Identifying key component in high-speed train based on complex network

Baojin Wang¹,a,Yanhui Wang²,b,Yucheng Hao³,c

¹Qingdao Sifang Rolling Stock Co. Ltd.CRRC Qingdao, China
²State Key Laboratory of Rail Traffic Control and Safety Beijing Jiaotong University
Beijing, China
³School of Traffic and Transportation Beijing Jiaotong University Beijing, China

a714260578@qq.com, b* Corresponding author: 18801007426@163.com,
c18114021@bjtu.edu.cn

Abstract—The identification of key components is very important for the real-life system. Because the complex network is a powerful tool to analyze the behavior of the complex network, we explore the key nodes by using the complex networks. First, we utilize the degree, the betweenness, the closeness centrality, the eigenvector centrality and the Pagerank to reflect the importance of the node. Then the machine components are abstracted as a node, and the interactions are abstracted as an edge. According to the topology structure of the high-speed train, it is found that the high-speed train has the scale-free feature. By removing the nodes in the light of the sequence of nodes, we can obtain the effective measure and analyze the key nodes of the high-speed train performance degradation caused by the failure. The simulation results show that the node with the high PageRank value and degree is critical for the network. This work may contribute to the discovery of the key components in the high-speed train.

1. INTRODUCTION

The real-world complex system plays a key role in the modern society. As a critical system in high-speed railway, the reliability of the high-speed train decides whether the transportation mission can be finished safely. If a key component fails to works due to the failure, it is likely to have a serious impact on the operation of the high-speed train, such as the reduction of the speed, fire. Therefore, finding the key component is meaningful for improving the reliability of the high-speed train and decreasing the cost of the maintenance.

As we all know, there are a lot of theories and models for the assessment reliability, such as a fault tree, prognostics and health management. However, these methods hardly find the key component and cannot evaluate the whole performance of the high-speed train. With the development of network science, more and more researchers have started to use the complex network to assess the reliability. First, Albert et al. attacked some nodes in networks and found that the failures of server nodes tend to the paralysis of the whole network [1]. Inspired by this work, the vulnerability of complex networks was studied subject to attacks on nodes and edges [2]. After these works, many studies concerning the vulnerability and robustness were performed.

In terms of real-life systems, Xu et al. adopted complex network theory to analyze the connectivity of railway networks, finding that the physical connectivity hardly reflects its connectivity while the
logical connectivity is able to capture the characteristic of the railway networks [3]. Qian et al investigated the classic road traffic network and the impact of some characteristics (including the time delay, restorative and so on) on the network robustness [4]. Similarly, considering the phenomenon of cascading failures, Wu et al. also put forward a more practical traffic flow assignment mechanism to explore the weighted urban traffic equilibrium networks based on the complex network [5]. In addition, the ability of the public transportation system in Beijing against random failures and intentional attacks was assessed [6]. The results revealed that Beijing subway system has the feature of the classic scale-free network and its dynamic process also was provided. In the field of the air transport, Du et al. studied the Chinese Airline Network by the means of “k-core decomposition” method [7]. Their results showed that the network keeps the most robust when nodes with the low degree or links with high flight flow are removed. Lin et al. proposed a reliability assessment framework concerning the complex electromechanical systems from a network and the percolation theory [8, 9]. Based on the proposed approaches, the reliability of the high-speed train was assessed and the effectiveness of the proposed method was verified. In addition, the reliability of power grid [10, 11], water distribution network [12, 13], supply chain network[14-16] also has been attracted a lot of attentions in the light of complex network.

Although there are many researches on the assessment performance of the complex system, few works paid attention to the evaluation safety and the identification of key component in the high-speed train [17-20]. As a key system, the high-speed train undertakes the mission to achieve the transport of passengers. Therefore, from the perspective of the network, this paper investigates the characteristic of the high-speed train. Then five measures that quantify the importance of components are adopted and compared in the network. Based on the order of the measures, we attack the corresponding nodes and analyze the impact of the failure of the node on the network. According to the simulation, we can obtain the effective measure and the key nodes.

