Assessing Network Vulnerability of Rail Transport Networks – Nodal Connectivity, Partial Node Failure, and Shortest Path Network Problems

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Qian Ye

August 2018

This dissertation is dedicated to

my parents, grandparents,

love, and life.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude and special appreciation to my advisor Dr. Hyun Kim. His considerate guidance has continuously supported me through all my time of Ph.D. study, academic research, and writing of this dissertation. He cares about his students with warm patience. He has stimulated my potential and passion in network modelling and optimization with his tremendous knowledge. He is such a fabulous mentor that his motivation and inspiration have been priceless to me in various aspects.

I would like to thank all my dissertation committee members – Drs. Shih-Lung Shaw, Bruce Ralston, Lee Han, and Russell Zaretzki – for their courses that I gained a lot from and their insightful comments to this dissertation. In Dr. Shaw's class, I learned knowledge and skills on mining big data and modeling human dynamics. I greatly appreciate his support in connecting me with Oak Ridge National Laboratory and offering me the Research Assistant opportunity with National Transportation Research Center. Through Dr. Ralston's class, I built up my knowledge and techniques in web-based GIS. His enthusiasm for the GIS related subjects and dedication for lifetime professor career inspires me and my academic research. I appreciate his thoughtful and detailed suggestions to my dissertation and the time and effort he has spent in making this dissertation a better one. Through Dr. Lee Han's class, I gained knowledge on transportation characteristics from the perspective of civil engineering. He presented theories and practice in such an easy-to-understand and humorous way. I would like to thank Dr. Russell Zaretzki, for being my advisor for the M.S. in Statistics program and serving on my doctoral committee as well. His well-organized and easy-to-follow classes have contributed substantially to my deep understanding in statistics and quantitative methods.

I would like to thank Dr. Liem Tran for providing me the opportunity of Graduate Research Assistant and involving me in the Joint Directed Research and Development research project collaborated between The University of Tennessee, Knoxville and Oak Ridge National Laboratory; thank the Department of Geography for offering me the Graduate Teaching Assistantship and continuously supporting my doctoral study financially till my graduation; thank Ms. Norma Galyon and Ms. Pamela Hope Sharpe for their administrative work taking good care of the entire department including me.

Last but not the least, my special thanks go to my family, especially my parents and grandparents. Life has been so much easier and enjoyable with their endless love and unconditional support.

ABSTRACT

Plenty of literature has examined the vulnerability of transportation networks. To identify appropriate measures of connectivity for heavy rail systems, this research presents a comprehensive measure named *Degree of Nodal Connection* (DNC) index along with a new classification of transfer stations – *Mandatory Transfer* (MT), *non-Mandatory Transfer* (non-MT), and *End Transfer* (ET). The DNC index reevaluates nodal connectivity among various types of transfer stations in heavy rail networks with multiple lines. The concept of *partial node failure* is proposed and addressed in network modeling, and the disruption results are compared between partial and complete node failures using four local and global DNC indexes. Numerical assessment is presented with a case study of major heavy rail networks in the United States. To incorporate network flow in addition to nodes and links, the research proposes two network optimization models in order to identify and evaluate critical components of a flow-based network from the perspective of shortest paths. The Shortest Path Network model (SPN) identifies the optimal flow distribution across the network where all flows find their shortest paths from origin to destination. The Assessing Nodal Disruption in Shortest Paths Network model (AND-SPN) assesses the influence of *r* nodes' disruption on the network flow pattern. Supplementary criticality indicators are provided to reflect average arc use, average arc cost, and average arc flow. In the case study of the Amtrak rail system, the criticality of stations is evaluated comprehensively in terms of objective function and criticality. To accommodate partial node failure, the SPN model is further expanded to the SPN for Partial node failure model (SPN^{Pr}) by introducing link attribute index and distance update constraints. A case study on the Washington Metropolitan Area Transit Authority (WMATA) network is carried out using daily and monthly station-to-station flow to assess nodal criticality of MTs and thus network vulnerability. From different perspectives of concern, the indexes and optimization models proposed in the dissertation can help decision makers and planners to decide which rail stations are the most important to protect during special events or disasters or when seeking to reduce transportation network vulnerability.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Research Background and Research Questions

Proper management of critical infrastructures is a paramount issue in the fields of homeland security and disaster preparedness (Presidential Decision Directive, 1998). Concerns about homeland security have been prompted by major disasters during the past several decades, such as the *Galaxy IV* satellite failure in 1998, the 9/11 attacks on the World Trade Center in 2001, the blackout in North America in 2003, and Hurricane Katrina in 2005. These are mirrored by the concerns of those working in research fields that emphasize and identify critical parts of infrastructure and their influence on the vulnerability of spatially networked systems, including those supplying water, energy, telecommunications, and transportation (Church et al., 2004; Taylor et al., 2006; Murray et al., 2008; Nakanishi, 2009; Akgun et al., 2010; Strandberg, 2013).

In the literature of vulnerability and resilience of transportation networks, vulnerability generally depicts the decrease of network performance or efficiency when malfunction occurs at specific network elements; resilience generally describes the ability of maintaining network performance or efficiency in response to the malfunction of specific network elements. Examples of the malfunction of network elements include the removal of nodes or arcs, the lengthening of arcs, and the degradation of arc capacity or node capacity. The malfunction of network elements could be caused by special events such as random or targeted attacks, natural disasters, traffic control, facility maintenance, and managerial purposes.

Specifically, in the context of chapter 2 of this dissertation, *vulnerability* at local scale is defined as the decrease of nodal connectivity due to complete or partial node failures; vulnerability at global scale is defined as the decrease of network size caused by complete or partial node failures. Correspondingly, *resilience* at local scale is defined as the ability to maintain nodal connectivity in complete or partial node failures; resilience at global scale is defined as the ability to keep the network as a whole in complete or partial node failures. In the context of chapters 3 and 4 of this dissertation, a vulnerable network is depicted by a substantial increase in total network disruption cost or flow reroute (and thus a substantial decrease of network efficiency) after specific nodes fail completely or partially. In addition, a node is considered as critical if it possesses a high value in nodal connectivity, or its failure significantly increases total network disruption cost or flow reroute cost in a shortest path network.

Network connectivity is an important set of measures of the network's vulnerability and resilience.

Basic concepts in network connectivity include connectivity, subnetwork, directed and undirected networks, and adjacency list. Network connectivity is consisted of two major components – arc connectivity and node connectivity. Specifically, this dissertation focuses on node connectivity. A network is connected if very pair of its nodes is connected; otherwise, the network is disconnected. A disconnected network contains two or more subnetworks. Two nodes *i* and *j* are connected if the network contains at least one path from node *i* to node *j*. A confusable term of connectivity is accessibility. *Accessibility* refers to the ease of reaching destinations, or in other words, how well two nodes are connected; meanwhile, *connectivity* addresses whether or not two nodes are connected. A directed network consists of a set of nodes and a set of directed arcs (i.e., the arcs are ordered pairs of distinct nodes) (Ahuja et al., 1993). Similarly, an undirected network consists of a set of nodes and a set of undirected arcs (i.e., the arcs are unordered pairs of distinct nodes). The arc adjacency list of a node is the arcs emanating from that node, while the node adjacency list is the set of nodes adjacent to that node.

With a rail system, disruptive effects on the system can be explained by changes in availability of nodes or linkages, total length of travel routes, amounts of passengers, etc., and thereby judging the degree of disruption suffered by stations and linkages (Matisziw et al., 2010; Kim et al., 2016). In the context of this dissertation, *disruption* refers to complete or partial node failure that impacts network structure or operation to a certain extent, where *complete* node failure refers to the removal of a node completely from the network, and *partial* node failure refers to the loss of the nodal function in transferring flows.

Network connectivity is a classical measure in assessing network vulnerability. However, each type of network has its own special characteristic, and thus a measure that works on one network is not guaranteed to perform well on another network. Especially for heavy rail systems, as more complicated rail networks, most existing indexes cannot well capture nodal connectivity of stations, since multiple links present among node pairs.

Most existing literature carry out disruption scenarios based on complete failure of network infrastructures that assume complete shutdown of stations or cutoff of linkages. Nevertheless, a facility failure does not necessarily cause the loss of entire operation, especially for rail systems with multiple transfer stations connected by multiple lines. In such cases, part of the function performed by nodes or links might be preserved so that the nodes or links remain operable despite some level of degradation of network operation.

In addition, transportation systems are not merely networks with nodes and links. Network flow – passenger or freight – is another, and usually crucial, component of transportation networks. A spatially networked transport system is a facility to transfer people and goods between demand and supply. Therefore, traditional optimization models, such as shortest path model, should be upgraded to embrace network flows so that all the three essential network components (i.e., nodes, links, and flow) can be simulated simultaneously when optimizing objective functions.

To address the problems represented above, this dissertation will answer the following research questions:

1. Chapter 2: How to re-classify transfer stations in rail networks to address and distinguish their different transferring functions? What can we do to improve existing nodal degree indexes to well capture the comprehensive nodal connectivity across different types of transfer stations? To what extent partial node failure scenarios are different from complete node failures when assessing network vulnerability at both local and global scales, based on the re-classification of transfer stations and the nodal degree indexes?

2. Chapter 3: How to address the issue of flows to the shortest path problems and obtain the best network flow distribution? Which node(s) is most critical in impacting network vulnerability in a shortest path network when removed from the network?

3. Chapter 4: What are the effects of partial node failures on assessing network vulnerability in a shortest path network?

The goals and scope of this dissertation are fourfold. Firstly, re-classify transfer stations in rail networks, and build a comprehensive index to fully capture the characteristic of nodal connectivity in order to quantitatively compare the criticality of rail stations. Secondly, analyze and address partial node failure in rail networks at both local level and global scale so that results from the more realistic scenarios (i.e., partial node failures) can be applied in guiding the practice of rail network management. Thirdly, incorporate network flow into traditional shortest path model so that all three network components (i.e., nodes, links, and flow) can be analyzed simultaneously in order to obtain the best network flow distribution in a shortest path network; extend the proposed Shortest Path Network model, and conduct complete node failure scenarios to assess network vulnerability in a shortest path network. Fourthly, accommodate partial node failure in the proposed Shortest Path Network model so that not only complete node failure but also partial node failure scenarios can be simulated to assess network vulnerability.

This dissertation is theoretically informed by several related literatures: studies of network vulnerability, network connectivity and nodal degree, shortest path problems and optimization, and network flows. The proposed research will draw from recent inquiries in these literatures, contributing methodologically, theoretically, or practically to each. Specifically, this dissertation will contribute to the literature in three ways: [1] methodologically, as we present advanced connectivity measures which demonstrate a better ability to capture networks' vulnerability compared to existing connectivity indexes, and we also construct advanced shortest path problems models to evaluate disruption scenarios of flow-based networks, [2] theoretically, we demonstrate the differential impact of complete node failure and partial node failure to network performance, and we propose an efficient and expanded shortest paths models for disruption, and [3] practically, as the index-based measures offer an simple but effective way to compare networks' performance, and the optimization-based models provide a general framework of criticality of network components, helping policy makers of network planning to assess the network which needs fortification of resilience and enhancement of transportation security.

1.2 Organization of the Dissertation

This dissertation consists of five chapters. Chapters 2, 3, and 4 are three independent manuscripts that have been published or are ready for submission to academic journals. Each of them solves one set of research questions described in section 1.1.

Chapter 2 evaluates nodal connectivity for station criticality of heavy rail systems and assesses network vulnerability of heavy rail systems with the impact of partial node failures. The contributions of this chapter are threefold. Firstly, the notion of three types of transfer stations are addressed in heavy rail networks – mandatory transfer, non-mandatory transfer, and end transfer stations. Secondly, a comprehensive index named Degree of Nodal Connection (DNC) is introduced and applied in calculating nodal connectivity of stations in the eleven U.S. heavy rail systems. Thirdly, the concept of partial node failure is address in network modeling, in contrast of complete node failure where only all-or-nothing scenarios are assumed. DNC loss indexes are adopted to evaluate the loss of nodal connection caused by partial node failure at local level, and

DNC resilience indexes are employed as global measures to assess the impact of partial node failure. Corresponding results of complete node failure are provided as well for comparing purpose.

Chapter 3 assesses network vulnerability using shortest path network problems. Based on the traditional shortest path problems, this chapter takes a further step by incorporating network flows into shortest path model. Two network optimization models are proposed in order to identify and evaluate critical network components – Shortest Path Network model (SPN) and Assessing Nodal Disruption in Shortest Paths Network model (AND-SPN). The SPN is an extended version of a single shortest path problems to apply for all origin–destination pairs in a network while considering passenger flows. The AND-SPN assesses the influence of *r* nodes' disruption of the network flow pattern based on the SPN. Moreover, three criticality indicators are supplemented to reflect average arc use, average arc cost, and average arc flow.

