proposed by Yan *et al.*[4] Lin *et al.*[6] proposed these improved algorithms based on conventional routing strategies, namely a node duplication avoidance algorithm, a next-nearest-neighbor algorithm and a restrictive queue length algorithm. After applying them to typical local routing strategies, the critical generation rate of information packets R_c increased by over 10-fold and the average transmission time $\langle T \rangle$ decreased by 70%. Tan *et al.*[7] presented a novel hybrid routing on scale-free networks, combining static structural properties and dynamic traffic conditions together, thus balancing the traffic between hubs and peripheral nodes more effectively. Zhang *et al.*[8] proposed a Multi-priority Efficient Path (MPEP) strategy through which packets with different priorities were routed through different paths with different exponents of nodes' degree in the efficient path.[4] Compared with the efficient path[4] its routing effect was marginally improved, while the improvement was not distinct due to random packet classification without considering the network characteristics.

The effective routing strategy avoids the use of nodes with large degree value, which leads to low utilization of nodes and failure to consider the priorities of different packets. On this basis, the classification algorithm that is based on importance of source node is proposed. This strategy introduces an adjustable parameter. By adjusting this parameter, the data packets generated at the important nodes are transferred to the nodes with higher degree, which is to say they can reach the destination faster. The data packets generated at the sub-important nodes are transmitted by nodes with smaller degrees, thus reaching the destination relatively slowly. Packets generated at different nodes adopt different routing strategies to make full use of each node in the network.

Since the revelation of small-world characteristic of complex networks[10] by Watts and Strogatz and discovery of scale-free characteristics of complex networks[9] by Barabasi and Albert, many actual networks have been proved to have these two characteristics through large numbers of researches, such as Internet network,[11,12] communication network,[13–15] aviation network,[16] highways and urban transportation network,[17] etc. Therefore, it is of practical significance to establish and analyze the theoretical model of actual network structures with the complex network model in which information transmission in a network is to send messages from one node to another in practice.

This paper is organized as follows. The first part introduces traffic model, network topology and evaluation criteria for a routing strategy in the network, including packet threshold, order parameter, average routing time and node utilization. In the second part, a new routing strategy is proposed, through which nodes in the network are divided into important nodes and secondary nodes according to the network topology and different routing paths are adopted according to the importance of the generated packet nodes. In the third part, the proposed strategy are evaluated and compared with other strategies in different aspects in BA scale-free network model and real networks. Finally, conclusions of this research are drawn in the fourth part.

2. Traffic Model, Network Topology and Evaluation Criteria

Data or information in a communication network is presented as packets and transmitted through a specific routing strategy in real life.

2.1. *Traffic model*

Each node in the network can generate, receive and deliver packets. After the generation of a packet, it is transmitted between network nodes until its arrival at the destination node and then is removed from the network. In general, it goes through the following three procedures in the network:

Generation: At each time step, a number of R packets are generated in the network, the source and destination nodes of which are random.

Transmission: At each time step, packets in the network are transmitted to the next nodes according to the routing strategy adopted until their arrival at the destination nodes.

Removal: Upon their arrival at the destination nodes, the packets are removed from the entire network.

2.2. *Network topology*

In this paper, the scale-free network model is utilized for simulation, the degree of which is distributed according to power-law distribution $p(k) = k^{-\gamma}$, where k is node degree and γ is power exponent. The construction strategy of this model is stated as follows:

- Growth: Starting with a connectivity diagram with a number of m_0 nodes, one new node is added and connected to existing nodes with a number of m each time ($m < m_0$).
- Priority connection: The probability of connection between the newly added node and the existing node i is defined as $p_i = \frac{k_i}{\sum_{j \in V} k_j}$, where k_i is the degree of node i and V is the collection of all the nodes in the network.

2.3. *Evaluation criteria of routing strategies*

Specific measurement parameters are required as evaluation criteria of routing strategies to compare the advantages and disadvantages of various routing strategies. In this section, four evaluation parameters are illustrated, namely packet threshold, order parameter, average routing time and node utilization, which, respectively, represent the advantages and disadvantages of routing strategies from different aspects.

