ABSTRACT

Title of Thesis:	TOPOLOGICAL ANALYTICS FOR VULNERABILITY ENHANCEMENT AND RECOVERY STRATEGY AFTER DISRUPTIONS OF RAIL NETWORKS IN THE UNITED STATES
	Siqi Cao, Master of Science, 2020
Thesis Directed By:	Professor Bilal M. Ayyub, Department of Civil and Environment Engineering

Rail networks are real-life examples of complex networks and critical logistic and economic contributors to the wellbeing of society. Natural or human-caused hazards leading to the disruptions of rail network's components can cause severe consequences including significant economic impacts. Therefore, analyzing rail networks and further reducing the impacts of potential disruptions are critical in order to manage risks to the performance of rail networks. Based on existing research on rail networks, this thesis proposes a methodology to analyze the rail networks with a large number of nodes, links, and complex connectivity from topological perspectives. Additionally, topology enhancement prior to failures and recovery strategies post to failures are used to reduce the impacts of potential failures based on vulnerability and resilience assessments. The analysis results of two case studies, the Amtrak and Class I rail networks, indicate that the proposed methodology is well suited to analyze and enhance the topology, vulnerability, and resilience of complex rail networks effectively and efficiently.

TOPOLOGICAL ANALYTICS FOR VULNERABILITY ENHANCEMENT AND RECOVERY STRATEGY AFTER DISRUPTIONS OF RAIL NETWORKS IN THE UNITED STATES

by

Siqi Cao

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2020

Advisory Committee:
Professor Bilal M. Ayyub, Chair
Professor Amde M. Amde
Professor Michelle (Shelby) Bensi

Acknowledgements

Firstly, I wish to express my sincere gratitude to my advisor, Professor Bilal Ayyub, for his continuous support of my thesis research at the University of Maryland, College Park, for his guidance in academy and writing, and for his patience. I would not been able to complete my thesis research without his help. I hope that all the work I do can contribute to his future research.

I would also like to express my deepest appreciation to my thesis committee, Professor Amde and Professor Bensi, for taking the time to review my work. I want to thank every professor who instructed me in my graduate study. They taught me to deal with every academic issue seriously. I also wish to express my sincere appreciation to an outstanding Ph.D. student, Yalda Saadat. Thanks to her valuable suggestions and comments on the technical and writing parts of my thesis.

I must express my profound gratitude to my parents and girlfriend for all their love, support, and encouragement. Thank you for making my dreams come true. My closest friends, Naiyi, Sally, Yujie, and Zhihao, thanks for being there for me.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Railroads serve as an effective and reliable transportation mode that is a critical logistic, social, and economic contributor to the wellbeing of society. Natural and human-caused hazards can seriously affect every component of a railroad, especially if it leads to the failure of rails and stations. This thesis mainly focuses on the analysis of rail networks composed of rails and stations to measure vulnerability and resilience from network topological perspectives. Such an analysis informs decisions on rail network enhancement and recovery strategies in order to reduce the impacts of disruptions on rail networks consisting of rails and stations, and more broadly on railroad systems. This chapter starts by highlighting the significance and criticality of railroads, and then defines the work breakdown structure of a typical railroad system. A rail network, a real-life example of complex networks consisting of two infrastructures component types of rails and stations, is also introduced. Finally, the needs in studying the rail network are identified as a basis for the research performed in this thesis.

1.1 Significance and Criticality of Railroads

1.1.1 Significance of Railroads

A key mode of transporting passengers and freight is by trains on rails. The railroad in the United States (U.S.) is recognized as one of the most complete and dynamic worldwide. In the 1990s, the ridership of U.S. railroad passenger transportation was more than 20 million people a year (Morrison 1990), while the ridership had increased more than 1.5 times by 2016, exceeding 30 million (Amtrak 2016). Increasingly Americans choose to travel by trains, and the annual

ridership is still growing. The transportation of goods relies more on freight railroads than other transportation modes. For example, in 2019, U.S. freight railroads transported more than 4 million carloads of coal, which exceeded 70 percent of U.S. coal to its destination. This coal transported by freight railroads was used to power 78 percent of the electricity in American homes (Association of American Railroads 2019). Hence, railroads, which have undertaken the majority of passenger and freight transportation, are an important economic and social contributor to the wellbeing of a society. In 2019, the U.S. railroad supported more than 1 million jobs, earned more than $91 billion in revenue, nearly $71 billion wages, and paid almost $26 billion in taxes (Association of American Railroads 2019).

1.1.2 Criticality of Railroads in Case of Disruptions

Railroads are vulnerable to natural and human-caused hazards. For example, hurricane Alicia in 1983 caused $3 billion in damage in southeastern Texas. The greatest impact of this hurricane was the inability of the railroad to perform its normal functions, i.e., the failure of the railroads. The storm surge caused the failure of 6km freight railroads in the area between Texas City Junction and Virginia Point. Additionally, railroads in other parts of Texas were out of service for four days (Byers 2011). Railroads are vulnerable to disruptions caused by natural hazards, such as hurricanes and earthquakes. These natural events can easily lead to the failure of rails, stations, power supply, trains, and operation systems. Once railroad components fail at a location, all railroad operations connected to this location are affected. As a consequence, a large area of railroad passenger transportation and the flow of cargo are delayed or even stopped. Chinowsky et al. (2019) indicated that in the U.S., increased temperatures due to climate change could cause rail deformation, causing delay-minute costs between $35 and $60 billion (in 2019

values) accumulated through 2100 under a high greenhouse gas emission. In China, the 2008 ice storm caused severe disruptions of railroads in southern China. The railroad passenger transportation throughout China almost shut down. Over 5 million travelers were delayed; the direct economic losses totaled $22.3 million in the fiscal year 2008 (Chen and Wang 2019). Janić (2018) reported that the 2011 Tohoku Earthquake caused the failure of Japan Shinkansen railroads in the northeast region. The entire transportation service of Japan Shinkansen railroad was delayed, causing a total cumulative cost for 92 days (from 11/04/2011 to 08/07/2011) of nearly $12.2 million (estimated values in June 2012).

1.2 Railroads in the United States

1.2.1 Passenger Railroads: Amtrak

The National Railroad Passenger Corporation, usually called Amtrak, is a for-profit, quasi-public corporation founded in 1971. Amtrak railroads contain three types of rails: long-distance, state-supported, and the northeast corridor, which cumulatively serve more than 510 stations in North America, including 46 U.S. states, Washington D.C and three Canadian provinces. The entire Amtrak route is more than 21,400 miles in length. As an important transportation mode for the American people, customers made more than 87,000 trips on an average day in 2018. Over 300 trains were on the rail every day (Amtrak 2018). Increasingly people are choosing to travel by train. According to data from Amtrak, the ridership had a straight five-year increase from 2015 to 2019.

Railroad passenger transportation plays a critical role in the U.S. economy by promoting economic activities in regions, especially between big cities. For example, more than two-thirds of the trip occurs between Boston-New York-Washington D.C., the most important political and economic regions in the U.S. (Morrison 1990). Additionally, the Amtrak enterprise employs more than 20,000 workers, with additional jobs are being created as railroads develop and expand (Amtrak 2018). Depending on the GAAP (Generally Accepted Accounting Principles), Amtrak generated $3.4 billion in revenue in the fiscal year 2018, an increase of 4.5 percent over the fiscal year 2016 (Amtrak 2018). Figure 1.2 shows the revenue of Amtrak in the last five years. In 2012, the federal government invested $1.42 billion dollars on Amtrak (Peterman 2017), representing 0.4 percent of all federal nondefense investments. In short, the Amtrak railroad is vital for public services, economic development, and government operations.

Amtrak, as a railroad covering the entire United States, is vulnerable to natural and human-caused hazards. For example, the 2005 Hurricane Katrina caused the failure of the Amtrak railroad in New Orleans. Important Amtrak operations through New Orleans were forced to stop, such as the operation of routes from New Orleans to Chicago, from New Orleans to New York, and from Orlando to Los Angeles. These nationwide operations were recovered after completing repairs to rails, bridges, and other infrastructures necessary for travel (DesRoches 2006). In addition, over 60 railroad failures happened in the Seattle-Vancouver's Amtrak operation between 2009 and 2013. The Amtrak operation from the northwestern United States to Canada was canceled or delayed more than 15 times a year (Azad et al. 2016). Amtrak disruptions cost over $41 million (in 2019 values) losses per year.

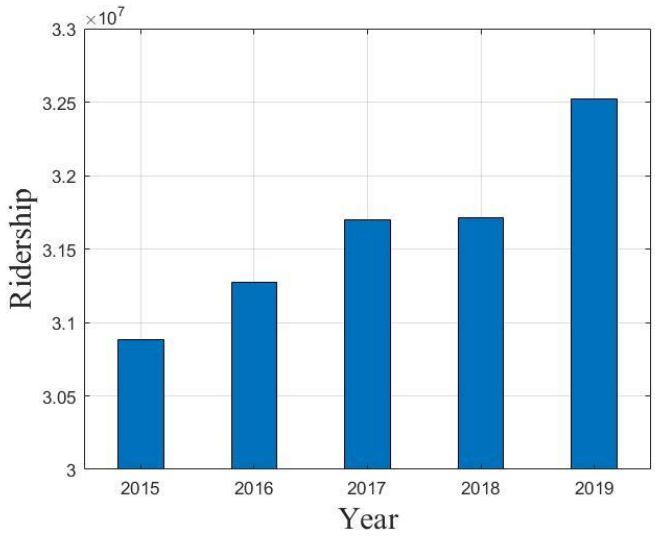

Figure 1.1 The Amtrak ridership from 2015 to 2019.

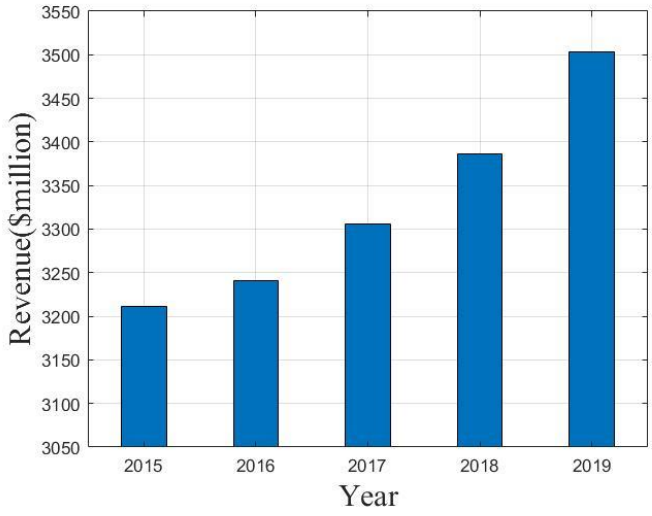

Figure 1.2 The Amtrak revenue from 2015 to 2019.

1.2.2 Freight Railroads: Class I

The entire freight railroad system in the U.S. consists of 136,898 rail miles, including more than 500 local railroads and 21 regional railroads. The freight railroad can be divided into three

categories as Class I, II, and III, depending on the annual gross revenue of the entity to which the freight railroads belong. In 1991, the Surface Transportation Board (STB) described the Class I railroad as "having the annual carrier operating revenue of $250 million or more in 1991 dollars." However, STB published in 2011 (Federal Register 2011) the annual inflation-adjusted index factors to update the annual gross revenue for classification resulting in increasing the $250 million threshold to $467.1 million by 2013 (Lawrence 2015). Currently, seven railroad entities in the U.S. are designated as Class I: (1) CSX Transportation, (2) Grand Trunk Corporation (held by Canadian National Railroad), (3) Kansas City Southern Railroad, (4) Norfolk Southern Railroad, (5) BNSF Railroad, (6) Soo Line Railroad (held by Canadian Pacific Railroad), and (7) Union Pacific Railroad. Figure 1.3 shows the annual revenue of Class I railroads. The entire industry created over $90 billion in revenue in the fiscal year 2018, which is close to 30 times that of the Amtrak railroad.

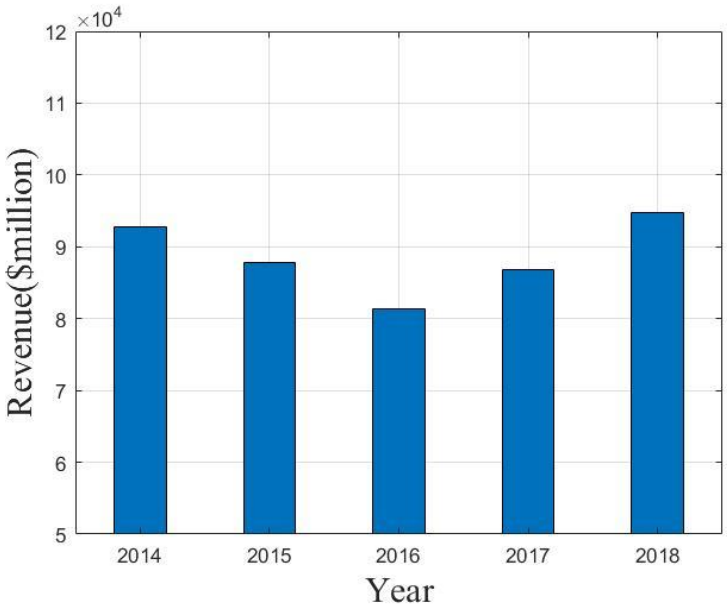

Figure 1.3 The annual revenue of Class I freight railroads from 2014 to 2018.

The significance of the Class I railroad is not only reflected by its revenue but more importantly, it is the primary transportation mode for freights. Five principal modes are used to transport freights in the United States: Class I railroads, trucking, pipelines, waterways, and airfreight. According to the National Rail Plan Progress Report (Federal Railroad Administration 2010), in 2010, Class I railroads carried nearly 40 percent of the United States freight by ton-miles. Figure 1.4 shows that Class I railroads can be considered the most crucial transportation mode. Class I railroads transport some of the most needed daily commodities, and the necessities of industrial manufacturing. For example, in 2019, U.S. Class I railroads moved 1.6 million carloads of agricultural and food products, 4 million carloads of coal and 2.4 million carloads of chemicals (Association of American Railroads 2019).

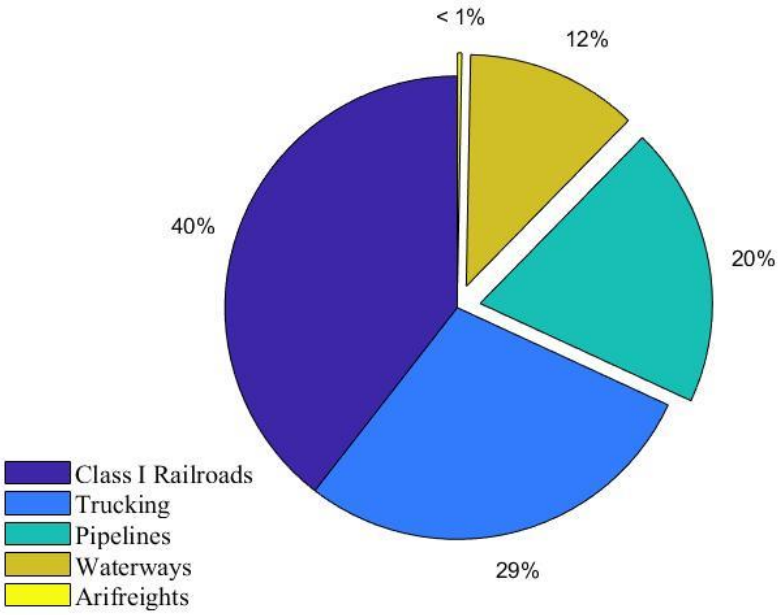

Figure 1.4 Percent ton-miles of each transportation mode in 2010.
Data Resource: National Rail Plan Progress Report (Federal Railroad Administration 2010).

The Class I railroad contributes the majority of the total railroad revenue and is one of the most critical transportation modes, with great economic and fiscal impacts on the United States. In the fiscal year 2017, Class I railroads generated almost $26 billion in tax (Irani et al. 2018). According to a report from Towson University (Irani et al. 2018), the economic impact can be measured in three types:

- The direct economic impact of generated jobs. In 2017, Class I railroad employed 90 percent of 1.1 million U.S. railroad workers and supported $71 billion in wages (Association of American Railroads 2017);
- The indirect economic impact of generated production and related services that can be purchased from other companies;
- The increased employment and increased income levels for households, which results in an increase in household purchases from local businesses.

Class I railroads, as a complex railroad system covering the entire United States, are also vulnerable to natural and human-caused hazards. In addition to the damage to Amtrak's rails in New Orleans, Hurricane Katrina in 2015 also caused the failure of Class I railroads. As a consequence, the Norfolk Southern railroad operations from Slidell, Louisiana to Shrewsbury, Massachusetts, from New Orleans to the Port Nickel and other Norfolk Southern railroad operations connected to New Orleans were forced to stop and switch to other railroads (DesRoches 2006). Class I railroad operations in the entire eastern region were greatly affected and delayed. The flooding of the Mississippi and Missouri rivers in 1993 caused Class I railroads to fail. The estimated losses of Class I freight railroad due to this flooding were more than $182 million (estimated values in 2014) (Gedik et al. 2014).

1.3 Railroad Systems and Rail Networks

1.3.1 Railroad Systems

Every component of railroads can be affected by natural and human-caused hazards (Batarlienė 2008). Therefore, the composition of the entire railroad must be understood first. A system can be described as an interdependent group of components building a unified whole (Ayyub 2014). The main components of a railroad system are trains, entities, infrastructures, users and environment. The most basic level of the railroad system is the trains, which are a set of wheeled vehicles moving along the rail. The powered vehicle that pulls the train is called a locomotive. And other wheeled vehicles used for the hauling of freight or passengers are called railroad cars. The entire railroad is owned, operated, and maintained by entities. Therefore, the railroad entity is also an important part of a railroad system. The most important component in a railroad system is infrastructure. Infrastructures in a railroad consist of rails, stations, train inspection, signaling, and electrification equipment. The users of the railroad mainly include passengers and producers, such as agricultural and food, coal, construction-related, and chemical producers. Finally, the environment where a railroad is located is composed of topography, geological conditions, climate, and weather. The work breakdown structure of a typical railroad system is shown in Figure 1.5.

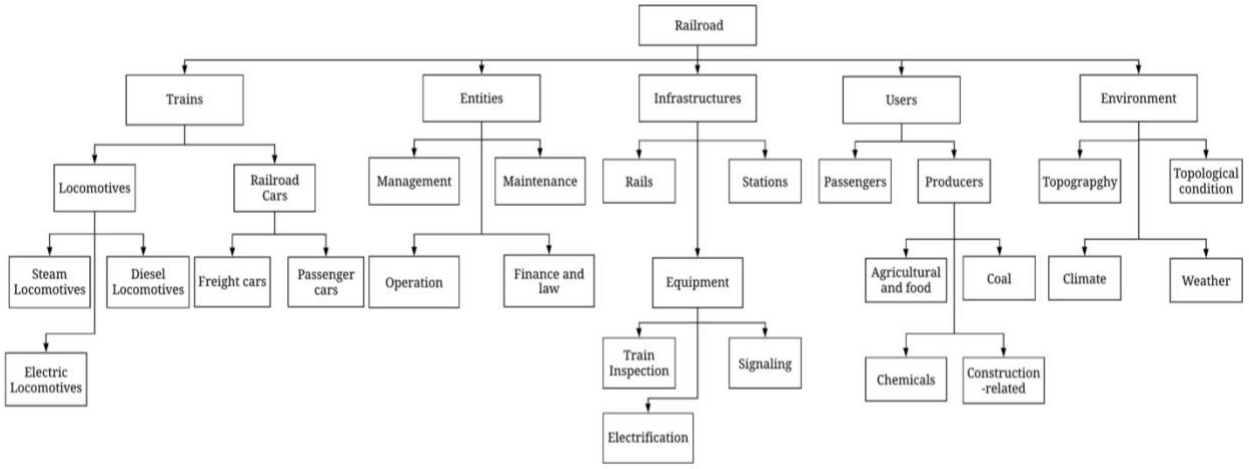

Figure 1.5 Work breakdown structure of a railroad system.

1.3.2 Rail Networks

The failure of rails and stations is a central and serious situation for a railroad system leading to disruptions. For example, the 2005 Hurricane Katrina destroyed most of Class I and Amtrak rails in New Orleans as well as the New Orleans Station. As a consequence, almost all railroad operations connected to New Orleans in the southeast U.S. were forced to stop for one week and were fully recovered after all damaged rails and stations were repaired (DesRoches 2006). In order to focus on the disruptions of rails and stations, the notion of a rail network is introduced.

A network includes two basic components: vertices or nodes, and the connections between them, called edges or links. Figure 1.6 shows a simple network in which red dots represent nodes, and black lines represent links. Most networks are defined and mapped from physical connections between a set of items (Barabási and Pósfai 2016). For instance, the nervous system in mammalian brains, consisting of more than 100 billion neurons, is modeled by scientists as a complex network. The cell called neuron is modeled as nodes, such as red dots in Figure 1.6,

while the axon connecting neighboring cells are modeled as links, such as black lines in Figure 1.6 (Barabási and Pósfai 2016).

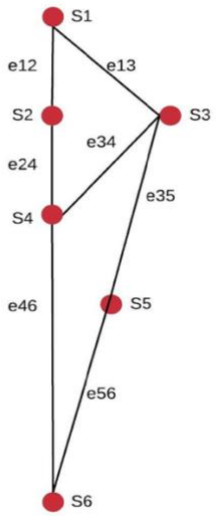

Figure 1.6 Graph of a network.

A transportation network, especially the rail network, is a real-life example of complex networks. Stations are infrastructures where passengers board or alight from trains, and freights can be loaded or unloaded. Stations can be directly mapped as nodes, such as the red dots in Figure 1.6. Rails that physically connect stations, including two parallel bars of rolled steel, to form the whole rail network can be mapped as links, such as the black lines in Figure 1.6. Although the rail network can be highly complex, the arrangement of nodes and links and connectivity patterns can be clearly graphed, modeled and expressed mathematically as provided in subsequent chapters.

1.4 Needs in Studying Rail Networks

In rail networks, any rail or station failure due to natural and human-caused hazards would lead to serious consequence, such as the performance loss of the entire or a portion of the rail network, which means that rail networks are vulnerable to disruptions of rails and stations. This network vulnerability is defined as the degree of susceptibility of a network due to the connectivity changes between nodes and links (Saadat et al. 2019). Higher network vulnerability means more susceptibility to disruptions. Therefore, connectivity and vulnerability of rail networks need to be analyzed and assessed to evaluate the consequence of disruptions in a risk analysis framework for managing risk and resilience. Resilience refers to the ability to prepare for and adapt to changing conditions, withstand hazards, and recover from disruptions (The White House Office of the Press Secretary 2013; Ayyub 2014). The resilience of rail networks also needs to be assessed and further strengthened to reduce the impact of disruptions on rails and stations. The following research needs of rail networks are summarized:

1. Modeling and analyzing the connectivity among stations by rails in order to assess the rail network vulnerability.
2. Assessing the rail network vulnerability due to the failure of rails and stations.
3. Assessing the rail network resilience in order to measure the ability of networks to prepare for and adapter to changing conditions, withstand hazards, and recover from disruptions of rails and stations; and
4. Enhancing rail networks and determining strategies for recovery based on network vulnerability and resilience assessment.

1.5 Organization of Thesis

Chapter 1 serves as an introduction of this thesis, including the significance and criticality of railroads, a definition of railroad systems and rail networks, the need to study rail networks, and an outline of the organization of each chapter.

Chapter 2 is the literature review of rail network topology, vulnerability and resilience. After which, the existing research on topology enhancement and recovery strategies are summarized. Based on the literature review, gaps are identified, and objectives are proposed in Chapter 2.

Chapter 3 presents theories and processes with illustrative examples of a methodology to analyze the rail network topology, assess and reduce the vulnerability and strengthen the resilience.

Chapter 4 shows how the methodology can be applied to specific case studies, such as Amtrak and Class I rail networks.

Chapter 5 provides a summary, conclusions and contributions from this thesis. Additionally, this chapter provides suggestions for future work related to rail networks.

Chapter 2: Topological Analytics of Rail Networks: A Literature Review

2.1 Rail Network Topology

Network connectivity among stations by rails can be quantified by analyzing network topology. The network topology analysis starts by mapping a rail network in the form of a graph. Garrison and Marble (1962) first proposed how to graph the components of rail networks, such as nodes and links. Additionally, they defined the connection matrix, structural patterns, and the cyclomatic number to demonstrate the network topology mathematically. Musso and Vuchic (1988) defined indicators used to analyze the rail network topology, such as path length and network density. They also analyzed the passenger flow and the population expression of metro rail networks. The network topology was assumed to be either completely random or completely regular. Watts and Strogatz (1998) first proposed the "small-world" characteristics of network topology. The “small-world” network is highly clustered with relatively small characteristic path lengths due to the presence of long-range links. Barabási and Albert (1999) proposed another important characteristic of network topology called “scale-free”. The probability $p(k)$ that the node in a complex network connects with *k* other nodes is supposed to decay as a power law, which means the degree of network connectivity decreases as the number of nodes increases. Latora and Marchiori (2002) first proved that the topology of the Boston metro rail network had the “small-world” characteristic. Derrible and Kennedy (2010) proved that most metro rail networks are either scale-free or small-world networks. Braha (2017) summarized characteristics and indicators of network topology.

Measuring and analyzing network topology are based on the calculation of topology indicators. However, as the number of nodes and links of networks increases, the size of matrices and the number of iterations in the topology indicator calculations can become extremely large. For example, the real Class I freight rail network contains more than 40,000 nodes and links, which means that in the calculation of topology indicators, the number of rows and columns of matrices exceeds 40,000. Using these matrices to perform over 40,000 iterative calculations is extremely difficult and inefficient. Therefore, the current analysis of rail network topology has a limitation based on its use for analyzing metro rail networks with 100 to 350 nodes and associated links (Derrible and Kennedy 2010; Zhang et al. 2018; Saadat et al. 2019).