Chapter 4 expands the SPN model to the SPN for Partial node failure model (SPN^{Pr}) so that partial node failure can be simulated as well. The link attribute index and distance update constraints are added for implementing partial node failure scenarios in rail systems with multi-lines. Two criticality indicators – Flow Reroute Cost and ViaFlow – are constructed to assess the influence of partial node failures using the SPN^{Pr}.

Chapter 5 concludes the dissertation with a summary of research and major contributions of this dissertation, and discusses the limitations of this dissertation and possible directions for future research.

CHAPTER 2

ASSESSING NETWORK VULNERABILITY

OF HEAVY RAIL SYSTEMS

WITH THE IMPACT OF

PARTIAL NODE FAILURES

A version of this chapter is published in *Transportation* (2018).
https://doi.org/10.1007/s11116-018-9859-6. The use of "we" in this chapter refers to co-author Dr. Hyun Kim and myself. As the first author, I searched and reviewed literature, performed data analysis and index modeling, and wrote the manuscript.

Abstract

Much of the literature in recent years has examined the vulnerability of transportation networks. To identify appropriate and operational measures of nodal centrality using connectivity in the case of heavy rail systems, this paper presents a set of comprehensive measures in the form of a *Degree of Nodal Connection* (DNC) index. The DNC index facilitates a reevaluation of nodal criticality among distinct types of transfer stations in heavy rail networks that present a number of multiple lines between stations. Specifically, a new classification of transfer stations—*Mandatory Transfer*, *non-Mandatory Transfer*, and *End Transfer*—and a new measure for linkages—*Link Degree (LD)* and *Total Link Degree (TLD)*—introduces the characteristics of heavy rail networks when we expose vulnerability of a node. The concept of *partial node failure* is also introduced and compared with the results of complete node failure scenarios. Four local and global indicators of network vulnerability are derived from the DNC index to assess the vulnerability of major heavy rail networks in the United States. Results indicate that the proposed DNC indexes can inform decision makers or network planners as they explore and compare the resilience of multi-hubs and multi-line networks in a comprehensive but accurate manner regardless of their network sizes.

Key words: Degree of Nodal Connection, Mandatory Transfer, partial node failure, network vulnerability, heavy rail networks

2.1 Introduction

Current literature highlights network vulnerability in transportation systems due to a degradation of service of network infrastructure, which occurs for a number of reasons, including the failure of critical nodes or links resulting from partial or complete malfunction of the system and leading to a significant loss in network service and operation (Murray, 2013; O'Kelly and Kim, 2007). The disruption and failure of network infrastructure can be caused by either unintended incident such as a tornado and earthquake, or human factors, such as a terrorist attack or even the normal operational degradation cycle of the system. Most of the literature aims to understand, assess, and prescribe ways of improving network performance and protecting critical facilities from unfavorable events. Early analyses of network vulnerability focused on identifying critical network components susceptible to possible interruptions to achieve effective management of infrastructural networks in case of a crisis (Demsar et al., 2008; Kim and Ryerson, 2017). However, the definition of network vulnerability includes some ambiguity, since the meaning differs depending on the field of application—for example, highway networks versus mass transit networks—and the methodology of application (Yates and Sanjeevi, 2012; O'Kelly, 2014). Thus, little consensus on the definition of network vulnerability exists in the fields of geography, engineering, and information science because of the distinct perspective of each domain. Perhaps, a common theme throughout the vulnerability literature is the link to coping and capacity to handle stress or perturbation in unfavorable events based on an identified resilience within the systems (D'Este and Taylor, 2003; Schoon, 2005; Holmgren, 2007).

Within the context of rail systems, Angeloudis and Fisk (2006) examined the network vulnerability of the networks by considering the factors of mean number of non-termini stations and lines as well as the total number of non-termini stations. Wang et al. (2013) focused on identifying the locations of vulnerable parts of the urban rail networks using the degree of trip failures. Rodríguez-Núñez and García-Palomares (2014) used the disruptions of riding times and the number of missed trips and then applied those elements to each link in a metro network under interdiction scenarios. As an advanced approach, Li and Kim (2014) investigated the network vulnerability of subway systems in terms of system degradation based on changes of network connectivity as well as passenger flows for the failures of targeted stations. Sun et al. (2015) used the traditional concepts of topological vulnerability and functional vulnerability to investigate urban rail transit networks in terms of the reciprocal of shortest paths and the subsequent reduction in transport ability (i.e., remaining network size). Recently, Sun and Guan (2016) defined network vulnerability of urban metro networks from the line-operation perspective, integrating network theory, such as betweenness centrality and passenger betweenness centrality, to draw a relative disruption probability of transit lines.

In this study, network vulnerability is defined as the ability of a rail network to maintain comprehensive nodal connectivity at local scale as well as network integrality at global scale to *partial* and *complete* node failure. To leverage the impact of different failure scenarios to the resilience of networks, this study employs three analytical elements in the assessment of network vulnerability of heavy rail networks: [1] partial node failure, [2] mandatory transfer stations, and [3] degree of nodal connection index. Specifically, the importance of partial node failure is highlighted, and the characteristics of transfer stations are re-conceptualized as *Mandatory Transfer* (MT), *non-Mandatory Transfer* (non-MT), and *End Transfer* (ET) stations. A comprehensive index of *Degree of Nodal Connection* (DNC) is constructed for multi-lines of heavy rail systems, and a set of vulnerability indexes based on DNC are presented to compare the vulnerability of networks at local and global level to the partial/complete node failure scenarios in a case study of heavy rail systems in the United States.

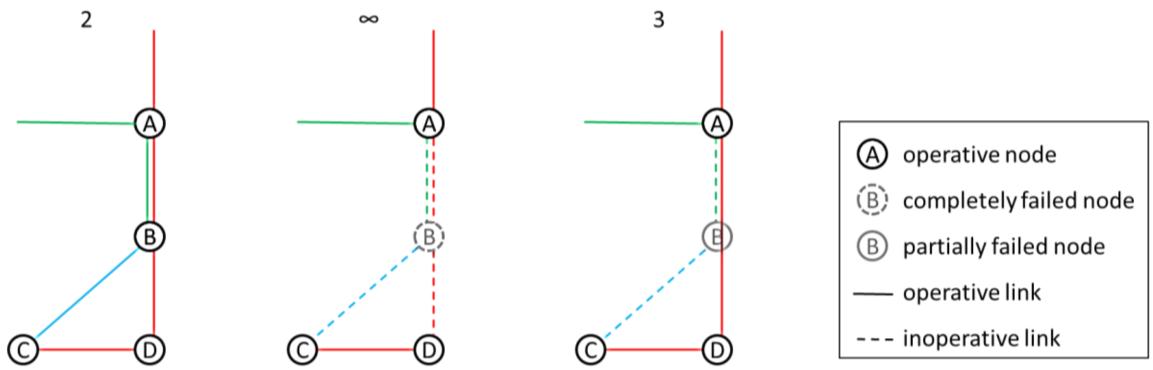
2.2 Background

The general problem of network vulnerability involves finding network components (nodes, links, or both) when they are removed or rendered inoperable, increasing the concerned attribute to the largest extent (Murray et al., 2007; Murray, 2013). The concerned attribute can include network connectivity (Albert et al., 2000), system flow (Myung and Kim, 2004), the maximum flow possible (Wood, 1993), the capacity on arc (Murray and Grubesic, 2007), and reliability of flows (Kim 2009; Kim et al., 2016). Given the literature, a common approach in modeling network vulnerability involves an index-based measure (Derrible and Kennedy, 2010). Subject to some restrictions such as flow capacity and resource availability, the approach adopts indicators to evaluate network vulnerability, which includes an index for network links to embed vulnerability attributes with different weights based on fuzzy logic (El-Rashidy and Grant-Muller, 2014), an indicator using the speed of a system to recover from disruptive events or shocks (D'lima and Medda, 2015), the difference of an objective function between *status quo* and an interrupted case (Kim and Ryerson, 2017; Kim et al., 2016; Zhang et al., 2015), and the operational loss or gain (Chiou, 2015; Rashidi, 2016). Because of the simple nature of index-based measures that allow easier computation, the failure of nodes inherently assumes a complete loss of its operation, and the transportation lines connecting nodes are represented as a single linkage. Even the different role of stations is largely ignored. In the real world, however, a node failure does not always bring about all-or-nothing results, assuming a failure of a railway station does not always mean it is completely inoperable, because of multi-line connections and a complex structure of transfer stations (Derrible and Kennedy, 2010). This is particularly an issue when previous simple connectivity-based measures are applied to heavy rail systems, where a number of transfer stations and multiple lines form the networks.

2.2.1 Partial node failure

Partial node failure is often a more realistic aspect of network operation, as public transport systems do not necessarily experience complete closure, and system degradation is typically progressive (Cats and Jenelius, 2016). This is particularly true when the operation of heavy rail systems ceases due to scheduled maintenance or temporary gate closures in a station rather than because of unexpected failures or extreme events, such as natural disasters. In our study, the notion of *partial node failure* is defined as the failure of a node that still allows traffic to bypass or go through, although the node cannot load, unload, or transfer passengers completely. The counter term, *complete node failure* refers to a situation in which no traffic can pass through the node when disrupted. Figure 2.1 illustrates the concept of partial node failure in a sample network with four stations and three rail lines (Red, Green, and Blue). For the sake of simplicity, the distance between nodes is set at 1. At *status quo* (Figure 2.1-a), the shortest path between nodes A and C is 2 (from A to B to C). In a scenario depicting complete failure of node B, all links incident to node B are rendered inoperative, resulting in no available path from nodes A to C, as represented by the symbol ∞ (Figure 2.1-b). However, if node B partially fails instead (Figure 2.1-c), the path from nodes A to C continues to be available and the shortest path between the two increases to 3 (from A across B to D to C). Traffic proceeds directly from A to D on the Red line without stopping at node B. The Green line between nodes A and B and the Blue line between nodes B and C are rendered inoperative, because the Green line appears in inbound links of node B while the Blue line does not (the order of inbound and outbound links is arbitrary). The Red line between nodes A and B and nodes B and D remains operative, since the Red line appears in both inbound and outbound links of node B so that traffic on the Red line can pass through node B. In these cases, the shutdown of a station does not necessarily mean all passengers are out of access to a station if a path remains available to transport passengers within the station. Some examples of partial node failure include ground construction (e.g., station renovation) and administrative closure (e.g., during a special event or scheduled maintenance). As an example, when a partial node failure occurred in Beijing from February 11 to 17, 2015, the railway temporarily closed Zhichunlu Station, a transfer point between two lines, due to the reconstruction of transfer paths necessitated by the overhaul (heavy maintenance) of old transfer stations. However, trains transited these stations without interruption because of the partial operation. In Paris, France, the partial closure of subway stations occurs frequently because of scheduled renovation or construction projects for certain periods (The Local fr, 2017). However, those partial closures do not cause complete station malfunctions. Other common partial node failures occur during scheduled events. To ensure a network remains operative during scheduled maintenance, most transit agencies strategically close lanes so as not to affect operation, though the process entails a degraded operation. Beijing railway shut down Tiananmen East Station on October 1, 2016, to restrict passenger volume in Tiananmen Square for security reasons during National Holiday. However, the system operation was not seriously affected by the shutdown.

Shortest path distance from A to C:



(a) *status quo* (b) complete node failure (c) partial node failure

Figure 2.1 Impact of node failure for paths from A to C: (a) *status quo*, (b) complete node failure, and (c) partial node failure

Note: Distance between nodes are assigned to 1 for simplicity.

2.2.2 Mandatory Transfer stations

Heavy rail systems are considered the most efficient massive transport mode for densely populated areas. Stations in heavy rail systems are categorized into two types: transfer stations and non-transfer stations (Kim and Song, 2015). Transfer stations usually manage larger passenger flows than non-transfer stations, and they connect non-transfer stations to assist passengers in completing trips at the local level. Transfer stations are viewed as hubs, because they connect several different rail lines, structuring a sub-network among lines within a node (O'Kelly, 2014). Because of higher connectivity among lines, transfer stations are highly vulnerable to potential disruptions and need to be protected with high priority and special attention (Nakanishi, 2009). The malfunction of a transfer station causes not only immediate loss of its own large passenger loadings, but also risks potential termination of interactions among rail lines, leading to overall degradation of network performance. This existing classification of transfer stations can be further divided into three categories, as illustrated in Figure 2.2:

Mandatory Transfer (MT), *non-Mandatory Transfer* (non-MT), and *End Transfer* (ET), which differ from each other in their roles within the heavy rail line. *Mandatory Transfer* (MT) is defined as a transfer station that connects more than two stations, or a station connects only two stations but hosts different inbound and outbound rail lines (i.e., the set of inbound rail lines are not the same as the set of outbound rail lines). MT_B and MT_A in Figure 2.2 are the two types of MTs as defined above, respectively. In contrast, *non-Mandatory Transfer* (non-MT) is defined as a transfer station that only connects one inbound station and one outbound station with the same set of rail lines (e.g., non-MT_C). *End Transfer* (ET) is defined as a transfer station that plays a transferring role for multi-lines at the end of lines (e.g., ET_D). For non-transfer stations, two types of stations exist: *end station* and *intermediate station*. Notice that MT, non-MT, and ET stations engage in multiple linkages if the network operates multiple lines. Also noteworthy is the fact that MT is not necessarily a central station and non-MT is not necessarily a peripheral station, since they are characterized by nodal connectivity, the number of multi-lines, and the degree of links by intermediate stations.