2.3.1. *Packet threshold and order parameter*

Packet threshold R_c refers to the number of packets generated per unit time increment when the network experiences a phase transition from a free-flow state ($R < R_c$) to a congested state ($R > R_c$). R_c is an important evaluation parameter for routing strategy, which directly reflects the advantages and disadvantages of a routing strategy. In general, it is less likely for a network to generate congestion when R_c is large, indicating a better routing strategy. In order to characterize the phase transition more directly and definitely, order parameter[18,20] H is introduced for quantitative analysis, which is defined as

$$H(R) = \lim_{t \to \infty} \frac{C\langle \Delta N_p \rangle}{R \Delta t}, \tag{1}$$

where $\langle \Delta N_p \rangle = N(t + \Delta t) - N(t)$ with $\langle \cdots \rangle$ indicates average over time windows of width Δt, and $N(t)$ is the total number of packets in the network at time t, C is the delivering capability of nodes within a time interval, R is the packet generating rate. For $H = 0$, the system is in a free-flow state, while for $H > 0$, the system is congested. A larger value of H indicates a severer congestion.

2.3.2. *Average routing time*

Routing time refers to the time it takes from the source node to the destination node or the path length from the source node to the destination node. For routing strategies of complex network, routing time is an important evaluation parameter which directly reflects the speed of routing strategy. However, the routing time of a few packets alone is not convincing in a typical large-scale network. Therefore, average routing time T with statistical properties is considered in this paper, which is defined as the average time of all packets that have reached the destination at a certain time in the network.

2.3.3. *Node utilization*

In previous studies, the routing load is often described with betweenness, which involves node betweenness and edge betweenness. Node betweenness is defined as the percentage of the shortest paths passing through the specified node in all the shortest paths in the network, while edge betweenness is defined as the percentage of the shortest paths passing through the specified edge in all the shortest paths. However, betweenness is merely able to evaluate routing load of the shortest path. In order to better analyze the situations of other routing strategies, node utilization U is introduced, which is stated as the proportion of the routing strategies going through the specified node in all the routing strategies adopted in the network:

$$U(i) = \frac{\sum_{\substack{s \to t \\ s \neq t \neq i}} \delta(i)}{\sum_{j \in V} \sum_{\substack{s \to t \\ s \neq t \neq j}} \delta(j)}, \tag{2}$$

where V is the collection of all the nodes in the network. If the path from node s to node t goes through node i, $\delta(i) = 1$, otherwise, $\delta(i) = 0$.

3. Routing Strategy and Methodology

3.1. *Routing strategy*

SR strategy is defined as $L_{s \to t} = \sum_i 1$, where s and t denote source node and destination node, respectively, i represents all the nodes along the specified routing. According to Yan,[4] an efficient routing strategy is presented as $L_{s \to t:\alpha} = \sum_i k_i^\alpha$, where k_i denotes the degree of node i, α is a parameter and the efficient routing from source node s to destination node t is defined as the routing minimizing $L_{s \to t:\alpha}$. Apparently, the SR corresponds to the case of $\alpha = 0$ and R_c of the packet is the largest when $\alpha = 1$ through simulation. To some extent, the efficient routing strategy alleviates the overdependence of SR on the nodes with larger degree which makes these nodes more susceptible to congestion. However, neither routing strategy takes into account the characteristics of the network itself. Through analysis, it is observed that only a few nodes have large degrees, while most nodes have small degrees in most actual networks. In the process of packet transmission, the few nodes with large degrees undertake a large number of transfer tasks, while the majority nodes with small degrees only undertake a small part of tasks. To some extent, the efficient path strategy redistributes traffic load in large-degree nodes to small-degree nodes, increasing the utilization frequency of nodes with medium degrees and reducing the utilization frequency of largest-degree nodes, which is also confirmed by simulation. However, the nodes with medium values are more susceptible to traffic congestion due to the uneven distribution of resources. This drawback of efficient path is even more evident in actual networks where nodes with largest degree tend to transmit packets to destination nodes more easily since those nodes have more neighbor nodes.