2.2 Rail Network Vulnerability

From network topological perspectives, Latora and Marchiori (2001) first defined a performance indicator, called network efficiency, for the Boston metro rail network. Derrible and Kennedy (2010) further investigated the efficiency of 33 metro rail networks in the world. Based on the network efficiency, the assessment of node and link vulnerability mainly follows the approach of complete enumeration to measure the loss of network efficiency due to the disruption of nodes and links (i.e. rails and stations) (Saadat et al. 2019; Bešinović 2020). Network vulnerability is the maximum value of node and link vulnerability (Zhang et al. 2011). Chang et al. (2006) calculated the efficiency of metro rail networks in Seoul, Tokyo, Boston, and Beijing. Additionally, Zhang et al. (2018) and Saadat et al. (2019) assessed the vulnerability of the Shanghai and Washington D.C. metro rail networks respectively.

The same limitation of network topology analysis also exists in the network vulnerability assessment. As the number of nodes and links of networks increases, the size of matrices and the number of iterations in the network efficiency and vulnerability calculations become large. Therefore, the network efficiency and vulnerability assessment focused on metro rail networks with a limited number of nodes and links, such as for metro networks (Derrible and Kennedy 2010; Zhang et al. 2018;).

2.3 Rail Network Resilience

The increasing number of disruptions caused by natural and human-caused hazards seriously affects the performance of rail networks (Bešinović 2020). Resilience refers to the ability to prepare for and adapt to changing conditions, withstand hazards, and recover from disruptions (The White House Office of the Press Secretary 2013; Ayyub 2014). Therefore, the demand for assessing and enhancing network resilience has greatly increased in order to manage risks to the performance of rail networks. Two primary methods of assessing the rail network resilience are the data-driven method and the topological method (Bešinović 2020).

2.3.1 Data-driven Methods

Data-driven methods rely on the recorded historical data, such as ridership, the time of arrival, and the number of kilometers traveled by trains, to develop statistical models which can reflect the change of rail network performances when different events occur. Network resilience can be further assessed based on the statistical model. Additionally, data-driven methods are mostly used to assess the resilience of rail networks affected by disruptive events due to natural hazards.

Janić (2018) defined social-economic related performance indicators and developed statistical models to assess the resilience of Japan's Shinkansen Rail network affected by the 2011 Tohoku Earthquake. Zhu et al. (2017) used the data-driven method to analyze the impact of storm surge caused by Hurricane Irene and Sandy on the New York City rail network and to assess the network resilience based on individual ridership data. Dawson et al. (2016) assessed the impact of sea-level rise on the England coastal rail network and to assess the network resilience in the event that the worst sea-level rise occurs. Other researchers use data-driven methods to assess the resilience of passenger or freight rail networks affected by different types of natural hazards, such as heat waves (Ferranti et al. 2016), snow and rainfalls (Chan and Schofer 2016).

2.3.2 Topological Methods

From network topological perspectives, the network resilience can be assessed based on topological performance indicators rather than the recorded historical data of rail networks. Adjetey-Bahun et al. (2016) proposed time-varying graphs and integrated operating conditions into topological performance indicators to assess rail network resilience. Additionally, they performed a case study of the Paris rail network, showing that some components of the Paris rail network are not related to topological performance indicators. However, when integrating operating conditions, these components become relevant to topological performance indicators. Chen and Miller-Hooks (2012) used a stochastic mixed-integer program to quantify network resilience based on topological indicators and also proved the significance of recovery strategies on the ability of networks to recover from disruptions. Saadat et al. (2020) used the resilience triangle proposed by Bruneau et al (2003) to demonstrate the loss and recovery of network

efficiency and to further assess the time-dependent network resilience. They also determined the best recovery sequence with respect to the value of a resilience index.

2.4 Topology Enhancement and Recovery Strategies

2.4.1 Topology Enhancement

From network topological perspectives, the network vulnerability assessment allows us to evaluate the impact of disruptions of nodes and links on rail networks (i.e. the loss of network efficiency), while network resilience can also be assessed based on topological performance indicators. After which, strategies to reduce the impact of disruptions by reducing the network vulnerability and strengthening network resilience can be determined.

Saadat et al. (2020) proposed a pre-failure strategy called topology enhancement to reduce the network vulnerability by adding additional links into rail networks. They added three hypothetical loop lines consisting of several links into the Washington D.C. metro rail network, creating topological redundancy and reducing network vulnerability. With topological redundancy, when nodes or links fail, other alternative nodes or links can be used to reduce the loss of network efficiency. Therefore, the network vulnerability can be further reduced. It should be mentioned that adding hypothetical loop lines is only used to enhance the network at the theoretical level from the topological perspectives. Designing and adding loop lines into a real rail network will have many limitations because of the actual situations and policy restrictions. However, the results in this thesis can provide additional insights to decision-makers in managing risks.

The topology enhancement method first proposed by Saadat et al. (2020) was only researched on the Washington D.C. metro rail network. Additionally, the Washington D.C. metro only has 91 nodes and 140 links, which means adding few links can create enough redundancy and reduce the network vulnerability effectively. This topology enhancement strategy needs to be verified in rail networks with more nodes and links than the Washington D.C. metro rail network.

2.4.2 Recovery Strategies

A recovery strategy focuses on strengthening the rail network resilience through identifying the best recovery sequence after the disruption of nodes and links (i.e. rails and stations). Henry and Ramirez-Marquez (2012) proposed five time-related transition states of network resilience and further measured the network resilience as a function of time. They used a road network as a case study, proving that designing a good recovery sequence is an effective way to increase the network resilience. Zhang et al. (2018) measured the network resilience using the resilience triangle proposed by Bruneau et al (2003). They assumed that only one component of rail networks can be repaired in a recovery stage and determined the best sequential recovery strategy for Shanghai metro network. Saadat et al. (2020) determined recovery strategies for four different hypothetical disruption cases in Washington D.C. metro: (1) one transfer station and its connected rails; (2) multiple stations with different node degrees; (3) multiple stations with the same node degrees; and (4) multiple rails.

When station disruptions caused by natural or human-made hazards occur, the entrances and exits must be closed, and these disrupted stations cannot be used as departure or destination stations before they are repaired (Yin et al. 2018). Therefore, the repair of disrupted stations is

important for the recovery strategy. However, for the disruption case of one transfer station and its connected rails, the recovery strategy only considers the repair of rails (Zhang et al. 2018; Saadat et al. 2020), while the repair of the station itself is ignored in the recovery sequence.

2.5 Gaps and Objectives

The literature review highlights the significance and needs in studying rail networks. On this basis, the following gap from the literature review helps to focus the objectives of this thesis:

- The network topology analysis is an effective tool to graph and analyze the arrangement and connectivity between nodes and links. However, as the number of nodes and links increases, the network topology analysis can become overly complicated. Therefore, how to graph and analyze the topology of more complex rail networks accurately and effectively requires further research;
- The vulnerability of metro rail networks can be well reduced through the topology enhancement strategy (Saadat et al. 2020). However, as the network connectivity becomes more complex and developed, adding a small number of links is difficult to reduce the network vulnerability effectively. Whether the same topology enhancement strategy is applicable to rail networks with more complex connectivity than metro rail networks requires further research; and
- For the disruption case of one transfer station and its connected rails, the repair of stations needs to be considered in the sequential recovery strategy.

Based on the literature review and gaps, the main objectives of this research are to:

1. Analyze the topology of rail networks containing more nodes and links than metro rail networks accurately and effectively;
2. From network topological perspectives, assess the rail network vulnerability and resilience; and
3. Determine the topology enhancement to reduce the network vulnerability and the sequential recovery strategy for the disruption of one station and its connected rails.

Chapter 3: Methodology and Illustrative Examples

Using complex network theory, a network is treated mathematically and represented by a graph with two components: nodes and links for the connectivity among nodes and links. Based on the complex network theory, a methodology is proposed to analyze rail networks and to reduce the impact of potential failures by examining: (1) network topology; (2) network efficiency and vulnerability; (3) network resilience; and (4) the impact of potential failures. The methodology as shown in Figure 3.1 consists of the following steps:

1. Defining nodes, links, and the connectivity pattern of rail networks, then, mapping rail networks in the form of a graph;
2. Analyzing network topology by calculating topological indicators;
3. Measuring network efficiency and further network vulnerability from network topological perspectives;
4. Identifying the characteristics of vulnerable nodes and links, and the critical areas of networks;
5. Evaluating the network resilience index based on the changes of topological performance indicators (i.e., network efficiency); and
6. Enhancing the network topology by adding loop lines and determining the best sequential recovery strategy based on resilience assessments.

Bollobás (1985) and West (1995) provides background information on complex network theory. Boccaletti et al. (2006) and Braha (2017) summarized methods on measuring the topological indicators and characteristics of general complex networks. Derrible and Kennedy (2010), Zhang et al. (2018), and Saadat et al. (2019) used the complex network theory to represent metro rail

networks as a graph and analyzed the rail network topology. Latora and Marchiori (2002) initially defined the network efficiency for the Boston metro rail network, which provided a basis to quantify the robustness and vulnerability (Zhang et al. 2018) of rail networks subjected to potential failures and attacks. Later, Saadat et al. (2020) proposed the topology enhancement strategy to reduce network vulnerability. Based on the resilience triangle proposed by Bruneau et al (2003), Zhang et al. (2018) and Saadat et al. (2020) provides methods on assessing the network resilience dynamically from the topological perspectives and determining the sequential recovery strategy regarding resilience restoration.

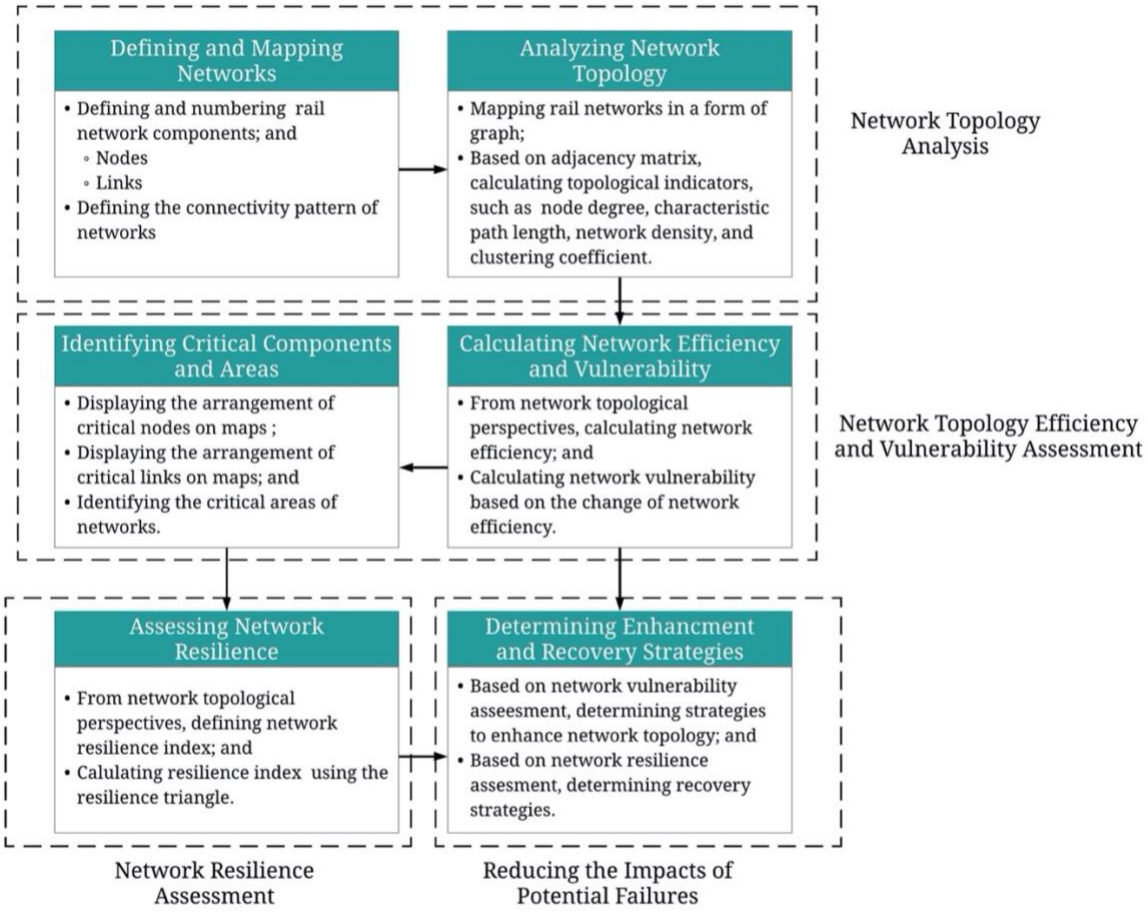

Figure 3. 1 Proposed methodology for analyzing rail networks and mitigating the impact of disruptions.

3.1 Analyzing Network Topology

Network topology is defined as the arrangement and connectivity among the components of a network. The analysis of rail network topology starts with defining and mapping stations and rails to nodes and links, respectively, in the form of a graph. Afterward, the topological indicators of rail networks are calculated to analyze network, which helps to quantify the connectivity of complex rail networks. Combined with an illustrative example, the analysis of network topology is introduced in detail in this section.

3.1.1 Defining and Mapping Network

In mapping rail networks as the form of a graph, the stations of a rail network can be mapped as nodes, such as the red dots in Figure 1.6, while links represent the rails of a rail network, such as the black lines in Figure 1.6. After numbering each node and link, the topology vector G is specified as Equation (3.1):

$$G = \{S, E\} \tag{3.1}$$

where,

G: the topology vector of a network

S: the collection of all nodes

E: the collection of all links

For example, the topology vector G of the network as shown in Figure 1.6 is:

$$G = \{6,7\} \tag{3.2}$$

Then, s_i denotes each node in the set *S.* For the network as shown in Figure 1.6:

$$S = \{s_i | i = 1,2,3,4,5,6\} \tag{3.3}$$

Also, e_{ij} in the set *E* represents a link that connects node *i* and node *j*. Referring to the network as shown in Figure 1.6:

$$E = \{e_{ij} | i,j = 1,2,3,4,5,6\} \tag{3.4}$$

Each link e_{ij} can also be represented as (i,j). Also, for an undirected network, the link from node *i* to node *j* is the same as the link from node *j* to node *i.*

3.1.2 Calculating Topological Indicators

A network's connectivity can be expressed mathematically by a $n \times n$ symmetric adjacency matrix, where *n* is the number of nodes in a network. Elements in the adjacency matrix are denoted as a_{ij}:

$$a_{ij} = \begin{cases} 1 & if\ node\ i\ connects\ with\ node\ j, \\ 0 & otherwise. \end{cases} \tag{3.5}$$

Three types of networks can affect the value of elements a_{ij} in the adjacency matrix. First, networks can be distinguished into weighted and unweighted. In a weighted network, some links

are more important or stronger than other links. Therefore, a_{ij} is not always equal to 0 or 1 but depends on the importance or strength of e_{ij}. Figure 3.2 shows a weighted and an unweighted network.

Figure 3. 2 Weighted and unweighted networks: (a) weighted; (b) unweighted.

For example, in the weighted network as shown in Figure 3.2 (a), dash lines indicate less important or weaker links, while the solid line indicates more important or stronger links. All links are considered equally important or strong in the unweighted network as shown in Figure 3.2 (b). As a result, the adjacency matrix for weighted and unweighted networks in Figure 3.2 can be represented as:

$$A_{weighted} = \begin{bmatrix} 0 & x & x \\ x & 0 & 1 \\ x & 1 & 0 \end{bmatrix} \tag{3.6}$$

$$A_{unweighted} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{3.7}$$

where,

x: the value of a_{ij} of less important or weaker links; $0 < x < 1$

The second type of network is the directed and undirected network. A directed graph is called digraph in which the link pointing from node *i* to node *j* is in a particular direction. For a directed link e_{ij}, the value of a_{ij} in the adjacency matrix is not equal to the value of a_{ji}. Therefore, the adjacency matrix of directed networks is asymmetric. Figure 3.3 shows a directed and an undirected network:

Figure 3. 3 Directed and undirected networks: (a) directed; (b) undirected.

The adjacency matrix for directed and undirected networks in Figure 3.3 can be represented as:

$$A_{directed} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.8}$$

$$A_{unweighted} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{3.9}$$

Also, networks can have multiple links among node *i* to node *j*. Additionally, a self-edge means that a link connects a node to itself. For the node *i* with a self-edge, the value of a_{ii} is equal to 2. Figure 3.4 shows a multi-edges network:

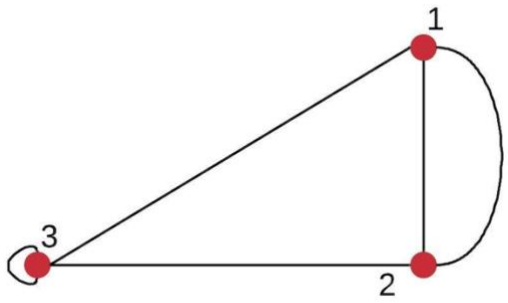

Figure 3. 4 Multi-edges network.

The adjacency matrix for the multi-edges network as shown in Figure 3.4 is specified as:

$$A_{multi} = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{3.10}$$

In the network studied in this thesis, every link among each pair of nodes is undirected and unweighted. Additionally, multiple links and the self-edge are not considered. The adjacency matrix is the most critical mathematic expression of graphing a network. All the calculations of topological indicators are based on the adjacency matrix. A comprehensive definition of the adjacency matrix for rail networks was proposed by Zhang et al. (2018), which applies in both simple and non-simple (i.e., weighted, directed, and multi-edges) networks. For the adjacency matrix defined by Zhang et al. (2018), if node *i* and *j* are connected, the value of a_{ij} equals to

one. If node *i* and *j* are not connected, the value of a_{ij} is equal to infinity. Otherwise, the value of a_{ij} equals to zero for the self-edge. Hence, for the illustrative example in Figure 1.6, a 6×6 symmetric adjacency matrix is represented as follows:

$$A_{ij} = \begin{bmatrix} 0 & 1 & 1 & \infty & \infty & \infty \\ 1 & 0 & 1 & 1 & \infty & \infty \\ 1 & 1 & 0 & 1 & 1 & \infty \\ \infty & 1 & 1 & 0 & \infty & 1 \\ \infty & \infty & 1 & \infty & 0 & 1 \\ \infty & \infty & \infty & 1 & 1 & 0 \end{bmatrix} \tag{3.11}$$

Node degree is a topological indicator that can be used to demonstrate the centrality of networks. The node degree of node *i* is equal to the number of links connected to it. Based on the adjacency matrix, the node degree denoted as K_i can be calculated as follows:

$$K_i = \sum_{j=1}^{n} a_{ij} \ (a_{ij} \neq \infty) \tag{3.12}$$

where,

n: the number of nodes

When using Equation (3.12) to calculate the node degree, infinity elements a_{ij} in the adjacency matrix Equation (3.11) need to be eliminated first. Table 3.1 shows the node degree of the network as shown in Figure 1.6:

Table 3. 1 Node degree of the network as shown in Figure 1.6.

Node	Node Degree
1	2
2	2
3	3
4	3
5	2
6	2

The average node degree denoted as $\overline{K}$ can be calculated by Equation (3.13). The average node degree of the network as shown in Figure 1.6 is equal to 2.33.

$$\overline{K} = \frac{1}{n}\sum_{i=1}^{n} K_i \tag{3.13}$$

In order to describe the process of moving from one node to another node in rail networks, a path is defined as a node sequence that each consecutive pair of nodes in the path is connected by links. In an unweighted rail network, the path length is the number of links between all consecutive pairs of nodes. Another critical network topological indicator is the minimum number of links moving from node *i* to node *j*, called the shortest path length or geodesic path length d_{ij}. For instance, the shortest path from node 1 to node 6 of the network as shown in Figure 1.6 is $s_1 s_2 s_4 s_6$, and the shortest path length is equal to 3.

In a complex network, finding the shortest path length between each pair of nodes is a challenging problem. Several shortest path algorithms have been formulated to solve this problem (Cherkassky et al. 1996). One of the algorithms called Shortest Path Faster Algorithm

(SPFA) based on the Breadth-first search, is a refinement of the Bellman-Ford Algorithm. The main philosophy of this algorithm is to start the procedure with "root nodes," then to improve the shortest path by examining all paths from this "root node" to its neighbors (Ding 1994). The neighbor of a node *i* are defined as the nodes that directly connect to node *i* by links. The procedure of Shortest Path Faster Algorithm is laid out as follows in Step1 through Step3:

1. Given a network $G = \{n, e\}$. Defining a subset of "root nodes" *r* ($r = \{1,2,3 \dots n\}$). Set all the distance of root node as zeros (i.e. $Dist(r) = 0$), while set all the distance of non-root node *i* as infinity (i.e. $Dist(i) = inf$). Continue Step 2;
2. Among the "root node" subset, set a "root node" *i* as an "initial node." Then, select a neighbor *j* of node *i* as the "marked" node. If no node is marked, the algorithm ends, otherwise, continue to Step 3; and
3. For each length of links between node *i* and node *j* denoted as $L(i,j)$, compare the distance $Dist(j)$ with the sum of $Dist(i)$ and $L(i,j)$. Whenever the sum is less than the $Dist(j)$, update $Dist(j)$ equals to the sum of $Dist(i)$ and $L(i,j)$. Also, set the node *j* as the next "initial node". After all neighbors of initial node *i* have been analyzed, back to Step 2.

For example, in the unweighted network as shown in Figure 1.6 (i.e. the length of all links equal to 1), node 1 is set as root node and the first "initial node." Therefore,

$$Dist(i) = \begin{cases} 0 & i = 1 \\ \infty & otherwise. \end{cases} \tag{3.14}$$

Node 2 is one of node 1's neighbors. Node 2 is the "marked" node that the distance $Dist(2)$ is equal to infinity in this stage. Therefore:

$$Dist(2) < Dist(1) + L(1,2) = 1 \tag{3.15}$$

The distance of node 2 is updated to one and set as the next "initial node." By repeating the procedure of Shortest Path Faster Algorithm, the following results can be obtained:

Table 3. 2 The shortest path length between node 1 and others.

Node	Dist(i)
1	0
2	1
3	1
4	2
5	2
6	3

where $Dist(i)$ is the shortest path length d_{1i} from node 1 to node i. The shortest path length between any pair of nodes can be obtained by changing the root node, which can be further used to calculate the diameter of a network denoted as D. The concept of network diameter is defined as the maximum value of shortest path lengths between all pairs of nodes.

$$D = Max\{d_{ij}\} \tag{3.16}$$

For the network as shown in Figure 1.6, the network diameter is equal to 3. Additionally, average path length or characteristic path length denoted as L represents the mean of all shortest path lengths:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij} \tag{3.17}$$

For the network as shown in Figure 1.6,

$$L = \frac{1}{6 \times (6-1)} \times 50 = 1.67 \tag{3.18}$$

Another topological indicator, called network density denoted as ρ is the fraction of the number of links to all possible number of links between:

$$\rho = \frac{E}{E_{max}} \tag{3.19}$$

where,

E: the number of links

E_{max}: all possible number of links that can be calculated by Equation (3.20):

$$E_{max} = \binom{n}{2} \tag{3.20}$$

Using Equation (3.19) and Equation (3.20), the density of the network as shown in Figure 1.6 can be obtained:

$$\rho = \frac{E}{E_{max}} = \frac{7}{\binom{7}{2}} = \frac{7}{21} = 0.33 \tag{3.21}$$

Characteristic path length *L* is a topological indicator measuring the separation degree of nodes in networks, whereas the clustering coefficient *C* is an indicator representing the aggregation degree of nodes in network. The clustering coefficient can be divided into two levels: local and global. The global clustering coefficient denoted as C_G is used to measure the degree of node aggregation among the overall network, which can be calculated by the ratio of the number of closed triplets to the number of all closed and open triplets (Prokhorenkova et al. 2015). A closed triplet is a set of three nodes that any pair of nodes are connected with, whereas an open triplet means three nodes are connected only by two links. For instance, zero closed triplets and 18 open triplets are in the network as shown in Figure 1.6. Therefore, the global clustering coefficient is zero.

The local clustering coefficient C_i of node i can be measured by the ratio of the number of links between its neighbors to the number of all possible links between its neighbors. For a node i with K_i node degree, the number of all possible links between its neighbors is given as a binomial coefficient $\binom{K_i}{2}$. Besides, denoting e_{ni} as the real number of links between the neighbors of node i, the local clustering coefficient can be obtained:

$$C_i = \frac{e_{ni}}{\binom{K_i}{2}} \tag{3.22}$$

Also, the average clustering coefficient $\bar{C}$ of a network is depicted as Equation (3.23):

$$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i = \frac{1}{n}\sum_{i=1}^{n} \frac{e_{ni}}{\binom{K_i}{2}} \tag{3.23}$$

Using Equation (3.23) to calculate the average local clustering coefficient of the network as shown in Figure 1.6:

$$\bar{C} = \frac{1}{6} \times (0 + 0 + 0 + 0 + 0 + 0) = 0 \tag{3.24}$$

3.2 Assessing Network Efficiency and Vulnerability

The network efficiency used to measure the efficiency of information exchange within a network was initially introduced by Latora and Marchiori (2001). The efficiency between node *i* and node *j* is inverse proportional to the shortest path length d_{ij}, i.e., the smaller the shortest path length from node *i* to node *j*, the more efficient information exchange between them. Therefore, the network efficiency represented by E_G can be calculated as follows:

$$E_G = \frac{1}{n(n-1)} \sum\nolimits_{i \neq j} \frac{1}{d_{ij}} \tag{3.25}$$

Using Equation (3.25), the efficiency of the network as shown in Figure 1.6 is calculated as:

$$E_G = \frac{1}{6 \times (6-1)} \times 21.33 = 0.711 \tag{3.26}$$

The potential failures of nodes and links might cause part of networks disrupted, affecting the shortest path length between each pair of nodes and further affecting the network efficiency. The vulnerability of nodes and links can be defined as the degree of the network efficiency loss after removing one node or link (i.e., disruptions of a node and link), which can be quantified as follows:

$$V_i = \frac{E_G - E_{Gi}}{E_G} \tag{3.27}$$

where,

E_G: initial network efficiency

E_{Gi}: network efficiency after removing a node or link

V_i: the vulnerability of one network component

The vulnerability of a network is the maximum value of node and link vulnerability as follows:

$$V = \max V_i \tag{3.28}$$

where,

V: network vulnerability

For example, if node 4 in the network of Figure 1.6 is removed, the network is changed as Figure 3.5:

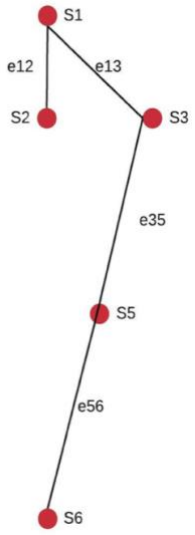

Figure 3. 5 Network graph after node 4 disrupted.