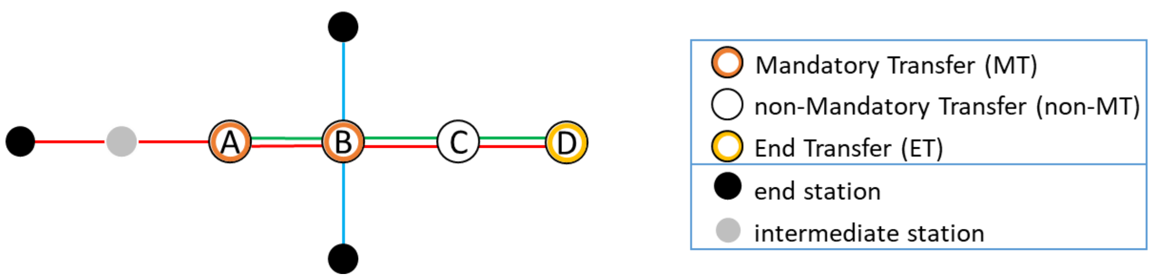


Figure 2.2 Types of stations in a heavy rail system

Figure 2.3 identifies the MT and ET stations of the Bay Area Rapid Transit (BART) in San Francisco, CA. Seven stations are identified as MTs, (stations #1, #10, #15, #17, #24, #26, and #28) and three stations are ETs (stations #14, #23, and #33). From a passenger perspective, MT likely serves as a final opportunity to make a transfer if a traveler wants to use different lines. For example, if passengers travel from station #24 to station #14 starting from the Green or Blue line, they can switch at stations #1, #2, #3, #4, #5, #6, #7, #8, #9, or #10. Given this scenario, station #10 is the last station passengers "have to" (i.e., "mandatory") make a transfer to the Red or Purple line to reach their destination at station #14. Similarly, if passengers travel from node #24 to node #35, MT_{26} is the last and only choice to execute a transfer. The traditional dichotomy between transfer and non-transfer stations does not adequately reflect the complex characteristic of heavy rail systems. Clearly, the importance of MT is more significant than that of non-MT or

ET stations, because the loss of MT functionality would critically affect the resilience of a network in terms of nodal connectivity.

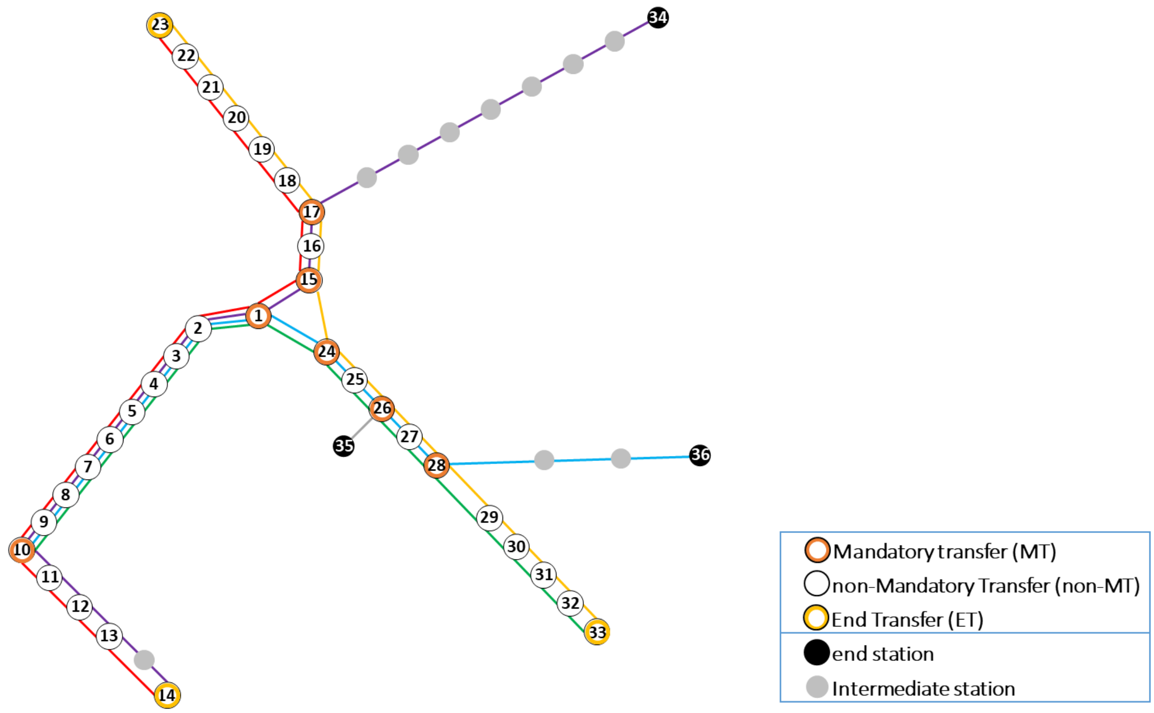


Figure 2.3 Graphic representation of a heavy rail network using the concept of MT, non-MT, and ET (BART. San Francisco, CA)

Note: The network was modified and redrawn by authors based on the BART network (Source: https://www.bart.gov/stations)

2.2.3 Degree of Nodal Connection (DNC) index

Nodal degree is popularized in *graph theory* as a theoretical basic and core concept. In transport systems, nodal degree is a simple but effective measure of nodal importance (Rodrigue et al., 2017) that has been used in modeling network vulnerability. The measure is modified to reflect different aspects of network connectivity, often translated as survivability, to define the capability of communications to withstand the malfunction of selected hubs in a network (Kim, 2012). Grubesic et al. (2003) applied the concept of survivability when calculating the change of nodal degrees by simulating node failures and comparing the resilience of Internet providers in the United States. The survivability of commercial Internet networks was gauged based on two indexes, the rate of loss of nodal degrees of the network and the remaining availability of network paths. Using spatial weight, Wan et al. (2011) proposed a nodal degree model in which both the functional class and length of network links are crucial factors in determining nodal degrees of a network. Other measures of connectivity have been used to assess network vulnerability, including connectivity loss and change of total network connectivity in characterizing habitat connectivity associated with habitat networks (Matisziw and Murray, 2009), measuring the resulting minimum pair-wise connectivity among the remaining nodes after deletion of critical nodes (Arulselvan et al., 2009), and measuring the number of distinct paths with acceptable travel time between each origin-destination pair (Kurauchi et al., 2009).

According to Murray (2013), the most challenging facet of assessing network vulnerability in terms of these connectivity-based measures lies in how to design the operational measure of connectivity given networks' distinctive characteristics. A measurement of connectivity applied to a certain network system is not necessarily appropriate when applied to other systems. For example, traditional simple nodal degree measures using a dichotomy (connected/not connected) do not capture the comprehensive nodal connectivity in which multiple links present among nodes such as current subway systems or heavy rail systems, thus necessitating a different index of nodal connectivity. Given this context, Derrible and Kennedy (2010a, 2010b) proposed an improved measure that considers multiple links between a pair of nodes when calculating nodal degree. In Derrible's measure, a single connection between two nodes is called *single edge*, while *multiple edge* defines two nodes connected by two or more lines. The nodal degree at node i (e_i) is defined as the sum of single edges e_i^s (i.e., k_i) and multiple edges e_i^m. The concept is effective in evaluating a network's form and complexity, because the role of multiple links of network organization (form, states) is reflected in the developed index.

Based on the concepts of multiple edges in connectivity measure, we propose the *Degree of Nodal Connection* (DNC) index to remedy conventional nodal degree methods in evaluating nodal connection for transfer stations with multiple edges, particularly to discern the criticality of different types of transfer stations in heavy rail systems. The DNC consists of two sub-measures: *Link Degree* (*LD*) and *Total Link Degree* (*TLD*). The concept of *Link Degree* (*LD*) reflects the level of nodal connectivity by summing the number of links a node directly connects. For instance, in Figure 2.3, the *LD* values for stations #9, #10, and #11 are 8, 6, and 4, respectively. Note that a traditional connectivity index suggests the nodal degree for each of the three nodes is 2 because it only accounts for the existence of links incident to the node. As an advanced nodal degree measure, Derrible's index only recognizes a "multiple edge" if more than one link exists between two nodes; thus, the connectivity of stations #9, #10, and #11 is 4 in each instance.

Another measure, the *TLD*, is formulated as an additive relationship between the sum of the *LD* and the number of links counted via the number of intermediate stations when each incident link to a node is extended via intermediate stations until transfer or end station appears. The main reason to develop this measure stems from the fact that the traditional degree of node or Derrible's index only counts the directly incident links but does not account for the importance of intermediate stations. For instance, stations #17 and #28 in Figure 2.3 register the same *LD* value of 6. However, if we consider the number of links counted via intermediate stations (grayed circles), the *TLD* for station #17 increases to 13 (= 6+7) while that of station #28 shifts to 8 (= 6+2), indicating that station #17 is more important than station #28, though they present very similar patterns of linkages. From the perspective of nodal criticality, the loss of station #17 stops all passenger flow at intermediate stations, causing a greater impact than the loss of station #28 does to its connected intermediate stations. The *TLD* is unique because its value captures the different levels of criticality of transfer stations to reflect the number of intermediate stations vulnerable to the loss of transfer stations. Based on k, e^m, *MT, LD,* and *TLD,* a comprehensive index, the Degree of Nodal Connection (DNC) index, is drawn using equations (2.1) and (2.2) below.

$$DNC_i = S \cdot [\sqrt{k_i * LD_i} + MT_i * \sqrt{TLD_i * (e_i^m + 1)}] \tag{2.1}$$

where

$$MT_i = \begin{cases} 0, & \text{if } L_i^O * L_i^I * \left[\left|L_i^O - L_i^I\right| + (k_i - 2)\right] = 0 \\ 1, & \text{if } L_i^O * L_i^I * \left[\left|L_i^O - L_i^I\right| + (k_i - 2)\right] \neq 0 \end{cases} \tag{2.2}$$

where

S = scaling factor (constant number)

i = index of a station,

k_i = nodal degree of station i,

LD_i = link degree of station i,

MT_i = 1: if station i is a Mandatory Transfer, 0: otherwise,

TLD_i = total link degree of station i, which is the sum of the *LD* and the number of links extended via intermediate stations,

e_i^m = number of multiple edges of station i,

L_i^O = outbound link(s) of station i,

L_i^I = inbound link(s) of station i.

Equation (2.1) calculates the comprehensive connection level at station i, and equation (2.2) determines whether a transfer station is MT or not by using the binary variable of MT_i. In detail, if both values of inbound (L_i^I) and outbound links (L_i^O) to station i are not zero, and $\left[\left|L_i^O - L_i^I\right| + (k_i - 2)\right] \neq 0$, then it is identified as MT_i (=1). For non-MTs and ETs, $MT_i = 0$, and thus the formulation reduces to $DNC_i = S \cdot \sqrt{k_i * LD_i}$ whereby k_i is the traditional nodal degree, and LD_i is multiplied by k_i to elevate the importance of total links incident to the transfer station. The square root of $k_i * LD_i$ is taken to represent a comparative magnitude with k_i. Specifically, when applied to non-transfer stations (i.e., intermediate stations and end stations), DNC downgrades to the traditional nodal degree k_i leveraged by the scaling factor S, since the value of k_i equals the value of LD_i (i.e., for intermediate stations and end stations, $S* \sqrt{k_i * LD_i} = S \cdot \sqrt{k_i * k_i} = S \cdot k_i$). Note that scaling factor S is determined appropriately considering the range of terms in DNC to create cut-off points of DNC so that the range of DNC values can be easily recalled and used. In our study, S = $\frac{\sqrt{2}}{4}$ is applied, creating the cut-off at 1.67, 1.0, 0.5, and 0.0. However, the range can be adjusted by changing the scaling factor in equation (1). Note that DNC is equal to or greater than 1 for non-MTs; for ETs, DNC is equal to or greater than 0.5 because $k_i = 1$, $LD_i \geq 2$, and $MT_i = 0$. Notice that e_i^m and TLD_i values appear in the DNC index when MTs are involved in the computation. The purpose of $(e_i^m + 1)$ in the second term is to ensure the MT is leveraged by TLD_i, so the importance of the links on the intermediate stations is reflected in calculating the index. In other words, if $e_i^m \neq 0$, MT is leveraged by both TLD_i and e_i^m, because both the importance of intermediate stations and multiple edges should be engaged. Here, the square root of $TLD_i * (e_i^m + 1)$ is taken in order to be an addable magnitude to the first term $\sqrt{k_i * LD_i}$. Note that the units of both terms are associated with link-based metrics, and they are formulated as an index with an additive relationship. If no multiple edges are incident to the transfer station (i.e., e_i^m=0), the second term becomes a simple form, $\sqrt{TLD_i}$, which reduces the weight of links with the intermediate stations. Interpretation is intuitive. The greater the number of intermediate stations and multiple edges incident to the MT, the more important the MT is.