Therefore, in this paper, nodes in the network are divided into two categories according to their degrees, namely important nodes whose degrees are greater than the average and secondary nodes whose node degrees are smaller. Correspondingly, packets are also divided into the packets whose source nodes are important nodes and the packets whose source nodes are secondary nodes according to the importance of different nodes. For different packets, different routing strategies are adopted. On the basis of the routing strategy $L_{s\to t:\alpha} = \sum_i k_i^{\alpha}$,[4] an improved strategy is presented by the authors to minimize the value of L along the path for packet i $(i = 1, 2)$ as follows:

$$L_{s\to t:\alpha_i} = \sum_{j}^{n} k_j^{\alpha_i}. \tag{3}$$

Corresponding to different packet i, there are also different $\alpha_i (i = 1, 2)$, where $i = 1$ symbolizes the case when source nodes of the packets are important nodes and $i = 2$ symbolizes the case when the source nodes of the packets are secondary nodes. According to Ref. 4, in the case that the packets are not classified, it is optimal if $\alpha = 1$, when the packet avoids the nodes with large degree and are biased towards the nodes with small degree for transmission. It is observed from simulation that this approach improves the utilization frequency of nodes with medium degree to some extent, yet reducing the utilization of nodes with larger degree to a low level. According to the equation, the nodes with larger degree are more preferred if α is smaller when the transmission path and the corresponding transmission time are both shorter. In actual networks, the packet information of important nodes is generally more important and more urgent, thus being expected to be transferred to the destination more rapidly. Therefore, it is assumed that $\alpha_2 = 1$ for the packet source nodes as secondary nodes, and the target is to search for the optimal value of α_1 within the range $(0, 1)$.

3.2. *Theoretical derivation*

According to the definition of node utilization, which represents the proportion of the paths passing through the specified node in all paths, the average utilization of nodes with the same degree can then be calculated as

$$\bar{U} = \frac{1}{N_k} \sum_{k_i = i} U(i), \tag{4}$$

where N_k is the number of nodes with degree of k.

The average distance between packets can be estimated as

$$\bar{D} = \frac{\sum_{j\in V} \sum_{\substack{s \to t \\ s \neq t \neq j}} \delta(j)}{N(N-1)}, \tag{5}$$

where V is the collection of all the nodes and N is the total number of nodes in the network. If the path from node s to node t passes through node i, $\delta(i) = 1$, otherwise, $\delta(i) = 0$.

Since congestion usually occurs at nodes with high utilization, packet threshold R_c can be estimated as

$$R_c = \frac{N(N-1)}{\overline{U_{\max}(k)}}, \tag{6}$$

where N is the total number of nodes and $\overline{U(k)}_{\max}$ is the largest value of node utilization in the network.

4. Simulation and Empirical Analysis

4.1. *Simulation network*

A scale-free network of $N = 1000$ and $\langle k \rangle = 4$ is constructed through the algorithm of the BA network with 303 important nodes and 697 secondary nodes. The processing capacity of each node C equals to 1. Simulation is then performed.

Figure 1 depicts the variations of packet threshold R_c in the case of random classification of packets and in the case of packet classification according to source node importance under different values of parameter α_1. The simulation found, in the case of random division, the maximum value of R_c is 42 when $\alpha_1 = 0.5$, while the maximum value of R_c is 50 when $\alpha_1 = 0.4$ in the case of packet division according to source node importance. In the case of nonclassification of packets which corresponds to $\alpha_1 = \alpha_2 = 1.0$, the maximum value of R_c is 36. By comparison, the improvement

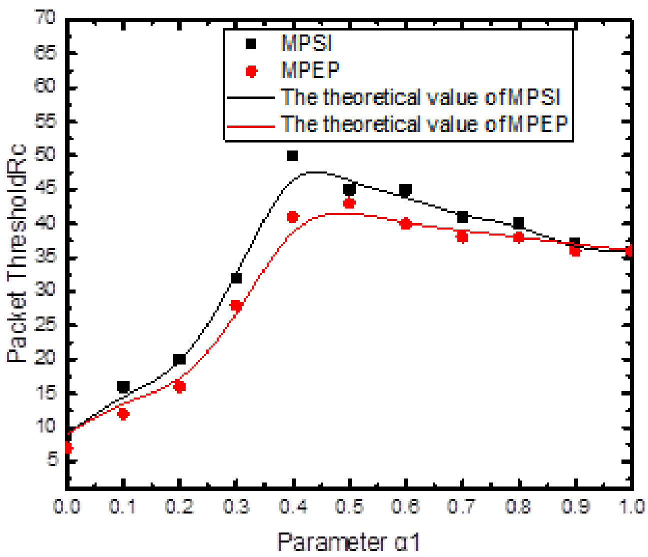

Fig. 1. (Color online) Simulated and theoretical R_c under different values of α_1.