The network efficiency E_{Gi} can be obtained using Equation (3.25), and the vulnerability of node 4 can be further calculated using Equation (3.27):

$$E_{G4} = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}} = \frac{1}{5 \times (5-1)} \times 12.83 = 0.64 \tag{3.29}$$

$$V_4 = \frac{E_G - E_{G4}}{E_G} = \frac{0.711 - 0.64}{0.711} = 0.1 \tag{3.30}$$

3.3 Identifying Critical Components and Areas

Following the approach of complete enumeration, the network vulnerability due to the failure of each component in a rail network can be measured. Then, the critical components of a network are those nodes or links whose failures lead to relatively high network vulnerability. Additionally, for rail networks covering a wide area and containing a large number of nodes and links, such as the Amtrak and Class I rail networks, some critical components will be concentrated in specific areas. The arrangement of critical components can be displayed on maps. Then, the critical area where critical components are concentrated can be identified. Enhancing network topology and resilience in the critical area provides a basis to reduce the impacts of potential failures.

3.4 Evaluating Resilience Index

Network resilience can be quantified and assessed by using the resilience triangle (Bruneau et al 2003) as shown in Figure 3.6 to calculate the resilience index. In the resilience triangle, the network performance changes over time, including the performance loss stage at time t_0 and the performance recovery stage during t_0 to t_0+t_h. The resilience index can be calculated as Equation (3.31):

$$R_e = \frac{\int_{t_0}^{t_0+t_h} Q(t)dt}{t_h Q_o} \tag{3.31}$$

where,

R_e: resilience index

t_0: the time when performance loses

t_h: the period when the network performance recovers to the initial condition

$Q(t)$: time-dependent network performance function

Q_o: initial network performance

The resilience index of rail networks can be calculated based on the change of topological performance indicators, such as network efficiency E_G. For example, if the disruptions of link e_{12} and e_{34} of the network in Figure 1.6 occur and the recovery sequence is e_{12}-e_{34} (i.e., repair link e_{12} first, then repair link e_{34}), the resilience triangle is shown in Figure 3.7, where link e_{12} and e_{34} disrupt at the recovery stage 1, then, link e_{12} is repaired at recovery stage 2 and link e_{34} is repaired at recovery stage 3. The resilience index can be calculated as Equation (3.32):

$$R_e = \frac{\int_{t_0}^{t_0+t_h} E_G(t)dt}{t_h E_{Go}} = \frac{1.31225}{1.4222} = 0.923 \tag{3.32}$$

where,

$E_G(t)$: time-dependent network efficiency function

E_{Go}: initial network efficiency

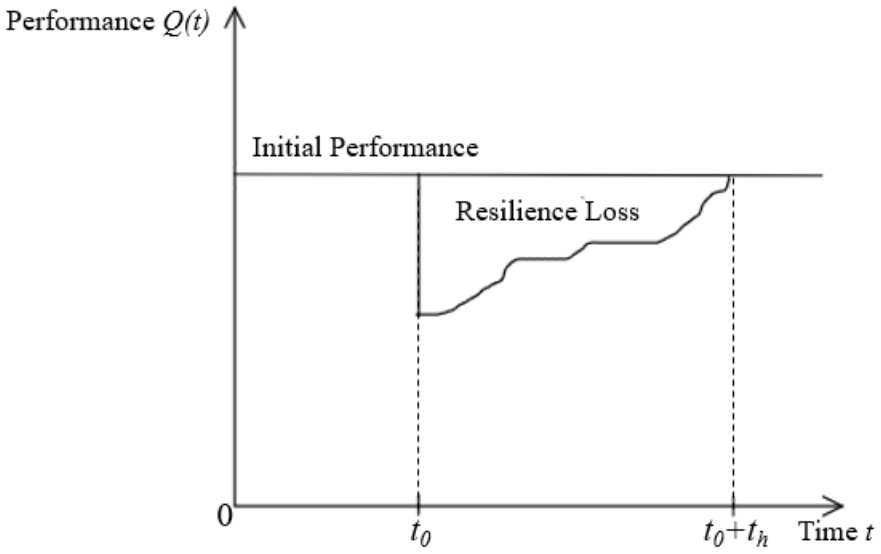

Figure 3. 6 Resilience Triangle.

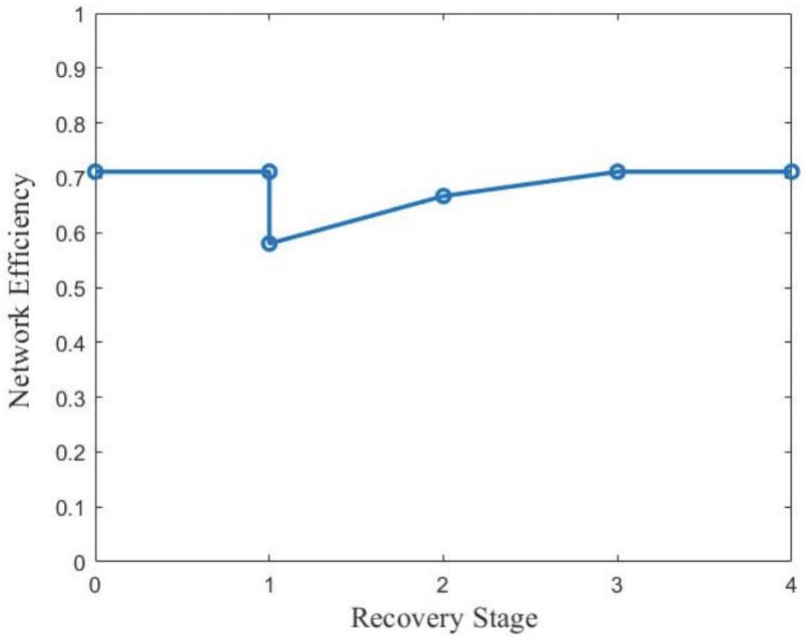

Figure 3. 7 Resilience triangle of the recovery sequence e_{12}-e_{34}.

3.5 Reducing the Impacts of Potential Failures

3.5.1 Enhancing Network Topology

Critical areas in which vulnerable nodes or links concentrate can be identified based on the vulnerability assessment of rail networks. Enhancing the network topology of these critical areas can significantly reduce the network vulnerability. In this study, the strategy of adding loop lines in critical areas is used to reduce the network vulnerability and further reduce the impacts of potential failures (Saadat et al. 2020). For example, a hypothetical loop line passing through nodes 1, 6, and 3 is added in the network as shown in Figure 1.6, then, a new network is created as Figure 3.8. This loop line decreases the characteristic path length of the network, increasing network efficiency. Additionally, the loop line creates topological redundancy, meaning when any node fails in the network, alternative nodes or links can be used to reduce the loss of network efficiency and the vulnerability. Figure 3.9 demonstrates the change of node 4 and network vulnerability after adding a hypothetical loop line, indicating that the vulnerability of node 4 and the network has been significantly reduced.

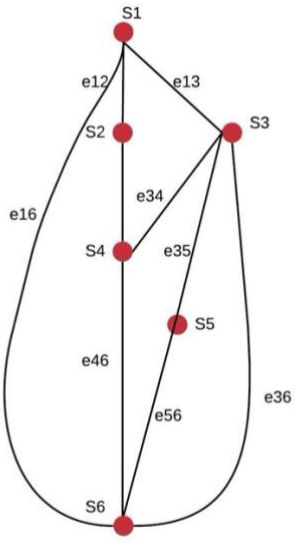

Figure 3. 8 Illustrative network after adding a loop line.

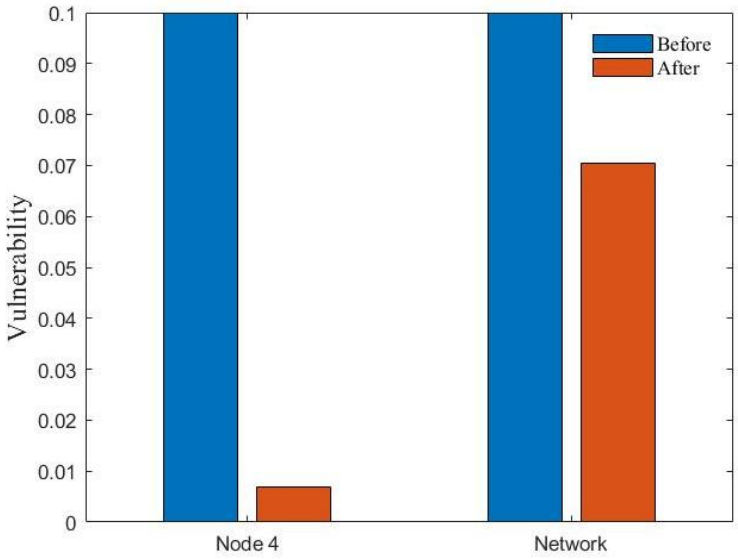

Figure 3. 9 The change of node 4 and network vulnerability.

3.5.2 Sequential Recovery Strategies

After the disruptions of nodes and links caused by potential failures, different recovery sequences lead to different network resilience, which can be measured and compared by the value of the resilience index. In order to reduce the impacts of potential failures, the best recovery sequence that can result in the largest value of resilience index needs to be determined to enhance the network resilience.

For the network as shown in Figure 1.6, after node 4 and its connected links are disrupted, 1 node and 3 links need to be repaired to recover the network fully. Assuming only one component can be repaired in a recovery stage, the number of all possible recovery sequences is equal to 24, i.e., the permutation of 4. In this study, the network efficiency is assumed not to be restored by the repair of links before the node repair. Because, in reality, sometimes even if rails are repaired, the route will not be re-operated until the disrupted station is repaired.

The initial network efficiency is equal to 0.7111, and the network efficiency decreases to 0.6417 after the disruption of node 4. The resilience index of different recovery sequences can be calculated using Equation (3.31). Table 3.3 shows the resilience index values of different recovery sequences. Figure 3.10 shows the comparison of recovery triangles. Recovery sequence e_{24}- s_4- e_{46}- e_{34} is identified as the optimal one with the largest value of resilience index 0.9234, which means if node 4 is disrupted, link e_{24} should be repaired first, then node 4, followed by link e_{46}, and finally link e_{34}.

Table 3. 3 Resilience index for the seven sequential recovery strategies of node 4.

Ranking	Recovery sequence	R_e
1	e_{24} - s_4 - e_{46} - e_{34}	0.9234
2	e_{24} - s_4 - e_{34} - e_{46}	0.9175
3	s_4 repaired in the third order	0.9146
4	s_4 - e_{34} - e_{24} - e_{46}	0.9048
5	s_4 repaired in the fourth order	0.9024
6	s_4 - e_{24} - e_{46} - e_{34}	0.9017
7	s_4 - e_{24} - e_{34} - e_{46}	0.8958

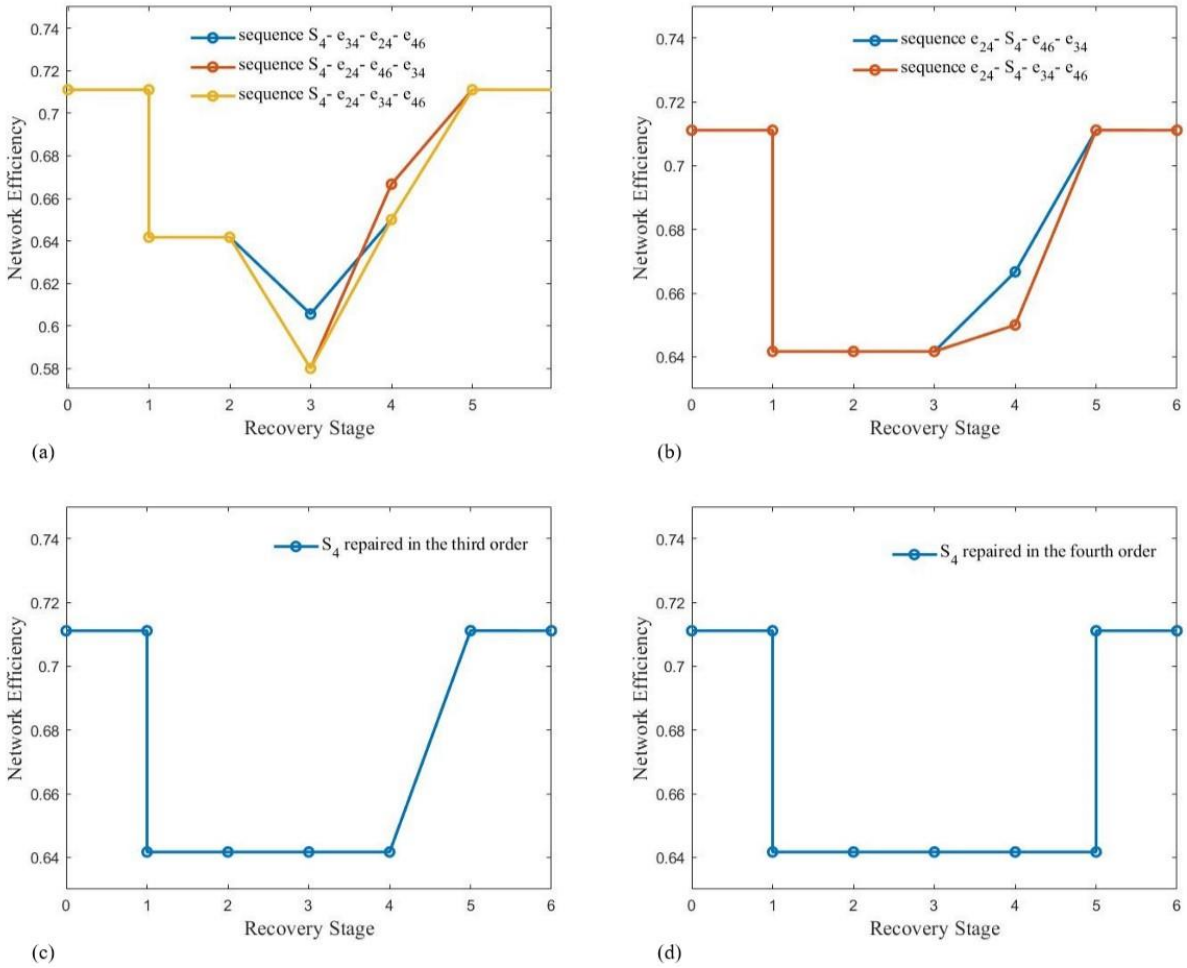

Figure 3. 10 (a) Node 4 repaired in the first order; (b) Node 4 repaired in the second order; (c) Node 4 repaired in the third order; (d) Node 4 repaired in the fourth order.

Chapter 4 Case Studies: Amtrak and Class I Rail Networks

Two critical rail networks in the United States are modeled and analyzed in this chapter. The first network is the Amtrak passenger rail network and the second is the Class I freight rail network. Both rail networks cover the entire United States, consisting of the greater number of nodes and links than metro networks. The methodology proposed in the Chapter 3 is used to analyze the topology of Amtrak and Class I rail networks to enhance the network vulnerability and assess the recovery strategies for the disruption case of one node and its connected links.

4.1 Case Study 1: Amtrak Rail Network

Most of the Amtrak's rails belong to the Class I railroad entities. The Amtrak rail network can be regarded as a subset of Class I rail network for independent research, which serves 529 stations and more than 21,400 miles rails in North America, including 46 states, Washington D.C. in the United States and three Canadian provinces. Figure 4.1 displays the entire Amtrak rail network in North America, where red lines belong to the Amtrak corporation and the yellow lines are commuter rails.

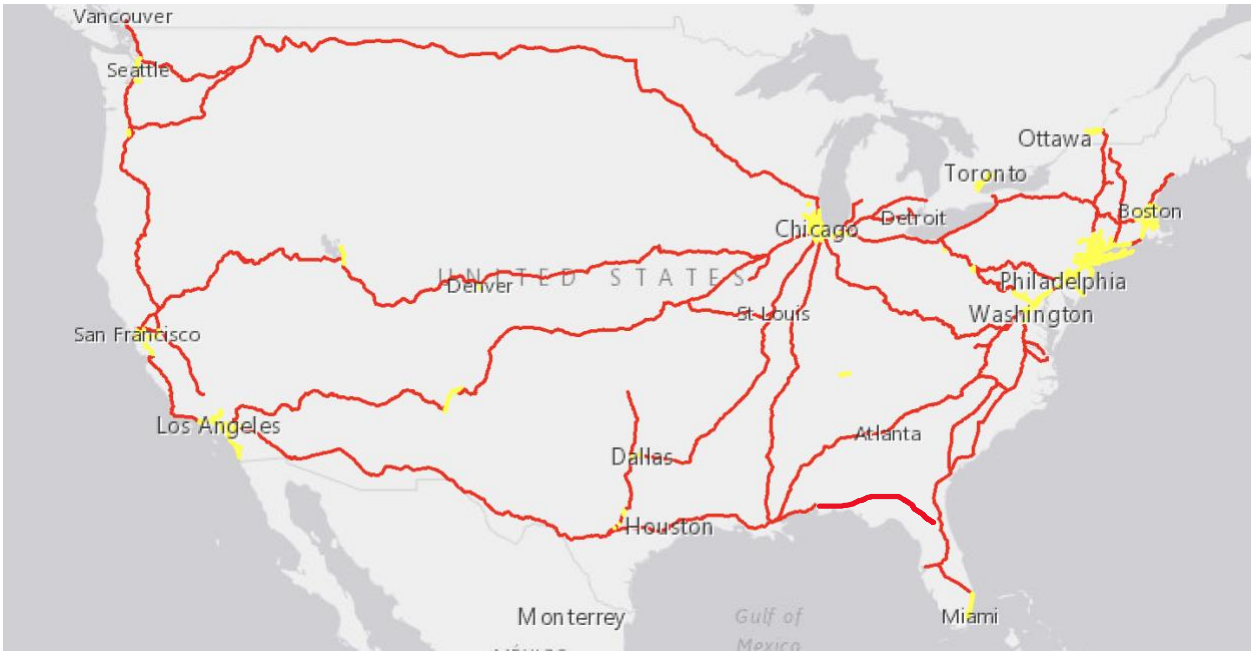

Figure 4. 1 Map of the Amtrak passenger rail network.
(Reproduced and edited from Federal Railroad Administration - Safety Map 2020.)

4.1.1 Mapping the Amtrak Rail Network

The first step of topology analysis is mapping the Amtrak rail network in the form of a graph, which can be divided into five areas depending on geographic locations: Northwest (NW), Southwest (SW), the Great Lakes area (GL), Southeast (SE), and Northeast (NE). Nodes represent the stations of the Amtrak rail network, while links represent the rails that directly connect stations. Additionally, the Amtrak rail network is modeled as an unweighted and undirected network. At the same time, each node or link is numbered for further analysis. Figure 4.2 to 4.6 show the graph and numbering of the Amtrak rail network in each area.

Figure 4. 2 Northwest area.

Figure 4. 3 Southwest area.

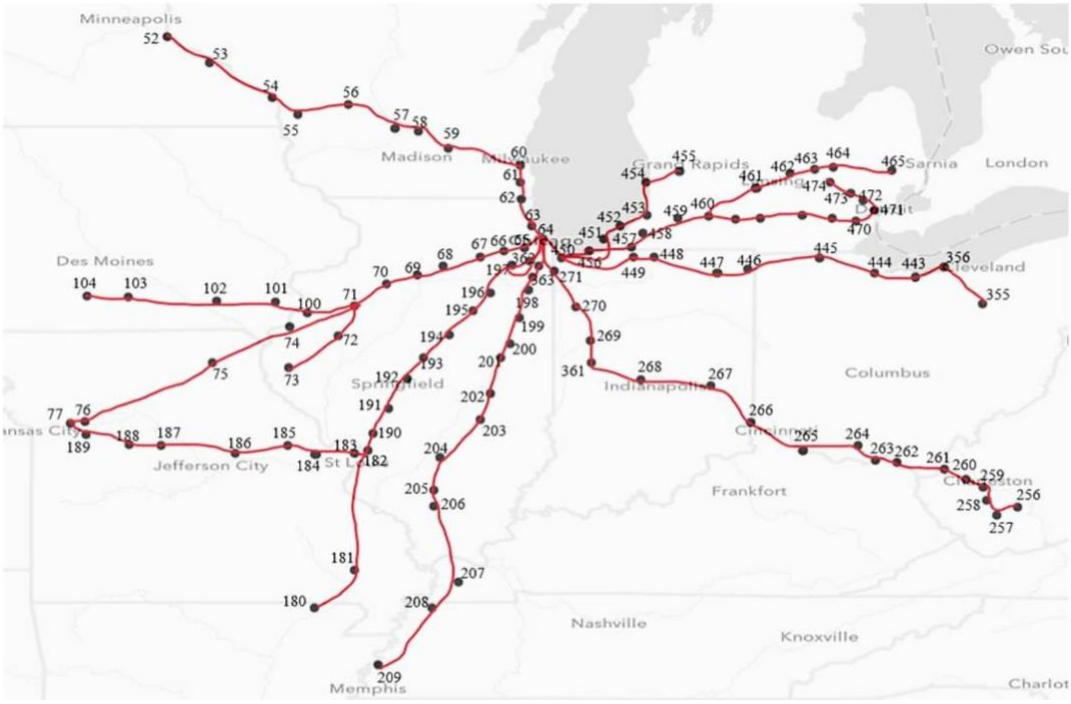

Figure 4. 4 The Great Lakes area.

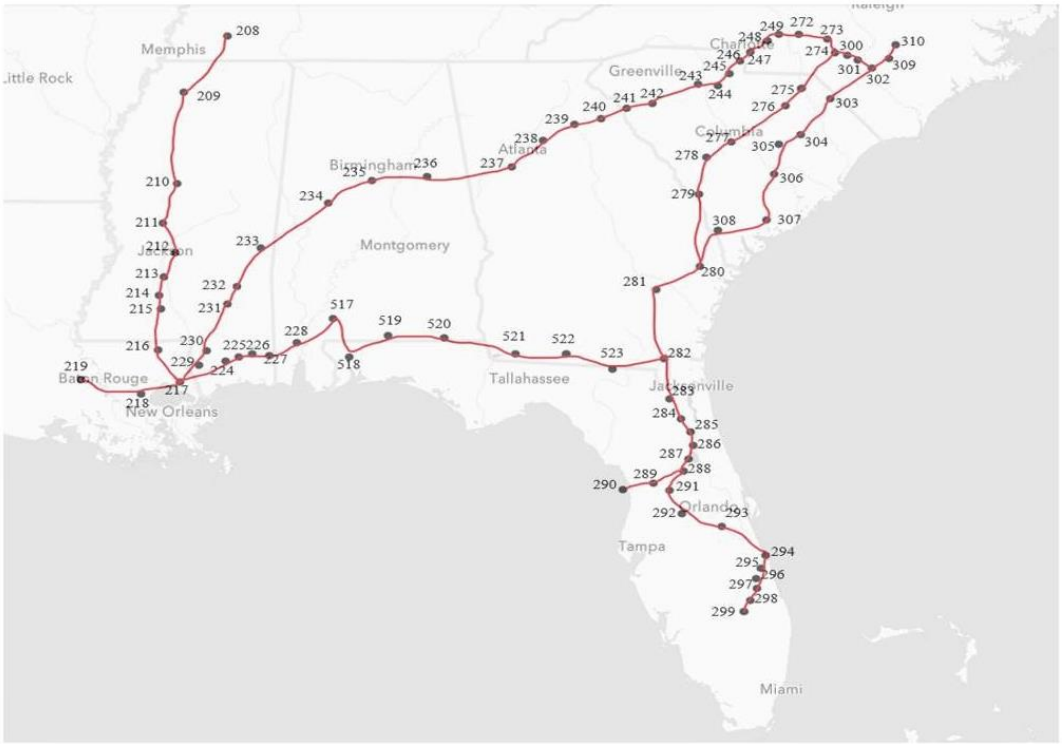

Figure 4. 5 Southeast area.