The DNC also provides information about the types of stations. Equation (2.3) summarizes the ranges of the DNC index for the different types of stations with four DNC cut-off points at 1.67, 1.0, 0.5, and 0.0. Within each range, we can compare their differences in nodal criticality precisely with respect to comprehensive nodal connection level. For ETs, the upper bound can be 1, as long as an ET maintains 7 or fewer links (i.e., $LD_i \leq 7$). Likewise, for non-MTs, the DNC index is less than 1.67 (the minimum DNC value for MTs), as long as $LD_i \leq 11$ of a non-MT. The upper bound of DNC_i for MT is theoretically infinite, but the results for the eleven U.S. heavy rail networks, the largest DNC_i value is found to be 7.50.

$$DNC_i \in \begin{cases} [1.67, \infty) & \text{for MT} \\ [1.0, 1.67) & \text{for non-MT} \\ [0.5, 1.0) & \text{either ET or intermediate station} \\ [0.0, 0.5) & \text{for end station} \end{cases} \tag{2.3}$$

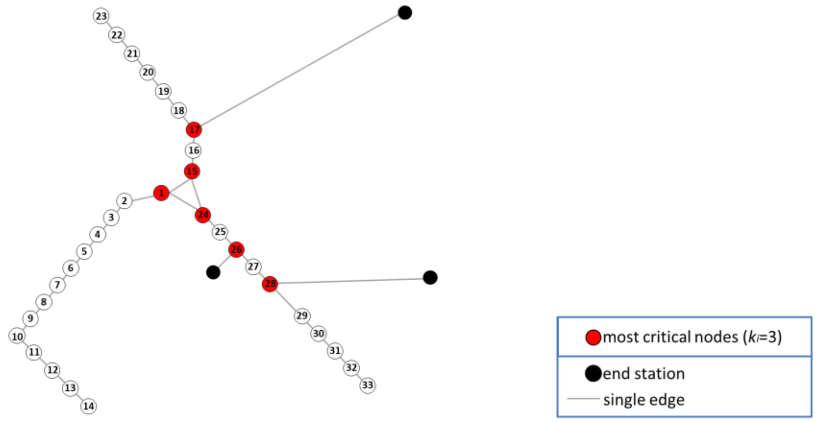
Although the DNC index is designed to address the discrepancy of nodal connection among transfer stations, it is applied to intermediate stations and end stations, too. For intermediate and end stations, the DNC index reduces to the value of traditional nodal degree by scaling factor and thus is less than 1.0, since $LD_i = k_i$ and $MT_i = 0$.

Table 2.1 and Figure 2.4 together indicate how well the DNC indexes capture the differences in nodal criticality compared to the traditional nodal degree (k_i) and Derrible's measure (e_i) for the thirty-three transfer nodes of the BART. Note that both traditional and Derrible's measures are perhaps less effective in describing degrees of importance among MTs in cases when a network has a number of MTs. Given the rankings in Table 2.1, traditional nodal degree categorizes the degrees using three values (1, 2, 3), Derrible's measure uses four groups, but DNC classifies the degrees with ten groups. As shown in Figure 2.4(a), six MTs (i = 1, 15, 17, 24, 26, 28) have identical values in traditional nodal degree ($k_i = 3$). In Figure 2.4(b), Derrible's measure reveals that MT_1 is more critical than the other five MTs. However, the other five MTs are still depicted with an identical value. In Figure 2.4(c), DNC index discerns the criticality of the six MTs into five groups (stations #1 / #17 / #26 / #28 / #15 and #24). In addition, the DNC index captures the MT_{10} stands out with greater DNC index than non-MTs such as stations #9 and #11. Neither traditional nodal degree nor Derrible's measure recognizes MT_{10} in this manner. Results indicate that the DNC index offers an enhanced method of measurement compared to existing measures in the sense that it fully captures the characteristics of transfer stations, including those presented by multiple links between stations, across heavy rail networks of the United States. Thus, the DNC index is a way of criticality assessment in terms of nodal centrality at *status quo* rather than nodal vulnerability to failure condition.

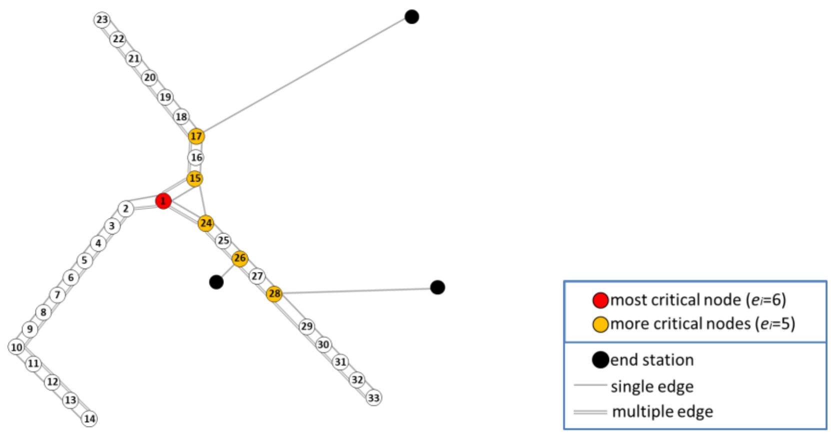
Table 2.1 Comparison of stations' criticality: traditional nodal degree, Derrible's measure, and DNC with a case of the BART (San Francisco, CA)

Station ID	***Traditional nodal degree*** (k_i)	***Derrible's measure*** (e_i)	***DNC index***	***LD***	***TLD***	***MT***
1	3	6	3.73	8	8	1
17	3	5	3.71	6	13	1
26	3	5	3.24	7	7	1
28	3	5	3.23	6	8	1
15	3	5	3.00	6	6	1
24	3	5	3.00	6	6	1
10	2	4	2.72	6	6	1
2	2	4	1.41	8	8	0
3	2	4	1.41	8	8	0
4	2	4	1.41	8	8	0
5	2	4	1.41	8	8	0
6	2	4	1.41	8	8	0
7	2	4	1.41	8	8	0
8	2	4	1.41	8	8	0
9	2	4	1.41	8	8	0
16	2	4	1.22	6	6	0
25	2	4	1.22	6	6	0
27	2	4	1.22	6	6	0
11	2	4	1.00	4	4	0
12	2	4	1.00	4	4	0
13	2	4	1.00	4	5	0
18	2	4	1.00	4	4	0
19	2	4	1.00	4	4	0
20	2	4	1.00	4	4	0
21	2	4	1.00	4	4	0
22	2	4	1.00	4	4	0
29	2	4	1.00	4	4	0
30	2	4	1.00	4	4	0
31	2	4	1.00	4	4	0
32	2	4	1.00	4	4	0
14	1	2	0.50	2	3	0
23	1	2	0.50	2	2	0
33	1	2	0.50	2	2	0

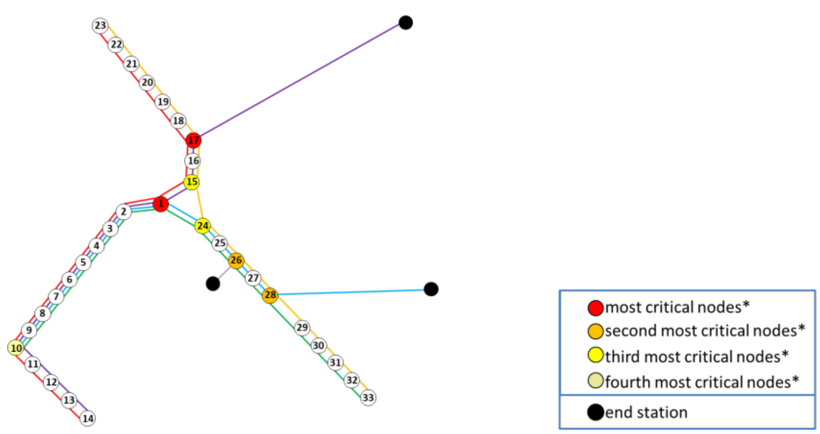
Note: The order of station ID follows the value of DNC. Different number of border lines are inserted based on the values to demonstrate how each measure captures the criticality of nodal degree.



(a) Traditional Nodal Degree



(b) Derrible's measure



(c) DNC

Figure 2.4 Comparison of traditional nodal degree, Derrible's measure, and DNC in identifying critical stations with an example of the BART (San Francisco, CA)

*: Refer to Table 2.1 for values of DNC index.

2.3 Analytical Framework

2.3.1 Network vulnerability with complete and partial node failure

Previous network vulnerability research often implemented disruption scenarios assuming c*omplete node failure,* which assumes all incident links and traffic close when a node is damaged. In this study, we compared the results of the complete node failure scenarios to those of partial node failures. Our numerical experiment of disruption scenarios only focused on MTs, since they represent more important nodes than others within heavy rail networks. The malfunction of MTs yields more considerable network loss than other non-MTs; and if impact occurs, the vulnerability of a network can be explored at local and global levels by [1] the decrement of DNC indexes at individual station levels and [2] the size of the separated sub-networks at a global level. Given this frame, four extended DNC indexes are formulated as two local DNC loss indexes (equations 2.4 and 2.5) and two global DNC resilience indexes (equations 2.6 and 2.7).

$$DNC_{Loss}^{Local-W} = max_r \left[\sum_j DNC_j^{MT} - \left(\sum_{r \in N^{MT}} DNC_{j'}^{MT} \right) \right] \qquad \forall\, j, j', r \in N^{MT} \tag{2.4}$$

$$DNC_{Loss}^{Local-M} = \frac{1}{n^{MT}} \sum_{r \in N^{MT}} \left[\sum_j DNC_j^{MT} - \left(\sum_{r \in N^{MT}} DNC_{j'}^{MT} \right) \right] \qquad \forall j, j', r \in N^{MT} \tag{2.5}$$

$$DNC_{Resil}^{Global-W} = max_r \left[\frac{n_r^{left}}{n^{net}} \right] \qquad \forall\, r \in N^{MT} \tag{2.6}$$

$$DNC_{Resil}^{Global-M} = \frac{1}{n^{MT}} * \sum_{r \in N^{MT}} \left(\frac{n_r^{left}}{n^{net}} \right) \qquad \forall\, r \in N^{MT} \tag{2.7}$$

where

$DNC_{Loss}^{Local-W}$ = Local DNC loss index for the worst scenario to disrupted station *r* at local scale,

$DNC_{Loss}^{Local-M}$ = Local DNC loss index for the average of disrupted station *r* at local scale,

$DNC_{Resil}^{Global-W}$ = Global DNC resilience index for the greatest value among the worst scenario of station failures *r* calculated by [left network size / total network size],

$DNC_{Resil}^{Global-M}$ = Global DNC resilience index for the averaged value of station failures *r* calculated by [left network size / total network size],

where

r = index of a disrupted MT,

N^{MT} = the set of all MTs in the network,

j = index of a MT at *status quo*,

j' = index of a MT after station failure,

DNC_j^{MT} = DNC index for MTj,

n^{MT} = total number of MTs at *status quo*,

n^{net} = whole network size at *status quo*,

n^{left} = size of the *left network* after station failure.

The local DNC loss measure, $DNC_{Loss}^{Local-W}$, identifies the greatest value of the DNC index, reflecting the worst-case scenario of the degree of nodal connection of the network when MT station r ($r \in N^{MT}$) is disrupted. $DNC_{loss}^{local-M}$ is the averaged value of DNC index considering all scenarios of MT station failures. As global measures, $DNC_{Resil}^{global-W}$ and $DNC_{Resil}^{global-M}$ examine the size of left-network when a whole network is broken into sub-networks following the failure of MT. The measure identifies the ratio of the largest sub-network at the failure of MT to the whole network at *status quo*. Theoretically, if the global measure is close to '1.0', it indicates the network resilience is robust in the worst scenario of MT failure; if the value is near '0', then it indicates the network's resilience is poor.

Table 2.2 DNC loss and resilience indexes for a sample network (Figure 2.2).