Table 1. The average routing time required to transmit packets in a free state of the network in three different situations, namely nonclassification, random classification and classification based on source node importance.

Routing strategy	Nonclassification	Random classification	Classification based on source node importance: Packets generated at important nodes	Classification based on source node importance: Packets generated at secondary nodes
Average routing time required to transmit packets in a free state	6.2452	5.6997	4.4756	6.4037

of routing efficiency with source node importance approach is 39% when compared with the efficient path strategy[4] and 19% when compared with the random classification approach, verifying the effectiveness of the proposed routing strategy from the aspect of raising the value of R_c. Besides, the black line is the theoretical value of the strategy proposed in this paper, the red line is the theoretical value of the MPEP routing algorithm.[8] It is observed that the simulation results are quite similar to the theoretical ones.

Moreover, the average routing time required to transmit packets in a free state of the network is considered in three different situations, namely nonclassification, random classification and classification based on source node importance as listed in Table 1. It is noted that the average routing time is longer in the nonclassification approach, indicating that packets require longer time to arrive at the destination node. In the case of packet classification based on source node importance, packets generated at important nodes take less time to reach their destination and then disappear in the network, reducing the possibility of system congestion. It is also concluded from theoretical analysis that packets from the source tend to select nodes with larger degree if the value of α_1 is smaller, arriving at the destination with fewer hops. In real life, the data packets generated at the nodes such as base station and traffic hub are of more significance, which are expected to reach the destination more effectively, safely and rapidly. Fortunately, this requirement can exactly be satisfied with the strategy proposed in this paper.

The utilization of nodes with different degrees in the network under efficient path strategy is depicted in Fig. 2(a). It is noted that the utilization of nodes with medium degrees are relatively higher, while that of nodes with larger and smaller degrees are lower. Compared with the shortest path strategy, the efficient path strategy alleviates the congestion of hub nodes to some extent. However, the advantages of nodes with larger degrees are not fully utilized, resulting in the over-lengthening of the path from the source to the destination node. Figure 2(b) shows the utilization of nodes with different degrees in the network under MPEP strategy. It is observed that the utilization of nodes with larger degrees is higher compared with the case in Fig. 2(a). Figure 2(c) shows the utilization of nodes with different degrees in the network under