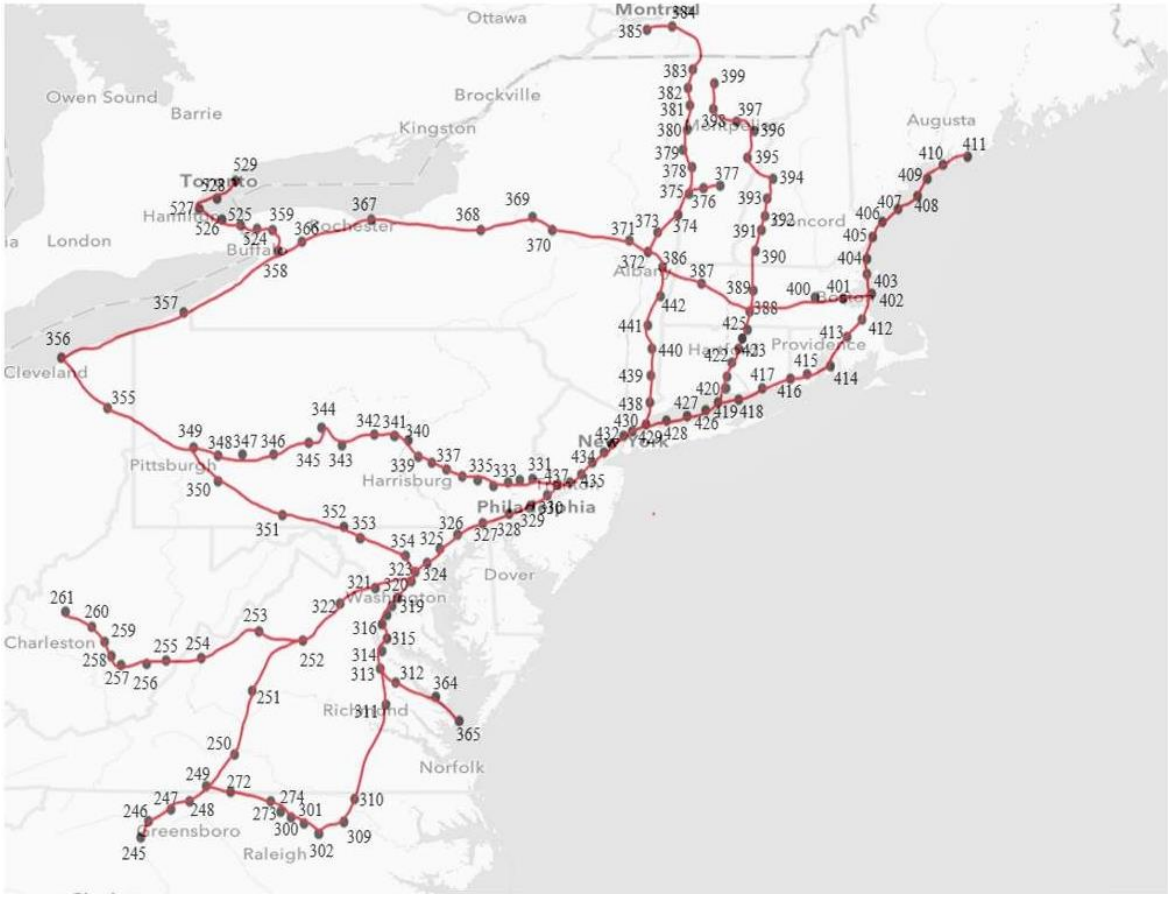

Figure 4. 6 Northeast area.

In summary, the Amtrak rail network consists of 529 nodes and 552 links. Therefore, the network vector is specified as follows:

$$G = [529,552] \tag{4.33}$$

4.1.2 Results and Analysis

4.1.2.1 Topology Analysis of the Amtrak Rail Network

The analysis of the Amtrak rail network is based on the adjacency matrix A_{ij}. According to the definition of simple adjacency matrix as used by Zhang et al. (2018):

$$a_{ij} = \begin{cases} \infty & node\ i\ are\ not\ directly\ connected\ with\ \ j. \\ 1 & node\ i\ are\ directly\ connected\ with\ j \\ 0 & self-edge \end{cases} \tag{4.34}$$

For the Amtrak rail network, the size of adjacency matrix is 529×529. Afterward, topological indicators can be measured one by one, including the average node degree $\overline{K}$ (using Equations (3.12) and (3.13)), network density ρ (using Equations (3.20) and (3.21)), characteristic path length L (using the shortest path faster algorithm (SPFA) and Equation (3.18)), diameter of the network D (using Equation (3.17)), local clustering coefficient $\bar{C}$ (using Equations (3.23) and (3.24)), and global network efficiency E_G (using Equation (3.26)). Table 4.1 summarizes the results of the Amtrak rail network's topological indicators:

Table 4. 1 Topological indicators of the Amtrak rail network.

No.	Topological Indicators	Values	Notes
1	Average node degree	2.087	/
2	Network density	0.004	/
3	Characteristic path length	35.404	/
4	Diameter of the network	91	/
5	Local clustering coefficient	0.0124	Ignore nodes with node degree 1
6	Global network efficiency	0.0463	/

4.1.2.2 Vulnerability Assessment of the Amtrak Rail Network

The vulnerability assessment includes 2 aspects: the network vulnerability due to node and link failures, and critical area identification. The network vulnerability due to the failure of a node is defined as the degree of network efficiency loss after removing this node and its connected links. However, when calculating the network vulnerability due to the failure of a link, only this link needs to be removed. Therefore, after removing a link, the size of the adjacency matrix remains unchanged. Using Equation (3.29) and (3.31), the network vulnerability due to the failure of each node and link can be obtained. Table 4.2 shows the top 40 most critical nodes and Table 4.3 demonstrates the top 40 most critical links. As such, the vulnerability of the Amtrak rail network is equal to 0.161.

Table 4. 2 Top 40 critical nodes of the Amtrak rail network.

No.	Name of Station	Area[a]	Node Numbering	Node Degree	Vulnerability
1	Chicago (Union Station), Illinois	GL	64	7	16.10%
2	Hammond-Whiting, Indiana	GL	450	4	12.78%
3	Cleveland, Ohio	GL	356	3	7.28%
4	Schenectady, New York	NE	372	3	6.04%
5	South Bend, Indiana	GL	449	2	5.89%
6	Buffalo (Exchange St), New York	NE	358	3	5.81%
7	Elkhart, Indiana	GL	448	2	5.73%
8	Waterloo, Indiana	GL	447	2	5.61%
9	Bryan, Ohio	GL	446	2	5.53%
10	Elyria, Ohio	GL	443	2	5.50%
11	Toledo, Ohio	GL	445	2	5.48%
12	Sandusky, Ohio	GL	444	2	5.47%
13	Michigan City, Indiana	GL	456	2	5.41%
14	Galesburg, Illinois	GL	71	4	5.12%
15	Niles, Michigan	GL	457	2	4.99%
16	Dowagiac, Michigan	GL	458	2	4.61%
17	Erie, Pennsylvania	NE	357	2	4.57%
18	Jacksonville, Florida	SE	282	3	4.30%
19	Kalamazoo, Michigan	GL	459	2	4.26%
20	Glenview, Illinois	GL	63	2	4.22%
21	Springfield, Massachusetts	NE	388	4	4.17%
22	Battle Creek, Michigan	GL	460	3	4.05%
23	Sturtevant, Wisconsin	GL	62	2	3.87%
24	Washington, DC	NE	323	3	3.77%
25	San Bernardino, California	NW	99	4	3.63%
26	Buffalo (Depew), New York	NE	366	2	3.61%
27	Milwaukee (General Mitchell Intl Airport), Wisconsin	GL	61	2	3.57%
28	New Orleans, Louisiana	SE	217	4	3.56%
29	Alexandria, Virginia	NE	320	3	3.54%
30	Rochester, New York	NE	367	2	3.49%
31	Albany/Rensselaer, New York	NE	386	3	3.44%
32	La Grange, Illinois	GL	65	2	3.43%
33	Syracuse, New York	NE	368	2	3.40%
34	Rome, New York	NE	369	2	3.33%
35	Palatka, Florida	SE	283	2	3.33%
36	Milwaukee, Wisconsin	GL	60	2	3.31%
37	Utica, New York	NE	370	2	3.30%
38	Amsterdam, New York	NE	371	2	3.30%
39	Naperville, Illinois	GL	66	2	3.21%
40	Saratoga Springs, New York	NE	373	2	3.16%

[a]NW (northwest area), SW (Southwest area), GL (the Great Lakes are), SE (southeast area), and NE (northeast area).

Table 4. 3 Top 40 critical links of the Amtrak rail network.

No.	Starting Node	Ending Node	Area[a]	Vulnerability
1	69	70	GL	7.97%
2	400	401	NE	5.97%
3	388	400	NE	5.80%
4	402	516	NE	5.71%
5	398	399	NE	5.67%
6	397	398	NE	5.58%
7	315	316	NE	5.53%
8	396	397	NE	5.52%
9	394	395	NE	5.50%
10	395	396	NE	5.49%
11	406	407	NE	5.29%
12	407	408	NE	4.90%
13	314	315	NE	4.67%
14	408	409	NE	4.55%
15	316	317	NE	4.53%
16	63	64	GL	4.43%
17	409	410	NE	4.23%
18	62	63	GL	4.08%
19	61	62	GL	3.78%
20	318	319	NE	3.72%
21	238	239	SE	3.64%
22	324	325	NE	3.59%
23	64	65	GL	3.52%
24	60	61	GL	3.51%
25	325	326	NE	3.49%
26	330	437	NE	3.47%
27	326	327	NE	3.42%
28	240	241	SE	3.37%
29	327	328	NE	3.37%
30	329	330	NE	3.35%
31	328	329	NE	3.35%
32	70	71	GL	3.29%
33	59	60	GL	3.27%
34	277	278	SE	3.25%
35	79	80	GL	3.23%
36	312	364	NE	3.15%
37	313	312	NE	3.13%
38	71	72	GL	3.12%
39	241	242	SE	3.12%
40	58	59	GL	3.05%

[a] NW (northwest area), SW (Southwest area), GL (the Great Lakes are), SE (southeast area), and NE (northeast area).

Figure 4.7 shows the arrangement of the top 40 critical nodes, where blue circles indicate the top 20 critical nodes and green circles represent the remaining top 21-40 critical nodes. In this case, the majority of top 40 critical nodes are concentrated in the Great Lakes and the northeast areas.

For example, 22 nodes are located in the Great Lakes area, accounting for 55 percent, while 14 nodes are located in the northeast area, accounting for 35 percent. Figure 4.8 demonstrates the arrangement of the top 40 critical links, with similar indicators of blue squares for the top 20 critical links and green squares represent the remaining top 21-40 critical links. The top 40 critical links are still concentrated in the Great Lakes and northeast areas. All top 20 critical links, and 90 percent of top 40 critical links are located in these two areas. As a result, the critical areas of the Amtrak rail network are the Great Lakes and northeast areas. Additionally, the network vulnerability due to the failure of node 64 (Chicago (Union Station), Illinois) and node 450 (Hammond-Whiting, Indiana) are much higher than other components of the Amtrak rail network. The most critical link is also in the Great Lakes area. Therefore, we will focus on the Great Lakes as the most critical area.

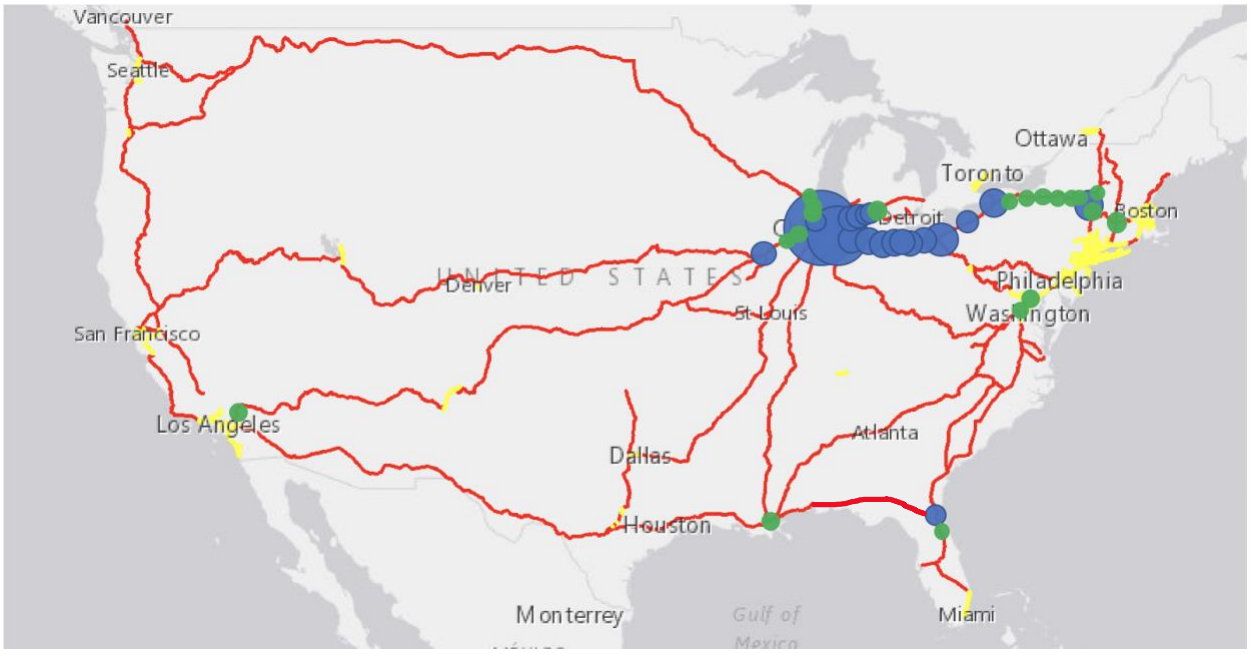

Figure 4. 7 Arrangement of the Amtrak rail network's top 40 critical nodes.
(Reproduced and edited from Federal Railroad Administration - Safety Map 2020.)

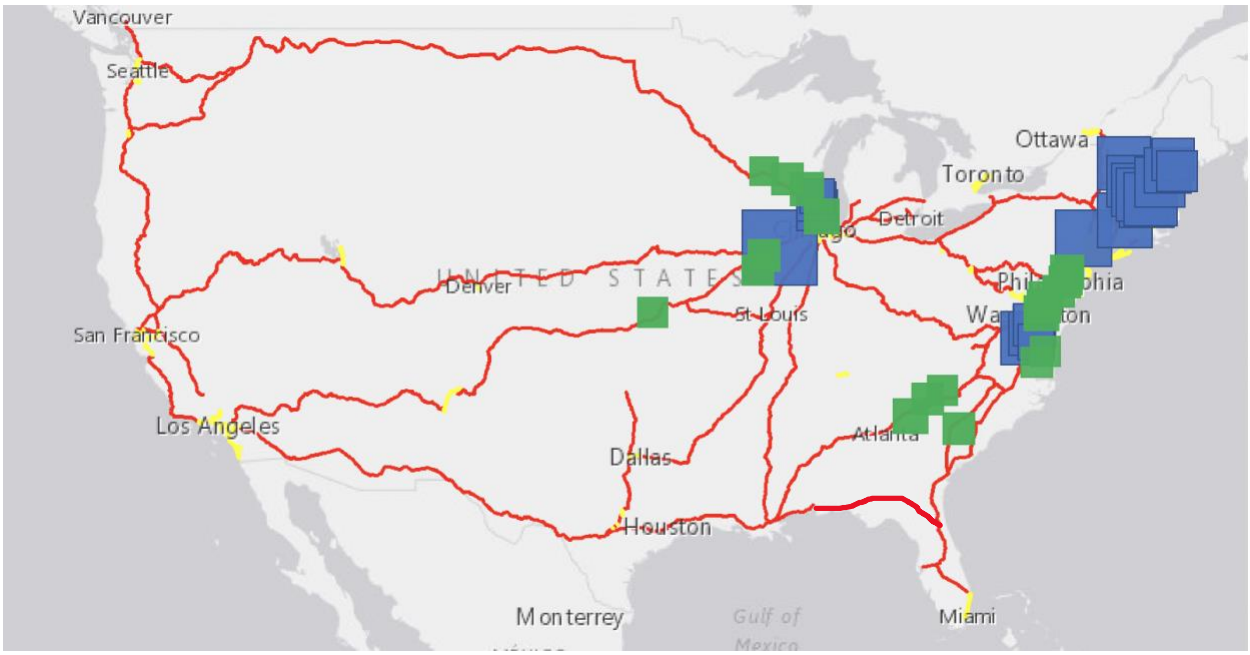

Figure 4. 8 Arrangement of the Amtrak rail network's top 40 critical links.
(Reproduced and edited from Federal Railroad Administration - Safety Map 2020.)

The failure of node 64 and node 450 leads to greater network vulnerability than other nodes in the Amtrak rail network. The characteristics of these two nodes need to be analyzed with emphasis. This thesis will analyze the characteristics of critical nodes from the perspective of the area where they are located, the node degree, and the connectivity pattern.

The top 40 critical nodes are mainly concentrated in the Great Lakes and the northeastern area. Table 4.4 demonstrates the comparison of the Amtrak network's five areas. Node density is the number of nodes per unit area. Diameter of an area is defined as the maximum value of the shortest path length between all pairs of nodes in this area. In this table, the value in each column is the ratio between any two areas' characteristics. The node density of Great Lakes and northeast are greater than that of other areas, while the difference of other characteristics between the Amtrak network's five areas are relatively small.

Table 4. 4 Comparison of the Amtrak network's five areas.

Area	Node Density	Average Node Degree	Characteristic Path Length	Diameter
Northwest	1.00	1.03	1.47	1.38
Southwest	1.99	1.04	1.82	2.00
Great Lakes	5.23	1.00	1.00	1.00
Southeast	2.10	1.04	1.09	1.16
Northeast	6.92	1.05	1.30	1.38

Node 64 (Chicago (Union Station)) and node 450 (Hammond-Whiting Station) are more critical than all other nodes. The most typical feature of these two nodes is that their node degrees are relatively large. Figure 4.9 shows the correlation between network vulnerability due to node failures and the node degree of failed node. For the top 20 critical nodes, network vulnerability is positive linear correlated to the node degree. The correlation continues to decrease as the number of critical nodes increases. However, node degree and network vulnerability still maintain a positive linear correlation. Thus, it can be determined that the greater the node degree, the greater node vulnerability might be. Because, when a node is disrupted, all links connected to it are considered disrupted. A node with the higher node degree means that when this node is disrupted, more links will also be disrupted, which leads to more severe network efficiency loss.

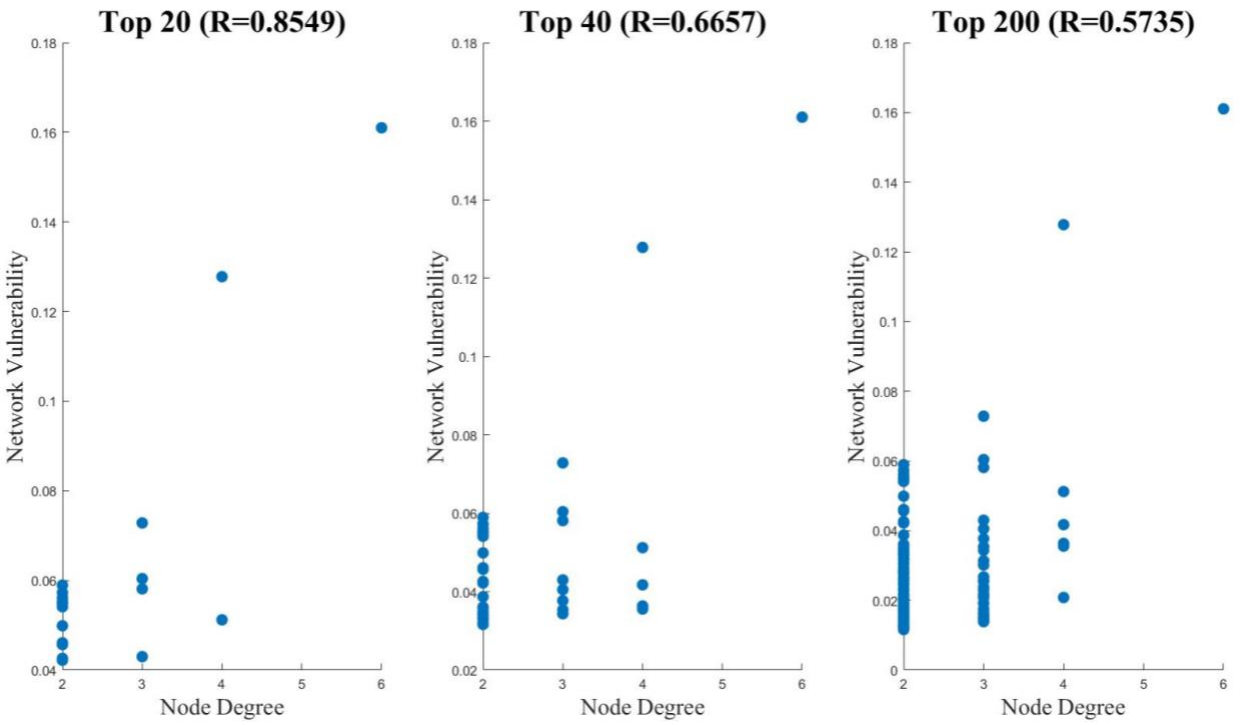

Figure 4. 9 Correlation between network vulnerability due to node failures and the node degree of the failed node.

The measurement of the vulnerability is based on the network efficiency loss that is related to the change in network connectivity. A connectivity pattern called the typical connectivity pattern 1 is defined and shown as Figure 4.10. In this connectivity pattern, a bridge link e_{12} connects two transfer nodes: s_1 and s_2, whose node degrees are equal to or more than 3. Also, several branch links connect the transfer nodes with their affiliated nodes. For example, the branch links e_{13} and e_{14} connect the transfer node 1 with its affiliated nodes 3 and 4. Additionally, a small number of intermediate nodes might be on the bridge link.

More than 90 percent of the top 40 critical nodes fall on the typical connectivity pattern 1, especially the transfer node and nodes falling on the bridge link. Because if these nodes disrupt,

the shortest path length between the nodes on two sides of the bridge link will greatly increase, which means the network efficiency will significantly decrease. In the typical connectivity pattern 1 of node 64 and 450 shown in Figure 4.11, over ten of the top 40 critical nodes fall on it, indicating that nodes on the typical connectivity pattern 1 are likely to produce larger node vulnerability than other nodes. In summary, for the typical connectivity pattern in high node density areas of the Amtrak rail network, the network vulnerability due to the failure of transfer nodes (i.e., large node degree) and nodes on the bridge link might be very large.

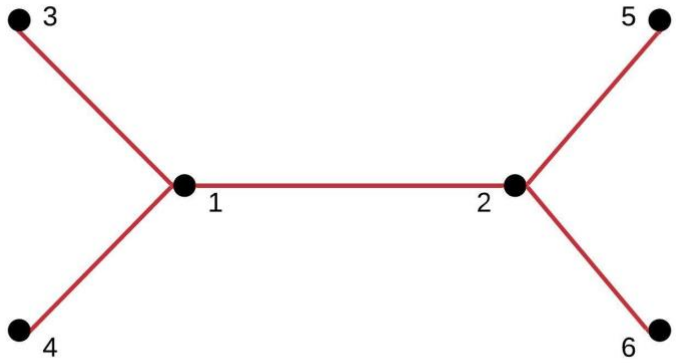

Figure 4. 10 Typical connectivity pattern 1.

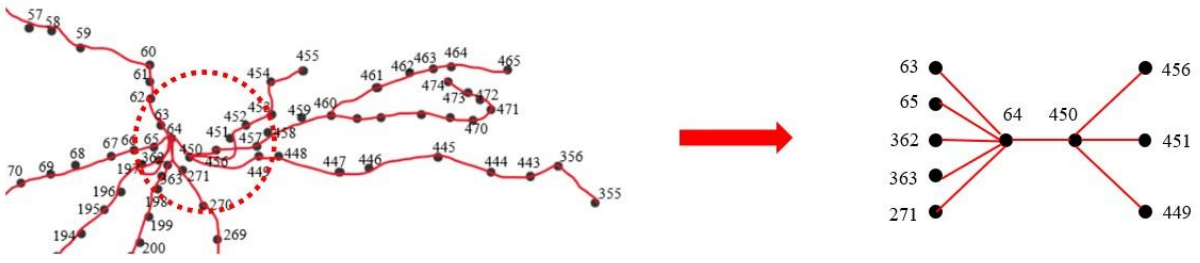

Figure 4. 11 The typical connectivity pattern 1 of node 64 and 450.

Table 4.3 and Figure 4.8 show the values and arrangement of the top 40 critical links. Compared with the node 64 and 450 that are more critical than other nodes, no link is significantly critical than other links. The top 40 critical links are still concentrated in the Great Lakes and northeast areas. However, in addition to some nodes that are still concentrated around the Chicago Union Station, more nodes are concentrated in the northeast corner instead of the area between Chicago and New York.

Links with high vulnerability concentrating around the Chicago Union Station and Washington D.C. are still in the typical connectivity pattern 1. However, more critical links are concentrated on the line that contains the "end-node" (i.e., a node with node degree of 1), for example, the line from node 388 to node 399 and node 402 to 411. The explanation of this phenomenon is that when a link in the line with "end-node" disrupts, several nodes on this line cannot be reached by other nodes in the network. Some shortest path length will be infinity, leading to the network efficiency significantly decreasing. For example, if link $e_{398,399}$ disrupts, the shortest path length between other nodes and node 399 will become infinity. Therefore, the network efficiency will be severely lost, which means the vulnerability of link $e_{398,399}$ is relatively high.

In summary, critical links of the Amtrak rail network are still concentrated in the areas with high node density. Also, most of the top 40 critical links are in the typical connectivity pattern 1 or the line with the "end-node."