2. Model

As mentioned in the previous discussion, various kinds of complex electromechanical systems have been defined by networks, which includes a set of nodes and edges. There are many node characteristics for measuring the node importance and a node has a wide variety of features. In this paper, the degree, the betweenness, the closeness centrality, the eigenvector centrality and the Pagerank are adopted. The degree \( k_i \) of node \( i \) is shown as

\[
k_i = \sum_{j=1}^{n} a_{ij}
\]

(1)

where \( a_{ij} \) represent the element of adjacency matrix. When the node \( i \) connects with the node \( j \) , \( a_{ij} = 1 \); otherwise \( a_{ij} = 0 \). \( n \) is the number of the nodes in the network.

The betweenness \( b_i \) of node \( i \) is shown as

\[
b_i = \sum_{pq} \frac{\sigma_{pq}(i)}{\sigma_{pq}}
\]

(2)

where \( \sigma_{pq}(i) \) denotes the number of the shortest paths from node \( p \) to node \( q \) through node \( i \) . \( \sigma_{pq} \) denotes the total number of the shortest paths from node \( p \) to node \( q \).

The closeness \( c_i \) of node \( i \) is shown as

\[
c_i = \frac{1}{\sum_{j=1}^{n} d_{ij} s_j}
\]

(3)

where \( d_{ij} \) denotes the shortest length form node \( i \) to node \( j \).

The eigenvector centrality \( e_i \) of node \( i \) is shown as
where $\Phi_i$ denotes the set of the adjacent nodes of node $i$. $\lambda$ is a parameter.

The PageRank $pr_i$ of node $i$ is shown as

$$pr_i = (1-d) + d \sum_{j \in \Phi_i} pr_j / k_i$$

(5)

where $d$ is a parameter and equals 0.85 usually.

On the basis of the above the definitions of measures for the node, we can find that these measures are capable of reflecting its importance from different perspectives. The degree takes into account the connections with other nodes. The betweenness and the closeness can reflect the importance of the node in the whole network. Additionally, by considering the impact of the adjacent nodes, the eigenvector centrality and the PageRank the local information on the node.

According to these measures, we sort the nodes and attack them so as to analyze the impact of nodes on networks. In general, the more important the node, the more serious the impact of its failure. Note that with the increase of the number of the attacked nodes, there may exist some single nodes. In this case, although these single nodes are not attacked, due to without the connection with other nodes, these nodes are also considered to be inactive. To this end, we adopt two measures to quantify the robustness of the network subject to attacks on nodes, i.e., the network efficiency and the relative size of the giant component. The network efficiency $ne$ is defined as follows

$$ne = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_l}$$

(6)

The relative size $r$ of the giant component is defined as follows

$$r = \frac{n_g}{n}$$

(7)

where $n_g$ represents the number of nodes in the largest connected cluster.

3. RESULTS

For the safety and the reliability, there are thousands of components to work, but considering the simplicity of the simulation model, we chose a machine component (i.e., a minimum maintenance unit) as a node. If two components are connected by the screw, there exists an edge between the corresponding nodes. According to the interaction between machine components, we can obtain the network with 142 nodes and 181 edges. By computing the degree of nodes, we can obtain the degree distribution of nodes in Fig.1.
It can be seen in Fig. 1, most of the nodes (close to 75% of nodes) have the small degree, while there are few nodes (close to 8% of nodes) with the degree over 12. This result shows that the network built by the high-speed train has an obvious scale-free feature. That is to say, there are a lot of nodes with the less connections, but there exist few high degree nodes. Many previous works have proved that a lot of real-world systems have the scale-free feature, such as the power grid, the communication networks, Internet, social networks and so on. In terms of the network with the scale-free feature, although we know that the number of the node with the high degree is small, if these nodes fail to work, the performance of the network will be greatly affected and the vulnerability of the network is obvious, which has been illustrated in Ref. 1. Therefore, it is found that few components are crucial for the high-speed train, and preventing their failures plays a key role in the maintenance of the reliability.