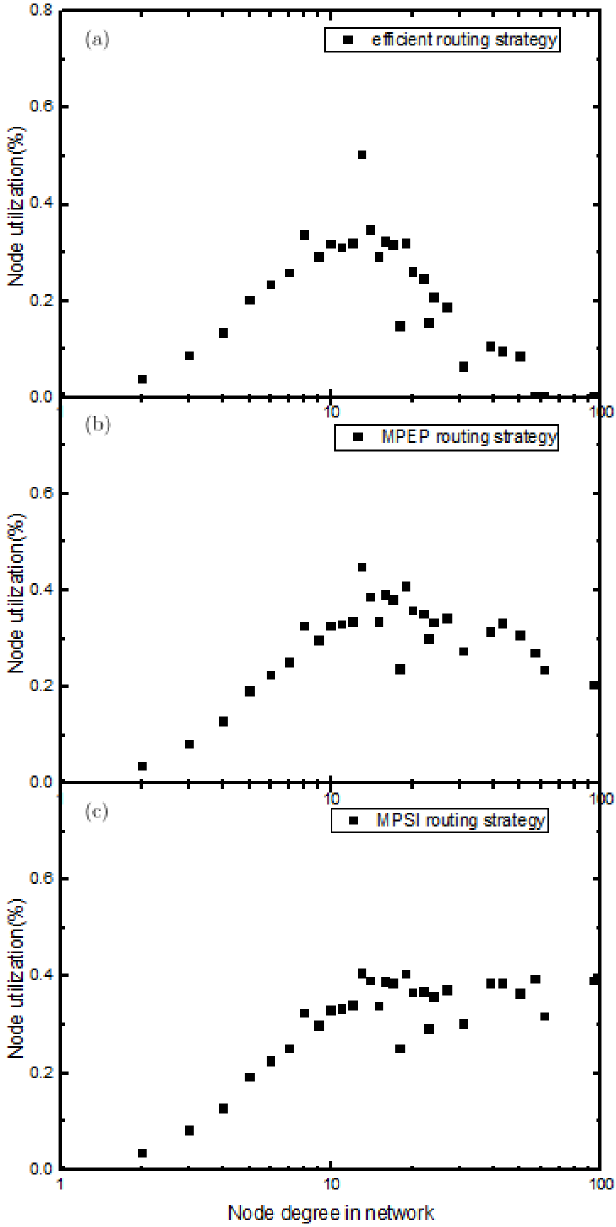

Fig. 2. (a)–(c) Utilization of nodes with different degrees under three different routing strategies: the efficient path strategy, the MPEP routing and the routing strategy proposed in this paper (MPSI: Multi-Priority routing algorithm based on Source node Importance).

the routing strategy proposed in this paper. It is noted that the utilization between nodes with different degree is quite uniform within the range of 0–0.4%. Compared with the case in Fig. 2(b) where node utilization is within 0–0.5%, the routing strategy proposed by the authors makes more uniform use of each node in

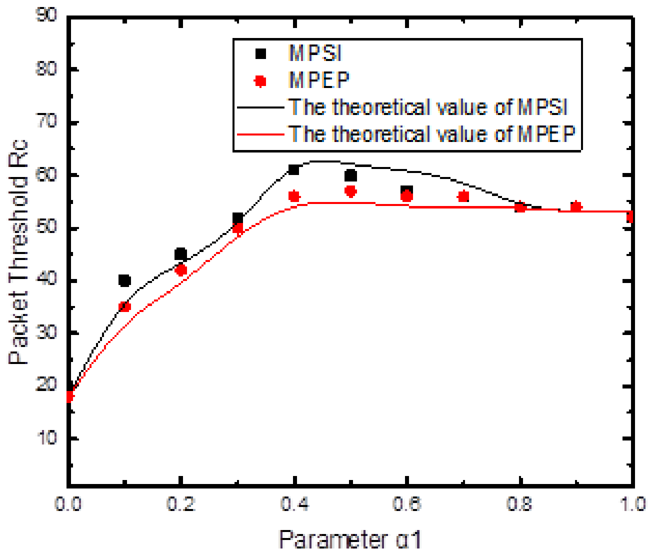

Fig. 3. (Color online) Simulated and theoretical R_c under different values of α_1 in web-EPA.

the network. Moreover, the routing strategy proposed in this paper makes full use of the utilization of intermediate and large degree nodes compared with effective routing and MPEP routing.

4.2. *Real-world networks*

In order to further verify the effectiveness of the routing strategy proposed in this paper compared with the shortest path, efficient path and random classification routing strategy in actual networks, network web-EPA[19] is selected for analysis, which contains 2221 nodes ($N = 2221$) representing routers or hosts in a communications network and 6802 edges ($K = 6802$) indicating direct packet transmission between two points. The average degree of the network is 6.1252 ($\langle k \rangle = 6.1252$).

Figure 3 describes the variations of order parameter with packet number per unit time interval R under different values of α_1 in the cases of source node importance

Table 2. The average routing time required to transmit packets in a free state of network in three different situations, namely nonclassification, random classification and classification based on source node importance.