4.1.2.3 Topology Enhancement of the Amtrak Rail Network

Depending on the vulnerability assessment, the Great Lakes area is identified as the most critical area of the Amtrak rail network. Two hypothetical loop lines are added around the most critical node 64 and 450 to create redundancy for the Great Lakes area, enhancing the network topology and reducing the vulnerability. In this section, the network vulnerability changes due to the failure of 10 nodes (i.e., node 64, 450, 356, 449, 448, 447, 446, 443, 445, and 444) and 10 links (i.e., $e_{69,70}$, $e_{63,64}$, $e_{62,63}$, $e_{61,62}$, $e_{64,65}$, $e_{60,61}$, $e_{70,71}$, $e_{59,60}$, $e_{79,80}$, $e_{71,72}$) are examined to verify if the topology enhancement is applicable to the Amtrak rail network.

The first hypothetical loop line connects existing nodes consisting of 7 added links: $e_{60,70}$, $e_{70,193}$, $e_{193,202}$, $e_{202,361}$, $e_{361,447}$, $e_{447,459}$, and $e_{459,455}$, which is the yellow line shown in Figure 4.12 (a). The second hypothetical loop line consists of 5 added links $e_{59,70}$, $e_{70,193}$, $e_{193,202}$, $e_{202,266}$, and $e_{266,355}$, again shown as the yellow line in Figure 4.12 (b).

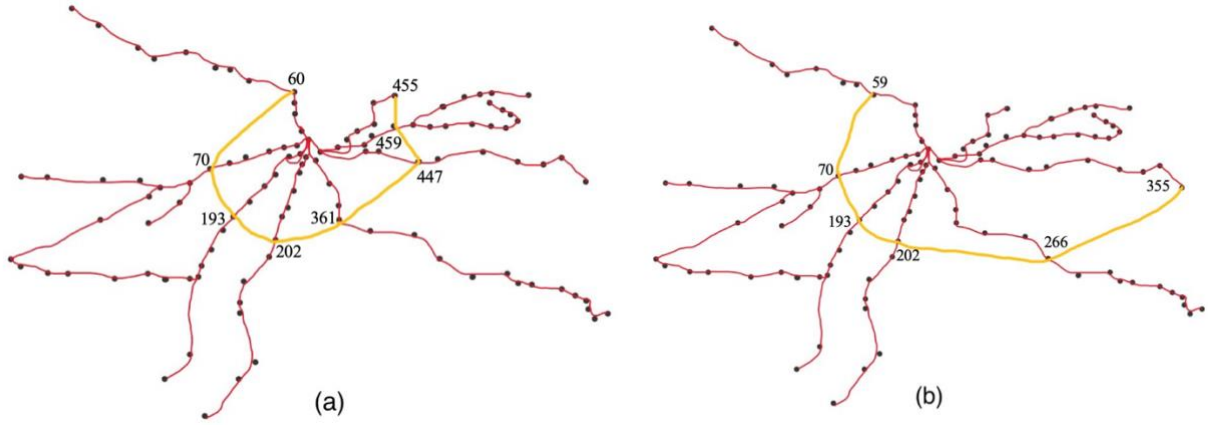

Figure 4. 12 Hypothetical loop lines added into the Amtrak rail network: (a) Loop Line 1; (b) Loop Line 2.

Table 4.5 and Figure 4.13 show the changes in the network vulnerability due to the failure of 10 nodes in the Great Lakes area after adding two loop lines. After adding the loop line 1, the network vulnerability due to the failure of nodes 64 and 450 has been significantly reduced to almost zero. The same phenomenon occurs on the node 449 and 448. However, the network vulnerability due to the failure of node 356, 443, 444, 445, 446 and 447 increases. These nodes can be regarded as falling outside the circle where node 64 or 450 is the center and loop line 1 is the arc. Therefore, loop line 2 is designed to cover more critical nodes and links inside the circle. After adding the loop line 2, the vulnerability due to the failure of all ten selected nodes have been reduced significantly to almost 0. However, the reduction magnitude of node 450 and 356 is smaller than other nodes, because the loop line 2 does not connect all the branch links of the transfer node 450 and 356 due to geographical reasons (i.e., the loop line is difficult to build across lakes). If the loop line 2 is extended to pass through node 355, 357, 370, 461, 455, the network vulnerability due to the failure of node 450 and 356 can be reduced significantly to 0.006 and 0.007, respectively, which demonstrates that adding a loop line is effective to reduce the network vulnerability due to the failure of nodes (Saadat et al. 2020).

Table 4. 5 Comparison of loop lines' impact on network vulnerability due to the failure of nodes.

Ranking	Station	Numbering[a]	Vulnerability		
			Original network	Loop line 1	Loop line 2
1	Chicago (Union Station), Illinois	64	16.10%	0.78%	1.83%
2	Hammond-Whiting, Indiana	450	12.78%	0.41%	7.15%
3	Cleveland, Ohio	356	7.28%	8.27%	6.51%
4	South Bend, Indiana	449	5.89%	0.10%	0.25%
5	Elkhart, Indiana	448	5.73%	0.17%	0.16%
6	Waterloo, Indiana	447	5.61%	8.62%	0.14%
7	Bryan, Ohio	446	5.53%	6.99%	0.16%
8	Elyria, Ohio	443	5.50%	6.63%	0.53%
9	Toledo, Ohio	445	5.48%	6.81%	0.23%
10	Sandusky, Ohio	444	5.47%	6.69%	0.35%

[a]These node numberings are provided in Figures 4.4.

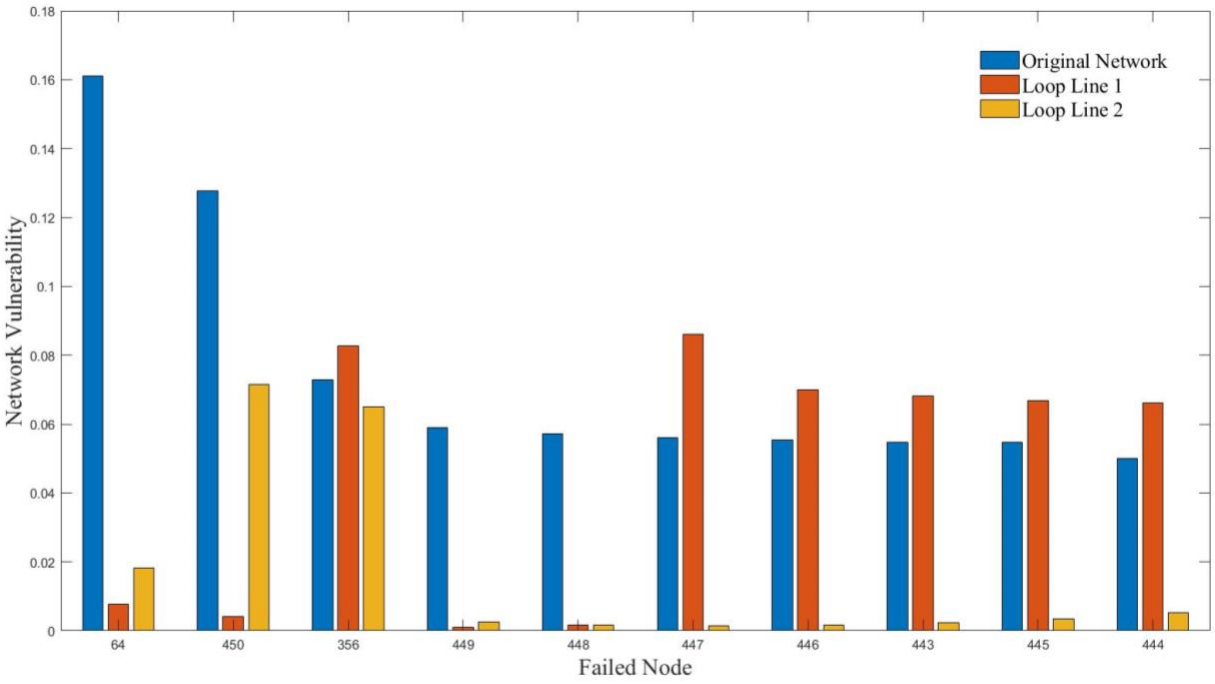

Figure 4. 13 Comparison of loop lines' impact on network vulnerability due to the failure of nodes.

Table 4.6 and Figure 4.14 show the impact of adding two loop lines on the 10 critical links in the Great Lakes area. After adding the loop line 1, the network vulnerability due to the failures of any links except $e_{59,60}$ has been greatly reduced, because the loop line 1 does not cover the link between node 59 and 60 inside the circle, which may cause negative impacts on this link. Adding the loop line 2 can enhance all selected critical links.

Table 4. 6 Comparison of loop lines' impact on network vulnerability due to the failure of links.

Ranking	Link	Numbering[a]	Vulnerability		
			Original network	Loop line 1	Loop line 2
1	(69,70)	71	7.97%	0.01%	0.07%
2	(63,64)	65	4.43%	0.03%	0.06%
3	(62,63)	64	4.08%	0.04%	0.06%
4	(61,62)	63	3.78%	1.36%	0.12%
5	(64,65)	66	3.52%	0.09%	0.13%
6	(60,61)	62	3.51%	0.13%	1.93%
7	(70,71)	72	3.29%	0.22%	1.36%
8	(59,60)	61	3.27%	3.98%	0.25%
9	(79,80)	80	3.23%	0.77%	0.79%
10	(71,72)	73	3.12%	0.05%	0.02%

[a]These link numberings are automatically generated by Matlab program.

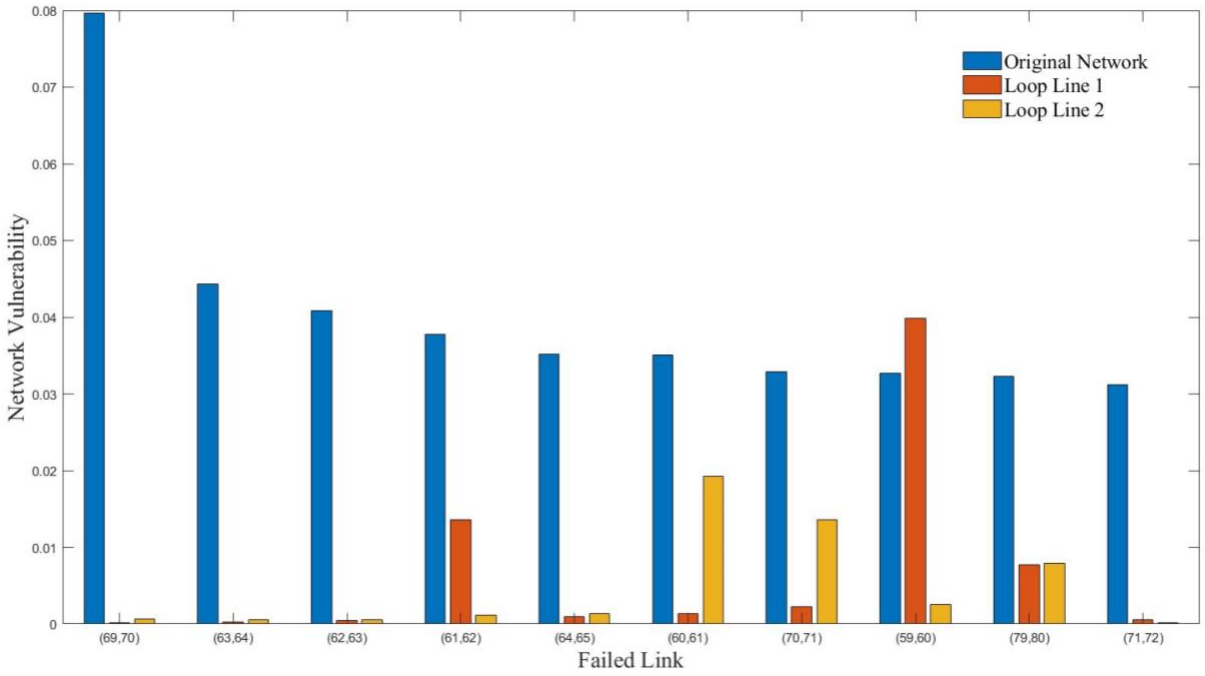

Figure 4. 14 Comparison of loop lines' impact on network vulnerability due to the failure of links.

Additionally, after adding hypothetical loop lines, the maximum network vulnerability of the Amtrak rail network needs to be assessed to examine the impact of loop lines on the entire network. Table 4.7 summarizes the change in the network efficiency, characteristic path length and vulnerability of the entire Amtrak network. As a result, adding the loop line 1, which

contains only 7, links improves the overall network efficiency by 9.1 percent and reduces the network vulnerability by 46.5 percent. Adding the loop line 2, which contains only 5 links, improves the overall network efficiency by 15.5 percent and reduces the network vulnerability by 46.6 percent. Therefore, adding a loop line can effectively enhance the topology of the Amtrak rail network.

Table 4. 7 The impact of adding loop lines on the entire Amtrak rail network.

Network type	Network efficiency	Characteristic path length	Network vulnerability
Original network	0.0463	35.40	16.10%
Network with loop line 1	0.0505	32.26	8.62%
Network with loop line 2	0.0535	30.12	8.60%

4.1.2.4 Recovery Strategies for Node 64

The most critical node 64 (i.e., Chicago (Union Station)) plays an essential role in the Amtrak rail network. Assuming a disruptive event of node 64 and its connected links, the optimal recovery sequence based on the resilience index needs to be determined. After node 64 and its connected links $e_{63,64}$, $e_{64,65}$, $e_{362,64}$, $e_{64,363}$, $e_{64,271}$, $e_{64,450}$ are disrupted, 1 node and 6 links need to be repaired to recover the network efficiency fully. Assuming only one component can be repaired in a recovery stage, the number of all possible recovery sequences is equal to 5,040, i.e., the permutation of 7. The resilience index can be calculated as Equation (3.31) by assuming that the time of each recovery stage is constant and that the network efficiency will not be recovered by the repair of links before the node is repaired.

The initial network efficiency is equal to 0.0464, and the network efficiency decreases to 0.0388 after the disruption of node 64 and its connected links. Figure 4.15 through Figure 4.18 (a) show the comparison of recovery triangles. Table 4.8 and Figure 4.18 (b) show the top 10 resilience index values of different recovery sequences. Recovery sequence s_{64}- (64,450)- (63,64)- (64,65)- (64,363)- (64,271)- (362,64) is identified as optimal, with the largest value of resilience index 0.9275, which means if node 64 and its connected links are disrupted, node 64 should be repaired first, then link $e_{64,450}$, $e_{63,64}$, $e_{64,65}$, $e_{64,363}$, followed by link $e_{64,271}$, and finally link $e_{362,64}$.

Table 4. 8 Top 10 Resilience index for the sequential recovery strategies of node 64.

Ranking	Recovery sequence	R_e
1	s_{64}- (64,450)- (63,64)- (64,65)- (64,363)- (64,271)- (362,64)	0.9275
2	s_{64}- (64,450)- (63,64)- (64,65)- (64,271)- (64,363)- (362,64)	0.9264
3	s_{64}- (64,450)- (63,64)- (64,363)- (64,65)- (64,271)- (362,64)	0.9263
4	s_{64}- (64,450)- (64,65)- (63,64)- (64,363)- (64,271)- (362,64)	0.9261
5	s_{64}- (64,450)- (64,65)- (63,64)- (64,271)- (64,363)- (362,64)	0.9251
6	s_{64}- (64,450)- (63,64)- (64,65)- (64,363)- (362,64)- (64,271)	0.9246
7	s_{64}- (64,450)- (63,64)- (64,271)- (64,65)- (64,363)- (362,64)	0.9244
8	s_{64}- (64,450)- (63,64)- (64,363)- (64,271)- (64,65)- (362,64)	0.9240
9	s_{64}- (64,450)- (64,65)- (64,363)- (63,64)- (64,271)- (362,64)	0.9235
10	s_{64}- (64,450)- (63,64)- (64,363)- (64,65)- (362,64)- (64,271)	0.9234

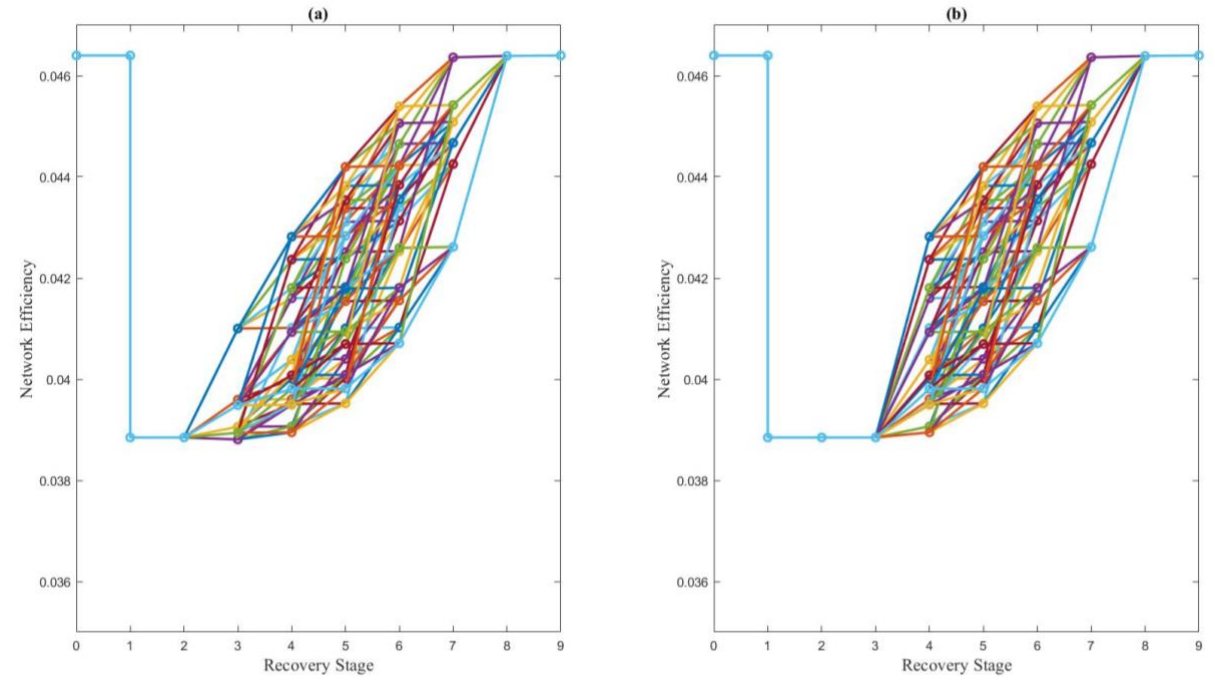

Figure 4. 15 (a) Node 64 repaired in the first order; (b) Node 64 repaired in the second order.

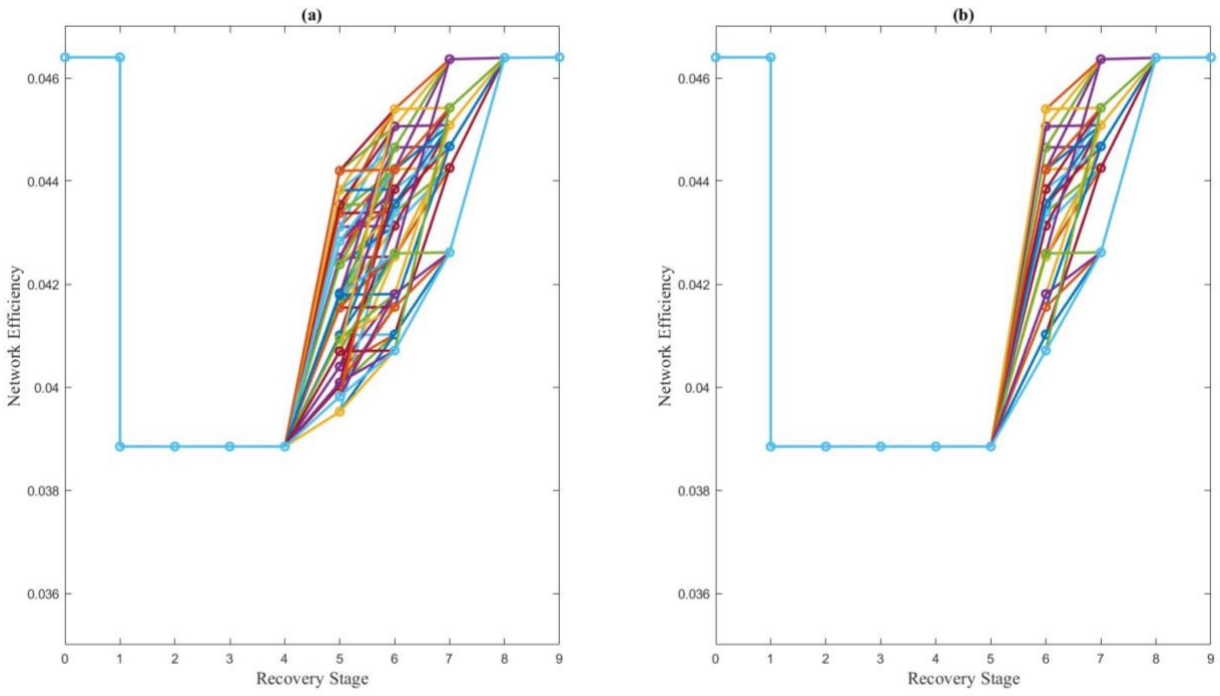

Figure 4. 16 (a) Node 64 repaired in the third order; (b) Node 64 repaired in the fourth order.

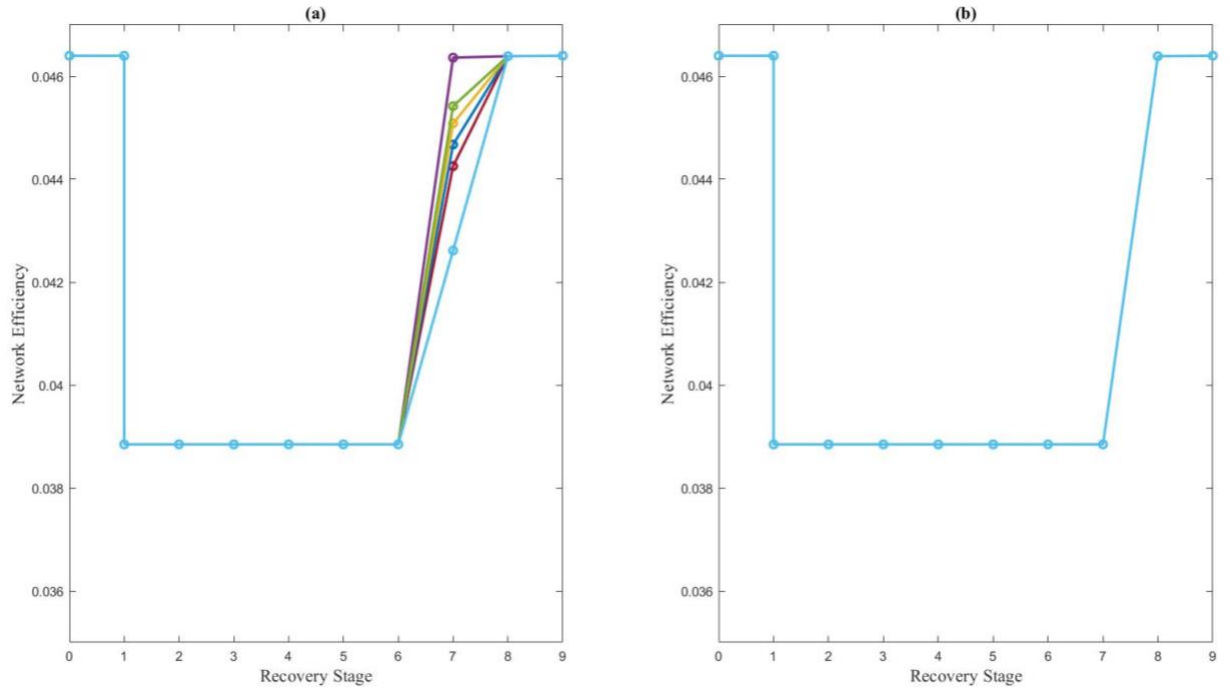

Figure 4. 17 (a) Node 64 repaired in the fifth order; (b) Node 64 repaired in the sixth order.

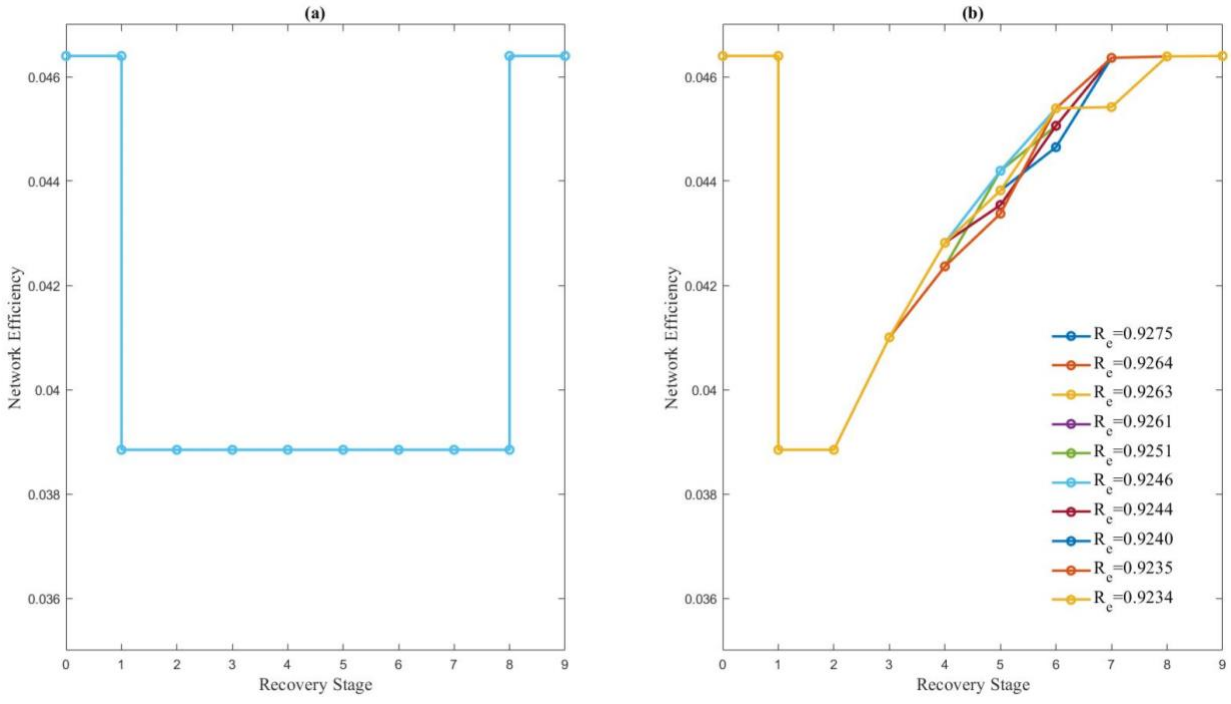

Figure 4. 18 (a) Node 64 repaired in the seventh order; (b) Top 10 resilience index.