Measure scale	Type of failure	Index	Index value
Local	Partial	$DNC_{Loss}^{Local-W}*$	3.23
		$DNC_{Loss}^{Local-M}$	2.82
	Complete	$DNC_{Loss}^{Local-W}*$	5.24
		$DNC_{Loss}^{Local-M}$	4.06
Global	Partial	$DNC_{Resil}^{Global-W}*$	0.63
		$DNC_{Resil}^{Global-M}$	0.75
	Complete	$DNC_{Resil}^{Global-W}*$	0.38
		$DNC_{Resil}^{Global-M}$	0.50

*Note: The worst cases for $DNC_{Loss}^{Local-W}$ and $DNC_{Resil}^{Global-W}$ were identified when r=MT$_B$.

Table 2.2 presents how these measures can identify distinctive network resilience to the different node failure scenarios in the case of a small network depicted in Figure 2.2. On a local scale, the impact of the worst scenario of complete failure, $DNC_{Loss}^{Local-W}$ (=5.24), is greater than that of a partial failure, $DNC_{Loss}^{Local-W}$(=3.23), which occurs at the failure of MT$_{B.}$ The averaged impact ($DNC_{Loss}^{Local-M}$) for all MT failure scenarios also reveals that the impact of complete failure is greater than the partial failures of MTs. Both global measures, $DNC_{Resil}^{Global-W}$ and $DNC_{Resil}^{Global-M}$, present the different levels of network resilience in the worst case and the averaged value of all failure scenarios, respectively. Notice that four measures can be used to make a comparative analysis across different sizes of networks as they are standardized or scaled.

The procedure to compute DNC indexes for all possible MT failures with complete and partial node failure scenarios for eleven heavy networks, an algorithm is designed as follows:

Initialization:

(a) Calculate DNC indexes of each transfer station at *status quo*;
(b) Sum up DNC indexes of all MTs at *status quo*;

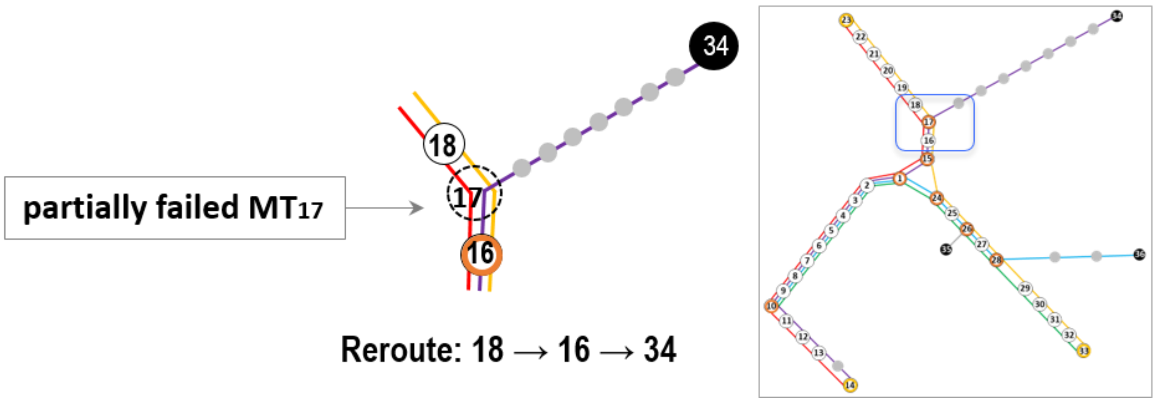
Simulating disruption:

(c) Select a Mandatory Transfer (MT) from the set of MTs ($\in N^{MT}$) for disruption;
(d) Remove all links incident to the selected MT for complete node failure; for partial node failure, keep the rail lines incident to the selected MT, which appear in both inbound and outbound lines of the MT, and remove other rail lines incident to the MT;
(e) Calculate DNC indexes to the disruption of the selected MT;
(f) Calculate the difference between DNC indexes using the values in steps (a) and (e);
  - Update local scale measures: $DNC_{Loss}^{Local-W}$ and calculate $DNC_{Loss}^{Local-M}$,
(g) Examine whether the network is divided into pieces of sub-networks;
  - If yes: record the number of resulting sub-networks and the size of each sub-network,
  - If no: record the number of resulting sub-networks ← Ø, and the size of each sub-network ← N/A,
  - Identify the worst-case scenario and update global scale measure: $DNC_{Resil}^{Global-W}$ and calculate $DNC_{Resil}^{Global-M}$,

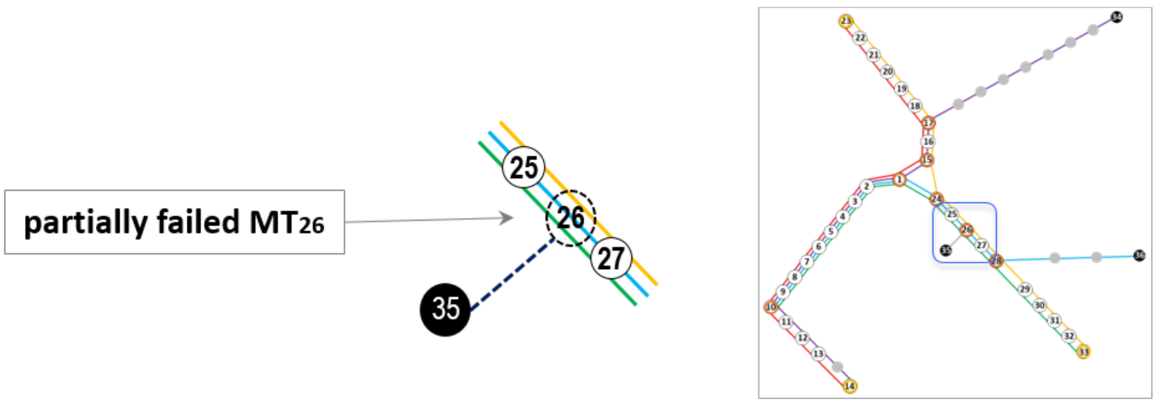
Termination:

(h) Terminate the procedure once all disruption scenarios of MTs have been evaluated (N^{MT}= Ø)

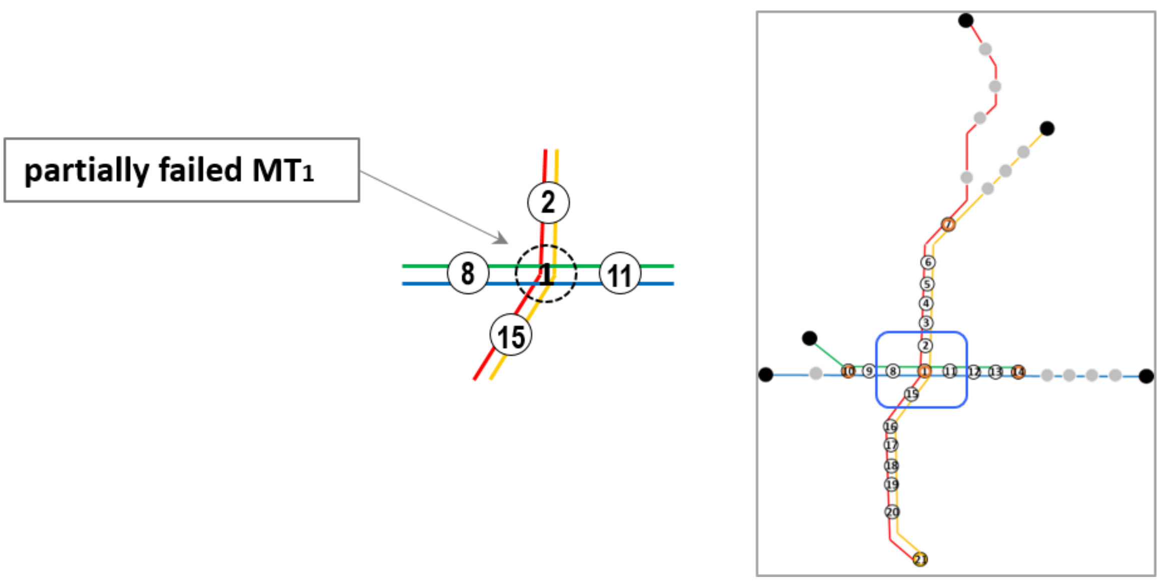
The key to differentiating network resilience to a partially failed MT is whether a neighboring transfer can substitute for the failed MT and continue the transferring function. For example, as illustrated in Figure 2.5(a), when MT_{17} partially fails in the BART, station #16 becomes an alternative MT, and no link is rendered inoperative. The Red and Yellow line passengers from



(a) Reroute to the destination via other line



(b) Separation from the network



(c) Split of the network

Figure 2.5 The different impact of partial node failure of MT with examples of the BART (San Francisco, CA) (a and b) and MARTA (Atlanta, GA) (c)

Note: The network in (c) is modified and redrawn by authors based on the MARTA network (Source: http://www.itsmarta.com/train-stations-and-schedules.aspx) station #18, who are traveling northeast to their end station #34 via the Purple line, must reroute using an additional two links to reach their final destination. However, in Figure 2.5(b), when MT_{26} partially fails, the Grey line is correspondingly triggered inoperative. None of its neighboring transfers (stations #25 and #27) emerge to serve as an alternative MT, causing the end node #35 to be separated from the network. In contrast, Figure 2.5(c) illustrates the Metropolitan Atlanta Rapid Transit Authority (MARTA) network (Atlanta, GA) during a partial failure of MT_1, which causes the entire network to split into two sub-networks (i.e., one sub-network serves the Red plus Yellow lines, and the other sub-network serves the Green plus Blue lines). In this worst-case circumstance, no passengers on the Red and Yellow lines can reach their final destinations via the Blue and Green lines, and vice versa. It is worth noting that the network will be separated into four sub-networks if a complete failure of MT_1 is assumed instead of a partial failure.

2.3.2 Data

According to the *Public Transportation FACT BOOK 2015*, fifteen transit systems in the United States are defined as heavy rail networks. Of these, the study focuses on eleven U.S. major heavy rail networks characterized as distinct systems with multi-transfer stations and multiple lines. Four networks are not included because they are simply single-line systems without any transfer station. Table 2.3 lists eleven heavy rail systems and their service coverage in the United States.

Table 2.3 Heavy rail agencies ranked by passenger miles, 2013*

Rank	Transit Agency	Area covered
1	MTA New York City Transit (NYCT)	New York, NY
2	San Francisco Bay Area Rapid Transit District (BART)	San Francisco, CA
3	Washington Metropolitan Area Transit Authority (WMATA)	Washington, DC
4	Chicago Transit Authority (CTA)	Chicago, IL
5	Massachusetts Bay Transportation Authority (MBTA)	Boston, MA
6	Southeastern Pennsylvania Transp. Authority (SEPTA)	Philadelphia, PA
7	Metropolitan Atlanta Rapid Transit Authority (MARTA)	Atlanta, GA
8	Port Authority Trans-Hudson Corporation (PATH)	New York, NY
9	Los Angeles County Metropolitan Transp. Auth. (LACMTA)	Los Angeles, CA
10	Miami-Dade Transit (MDT)	Miami, FL
11	The Greater Cleveland Regional Transit Authority (GCRTA)	Cleveland, OH

* Source: *2015 Public Transportation FACT BOOK.*

Using the rail route maps from the systems' official web pages, we created connectivity matrixes for analysis. Table 2.4 summarizes basic information of network properties (i.e., network size, the number of MTs, non-MTs, ETs, end stations, and intermediate stations) for each heavy rail system.

Table 2.4 Summary of the heavy rail networks for analysis

Heavy rail network	**Number of**					
	Network size	**MT**	**non-MT**	**ET**	**end station**	**intermediate station**
NYCT (New York, NY)	419	62	61	3	18	275
CTA (Chicago, IL)	140	12	12	0	11	105
MBTA (Boston, MA)	121	9	0	0	12	100
LACMTA (Los Angeles, CA)	93	5	3	0	8	77
WMATA (Washington, DC)	91	8	32	3	7	41
SEPTA (Philadelphia, PA)	73	2	1	0	5	65
GCRTA (Cleveland, OH)	50	3	9	1	4	33
BART (San Francisco, CA)	46	7	23	3	3	10
MDT (Miami, FL)	42	7	14	1	4	16
MARTA (Atlanta, GA)	38	4	16	1	5	12
PATH (New York, NY)	13	6	3	2	1	1

2.4 Numerical Results

2.4.1 DNC index for station criticality

Using the DNC index, we classify distinct categories of transfer stations—MTs, non-MTs, and ETs— with detailed numerical values, as well as the different levels of criticality in each group. Table 2.5 summarizes the results of the DNC index with an example of the BART network. The table is sorted in descending order according to DNC index. The type of station is noted in the last column to demonstrate how DNC index can determine the criticality of stations. The seven MTs with greater DNC indexes are clearly distinguishable from others, followed by non-MTs with smaller DNC indexes and, finally, ETs. The non-MTs are additionally classified into three tiers. Stations #2, #3, #4, #5, #6, #7, #8, and #9—the transfer stations for four rail lines— are grouped as non-MT_1 because of their larger DNC index to other non-MTs; the non-MT_2 are characterized as the second-tier because they host three lines; and twelve stations in the third-tier non-MT_3 connect only two rail lines.