			Classification based on source node importance	
Routing strategy	Nonclassification	Random classification	Packets generated at important nodes	Packets generated at secondary nodes
Average routing time required to transmit packets in a free state	5.8124	5.4983	4.0790	6.0367

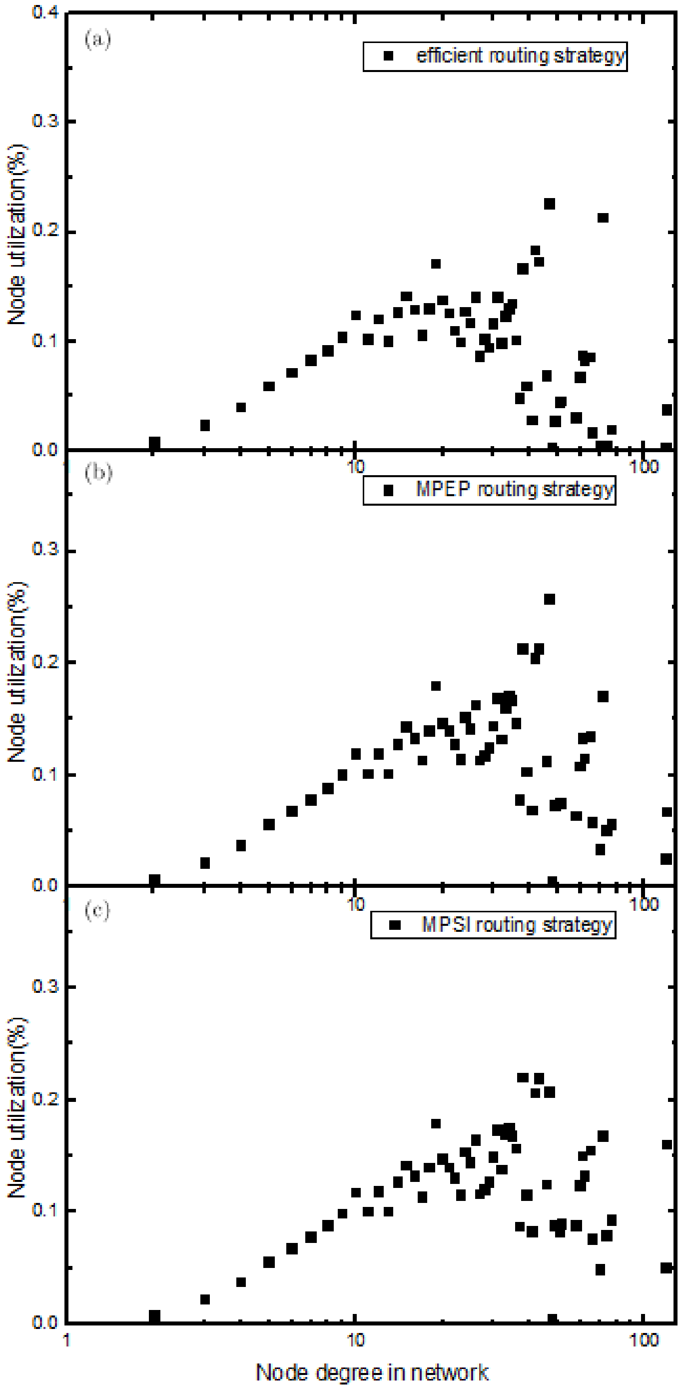

Fig. 4. (a)–(c) Utilization of nodes with different degrees under the efficient path strategy, MPEP and MPSI routing strategy in web-EPA.

classification and random classification routing in network web-EPA. It is observed that the values of R_c in the case of source node importance classification are all larger than those in random classification routing when the values of α_1 is between 0 to 0.6, with the maximum R_c being 61 and 57 at $\alpha_1 = 0.4$ under the case of source node importance classification and random classification, respectively. In the case of nonclassification of packets which corresponds to $\alpha_1 = \alpha_2 = 1.0$, the maximum value of R_c is 52. Besides, the black line is the theoretical value of the strategy proposed in this paper, the red line is the theoretical value of the MPEP routing algorithm.[8] It is observed that the simulation results are quite similar to the theoretical ones.

The results in Table 2 are quite similar to those in Table 1, also leading to the conclusions that the average routing time is longer in the nonclassification approach and the packets generated at important nodes take less time to reach their destination, relieving hub node congestion more rapidly.