4.2 Case Study2: Class I Rail Network

4.2.1 Mapping the Class I Rail Network

Each Class I railroad entity owns and operates its own rail network. However, the Class I entities can also share their rails with each other. Therefore, in this thesis, the Class I rail network is regarded as a whole for analysis and research, which serves more than 40,000 stations and 130,000 miles of rails in North America, including 46 states, Washington D.C. in the United States and several Canadian provinces. Compared to the metro and the Amtrak passenger rail networks, the Class I rail network has a larger number of nodes and links, and overly complicated connectivity patterns, which means the topology analysis and vulnerability assessment for the Class I rail network are very difficult. In order to improve efficiency while ensuring the accuracy of the analysis, the following method is proposed: selecting a part of nodes and links, then, reducing the size of the Class I rail network (i.e., reduce the number of nodes and links without changing the overall network connectivity):

1. The analysis of the Amtrak rail network indicates that the node density of areas has a great impact on the arrangement of critical nodes and links. Therefore, according to the actual number of Class I rail network stations, the number of nodes is selected in each state proportionally to ensure that the ratio of node density of each area remains unchanged;
2. All transfer nodes must be considered first during the process of reducing the network size, because they are the key components of network connectivity; and
3. Selecting more nodes on the longer rails can minimize the impact of link lengths, building an unweighted rail network.

Based on these three principles, the Class I rail network with 638 nodes and 860 links is defined and mapped. The network vector is specified as follows:

$$G = [638,860] \tag{4.2}$$

The Class I rail network can be divided into six areas depending on geographic locations: northwest (NW), southwest (SW), Great Lakes (GL), central south (CS), southeast (SE), and northeast (NE) areas. Figure 4.19 to 4.24 show the graph and numbering of the Class I rail network in each area.

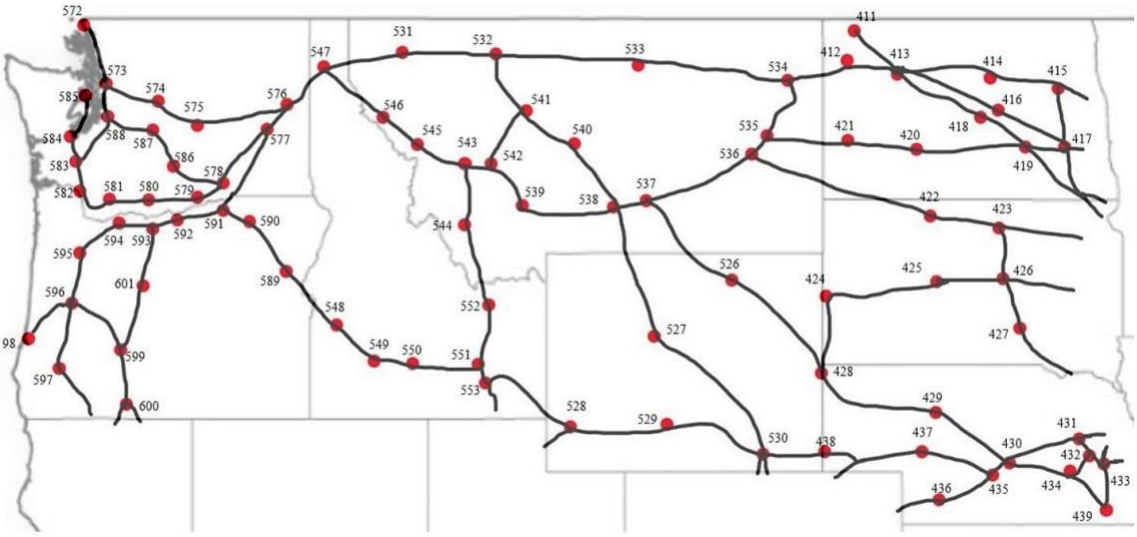

Figure 4. 19 Graph of the northwest area including Washington, Oregon, Idaho, Montana, Wyoming, North Dakota, South Dakota, and Nebraska.

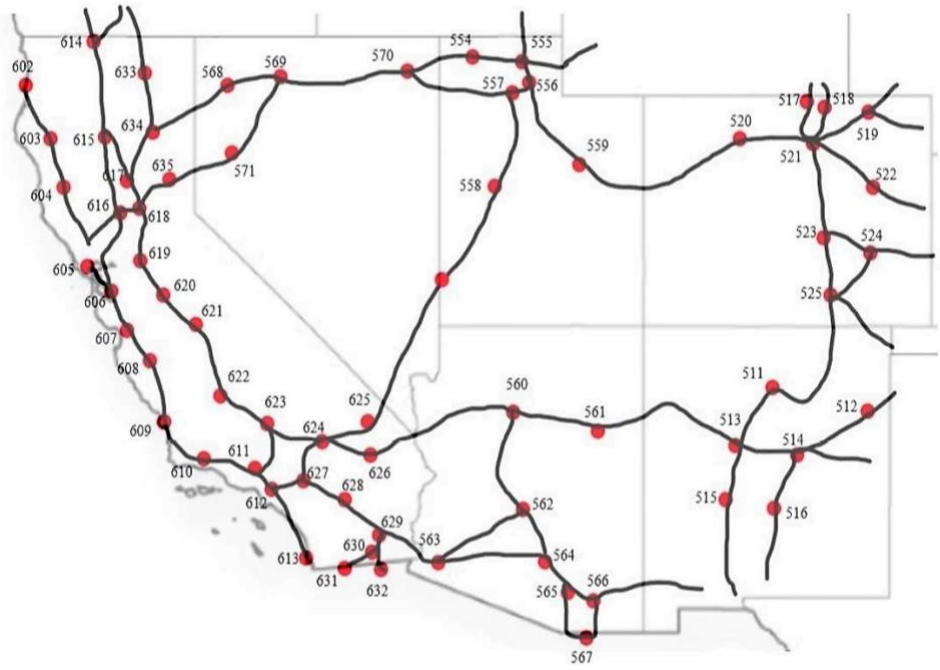

Figure 4. 20 Graph of the southwest area including California, Nevada, Utah, Arizona, Colorado, and New Mexico.

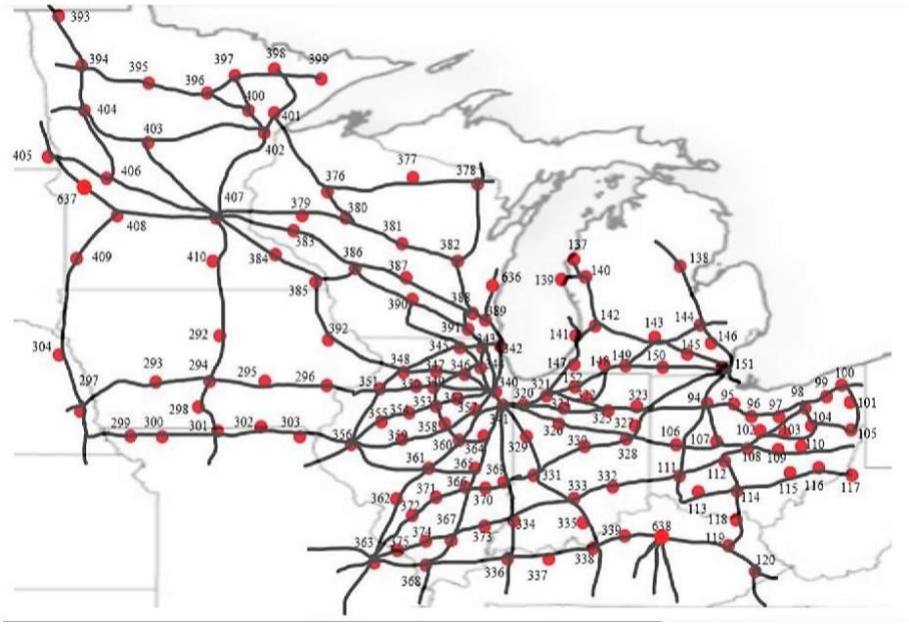

Figure 4. 21 Graph of the Great Lakes area including Minnesota, Iowa, Wisconsin, Illinois, Indiana, Michigan, and Ohio.

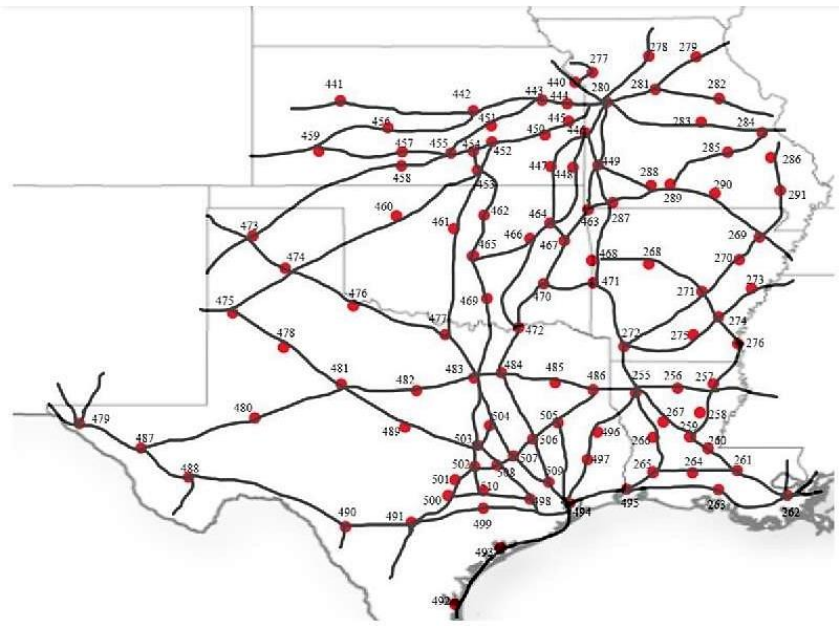

Figure 4. 22 Graph of the central south area including Kansas, Missouri, Oklahoma, Arkansas, Louisiana, and Texas.

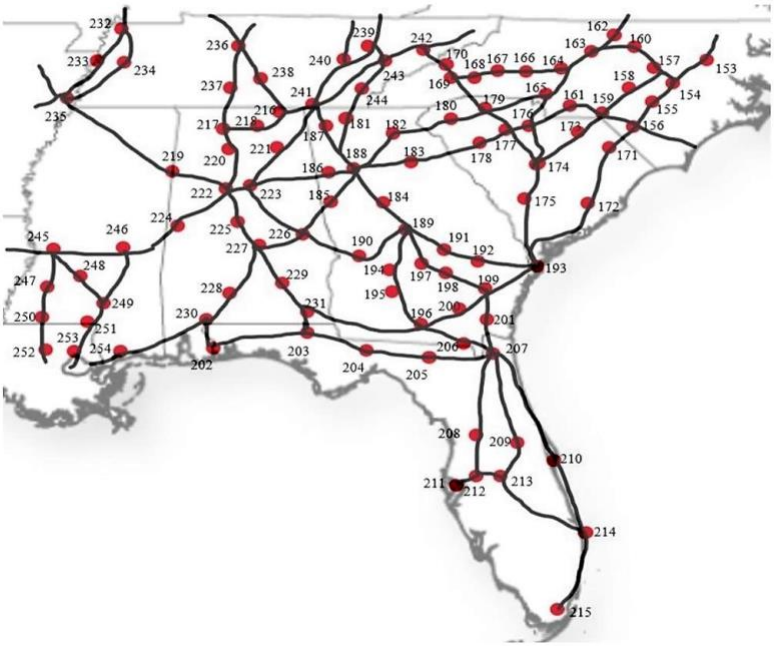

Figure 4. 23 Graph of the southeast area including Tennessee, North Carolina, South Carolina, Georgia, Alabama, Mississippi, and Florida.

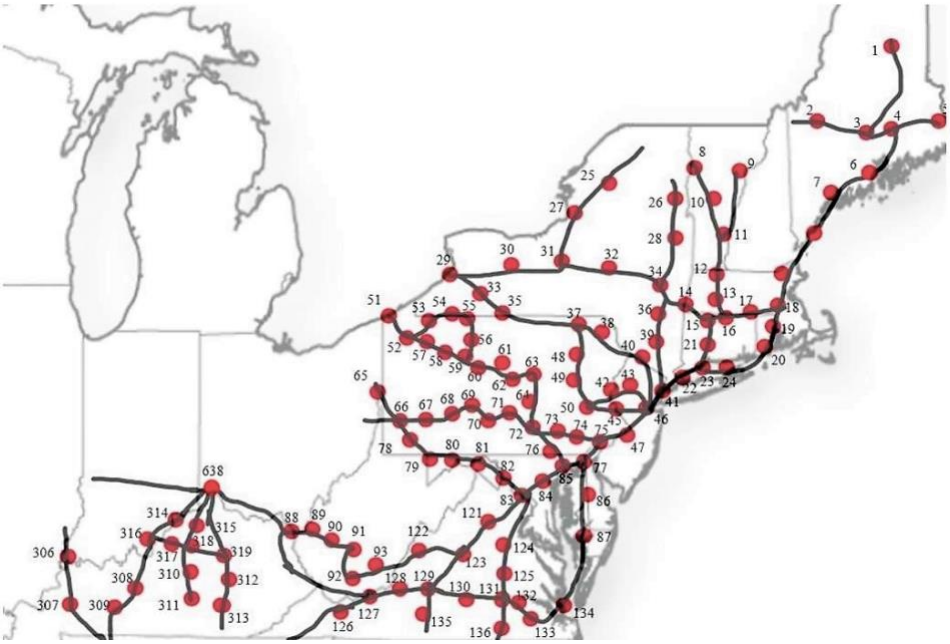

Figure 4. 24 Graph of the northeast area including Kentucky, West Virginia, Virginia, Maryland, Washington, D.C., Delaware, Pennsylvania, New Jersey, Connecticut, Rhode Island, New York, Massachusetts, Vermont, New Hampshire, and Maine.

4.2.2 Results and Analysis

4.2.2.1 Topology Analysis of the Class I Rail Network

The size of the adjacency matrix for the Class I rail network is 638×638. Table 4.9 summarizes the results of the Class I rail network's topological indicators. Compared with the Amtrak rail network, the characteristic path length and diameter of the Class I rail network have been lowered significantly, which means that the Class I rail network is more efficient and developed than the Amtrak rail network. Additionally, the local clustering coefficient indicates that the nodes of the Class I rail network are more clustered in local areas than the Amtrak rail network. The network topology characteristics reflected by topological indicators are consistent with the

Class I rail network as shown in Figure 4.25, indicating that the connectivity is accurately depicted after reducing the size of the Class I rail network.

Table 4. 9 Topological indicators of the Class I rail network.

No.	Topological Indicators	Values	Notes
1	Average node degree	2.696	/
2	Network density	0.004	/
3	Characteristic path length	16.726	/
4	Diameter of the network	46	/
5	Local clustering coefficient	0.033	Ignore nodes with node degree 1
6	Global network efficiency	0.085	/

4.2.2.2 Vulnerability Assessment of the Class I Rail Network

Table 4.10 shows the top 40 most critical nodes and Table 4.11 demonstrates the top 40 most critical links. Therefore, the vulnerability of the Class I rail network is equal to 0.0295. Compared with the vulnerability of the Amtrak network, the network vulnerability due to the failure of nodes and links in the Class I rail network has been dropped dramatically. This demonstrates that the Class I rail network is more developed and robust than the Amtrak rail network, because when any component is disrupted, more alternative paths in the Class I rail network can be used to reduce the loss of overall network efficiency (i.e., reduce the network vulnerability).

Table 4. 10 Top 40 critical nodes of the Class I rail network.

No.	State	Area[a]	Node Numbering	Node Degree	Vulnerability
1	New York	NE	29	5	2.95%
2	Missouri	CS	280	7	2.85%
3	lLLinois	GL	341	6	2.01%
4	Alabama	SE	222	5	1.80%
5	Florida	SE	207	6	1.63%
6	Tennessee	SE	235	5	1.51%
7	Minnesota	GL	407	8	1.42%
8	Michigan	GL	151	6	1.22%
9	lLLinois	GL	363	7	1.20%
10	New York	NE	31	3	1.19%
11	lLLinois	GL	340	8	1.18%
12	Alabama	SE	219	2	1.14%
13	Ohio	GL	119	3	1.10%
14	Colorado	SW	521	6	1.09%
15	Massachusetts	NE	16	3	1.02%
16	Ohio	GL	120	3	1.01%
17	Washington	NW	576	3	1.00%
18	New York	NE	34	4	0.98%
19	Virginia	NE	127	3	0.98%
20	Texas	CS	494	6	0.97%
21	New York	NE	30	2	0.96%
22	Indiana	GL	320	4	0.94%
23	Indiana	GL	336	4	0.93%
24	Massachusetts	NE	18	3	0.91%
25	Pennsylvania	NE	51	3	0.86%
26	Indiana	GL	338	4	0.82%
27	Virginia	NE	129	4	0.78%
28	lLLinois	GL	368	4	0.71%
29	New Hampshire	NE	7	2	0.70%
30	lLLinois	GL	356	6	0.67%
31	Michigan	GL	142	3	0.66%
32	Kansas	CS	443	3	0.66%
33	Kansas	CS	444	3	0.65%
34	Massachusetts	NE	13	2	0.64%
35	New York	NE	32	2	0.63%
36	Alabama	SE	223	4	0.58%
37	Virginia	NE	128	2	0.57%
38	Pennsylvania	NE	52	3	0.56%
39	Louisiana	CS	255	6	0.56%
40	Ohio	GL	114	4	0.55%

aNW (northwest area), SW (southwest area), GL (the Great Lakes area), CS (central south area), SE (southeast area), and NE (northeast area).

Table 4. 11 Top 40 critical links of the Class I freight network.

No.	Starting Node	Ending Node	Area[a]	Vulnerability
1	181	244	SE	1.15%
2	170	242	SE	1.13%
3	31	32	NE	1.09%
4	7	18	NE	0.99%
5	35	37	NE	0.97%
6	13	16	NE	0.94%
7	124	125	NE	0.89%
8	282	363	GL	0.88%
9	6	7	NE	0.81%
10	37	38	NE	0.74%
11	12	13	NE	0.74%
12	33	35	NE	0.73%
13	53	54	NE	0.71%
14	38	40	NE	0.68%
15	126	127	NE	0.65%
16	582	583	NW	0.65%
17	4	6	NE	0.65%
18	129	135	NE	0.64%
19	143	145	GL	0.63%
20	131	136	NE	0.59%
21	446	448	CS	0.58%
22	11	12	NE	0.55%
23	379	380	GL	0.53%
24	34	14	NE	0.53%
25	32	34	NE	0.50%
26	307	236	NE	0.50%
27	553	555	NW	0.48%
28	624	625	SW	0.48%
29	576	577	NW	0.47%
30	309	236	NE	0.46%
31	216	238	SE	0.44%
32	606	607	SW	0.44%
33	581	582	NW	0.43%
34	125	131	NE	0.42%
35	412	534	NW	0.42%
36	507	509	CS	0.42%
37	123	129	NE	0.41%
38	48	49	NE	0.41%
39	87	134	NE	0.40%
40	248	249	SE	0.40%

[a] NW (northwest area), SW (southwest area), GL (the Great Lakes area), CS (central south area), SE (southeast area), and NE (northeast area).

Figure 4.25 shows the arrangement of the Class I rail network's top 40 critical nodes, where blue circles indicate the top 20 critical nodes and green circles are the top 21-40 critical nodes. Compared with the Amtrak rail network, the top 40 critical nodes of the Class I rail network are scattered instead of being concentrated in any area. Figure 4.26 demonstrates the arrangement of

the Class I rail network's top 40 critical links, where blue squares indicate the top 20 critical nodes and green squares are the top 21-40 critical links. More than 75 percent of the top 20 critical links are concentrated in the northeast area. Additionally, the most critical node 29 is also located in this area. As a result, the northeast area of the Class I rail network is identified as the most critical area.

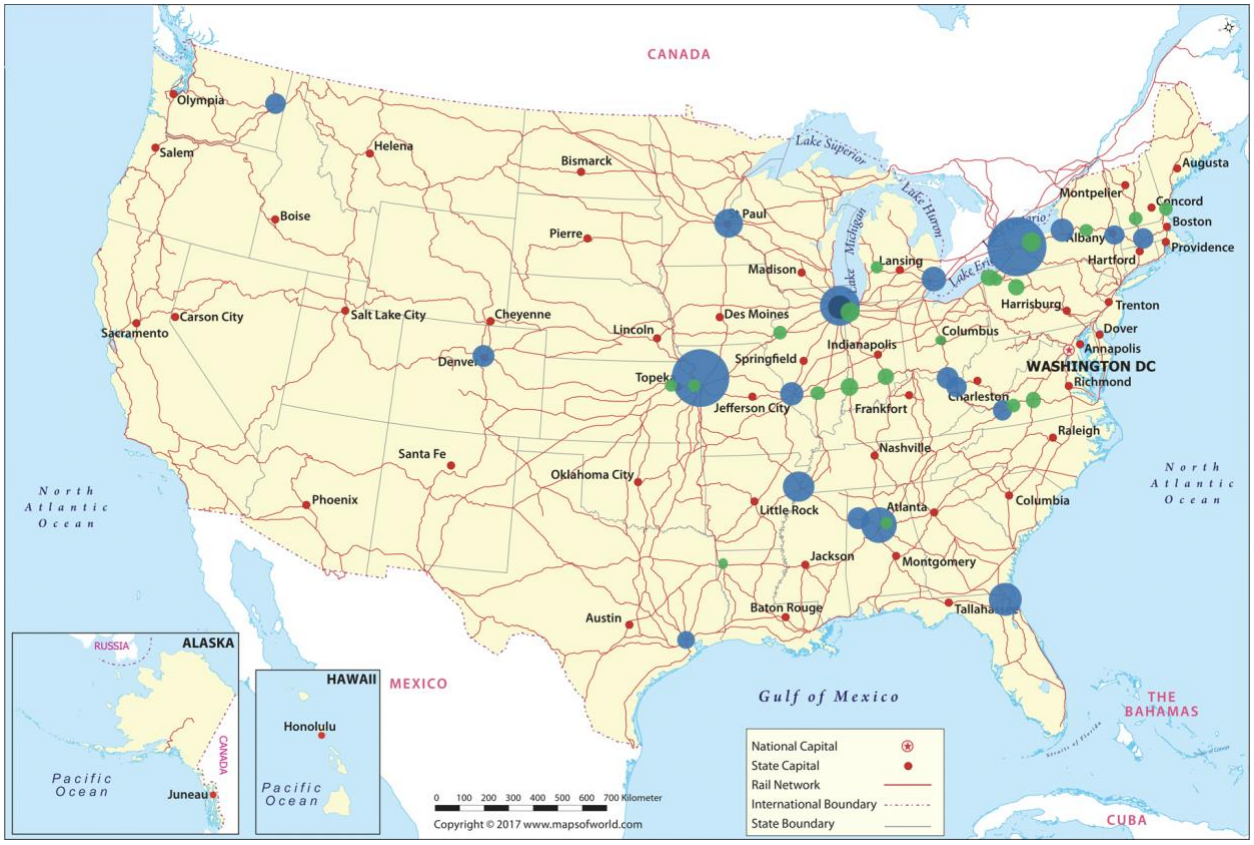

Figure 4. 25 Arrangement of the Class I rail network's top 40 critical nodes.

(Purchased, permitted and edited from Mapsofworld.com. (2020). US Railroad Map, US Railway Map, USA Rail Map for Routes. <https://www.mapsofworld.com/usa/usa-rail-map.html> (Accessed 31 May 2020).)