Table 2.5 Summary of DNC index for the BART (San Francisco, CA)

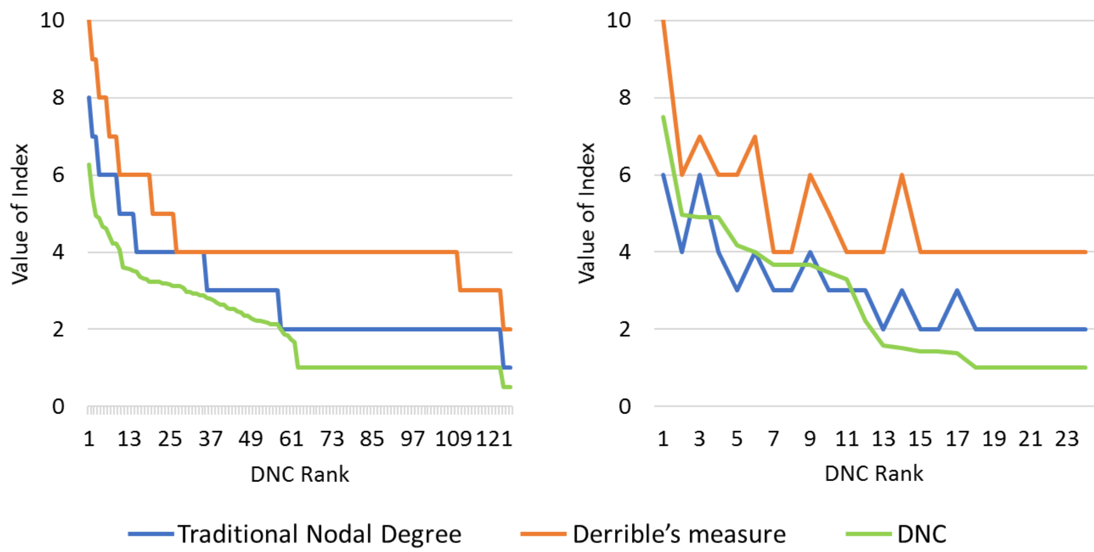
Station ID*	DNC index	Tier
1	3.73	MT – Tier 1
17	3.71	MT – Tier 2
26	3.24	MT – Tier 3
28	3.23	MT – Tier 4
15, 24	3.00	MT – Tier 5
10	2.72	MT – Tier 6
2, 3, 4, 5, 6, 7, 8, 9	1.41	non-MT – Tier 1
16, 25, 27	1.22	non-MT – Tier 2
11, 12, 13, 18, 19, 20, 21, 22, 29, 30, 31, 32	1.00	non-MT – Tier 3
14, 23, 33	0.50	ET

*: refer to Figure 2.4 for node ID.

Table 2.6 highlights the top three critical transfer stations for the eleven U.S. heavy rail networks using the DNC index. Noticeably, the top three transfer stations in each network are all MTs that host multiple rail lines and connect many non-MTs or intermediate stations. Given the DNC index, the different levels of criticality of transfer stations across networks are identified. For example, the Chicago Transit Authority (CTA) and the New York City Transit (NYCT) have higher levels of critical transfer stations, while the Port Authority Trans-Hudson (PATH) and the Southeastern Pennsylvania Transportation Authority (SEPTA) are relatively lower in their criticality of stations. The DNC index can provide more precise information regarding the criticality of stations than other existing connectivity measures through DNC index-based ranking distribution. Figure 2.6 compares the results from the three measures of connectivity along with the DNC index rankings with two examples of networks, the NYCT and the CTA. Note that the DNC results include decimal numbers in depicting the criticality of stations to allow more detailed comparisons of nodal connections. In the first case (Figure 2.6-a), the three methods of measure provide consistent ranking results for the transfer stations; however, the DNC index clearly discerns the rankings of stations for the rankings from 13 to 61, because the DNC expresses values of connectivity at the decimal level rather than integer. In the second case (Figure 2.6-b), DNC indexes are more appropriate than traditional indices in ranking transfer stations by presenting the change of ranking with nodal connectivity of heavy rail networks in a smoothed but detailed manner, particularly when a number of nodes are evaluated and compared.

Table 2.6 Top 3 transfer stations with respect to the DNC index of each heavy rail network

Network	Rank	Station Name	DNC
BART (San Francisco, CA)	1	West Oakland	3.73
	2	MacArthur	3.71
	3	Coliseum	3.24
CTA (Chicago, IL)	1	Clark / Lake	7.50
	2	Belmont	4.97
	3	Roosevelt	4.91
GCRTA (Cleveland, OH)	1	Tower City - Public Square	3.79
	2	Shaker Square	3.67
	3	E. 55th	3.53
LACMTA	1	Union Station	3.87
(Los Angeles, CA)	2	Pico	3.87
	3	Willowbrook/Rosa Parks	3.45
MARTA (Atlanta, GA)	1	Five Points	4.24
	2	Lindbergh Center	2.88
	3	Ashby	2.34
MBTA (Boston, MA)	1	Kenmore	3.81
	2	Downtown Crossing	2.83
	3	North Station	2.81
MDT (Miami, FL)	1	Government Center	5.12
	2	Knight Center	3.76
	3	Brickell	3.67
NYCT (New York, NY)	1	Jackson Hts Roosevelt Av - 74 St Broadway	6.27
	2	Atlantic Av-Barclays Ctr	5.46
	3	Times Sq-42 St	4.95
PATH (New York, NY)	1	Grove St	2.22
	2	Exchange PI	2.22
	3	Christopher St	2.22
SEPTA (Philadelphia, PA)	1	15th	3.59
	2	69 Street Terminal	2.68
	3	Girard	1.06
WMATA (Washington, DC)	1	L'Enfant Plaza	5.42
	2	Metro Center	4.81
	3	Fort Totten	4.10

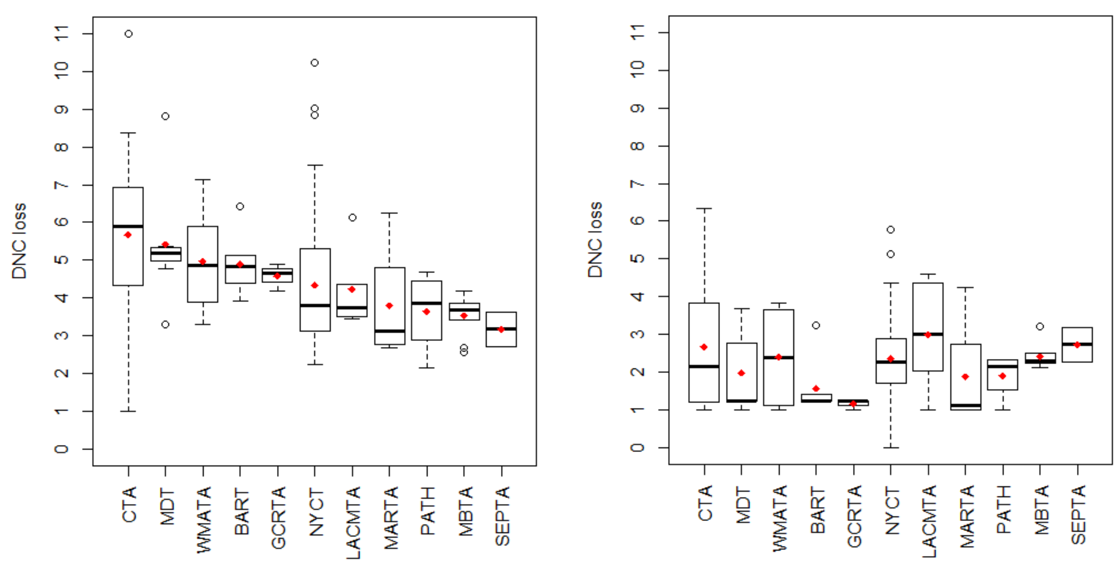


(a) the NYCT (b) the CTA

Figure 2.6 Comparison of connectivity indexes: Traditional nodal degree, Derrible's measure and DNC index with the cases of the NYCT (New York, NY) and CTA (Chicago, IL)

2.4.2 Local DNC loss index vs. Global DNC resilience index

In nodal disruption scenarios, the depression of a node initially causes a local effect but eventually triggers a loss of network operation resulting in the entire network splitting into sub-networks at a global level. Thus, the impact on a global network scale can be measured by examining the size of sub-networks (Grubesic et al., 2003). Figure 2.7 presents the results of local DNC loss indexes for complete (Figure 2.7-a) and partial node failure scenarios (Figure 2.7-b). The X-axis displays the names of heavy rail networks and the Y-axis represents the values of DNC loss. For both failure scenarios, heavy rail networks along the X-axis are ordered based on the $DNC_{Loss}^{Local-M}$ (red-dot in the box-plot) from the most vulnerable to the least vulnerable network in the complete node failure result. For comparison purposes, the order is reflected in the partial node failure result in Figure 2.7(b). Each box-plot captures the worst-case scenario of node failure (i.e., $DNC_{Loss}^{Local-W}$), range, mean (i.e., $DNC_{Loss}^{Local-M}$), and upper and lower quantiles among the DNC loss indexes of a network. The results imply that the impact on the network by the partial node failure is not necessarily in concordance with that of complete node failure. In both scenarios, the eleven networks generally experience decreased DNC loss indexes. However, within the GCRTA, the $DNC_{Loss}^{Local-M}$ drops significantly from 4.57 at complete node failure to 1.15 at partial node failure, while the SEPTA experiences a very small decrease in DNC loss from 3.16 to 2.72. It is worth noting that changes in rankings of heavy rail networks during complete and partial node failures offer valuable information on network resilience. In the cases of considerable change in DNC loss, the BART and the GCRTA, for examples, the resilience of the networks would not be as low as expected. In contrast, the SEPTA, which shows a minor change between two scenarios, would need a protection scheme to improve its resilience to the stations' partial failure. Interestingly, the LACMTA increases the range of quantiles at partial node failure, though the worst-case impact is reduced in a partial node failure scenario. In complete node failure, due to the complete shutdown of all links and traffic around the MT, thus the MT's incident non-MTs revert to ETs when no intermediate stations are between them. Similarly, the MT's incident ETs are eliminated as single nodes without any connection. In contrast, during partial node failure, the through-traffic is reserved, thus providing possibility and opportunity for the incident non-MTs to emerge as alternative MTs. When none of the nearby non-MTs can substitute for the disrupted MT, a portion of nodal connection remains to potentially allow the nodal connection of incident transfers to continue to function, even if weakened.



(a) Complete node failure (b) Partial node failure

Figure 2.7 Comparing DNC loss in complete (a) and partial node failure scenarios (b)

At global scale, the network integrality is also affected differently depending on the assumed failure scenario. Table 2.7 summarizes the most critical MT, which results in the worst-case scenario in terms of complete and partial node failure for each heavy rail network using global DNC resilience index ($DNC_{Resil}^{Global-W}$). Here, the "critical MT" refers to the MT that breaks the network into separate segments and most affects the network's resilience during a failure. Notice that the critical MT identified during complete node failure is not necessarily identified at partial node failure. For example, In the case of the MBTA, the most critical MT station is Park St with $DNC_{Resil}^{Global-W}$=0.50. However, in the case of partial node failure, the most critical one is Kenmore with $DNC_{Resil}^{Global-W}$= 0.78. Likewise, the network resilience indicator on a global scale can encompass all the other networks, allowing the comparison of resilience among them. The result indicates that the NYCT network is the most resilient in the cases of complete node failure; but the GCRTA, MDT, PATH, and WMATA networks are the most resilient in partial node failure because they are not completely divided into sub-networks. In contrast, networks such as the MARTA, SEPTA, and LACMTA are weakly resilient to both complete and partial disruption compared to other networks.

Table 2.7 The most critical MTs in complete and partial node failure scenarios for each heavy rail network

	Most critical MT station			
	Complete node failure		**Partial node failure**	
Network	**Station Name**	$DNC_{Resil}^{Global-W}$	**Station Name**	$DNC_{Resil}^{Global-W}$
MARTA	Five Points	0.39	Five Points	0.58
SEPTA	15th	0.42	15th	0.66
LACMTA	Pico	0.45	7th St/Metro Ctr	0.55
PATH	Newport	0.46	-*	1.00
MBTA	Park St	0.50	Kenmore	0.78
GCRTA	E. 55th	0.52	-	1.00
BART	12th St/Oak. City Center	0.63	Coliseum	0.98
MDT	Government Center	0.64	-	1.00
WMATA	L'Enfant Plaza	0.74	-	1.00
CTA	Fullerton	0.75	Howard	0.99
NYCT	161 St Yankee Stadium	0.95	Franklin Av.-Botanic Garden	0.99

*: No station causes network separation.