The node utilization results under the three different routing strategies for an actual network are quite similar to those results represented in Fig. 2. Figure 4(a) depicts the utilization of nodes with different degrees in the actual network under efficient path strategy. It is noted that the utilization of nodes with medium degrees is relatively higher than those with larger and smaller degrees. Figure 4(b) shows the utilization of nodes with different degrees in the actual network under MPEP strategy. It is observed that the utilization of nodes with larger degrees is higher compared with the case in Fig. 4(a). Figure 4(c) shows the utilization of nodes with different degrees in the actual network under the routing strategy proposed in this paper. It is concluded that the routing strategy proposed in this paper makes more uniform use of each node in the network with the maximum value smaller than 0.23%.

5. Conclusions

This paper presents research of the communication network and its performance with complex network theory. Complex network parameters containing network structure information are used to represent the communication network effectively. In detail, this paper proposes an improved routing algorithm based on the source node importance of data packets, which classifies the nodes in the network into important nodes and secondary nodes according to their degree values. Different routing strategies are adopted according to different importance of source nodes. Moreover, the algorithm is evaluated by four evaluation criteria, namely packet threshold, order parameter, average routing time and node utilization. Compared with the shortest path, efficient path and random classification routing strategies, the proposed routing algorithm is more effective in enhancing the system capacity and routing efficiency, as well as making node utilization more average in both artificial and actual networks. In this paper, the proposed classification approach is based on core/periphery structures, which has been presented in relative researches of complex networks.[21]

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China under Contract Nos. 11075057 and 11875133. We also acknowledges the National Key R&D Program of China under Grant No. 2018YFB2101302.

References

1. R. Guimer, A. Daz-Guilera, F. Vega-Redondo, A. Cabrales and A. Arenas, *Phys. Rev. Lett.* **89**, 248701 (2002).
2. V. Cholvi, V. Laderas, L. Lpez and A. Fernndez, *Phys. Rev. E* **71**, 035103 (2005).
3. A. Arenas, A. Daz-Guilera and R. Guimera, *Phys. Rev. Lett.* **86**, 3196 (2001).
4. G. Yan, T. Zhou, B. Hu, Z. Q. Fu and B. H. Wang, *Phys. Rev. E* **73**, 046108 (2006).
5. L. Chen, J. Chen, Z. H. Guan, X. H. Zhang and D. X. Zhang, *Physica A* **391**, 3336 (2012).
6. B. Lin, B. Chen, Y. Gao, C. K. Tse, C. F. Dong, L. X. Miao and B. H. Wang, *Plos One* **11**, e0156756 (2016).
7. F. Tan and Y. Xia, *Physica A* **392**, 4146 (2013).
8. X. Zhang, Z. Zhou and D. Cheng, *Plos One* **12**, e0172035 (2017).
9. A. L. Barabsi and R. Albert, *Science* **286**, 509 (1999).
10. D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
11. G. Siganos, M. Faloutsos, P. Faloutsos and C. Faloutsos, *IEEE/ACM Trans. Netw.* **11**, 514 (2003).
12. S. Zhou and R. J. Mondragn, *IEEE Commun. Lett.* **8**, 180 (2004).
13. J. Wu, K. T. Chi, F. C. M. Lau and I. W. H. Ho, *IEEE Trans. Circuits Syst. I Regul. Pap.* **60**, 3303 (2013).
14. J. Wu, K. T. Chi and F. C. M. Lau, *IEEE Trans. Circuits Syst. I Regul. Pap.* **62**, 1195 (2015).
15. M. Takayasu, H. Takayasu and K. Fukuda, *Physica A* **277**, 248 (2000).
16. D. D. Han, Q. Q. Yao, Q. Chen and J. H. Qian, *Acta Phys. Sin.* **66**, 267 (2017).
17. D. D. Han, J. H. Qian and Y. G. Ma, *Experimental Study of Topological and Dynamic Behaviors of Complex Networks* (Peking University Press, Beijing, 2012).
18. M. Tang and T. Zhou, *Phys. Rev. E* **84**, 026116 (2011).
19. Ryan A. Rossi and Nesreen K. Ahmed, The network data repository with interactive graph analytics and visualization, *AAAI*, 2015, http://networkrepository.com.
20. S. P. Borgatti and M. G. Everett, *Soc. Netw.* **21**, 375 (1999).
21. L. Zhao, Y. C. Lai, K. Park and N. Ye, *Phys. Rev. E* **71**, 026125 (2005).