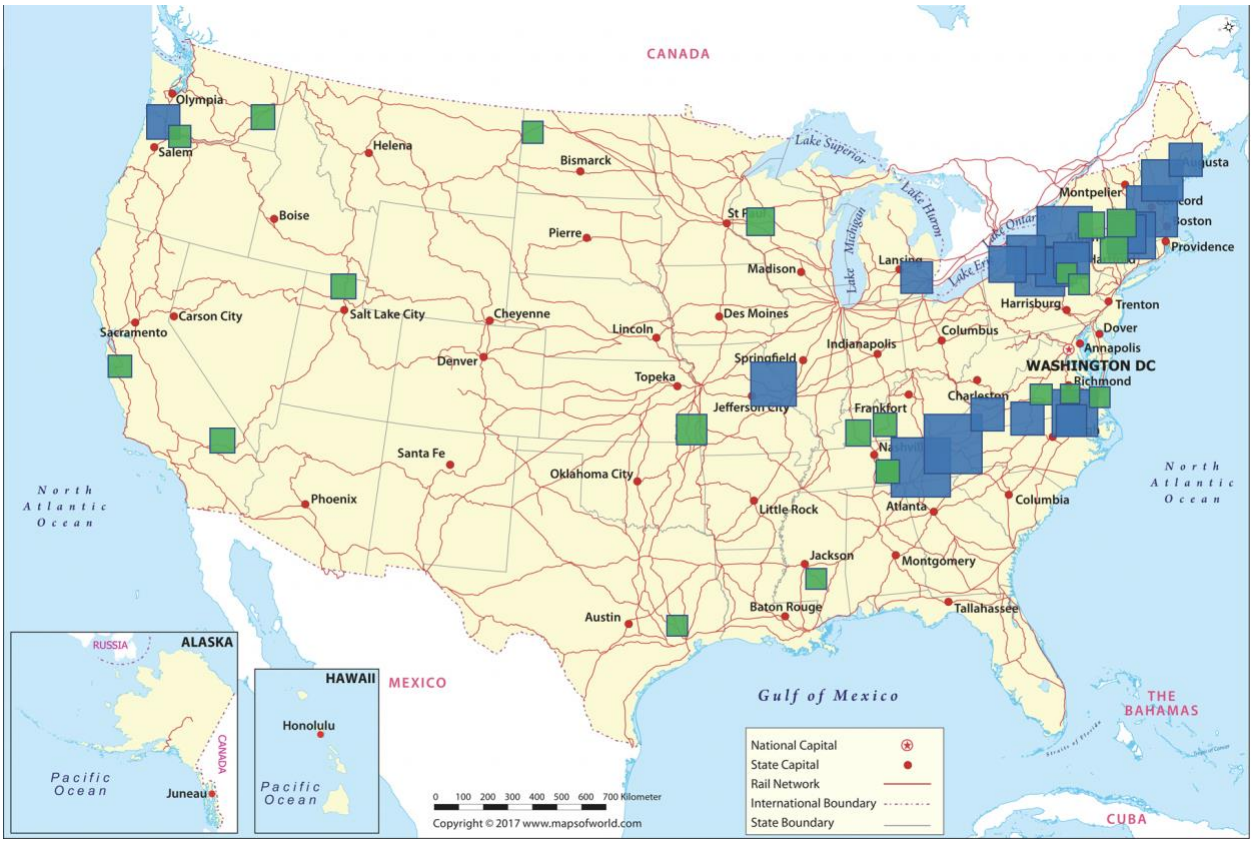

Figure 4. 26 Arrangement of the Class I rail network's top 40 critical links.

(Purchased, permitted and edited from Mapsofworld.com. (2020). US Railroad Map, US Railway Map, USA Rail Map for Routes. <https://www.mapsofworld.com/usa/usa-rail-map.html> (Accessed 31 May 2020).)

Based on the analysis of the Amtrak rail network, critical nodes are concentrated in the typical connectivity pattern 1 in areas with high node density. However, critical nodes of the Class I rail network are not concentrated in any areas but are scattered in almost all areas. Table 4.12 demonstrates the comparison of the Class I rail network's six areas. Node density is the number of nodes per unit area. Diameter of an area is defined as the maximum value of the shortest path length between all pairs of nodes in this area. In this table, the value in each column is also the ratio between any two areas' characteristics. From Table 4.12, the differences in characteristics between the Class I rail network's six areas are small with the exception of node density, which

is similar to the analysis result of the Amtrak rail network. Additionally, the differences in node density between the Class I rail network's areas are smaller than the Amtrak rail network, which may explain why the top 40 critical nodes of the Class I rail network are scattered in each area rather than being concentrated on one or two areas.

Table 4. 12 Comparison of the Class I rail network's six areas.

Area	Node Density	Average Node Degree	Characteristic Path Length	Diameter
Northwest	1.25	1.12	1.32	1.40
Southwest	1.00	1.11	1.08	1.07
Great Lakes	3.52	1.31	1.27	1.40
Central South	2.02	1.30	1.00	1.00
Southeast	3.01	1.16	1.07	1.13
Northeast	4.26	1.00	1.67	1.87

Figure 4.27 shows the correlation between network vulnerability due to node failures and the node degree of failed node for the Class I rail network. In this case, network vulnerability is not positively linear correlated with the node degree. Because in the Class I rail network, many loop lines are around the nodes with larger node degree, which greatly reduces the network vulnerability due to node failures, thereby reducing the correlation between node vulnerability and node degree. For example, nodes 340 in the Class I rail network has the same location and largest node degree as nodes 64 in the Amtrak rail network. However, the vulnerability due to the failure of node 340 in the Class I rail network has dropped significantly. From Figure 4.21, several loop lines are around nodes 340. One of the loop lines is $e_{342,343}$, $e_{343,345}$, $e_{345,348}$, $e_{348,351}$, $e_{351,356}$, $e_{356,361}$, and $e_{361,365}$. Another loop line is $e_{343,344}$, $e_{344,346}$, $e_{346,349}$, $e_{349,353}$, $e_{353,358}$, $e_{358,360}$, and $e_{360,364}$. These loop lines significantly reduce the criticality of nodes 340.

When node 340 and its connected links are disrupted, loop lines create redundancy and can be used as alternative routes leading to the reduction in network vulnerability. For example, if node 340 is disrupted, the shortest path length from node 378 to node 352 is 9 with loop lines, while the shortest path length becomes 17 without loop lines. Figure 4.28 shows the changes in the network vulnerability due to the failure of nodes 340 and 407 after losing loop lines around them. Afterward, the correlation coefficient between network vulnerability due to node failures and node degree for the top 20 critical nodes increases to 0.6814.

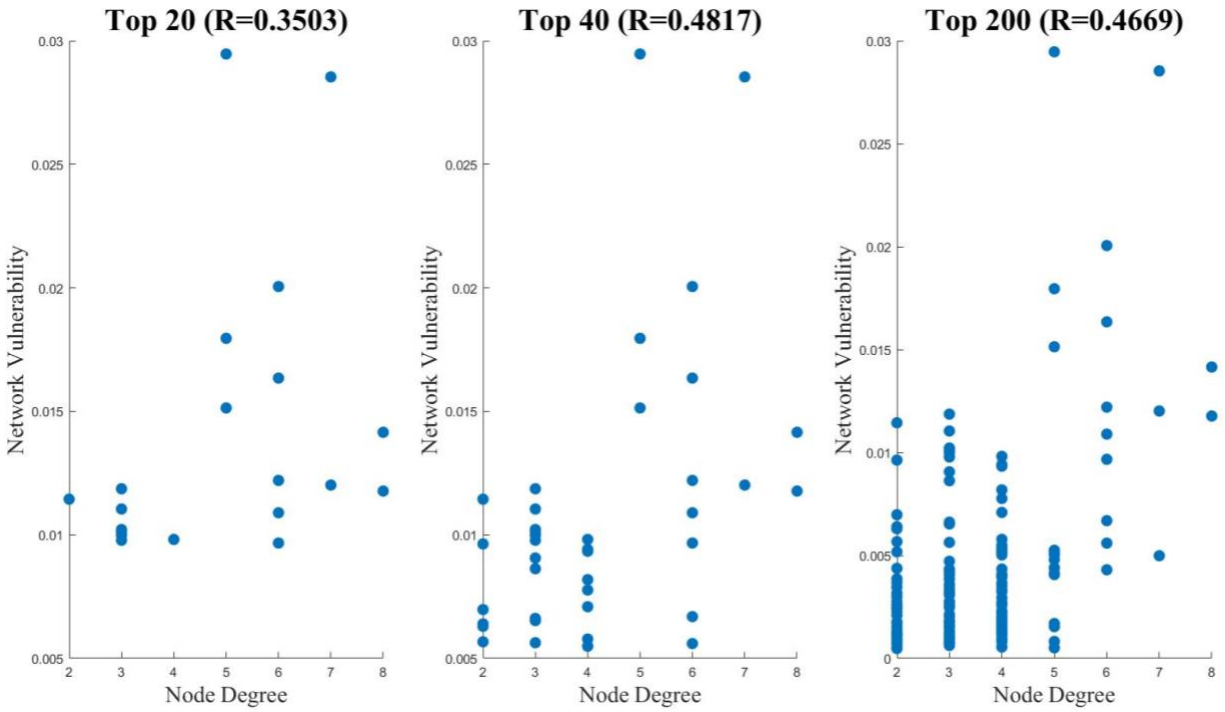

Figure 4. 27 Correlation between network vulnerability due to node failures and the node degree of the failed node.

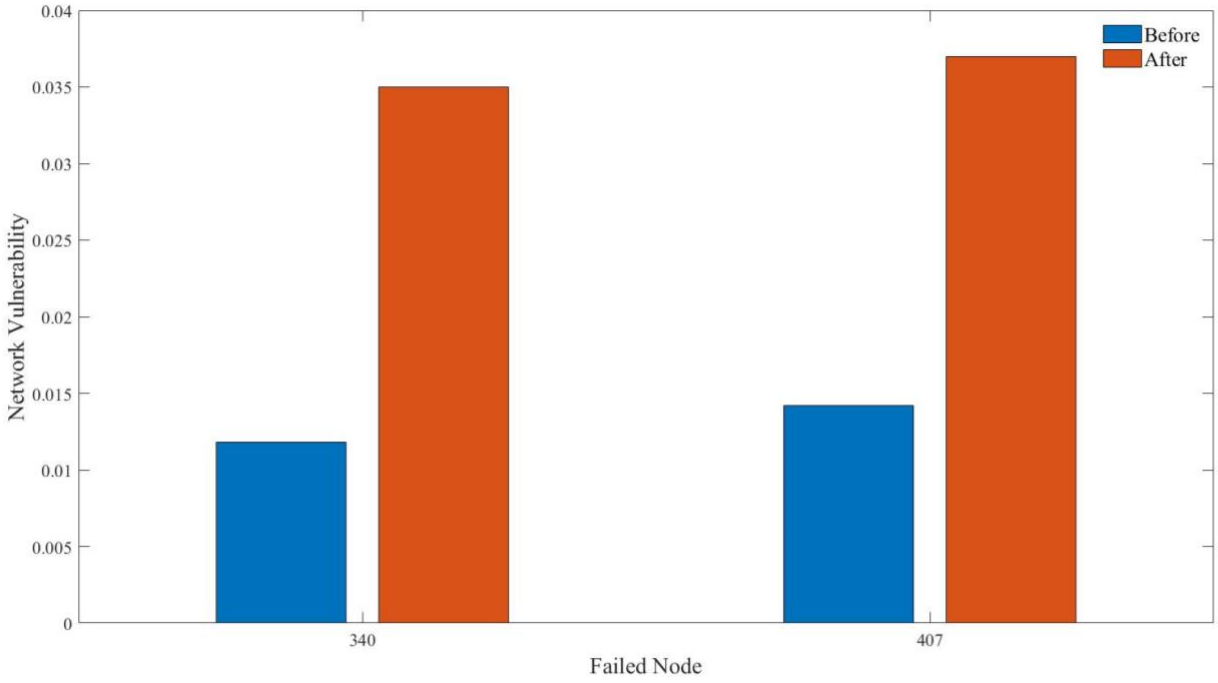

Figure 4. 28 Changes in the network vulnerability due to the failure of nodes 340 and 407 after losing loop lines.

Table 4.11and Figure 4.26 show the value and arrangement of the top 40 critical links. Compared with the Amtrak rail network, the network vulnerability due to the failure of links for the Class I rail network has dropped significantly, and 75 percent of the top 20 critical links are concentrated in the northeast area with the highest node density. Additionally, most of the top 40 critical links are located in the typical connectivity pattern 1 or the line with the "end-node," which is the same as the conclusion of the Amtrak rail network.

4.2.2.3 Topology Enhancement of the Class I Rail Network

Depending on the vulnerability assessment, the northeast area is identified as the most critical area of the Class I rail network. A hypothetical loop line is added around the most critical node

29 to create redundancy for the northeast area, enhancing the network topology and reducing the vulnerability. In this section, the vulnerability changes in 6 nodes (i.e., nodes 29, 30, 31, 32, 34, and 16) and 9 links (i.e., $e_{31,32}$, $e_{35,37}$, $e_{13,16}$, $e_{37,38}$, $e_{12,13}$, $e_{33,35}$, $e_{38,40}$, $e_{34,14}$, and $e_{32,34}$) in the northeast area are examined to verify if the topology enhancement is applicable to the Class I rail network.

The hypothetical loop line for the Class I rail network connects existing nodes consisting of 7 links: $e_{25,26}$, $e_{26,10}$, $e_{10,17}$, $e_{17,21}$, $e_{21,39}$, $e_{39,40}$, and $e_{40,49}$, which is the yellow line as shown in Figure 4.29.

Figure 4. 29 Hypothetical loop lines added into the Class I rail network.

Table 4.13 and Figure 4.30 show the changes in network vulnerability due to the failure of nodes in the northeast area after adding the hypothetical loop line. All examined nodes fall inside the

circle where node 29 is the center, and the loop line is the arc. Therefore, the vulnerability due to the failure of nodes 31, 16, 34, 30, and 32 has been reduced significantly to almost zero. Also, the vulnerability due to the failure of node 29 has been reduced by half. As a result, adding loop lines in the Class I rail network effectively reduces the network vulnerability due to the failure of nodes.

Table 4. 13 Changes in network vulnerability due to the failure of nodes after adding a loop line.

			Vulnerability	
Ranking	Numbering[a]	Area	Original network	Loop line
1	29	NE	2.95%	1.43%
2	31	NE	1.19%	0.087%
3	16	NE	1.02%	0.026%
4	34	NE	0.98%	0.092%
5	30	NE	0.96%	0.112%
6	32	NE	0.63%	0.036%

[a]These node numberings are provided in Figure 4.24.

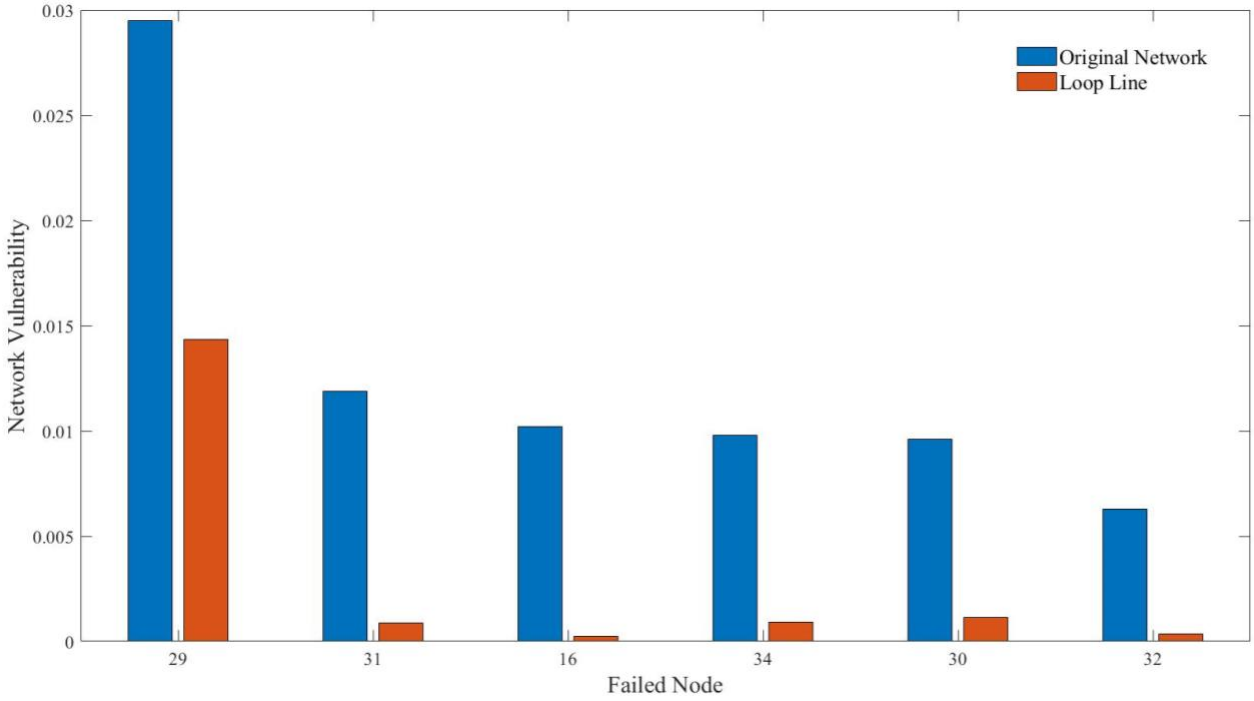

Figure 4. 30 Changes in network vulnerability due to the failure of nodes after adding a loop line.

Table 4.14 and Figure 4.31 show the changes in network vulnerability due to the failure of links in the northeast area after adding the hypothetical loop line. All examined links are located inside the circle. Therefore, adding the hypothetical loop line causes the positive impacts on the examined links, reducing network vulnerability significantly.

Table 4. 14 Changes in network vulnerability due to the failure of links after adding a loop line.

			Vulnerability	
Ranking	Link	Numbering[a]	Original network	Loop line 1
1	(31,32)	43	1.09%	0.10%
2	(35,37)	47	0.97%	0.32%
3	(13,16)	15	0.94%	0.024%
4	(37,38)	49	0.74%	0.099%
5	(12,13)	14	0.74%	0.009%
6	(33,35)	45	0.73%	0.390%
7	(38,40)	51	0.68%	0.091%
8	(34,14)	17	0.53%	0.055%
9	(32,34)	44	0.50%	0.083%
10	(71,72)	73	3.12%	0.05%

[a]These link numberings are automatically generated by Matlab program.

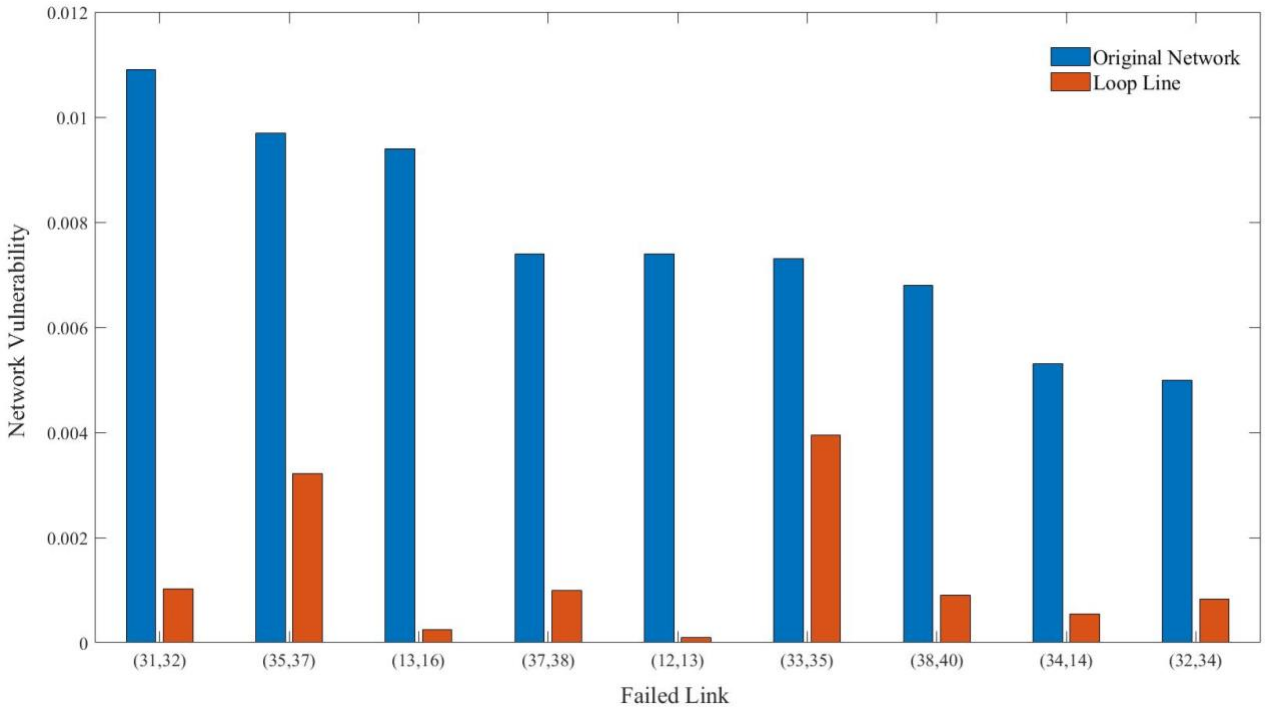

Figure 4. 31 Changes in network vulnerability due to the failure of links after adding a loop line.

Additionally, after adding loop lines, the maximum network vulnerability of the Class I rail network needs to be assessed to examine the impact of the addition of loop lines on the entire Class I rail network. Table 4.15 summarizes the changes in the efficiency, characteristic path length, and the vulnerability of the entire Class I rail network. The entire network has been enhanced slightly, indicating that adding a loop line in the northeast area cannot effectively enhance the topology of the entire Class I rail network. As a result, for the Class I rail network with complex and developed connectivity among nodes and links, more hypothetical loop lines need to be designed and added in several areas rather than the most critical area to enhance the topology of the entire network.

Table 4. 15 Impacts of adding the loop line on the entire Class I rail network.

Network type	Network efficiency	Characteristic path length	Network vulnerability
Original network	0.085	16.73	2.95%
Network with the loop line	0.0863	16.54	2.84%

4.2.2.4 Recovery Strategies for Node 29

The most critical node 29 plays a critical role in the Class I rail network. Assuming potential failures lead to the disruption of node 29 and its connected links, the optimal recovery sequence that generates the largest value of resilience index needs to be determined to enhance the network resilience. After node 29 and its connected links $e_{29,30}$, $e_{29,33}$, $e_{29,51}$, $e_{29,144}$, and $e_{29,151}$ are disrupted, 1 node and 6 links need to be repaired to recover the network efficiency fully. Assuming only one component can be repaired in a recovery stage, the number of all possible recovery sequences is equal to 720, i.e., the permutation of 6. The resilience index can be

calculated using Equation (3.31) by assuming that the time of each recovery stage is constant and that the network efficiency will not be recovered by the repair of links before the node repair. The initial network efficiency is equal to 0.085, and the network efficiency decreases to 0.0828 after the disruption of node 29 and its connected links.

Figure 4.32 to Figure 4.34 show the comparison of recovery triangles. In Figure 4.32, repairing a link has negative impacts on the recovery of network efficiency. The reason for the occurrence of this phenomenon is that after repairing a node, the number of nodes is also increased by one when calculating network efficiency. Depending on Equation 3.29, the network efficiency decreases when the denominator increases more than the numerator. Table 4.16 and Figure 4.35 show the top 10 resilience index values of different recovery sequences. Therefore, if node 29 and its connected links are disrupted, link $e_{29,151}$ should be repaired first, then node 29, link $e_{29,30}$, $e_{29,51}$, followed by link $e_{29,33}$, and finally link $e_{29,144}$.

Table 4. 16 Top 10 resilience index for the sequential recovery strategies of node 29.

Ranking	Recovery sequence	R_e
1	*(29,151)- s_{29}- (29,30)- (29,51)- (29,33)- (29,144)*	0.9835
2	*s_{29}- (29,30)- (29,151)- (29,51)- (29,33)- (29,144)*	0.9833
3	*(29,151)- s_{29}- (29,33)- (29,30)- (29,51)- (29,144)*	0.9826
4	*s_{29}- (29,151)- (29,30)- (29,51)- (29,144)- (29,33)*	0.9825
5	*s_{29}- (29,33)- (29,151)- (29,30)- (29,51)- (29,144)*	0.9824
6	*s_{29}- (29,30)- (29,151)- (29,51)- (29,144)- (29,33)*	0.9823
7	*(29,151)- s_{29}- (29,33)- (29,51)- (29,30)- (29,144)*	0.9816
8	*s_{29}- (29,51)- (29,30)- (29,151)- (29,33)- (29,144)*	0.9815
9	*s_{29}- (29,33)- (29,151)- (29,51)- (29,30)- (29,144)*	0.9814
10	*s_{29}- (29,30)- (29,151)- (29,144)- (29,33)- (29,51)*	0.9813

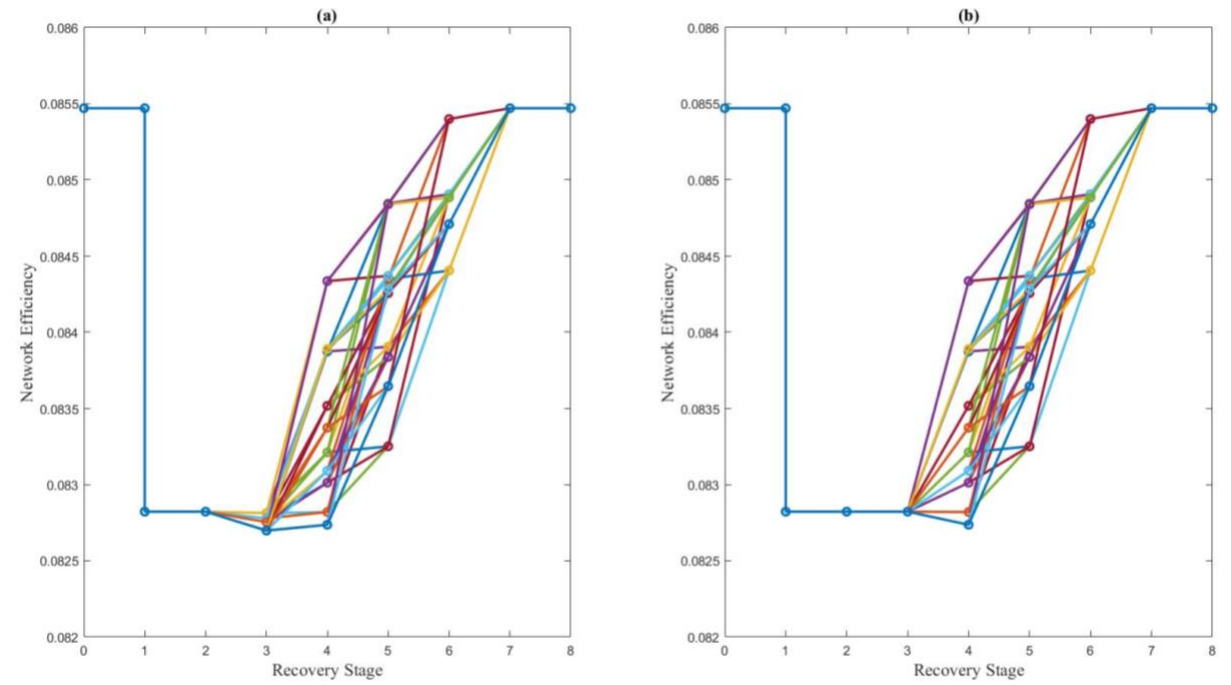

Figure 4. 32 (a) Node 29 repaired in the first order; (b) Node 29 repaired in the second order.