Figure 2.8 illustrates the positionality of heavy rail systems from the perspective of network resilience when considering complete and partial node failure. The positionality map is useful in characterizing the group of networks that share similar properties, such as resilience and accessibility (Kim and Song, 2015). Given four quadrants, each network is positioned within four classes: HR-P & HR-C (1st quadrant), HR-P & LR-C (2nd), LR-P & LR-C (3rd), and LR-P & HR-C (4th), where HR and LR represent High Resilience and Low Resilience, and -P and -C stand for partial node failure and complete node failure, respectively. Given this categorization, the network resilience of the CTA, NYCT, and WMATA networks are in good standing to both partial and complete node failures, as these heavy rail networks show high resilience in both HR-C and HR-P. Of the networks showing low resilience to complete node failure (LR-C), the MARTA, LACMTA, and SEPTA networks also appear to be weakly resilient to partial node failure (LR-P). Regarding network protection policies, the networks in this group should take steps to enhance their resilience. Despite their weak network resilience to complete node failure, five heavy rail networks perform much better in partial node failure: the MBTA, GCRTA, MDT, PATH, and BART networks. A few mid-size networks—the GCRTA, MDT, and PATH networks—maintain 100% network integrality (i.e., no sub-networks separation) in partial node failure despite their weak resilience in complete node failure. This indicates no partial failure of any MT will trigger a loss in network size. In this study, none of the eleven networks falls into the LR-P & HR-C category.

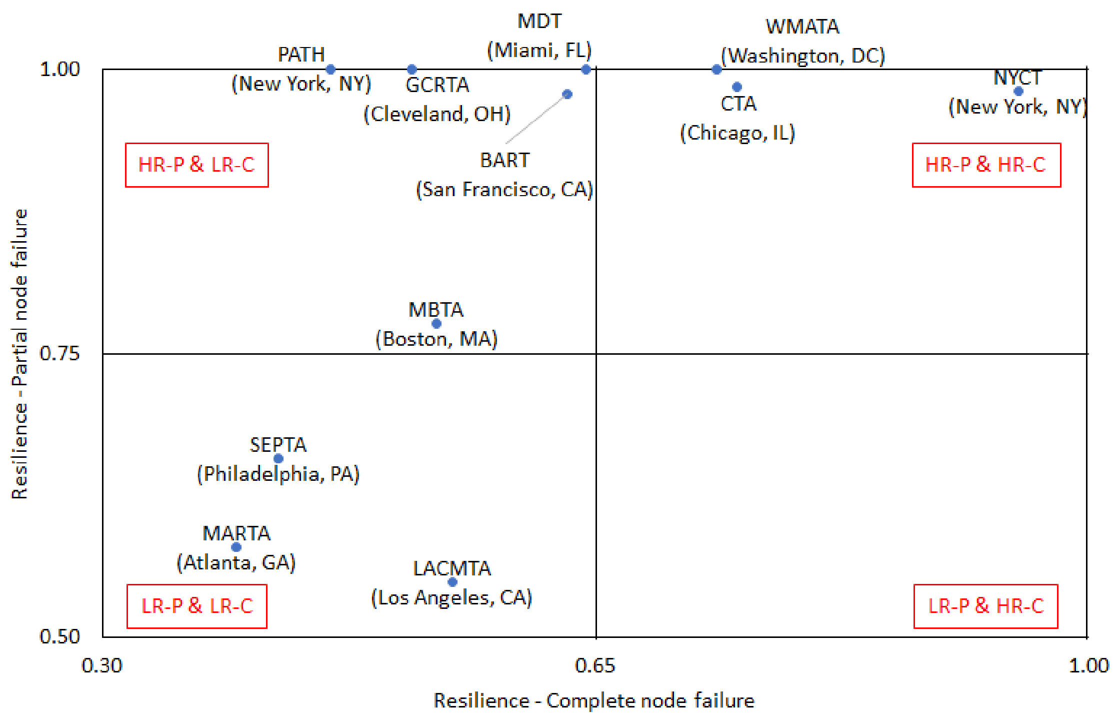
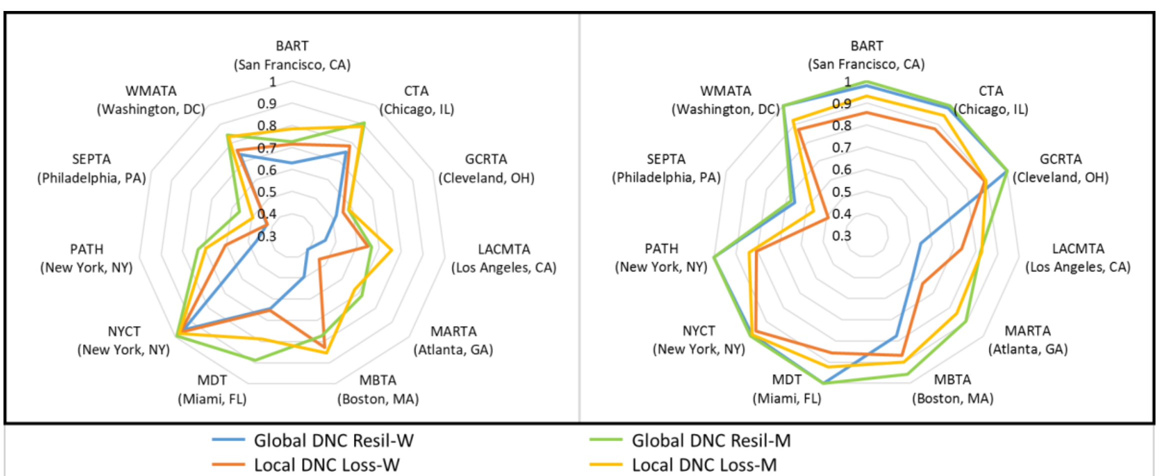


Figure 2.8 Network resilience in targeted interruption at global scale in complete and partial node failure scenarios

Finally, the network resilience of the eleven networks can be evaluated using four local and global DNC indexes, as Figure 2.9 indicates in performance radar plots. The radar plots clearly compare the networks' resilience in a multi-faceted dimension—complete or partial node failure, loss or resilience, the worst or averaged at local and global scales. The closer the point of each measure is to 1.0 (the outer rim), the more resilient the network is. Given Figure 2.9(a) and (b), each network behaves differently in the eight combinations. Figure 2.9(a) shows that in complete node failure, the SEPTA and MARTA are most vulnerable, as their four measures are closer to the center relatively. In particular, the worst local DNC index indicates that both the MARTA and SEPTA are least resilient because both networks follow a star network topology consisting of one central major hub, and the lines are connected to a few centralized hub stations. As a counter example, the NYCT network is the most resilient network among the eleven networks because of its complex network structure with many hubs and inter-hub connections. In Figure 2.9(b), the resilience of networks is enhanced as indicated by the expanded shape of the radar form. However, the SEPTA's resilience to partial node failure does not appear to have improved, and the LACMTA is relatively weak according to the global DNC resilience index. The LACMTA's weakness is largely due to the network topology that has only a few hubs connecting two different rail lines, making it difficult to reroute traffic when transferring traffic is not permitted at those hubs. In addition, the network integrality of four networks— the MDT, WMATA, GCRTA, and PATH—remains nearly unchanged according to the two global measures, which implies they have enough hubs to allow traffic to reroute through other transfers even when a MT is completely or partially failed. In terms of network resilience policy, the identified network with low resilience should increase the number of transfer stations to strengthen its resilience to complete and partial node failure.



(a) complete node failure (b) partial node failure

Figure 2.9 Network resilience in different situations for heavy rail networks

2.5 Concluding Remarks

The existing research on network vulnerability is based solely on complete node failure and is less appropriate or realistic in real world applications. Given that passengers choose other modes of transportation when a subway station fails completely or reroute to nearby transfer stations when a station partially fails, the network vulnerability considers both scenarios. The lessons from this research are summarized as follows. In complete node failure of a subway station, which is anticipated during a natural disaster (e.g., flood or tornado), it is very likely that even on-the-ground transportation cannot accommodate passengers as an alternative transport mode. In such a scenario, the area of destruction is usually large, typically larger than the subway station itself, and thus the entire surrounding area will experience such mass chaos so that no passenger can move within this area. This research, however, addresses the concept of partial node failure, which suggests that the impact on network resilience is limited to a section of the station lines or to the station itself, allowing passengers to potentially complete their trips by rerouting or changing lines with alternative stations near the impacted station. Specifically, if the disrupted MT is on the shortest path to a given destination, passengers may choose to reroute within the heavy rail system via the next shortest path when available. As a final note, when partial node failure happens, reroute within the network is more likely to occur in short-haul transit systems (such as metro systems) rather than in long-haul transport networks (such as the Amtrak network), since between-station distance is usually short and service frequency is usually high in short-haul transit networks and thus reroute within the network would not be too lengthy or time consuming.

CHAPTER 3

ASSESSING NETWORK VULNERABILITY

USING

SHORTEST PATH NETWORK PROBLEMS

A version of this chapter is revised and will be submitted to an academic journal. The use of "we" in this chapter refers to co-author Dr. Hyun Kim and myself. As the first author, I processed the data, performed the analysis and wrote the manuscript.

Abstract

This paper identifies and evaluates critical components of a flow-based network from the perspective of the shortest path problems. Two network optimization models are proposed: the first model, called the Shortest Path Network model (SPN), is an extended version of a single shortest path problems to apply for all origin-destination pairs in a network while considering passenger flows; the Assessing Nodal Disruption in Shortest Paths Network model (AND-SPN), which is based on the SPN, assesses the influence of *r* nodes' disruption of the network flow pattern. In this paper, network performance is assessed not only by the transport cost as an objective function but also through supplementary criticality indicators reflecting average arc use, average *s-t* pair cost, and average arc flow. In the case study of the Amtrak rail network, the criticality of stations is evaluated in terms of objective function and of criticality indicators—highlighting the effect of the disruption and allowing comprehensive exploration of it. Aided by the provided list of criticality rankings, decision-makers can decide which stations are the most important to protect during special events or disasters or when seeking to enhance transportation network security.

Key words: shortest path problems, network vulnerability, criticality, rail network, spatial optimization

3.1 Introduction

Although defining the term *critical infrastructure* is a challenge, a spatially networked transport system, which facilitates interactions between demand and supply, is recognized as being a critical part of infrastructure because it may face a risk for significant degradation of system functionality as well as the operation of the spatial economy. With a rail system, for example, disruptive effects on the system can be measured explicitly by flow-based metrics such as change in amounts of passengers, increased length of travel routes, availability of linkages, or partial node failure, thereby judging the degree of disruption suffered by stations and linkages (Matisziw et al., 2010; Kim et al., 2016). Hence, of particular interest in a flow-based system is the assessment of the system's response to various system-inoperable scenarios, exploring potential changes in the resilience of the system and deriving a protection strategy for mitigating the effects of such a scenario (Caschili et al., 2015). In the class of routing problems, great efforts have been made using a mathematical approach, but most of the literature focuses on a single aspect of the modeled response—for example, identifying locations for prepositioning resources available to given networks (Ukkusuri and Yushimito, 2009; Galindo and Batta, 2010), measuring the effects based on disruption scenarios (Murray et al., 2007; Rawls and Turnquist, 2010), or providing a strategy for protecting the system (Matisziw et al., 2010; Kim, 2011).

In this paper, we propose a set of mathematical models to identify and evaluate critical components of a flow-based network from the perspective of the shortest path problems. The assessment of a network's performance is made not only of the single objective function but also of supplementary criticality indicators: (1) average arc use, (2) average *s*-*t* pair cost, and (3) average arc flow. As a first step, we formulate the Shortest Path Network model (SPN), which is an extended version of the single shortest path problems. The SPN considers the shortest path of all origin–destination pairs in order to distribute flow demands optimally in the network. Based on the SPN, we further expand the model to the Assessing Nodal Disruption in Shortest Paths Network model (AND-SPN) to assess the influence of *r* nodes' disruption of the network flow pattern. This model aims to help network planners evaluate network vulnerability in the case of a targeted attack on a node or a set of nodes. To allow effective assessment, we prescribe three criticality indicators as well.

In the following sections, we provide a brief literature review of shortest path models in the context of network vulnerability. Moreover, to address how the SPN and AND-SPN models function in the case of a directional network, we define a general problem of nodal disruption that can be used to identify critical node(s) and determine protection priority, using a small sample network. To demonstrate the models' capability, we present numerical experiments using the SPN and AND-SPN models in the context of a case study of the Amtrak rail system in the United States, followed by concluding remarks and recommendations for further research.