The degree is able to reflect the local information as an important index. In order to explore the relationship between the other four measures and the degree, we plot the degree-betweenness correlation, the degree-closeness correlation, the degree-eigenvector centrality correlation, and the degree-Pagerank correlation.
From Fig. 2, we can find that the betweenness, eigenvector centrality and PageRank have a positive correlation with the degree, which means that for most of the nodes, the larger the value of the degree, the larger the value of the betweenness, eigenvector centrality and PageRank. It must be noted that with the increase of the degree, the value of the PageRank also increases steadily. Namely, there exists a strong linear relationship between the degree and the PageRank. These results indicate that the sequence of the attacked nodes in the light of the degree is similar to the one of attacked nodes in the light of the Pagerank. Then, the betweenness also has an obvious linear relationship with the degree. On the contrary, in terms of the closeness and eigenvector, it is clear that there exists a slight linear relationship with the degree and the closeness, the eigenvector centrality. For some nodes with larger values of the closeness and the eigenvector centrality, their degree are between 10 and 20.

On the basis of the values of the degree, the betweenness, the closeness centrality, the eigenvector centrality and the Pagerank of nodes, we attack a node and calculate the network efficiency and the relative size of the giant component under different numbers \( f \) of attacked nodes, which is shown in Fig. 3.

![Figure 3. \( ne \) as a function of the number of attacked nodes under different attack strategies](image)

As shown in Fig. 3, we can find that in terms of \( ne \), attacking nodes by the degree and the PageRank has a great impact on the network in the case of \( f < 30 \). This is because that the failure of the node with the high degree leads to the malfunctions of other nodes that only have an edge due to without the edge. When the value of \( f \) increases to over 30, the PageRank makes the value of \( ne \) lower. In this case, the only attack on 50 nodes with the higher PageRank value makes the entire network disconnected. This result shows that the node with the higher PageRank value is the key node for the whole network. The linear relationship between the degree and the PageRank also is proved based on the finding that the malfunction of the node with the high degree and the high PageRank results in the significant reduction of \( ne \). If these nodes fail to operate, the function of the whole high-speed train will be significantly impaired. Therefore, the reliability of these components should be enhanced for the examination, the repair and the design in case of their failures. In addition, the component that has many interactions with other components also should be drawn attention.

In terms of five measures, the node with the high closeness is not important because the value of \( ne \) in the network slightly decreases after it is removed. Until the value of \( f \) increases to near 140, the value of \( ne \) decreases to zero, which means that the high closeness node is not very critical. This is because that the high closeness node is near the center of the network. If it is removed, the adjacent nodes also have the connection to keep their functions.
Figure 4. $r$ as a function of the number of attacked nodes under different attack strategies

According to different attack orders, we can calculate the value of $r$, whose result is shown in Fig.4. It can be seen that attacking the nodes with the high degree and the PageRank values can make the value of $r$ greatly decreased in Fig.4, which means that the number of the functional nodes reduces significantly. On the contrary, the node with the high betweenness and the eigenvector centrality cannot have an obvious effect on most of the active nodes in the network. In addition, attacking the nodes by means of the closeness is not an effective attack strategy. The reason is that the adjacent node of the high closeness node has the more connections. Even though the node with the high closeness is removed, its adjacent nodes still keep operations by their other edges. It is found that the analysis of Fig.4 is in agreement with that of Fig.3, which also indicates that our simulation result is reliable.

4. CONCLUSION

In this paper, we adopt the complex network to model the high-speed train by considering the machine components and the interactions between them. The results regarding the topology structure show that the high-speed train exhibits the scale-free feature. In the case of five measures including the degree, the betweenness, the closeness centrality, the eigenvector centrality and the Pagerank for nodes, it is found that the node with the high degree and the high PageRank value is critical for the network. However, the high closeness node has little impact on the network when it fails to work. Our result is likely to be helpful for the design, and maintenance for the high-speed train.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support provided by the National Key R&D Program of China under grant (2016YFB1200402). We would like to thank the anonymous reviewers for their constructive comments and suggestions, which helped us to improve the paper.

REFERENCES

[1] Albert R., Jeong H., Barabasi A.L., "Error and attack tolerance of complex networks," Nature, vol. 340, no. 1(2000), pp. 378-382.