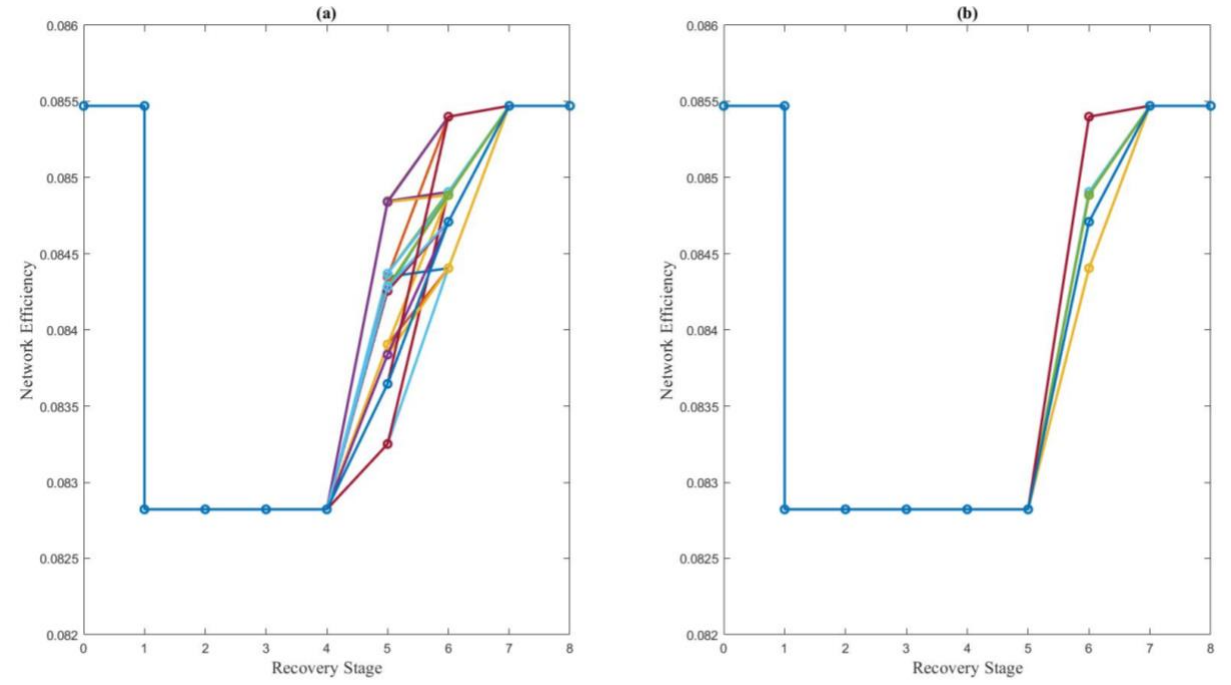

Figure 4. 33 (a) Node 29 repaired in the third order; (b) Node 29 repaired in the fourth order.

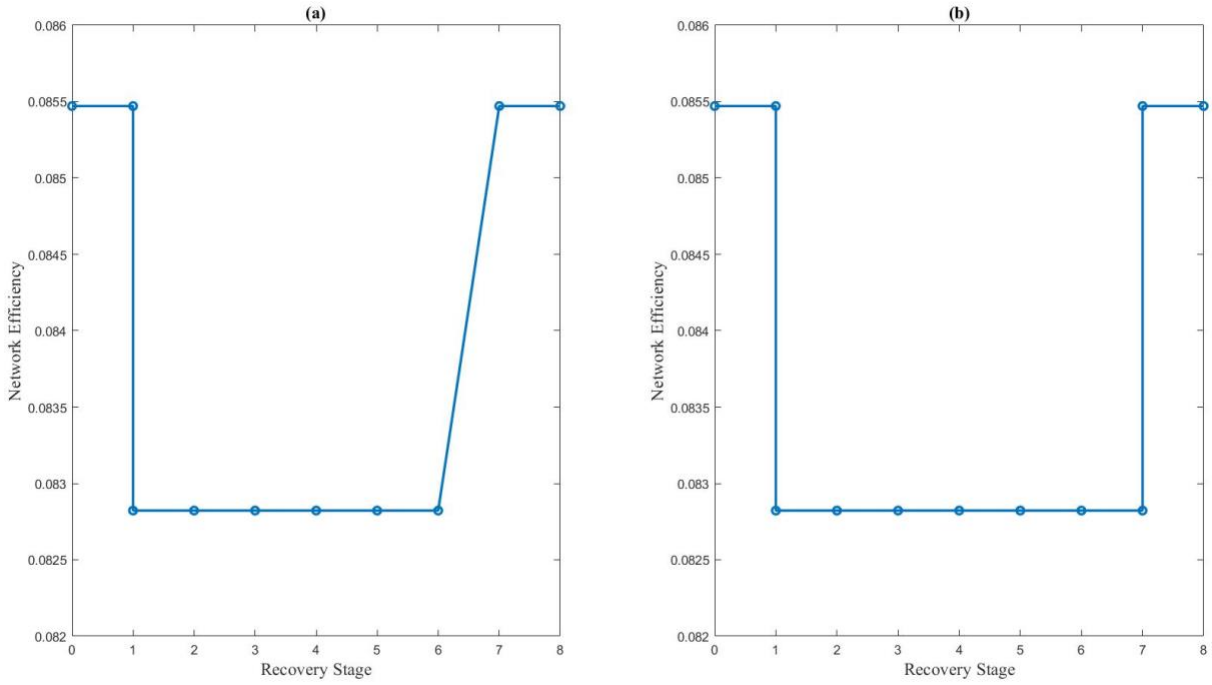

Figure 4. 34 (a) Node 29 repaired in the fifth order; (b) Node 29 repaired in the sixth order.

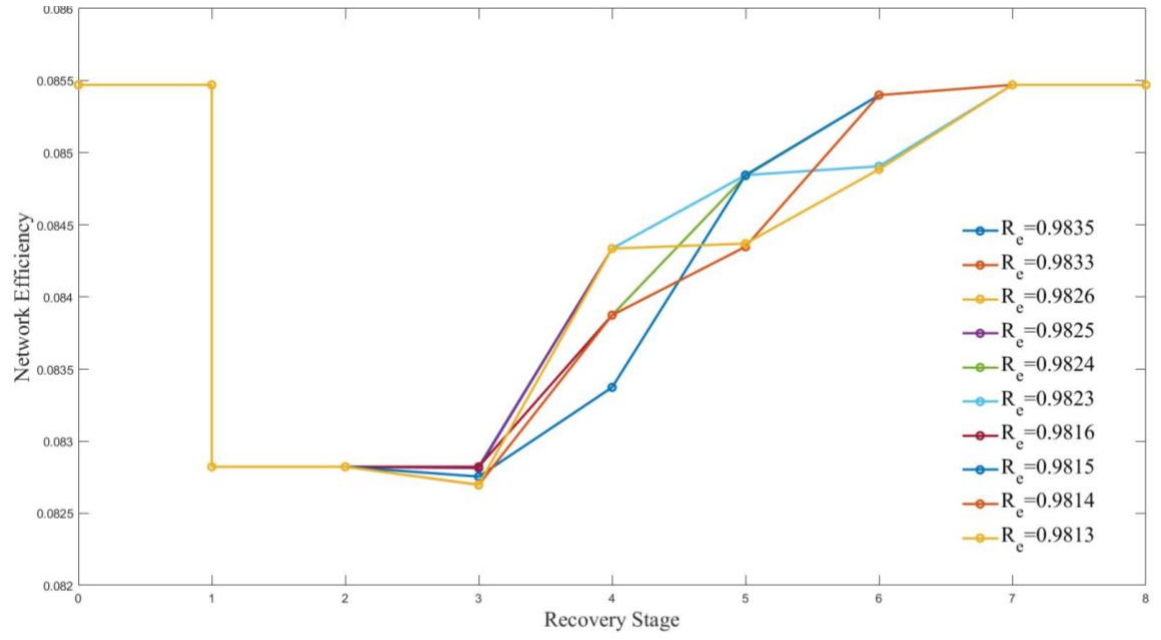

Figure 4. 35 Top 10 resilience index.

4.3 Discussion of the Proposed Methodology Results

The methodology proposed in Chapter 3 is applied to analyze the Amtrak and Class I rail networks from the topological perspectives, and to further reduce the impacts of potential failures. The calculation of topological indicators, such as average node degree, characteristic path length, and local clustering coefficient, can well describe the arrangement of and the connectivity among these two networks' components. The topology enhancement and sequential recovery strategies succeed in reducing the vulnerability and enhancing the resilience of the Amtrak and Class I rail networks based on the vulnerability assessment to identify the critical components and area.

4.3.1 Summaries of the Amtrak Rail Network

The Amtrak rail network contains 529 nodes and 552 links and can be divided into five areas as shown in Figure 4.2 to Figure 4.6. The size of the Amtrak rail network's adjacency matrix is 529×529. Using the equations in Chapter 3 to calculate the indicators can effectively analyze the topology of the Amtrak rail network. Table 4.1 summarizes the results of the Amtrak rail network's topological indicators.

The size of matrices and the number of iterations in the vulnerability assessment of the Amtrak rail network is also not overly large. Table 4.2 shows the top 40 most critical nodes and Table 4.3 demonstrates the top 40 most critical links. These critical nodes and links are identified as the critical components of the Amtrak rail network. Based on the arrangement of the critical components as shown in Figure 4.7 and 4.8, the Great Lakes area is identified as the critical area of the Amtrak rail network. Additionally, most of the critical nodes of the Amtrak rail network

are the nodes with a large node degree located in the typical connectivity pattern 1 of high node density areas. Most critical links of the Amtrak rail network are located in the typical connectivity pattern 1 or the line with the “end-node” of high node density areas.

The topology enhancement and sequential recovery strategies are used to reduce the impacts of potential failures on the Amtrak rail network by reducing vulnerability and enhancing resilience. Table 4.5 to 4.7 and Figure 4.13 to 4.14 indicate that adding a proper loop line in the critical area of the Amtrak rail network can effectively reduce the network vulnerability and improve the network efficiency. Additionally, the resilience of the Amtrak rail network can be well measured based on the changes in network efficiency. After node 64 and its connected links are disrupted, recovery sequence s_{64}- (64,450)- (63,64)- (64,65)- (64,363)- (64,271)- (362,64) is identified as the optimal one with the largest value of resilience index 0.9275.

In summary, the proposed methodology can be used effectively to analyze the topology, vulnerability, and resilience of the Amtrak rail network. The topology enhancement and sequential recovery strategies are also applicable to the Amtrak network to reduce the impacts of potential failures.

4.3.2 Summaries of the Class I Rail Network

The real Class I rail network contains more than 40,000 nodes and links. The size of the Amtrak rail network’s adjacency matrix is beyond 40000×40000. In addition, when calculating the network vulnerability, the number of iterations exceeds 128 trillion. Therefore, measuring the topological indicators and vulnerability of the Class I rail network is extremely difficult and

inefficient. In this case, the size of the Class I rail network is reduced to 638 nodes and 860 links using the method described in section 4.2.1, while keeping the connectivity of the Class I rail network unchanged. Table 4.9 summarizes the results of the Class I rail network's topological indicators, indicating that the Class I rail network is more efficient and developed than the Amtrak rail network.

Table 4.10 shows the top 40 most critical nodes and Table 4.11 demonstrates the top 40 most critical links. These critical nodes and links are identified as the critical components of the Class I rail network. Based on the arrangement of the critical components as shown in Figure 4.25 and 4.26, the northeast area is identified as the critical area of the Class I rail network. Compared with the Amtrak rail network, the critical nodes of the Class I rail network are more scattered instead of being concentrated in any area. Meanwhile, critical nodes with greater vulnerability may not have a larger node degree. Most of the critical links of the Class I rail network are still located in the typical connectivity pattern 1 or the line with the "end-node" of high node density areas.

Table 4.13 to 4.15 and Figure 4.30 to 4.31 indicate that adding a proper loop line in the critical area of the Class I rail network can still effectively reduce the vulnerability due to the failure of nodes and links in the critical area. However, the network efficiency, characteristic path length, and the maximum network vulnerability of the Class I rail network has been enhanced slightly. Additionally, the optimal recovery sequence for node 29 can also be determined with respect to the largest resilience index.

In summary, after reducing the size of the Class I rail network, the proposed methodology can be well used to analyze the network topology, vulnerability, and resilience accurately and effectively. Also, the topology enhancement and sequential recovery strategy is applicable to the Class I rail network to reduce the impacts of potential failures.

Chapter 5: Conclusions

Rail networks are real-life examples of complex networks and critical logistic and economic contributors to the wellbeing of society. The potential failures of rail networks due to natural or human-caused hazards can cause disruptions, leading to severe consequences including significant economic impacts. This thesis proposes a methodology to analyze rail networks with a large number of nodes, links, and complex connectivity from the topological perspectives and further evaluate and reduce the impacts of potential failures on rail networks. Based on the analysis of two case studies: the Amtrak and Class I rail networks, it can be concluded that the proposed methodology is appropriate and well suited to analyze the topology, vulnerability, and resilience of complex rail networks effectively and efficiently. The results also indicate that topology enhancements and recovery strategies are applicable to rail networks in order to reduce vulnerability and enhance resilience.

The network topology, defined as the arrangement and connectivity among components, can be clearly depicted, quantified, and assessed using complex network theory. Additionally, under the premise of ensuring network connectivity remains unchanged, the size of rail networks with an overly large number of components were reduced, successfully increasing the efficiency of analysis, and breaking through the current limitations on rail network analysis. In comparing the two analyzed networks in this thesis, the Class I rail network is more complicated, efficient, and developed than the Amtrak rail network.

The vulnerability of rail networks is a critical factor in evaluating the impacts of node and link disruptions due to potential failures. The vulnerability of rail networks is measured to compare

the impact of all possible failures on the networks. The reduction in the size of the Class I rail network ensures the analysis of the network vulnerability is effective and efficient. The critical areas can be clearly depicted and identified depending on the arrangement of the most critical nodes and links.

This thesis provides and examines a pre-failure topology enhancement strategy to reduce the impact of potential failures by adding hypothetical loop lines in critical areas to reduce network vulnerability. The vulnerability of most critical components can be reduced to almost zero, indicating that inserting a proper hypothetical loop line can significantly reduce the vulnerability of components in the critical areas. Furthermore, the results also reflect that in a network, such as the Amtrak rail network, adding loop lines in an area can effectively increase the entire network efficiency and reduce the network vulnerability. However, in some networks with overly complicated connectivity, such as the Class I rail network, adding loop lines in an area may not effectively enhance the entire network. Other hypothetical loop lines need to be examined in future work by adding them in several areas rather than the most critical area.

Network efficiency assessment also provides a basis to measure the resilience index based on assumed recovery sequences. The resilience index is used to quantify the ability of rail networks to withstand and recover from the failure of components. In order to enhance network resilience due to the failure of the most critical node (i.e., station) and its connected links (i.e., rails), this thesis compares all possible recovery sequences and ranks them with respect to the corresponding resilience index values. The recovery sequence with the largest resilience index value represents the best recovery strategy. In this case, after a failure occurs, the overall loss of

network efficiency can be effectively minimized to reduce impacts. It should be emphasized that the repair of nodes (i.e., stations) is essential in the sequential recovery strategy, which may significantly affect the restoration of network efficiency.

Future work can focus on combining several kinds of data sources to analyze the rail networks, such as travel time, passenger or freight flow, distance, and train routes. By inputting such data, the railroad blocking or traveling salesman problem can be solved to optimize the railroad's operations. In addition, inputting the disruption-related historical data can help us to investigate the probability of disruption occurrences and to assess potential disruptions and their impacts on rail networks.

Finally, the reduced Class I rail network is intended to select and analyze a subset of nodes in the network while keeping the network connectivity unchanged. For different analysis purposes, additional nodes need to be added to the subset and analyzed. The recovery strategy in this thesis is only for the disruption case of one node and links connected to it. The same recovery analysis process can be used to determine the recovery strategies for practical situations according to different disruption cases and assumptions involving several nodes concurrently.

Reference

Adjetey-Bahun, K., Birregah, B., Châtelet, E., and Planchet, J. (2016). "A model to quantify the resilience of mass railway transportation systems". *Reliability Engineering & System Safety*, 153, 1-14.

Amtrak. (2018). "*FY 2018 Company Profile*." <https://www.amtrak.com/content/dam/projects/dotcom/english/public/documents/corporate/nationalfactsheets/Amtrak-Corporate-Profile-FY2018-0319.pdf> (accessed 5 April 2020).

Amtrak. (2016). "Amtrak FY2016 Ridership & Revenue." <https://media.amtrak.com/wp-content/uploads/2015/10/Amtrak-FY16-Ridership-and-Revenue-Fact-Sheet-4_17_17-mm-edits.pdf> (accessed 11 June 2020).

Association of American Railroads. (2017). "*Freight Railroads In United States*." <https://www.aar.org/wp-content/uploads/2019/01/AAR-United-States-Fact-Sheet.pdf> (accessed 5 April 2020).

Association of American Railroads. (2019). "*Coal: Freight Rail Powers The Nation*." <https://www.aar.org/article/freight-rail-coal/> (accessed 21 April 2020).

Association of American Railroads. (2019). "*The Economic Impact Of America'S Freight Railroads*." <https://www.aar.org/wp-content/uploads/2018/05/AAR-Economic-Impact-US-Freight-Railroads.pdf> (accessed 5 April 2020).

Ayyub, B. (2014). Risk Analysis in Engineering and Economics, 2nd Ed., CRC Press.

Azad, N., Hassini, E., and Verma, M. (2016). "Disruption risk management in railroad networks: An optimization-based methodology and a case study". *Transportation Research Part B: Methodological*, 85, 70-88.

Barabási, A., and Pósfai, M. (2016). *Network science*. Cambridge University Press, Cambridge.

Barabási, A., and Albert, R. (1999). "Emergence of Scaling in Random Networks". *Science*, 286(5439), 509-512.

Batarlienė, N. (2008). "RISK ANALYSIS AND ASSESSMENT FOR TRANSPORTATION OF DANGEROUS FREIGHT". *TRANSPORT*, 23(2), 98-103.

Bešinović, N. (2020). "Resilience in railway transport systems: a literature review and research agenda". *Transport Reviews*, 1-22.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. (2006). "Complex networks: Structure and dynamics". *Physics reports*.

Bollobás, B. (1985). *Random Graphs*. Academic Press, London.

Braha, D. (2017). "Complex Design Networks: Structure and Dynamics". *SSRN Electronic Journal*.

Bruneau, M., Chang, S., Eguchi, R., Lee, G., O'Rourke, T., Reinhorn, A., Shinozuka, M., Tierney, K., Wallace, W., and von Winterfeldt, D. (2003). "A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities". *Earthquake Spectra*, 19(4), 733-752.

Byers, W. (2011). "Railroad Damage from Two Hurricanes". *Natural Hazards Review*, 12(1), 6-8.

Chan, R., and Schofer, J. (2016). "Measuring Transportation System Resilience: Response of Rail Transit to Weather Disruptions". *Natural Hazards Review*, 17(1), 05015004.

Chang, K., Kim, K., Oshima, H., and Yoon, S. (2006). "Subway Networks in Cities". *Journal of the Korean Physical Society*.

Chen, L., and Miller-Hooks, E. (2012). "Resilience: An Indicator of Recovery Capability in Intermodal Freight Transport". *Transportation Science*, 46(1), 109-123.

Chen, Z., and Wang, Y. (2019). "Impacts of severe weather events on high-speed rail and aviation delays". *Transportation Research Part D: Transport and Environment*, 69, 168-183.

Cherkassky, B., Goldberg, A., and Radzik, T. (1996). "Shortest paths algorithms: Theory and experimental evaluation". *Mathematical Programming*, 73(2), 129-174.

Chinowsky, P., Helman, J., Gulati, S., Neumann, J., and Martinich, J. (2019). "Impacts of climate change on operation of the US rail network". *Transport Policy*, 75, 183-191.

Dawson, D., Shaw, J., and Roland Gehrels, W. (2016). "Sea-level rise impacts on transport infrastructure: The notorious case of the coastal railway line at Dawlish, England". *Journal of Transport Geography*, 51, 97-109.

Derrible, S., and Kennedy, C. (2010). "The complexity and robustness of metro networks". *Physica A: Statistical Mechanics and its Applications*, 389(17), 3678-3691.

DesRoches, R. (2006). *Hurricane Katrina*. Performance of Transportation Systems.

Ding, D. (1994). "关于最短路径的 SPFA 快速算法". *西南交通大学学报*, 29.

Federal Railroad Administration. (2010). “National Rail Plan Progress Report.” <https://railroads.dot.gov/elibrary/national-rail-plan-progress-report> (accessed 5 April 2020).

Ferranti, E., Chapman, L., Lowe, C., McCulloch, S., Jaroszweski, D., and Quinn, A. (2016). "Heat-Related Failures on Southeast England's Railway Network: Insights and Implications for Heat Risk Management". *Weather, Climate, and Society*, 8(2), 177-191.

Garrison, W.L. and Marble, D.F. (1962). The structure of transportation networks, No. TR-62-11, Northwestern Univ Evanston Il.

Gedik, R., Medal, H., Rainwater, C., Pohl, E., and Mason, S. (2014). "Vulnerability assessment and re-routing of freight trains under disruptions: A coal supply chain network application". *Transportation Research Part E: Logistics and Transportation Review*, 71, 45-57.

Henry, D., and Emmanuel Ramirez-Marquez, J. (2012). "Generic metrics and quantitative approaches for system resilience as a function of time". *Reliability Engineering & System Safety*, 99, 114-122.

Irani, D., Siers, M., Menking, C. and Nickey, Z. (2018). "*Economic And Fiscal Impact Analysis Of Class I Railroads In 2017*." <https://www.aar.org/wp-content/uploads/2018/11/AAR-Class-I-Railroad-Towson-Economic-Impact-October-2018.pdf> (accessed 11 June 2020).

Janić, M. (2018). "Modelling the resilience of rail passenger transport networks affected by large-scale disruptive events: the case of HSR (high speed rail)". *Transportation*, 45(4), 1101-1137.

Latora, V., and Marchiori, M. (2002). "Is the Boston subway a small-world network?". *Physica A: Statistical Mechanics and its Applications*, 314(1-4), 109-113.

Latora, V., and Marchiori, M. (2001). "Efficient Behavior of Small-World Networks". *Physical Review Letters*, 87(19).

Lawrence, S. (2015). "Freight Railroads Background." <https://railroads.dot.gov/elibrary/freight-railroads-background> (accessed 5 April 2020).

Morrison, S. (1990). "The Value of Amtrak". *The Journal of Law and Economics*, 33(2), 361-382.

Musso, A., and Vuchic, V. (1988). "Characteristics of metro networks and methodology for their evaluation". *TRANSPORTATION RESEARCH RECORD*.

Ostroumova Prokhorenkova, L. (2015). "Global Clustering Coefficient in Scale-Free Weighted and Unweighted Networks". *Internet Mathematics*, 12(1-2), 54-67.

Peterman, D. (2017). "Amtrak: Overview." <https://fas.org/sgp/crs/misc/R44973.pdf> (accessed 5 May 2020).

Saadat, Y., Ayyub, B., Zhang, Y., Zhang, D., and Huang, H. (2019). "Resilience of Metrorail Networks: Quantification With Washington, DC as a Case Study". *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, 5(4).

Saadat, Y., Ayyub, B., Zhang, Y., Zhang, D., and Huang, H. (2020). "Resilience-Based Strategies for Topology Enhancement and Recovery of Metrorail Transit Networks". *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 6(2), 04020017.

The White House Office of the Press Secretary. (2013). "Presidential Policy Directive -- Critical Infrastructure Security And Resilience." <https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil> (accessed 11 June 2020).

Watts, D., and Strogatz, S. (1998). "Collective dynamics of 'small-world' networks". *Nature*, 393(6684), 440-442.

West, D. (1995). *Introduction to graph theory*. Prentice Hall, Englewood Cliffs, N.J.

Yin, H., Wu, J., Sun, H., Qu, Y., Yang, X., and Wang, B. (2018). "Optimal Bus-Bridging Service under a Metro Station Disruption". *Journal of Advanced Transportation*, 2018, 1-16.

Zhang, D., Du, F., Huang, H., Zhang, F., Ayyub, B., and Beer, M. (2018). "Resiliency assessment of urban rail transit networks: Shanghai metro as an example". *Safety Science*, 106, 230-243.

Zhang, J., Xu, X., Hong, L., Wang, S., and Fei, Q. (2011). "Networked analysis of the Shanghai subway network, in China". *Physica A: Statistical Mechanics and its Applications*, 390(23-24), 4562-4570.

Zhu, Y., Xie, K., Ozbay, K., Zuo, F., and Yang, H. (2017). "Data-Driven Spatial Modeling for Quantifying Networkwide Resilience in the Aftermath of Hurricanes Irene and Sandy". *Transportation Research Record: Journal of the Transportation Research Board*, 2604(1), 9-18.