[2] Holme P., Kim B.J., Yoon C.N., Han S.K., "Attack vulnerability of complex networks," Physical Review E Statistical Nonlinear & Soft Matter Physics, vol. 65, no. 2(2002), pp. 56109.

[3] Xu Z, Zhang Q, Chen D, He Y, "Characterizing the Connectivity of Railway Networks," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 4(2020), pp. 1491-1502.
[4] Qian Y., Wang B., Xue Y., Zeng J., Wang N., "A simulation of the cascading failure of a complex network model by considering the characteristics of road traffic conditions," Nonlinear Dynamics, vol. 80, no. 1-2(2015), pp. 413-420.

[5] Wu J.J., Sun H.J., Gao Z.Y., "Cascading failures on weighted urban traffic equilibrium networks," Physica A, vol. 386, no. 1(2007), pp. 407-413.

[6] Yang Y., Liu Y., Zhou M., Li F., Sun C., "Robustness assessment of urban rail transit based on complex network theory: A case study of the Beijing Subway," Safety Science, vol. 79, no. (2015), pp. 149-162.

[7] Du W., Zhou X., Lordan O., Wang Z., Zhao C., Zhu Y., "Analysis of the Chinese Airline Network as multi-layer networks," Transportation Research Part E: Logistics and Transportation Review, vol. 89, no. (2016), pp. 108-116.

[8] Lin S., Wang Y., Jia L., Zhang H., "Reliability assessment of complex electromechanical systems: A network perspective," Quality and Reliability Engineering International, vol. 34, no. (2018), pp.

[9] Lin S., Wang Y., Jia L., "System Reliability Assessment Based on Failure Propagation Processes," Complexity, vol. 2018, no. (2018), pp. 1-19.

[10] Wang J.W., Rong L.L., "Cascade-based attack vulnerability on the US power grid," Safety Science, vol. 47, no. 10(2009), pp. 1332-1336.

[11] Kinney R., Crucitti P., Albert R., Latora V., "Modeling cascading failures in the North American power grid," The European Physical Journal B - Condensed Matter and Complex Systems, vol. 46, no. 1(2005), pp. 101-107.

[12] Shuang Q., Zhang M., Yuan Y., "Performance and Reliability Analysis of Water Distribution Systems under Cascading Failures and the Identification of Crucial Pipes," Plos One, vol. 9, no. 2(2014), pp. e88445.

[13] Shuang Q., Zhang M.Y., Yuan Y.B., "Node vulnerability of water distribution networks under cascading failures," Reliability Engineering & System Safety, vol. 124, no. 124(2014), pp. 132-141.

[14] Liang T., Jing K., He J., Stanley H.E., "Complex interdependent supply chain networks: Cascading failure and robustness," Physica A Statistical Mechanics & Its Applications, vol. 443, no. (2016), pp. S1043376969.

[15] Kim Y., Chen Y., Linderman K., "Supply network disruption and resilience: A network structural perspective," Journal of Operations Management, vol. 33-34, no. (2015), pp. 43-59.

[16] Wang Y.C., Zhang F.P., "Modeling and analysis of under-load-based cascading failures in supply chain networks," Nonlinear Dynamics, vol. 92, no. (2018), pp. 1-15.

[17] Lin S., Wang Y., Jia L., Zhang H., Li Y., "Intuitionistic Mechanism for weak components identification method of complex electromechanical system," Journal of Intelligent & Fuzzy Systems, vol. 34, no. (2017), pp. 1-16.

[18] Lin S., Jia L., Wang Y., Zhang H., "A New Function-topology-based Method for Assessing Passive Safety of Mechatronics Systems," IEEE Access, vol. PP, no. (2020), pp. 1.

[19] Li M, Wang Y, Jia L, "A Research Into the Reliability of Equipment-Integrated System Regarding High-Speed Train Based on Network Model," IEEE Access, vol. 7, no. (2019), pp. 186328-186339.

[20] Wang Y., Cui Y., Li M., Wang S., "On identification method of key components of mechatronics system based on network model," Journal of Intelligent & Fuzzy Systems, vol. 36, no. (2019), pp. 3273-3287