3.2 Background

Identifying the shortest path from origin to destination is one of the recognized network performance functions. The "shortest path" could be variously defined as that covering the least distance, taking the least travel time, or incurring the least travel cost, depending on the types of network cost structure. In general, there are two approaches to construct the shortest path problems – algorithmic approach and mathematical programming approach. First, algorithmic approach is widely used in practice because of its easy implementation, and computationally reasonable to have optimal solutions. Given this approach, three classical types of shortest path problems are mentioned: single-source shortest path problems, all-pairs shortest path problems, and *k*-shortest path problems.

3.2.1 Single-source, all-pairs, and k shortest path problems

Single-source shortest path problem refers to the problem that finding shortest paths from one node to all other nodes in a network (Ahuja et al., 1993). The Dijkstra's algorithm (Dijkstra, 1959; Whiting and Hillier, 1960; Dantzig, 1960) identifies the shortest path from a single source node to all other nodes in directional networks. With slight modification, the reverse Dijkstra's algorithm searches for the shortest path from all other nodes to a single sink node; the bidirectional Dijkstra's algorithm recognizes the shortest path from a single source node to a specific sink node. When arcs with negative length present, the Bellman–Ford algorithm (Ford, 1956; Bellman, 1958) is adopted to handle the single-source shortest path problems.

All-pairs shortest path problem refers to the problem that finding the shortest paths from every node to every other node in a network. Floyd (1962) presented an algorithm to find all the shortest paths in a network, where the arc lengths can be negative but no circuits could have negative length. The Floyd-Warshall algorithm returns a matrix of the shortest paths among all node pairs through comparing all possible paths from origin to destination especially in dense networks. Johnson's algorithm (1977) is faster and more efficient in solving all-pairs shortest path problems on sparse graphs.

K shortest path problem refers to the problem that finding the shortest path as well as the other *k-1* paths, such as the second shortest path and the third shortest path and etc. *K* specifies the total number of shortest paths to be identified. Bock et al. (1957) developed a method to find all possible paths from source to sink and then rank them systematically. Bellman and Kalaba (1960) described a method in searching *k* shortest paths for larger networks that finding *k* shortest paths from all origins to a given set of destinations. Clarke et al. (1963) found the *k* shortest paths using a branch-and-bound method.

3.2.2 Network optimization approaches in shortest path problems

Network optimization approaches to assessing vulnerabilities in system performance have received considerable attention during the search for extremely important assets whose loss will decisively affect network operation. For transportation networks, particular interest has been focused on understanding how the disabled crucial components—whether through disruption of links (e.g., roads or rails) or removal of nodes (e.g., stations)—affect network performance and, furthermore, on identifying which link or node is the most critical or vulnerable in the entire infrastructure system.

In pursuing their long-standing research interest in shortest path problems, Fulkerson and Harding (1977) applied shortest path problems in the context of network vulnerability while focusing on network disruption. Their work demonstrated how lengthening the arcs between two specific nodes could influence the shortest path length in a graph. Corley and Sha (1982), addressed a problem of the vital links in a network because the removal of the identified vital links from the network could cause the largest increase in the shortest distance between a node pair. Nardelli et al. (2003) studied the removal of which node in a network would produce the maximum shortest distance between nodes. Snyder et al. (2006) presented a set of disruption models in terms of supply chain networks where the goal is to prevent disruptions at some facilities and enhance network resilient to disruptions. Lim and Smith (2007) translated a network disruption problem into a multi-commodity flow network problem instead of following the effects of arc removal in a network—a traditional viewpoint taken into disruption approach. Scaparra and Church (2008) allocated protection resources in a network of facilities from an optimization modeling approach based upon the *p*-median service protocol, in order to minimize the disruptive effects of possible attacks to the network. Laporte et al. (2011) introduced a model with three flow conditions to minimize the effect of disruption on total trip coverage for designing robust rapid transit networks through providing alternative routes for specific node pairs.

The general issue underlying the shortest path problems associated with network vulnerability is that of identifying a set of network components, such as nodes or arcs, in the network whose removal increases the total cost of the shortest path between an origin and destination (Murray, 2013). Considering the traditional approach, three issues are worthy of note when describing the motivation for this research.

First, earlier research paid more attention to the impact of a disruption based upon simple shortest path models. In detail, those models focused on identifying a "single" network component that would feature in a worst-case failure of the network rather than on using multifaceted assessments to examine a collective set of network component failures and their effect on the network (Matisziw and Murray, 2009). However, as pointed by Kim et al. (2016), an identified network component that critically affects network performance in a particular scenario need not be the crucial element in other scenarios in which multiple malfunctions are assumed.

Second, in terms of modeling complexity, if a shortest path problem is formulated as a class of disruption problem, this entails increased model complexity, for the model should formulate the role of the "interdictor" and "defender" to fit the original structure of the shortest path problem, hence, the model's objective is necessarily constructed as a bi-level optimization problem such as a max-min problem where the maximization of effects at a higher level and minimization of shortest path costs at a lower level are required. This prescription results in strong NP-completeness (Wood, 1993). When heuristic approaches are considered, the effectiveness is only valid when finding critical network component is for a single origin–destination pair or in the case of a small network. The solution process for shortest path disruption problems involves the complicated steps in order to verify that the solution is optimal (Cormican et al., 1998; Israeli and Wood, 2002). In summary, identifying network components that most affect *all* pairs of nodes in the network is challenging (Zenklusen, 2010).

Finally, considering that planning decisions are made to suit specific disruption scenarios in a real network, formulating models to embrace diverse vulnerability metrics in a simple but expandable structure of mathematical programming, which is a substantial concern. The effects produced by the malfunction of a network component can be investigated in various ways based not only on transport cost but also on other metrics, such as changes in network link use, the attendant flow pattern, and the anticipated concentration of flow across particular links throughout the network. From a network efficiency perspective, changes in travel time or on-time performance for passengers between the status quo and affected scenarios can be considered (Bruun, 2014; Kim and Ryerson, 2017).

3.3 Methodology and Formulations

Solving the shortest path problems using an algorithm-based approach is known to be more efficient in terms of solution time than are models based on a mixed-integer programming (MIP) approach to solving. However, for special cases such as in all-pairs shortest path problems or with multi-objective functions, an MIP-based solving approach offers a more flexible and effective model structure because the specific conditions are easily translated as constraints or incorporated into part of an objective function (Kuby et al., 1997). In addition, the complexity of solving the problem based on MIP-based approach is only monotonically increased if the fundamental structure of the formulations is well preserved (Bellman, 1958).

As a first step, to develop the Shortest Paths Network model (SPN), the traditional single source shortest path (SP) model is addressed as baseline model. Let $G = (N, A)$ be a directed network of a transportation system with a set of nodes N and a set of arcs $A=\{(i, j)\ldots (m, n)\}$. Each arc (i, j) has an attribute C_{ij}, the cost of the arc—such as the physical length of the arc or the travel time between two nodes i and j. The shortest path from a particular pair of origin s to destination t is determined by the set of arcs, which generates the smallest travel cost based on the C_{ij} of arcs. The model is formulated as a form of mixed-integer programming.

Minimize $\sum C_{ij} X_{ij}$ (3.1)

Subject to

$$\sum_j X_{ij} - \sum_j X_{ji} = \begin{cases} 1, \forall i = s \\ 0, \forall i \neq s, t \\ -1, \forall i = t \end{cases} \qquad \forall i \neq j \tag{3.2}$$

$$X_{ij} \in \{0,1\} \qquad \forall i, j; i \neq j \tag{3.3}$$

where

s = index of origin in network G

t = index of destination in network G ($s \neq t$)

C_{ij} = travel cost of arc (i, j)

X_{ij} = 1: if arc (i, j) is the member of the shortest path from s to t, 0: otherwise.

The objective function 3.1 minimizes the total travel cost from origin to destination. Constraints 3.2 are flow-balancing constraints for origin s, destination t, and intermediate arc (i, j) that are the members producing the least travel cost for s–t based on the cost attributes of the arcs. Integer restrictions are set by constraints 3.3.

Notice that this original single shortest path problem does not take into account the volume of passengers to be delivered from s to t, although a certain amount of flow should be an important factor when the best route is arranged among nodes in real transportation networks. To reflect this factor, the single pair and non-flow-based model should be expanded to an all-pairs and flow-weighted shortest problem in which all sets of best paths are identified while considering a volume of flows to be delivered from s to t through a given network, which is named the Shortest Paths Network (SPN) model in this research.

3.3.1 The Shortest Paths Network (SPN) model

The SPN is suitable for application to transportation networks such as rail or highway systems along which passengers between s–t pairs are delivered over the best routes. Because the SPN attempts to identify the shortest paths for all s–t pairs for all flows, each arc (i, j) will have a different level of traffic load based on how frequently it is used for all s–t pairs when they complete the journey. The SPN is formulated as follows.

$$\text{Minimize} \quad \Omega^{sq} = \sum_{s}\sum_{t}\sum_{i}\sum_{j\in N_i} W_{st} C_{ij} X_{stij} \tag{3.4}$$

subject to

$$\sum_{j\in N_i} X_{stij} - \sum_{j\in N_i} X_{stji} = \begin{cases} 1, \forall i = s \\ 0, \forall i \neq s,t \\ -1, \forall i = t \end{cases} \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.5}$$

$$X_{stij} \in \{0,1\} \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.6}$$

where

Ω^{sq} = objective function at *status quo*

N_i = set of nodes connected directly to node i by an arc (i, j)

W_{st} = amount of flows (passengers or freight) to be delivered from s to t

X_{stij} = 1: if arc (i, j) is the member of the shortest path from s to t, 0: otherwise

C_{ij} is as defined in equation 3.1

The objective function 3.4 minimizes total network cost (e.g., total passenger (W)×kilometers(C)) for a set of origins (s) and destinations (t) on network G. Constraints 3.5 are extended from the flow balance constraints 3.2 to ensure that the shortest path is structured for each s–t so that the set of shortest paths for all s–t pairs is identified. Constraints 3.6 set integer restrictions. From a network design standpoint, it is worth examining how arc capacity affects the behavior of the model, including the level of use of the network, compared to the standard non-capacitated model, because appropriate levels of arc capacity could produce a better-balanced passenger flow distribution over the networks. Thus, constraints 3.7 can be added as an extension to impose a maximum amount of flows on arc (i, j), where the sum of flows on the arc cannot exceed a given amount of CAP_{ij}^{route}.

$$\sum_s \sum_t W_{st} X_{stij} \le CAP_{ij}^{route} \qquad \forall s, t; s \neq t \tag{3.7}$$

Notice that the result produced by the SPN model with constraints 3.7 portrays the ideal flow distribution among nodes in the network, assuming all flows take the shortest paths to reach their destinations. The SPN model is used for network performance at status quo without any malfunction of network components assumed. Hence, if of concern is to assess the impact by the disruption of any nodal disruption in unfavorable events of nodal disruption, the SPN model is extended to reflect the disruption scenarios.

3.3.2 The Assessing Nodal Disruption in Shortest Paths Network (AND-SPN) model

Certainly, network performance is degraded if disruption halts any node operation. However, the level of effect is manifested differently over network space to reflect its various aspects (flow distribution, arc use, and increased distance) depending on which node is more critical—and, furthermore, on what combination of nodes is disrupted (Kim et al., 2016). In this context, the AND-SPN model is useful to investigate the outcomes of disruption scenarios with reference to network performance per the status quo.

The AND-SPN is characterized as a special case of network disruption problem, distinct from the SPN, because it includes disruption constraints and decision variables within the general structure of the SPN. Notice that the AND-SPN can be distinguished from traditional disruption models associated with shortest path problems. For example, the AND-SPN is more focused on assessing network vulnerability when specific node(s) *r* are targeted for malfunction in the network. This model is particularly useful if the critical concern is that of exploring which set of nodes could result in the least critical or most critical degradation of network performance when the nodes are simultaneously disabled. Given a specific scenario for disruption of *r*-nodes, the AND-SPN identifies the best alternative paths for all *s*–*t* pairs so as to minimize total network disruption cost. The model formulation is as follows.

$$\text{Minimize} \quad \Omega^r = \sum_s \sum_t \sum_i \sum_{j \in N_i} W_{st} C_{ij} X_{stij} + (M \cdot \sum D_{st}) - \Omega^{sq}_{\text{-}D_{st}} \tag{3.8}$$

subject to

$$\sum_{j \in N_i} X_{stij} - \sum_{j \in N_i} X_{stji} = \begin{cases} 1 - D_{st}, \forall i = s \\ 0, \qquad \forall i \neq s,t \\ -1 + D_{st}, \forall i = t \end{cases} \quad \forall s,t,i,j; s \neq t, i \neq j \tag{3.9}$$

$$\sum_r \sum_{j \in N_r^R} (X_{strj} + X_{stjr}) \leq 2 \,|\, N_r^R \,|\, (1 - Z_r) \qquad \forall s,t; r \in R \tag{3.10}$$

$$X_{stij} \in \{0,1\} \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.11}$$

where

r = index of disabled node(s)

M = a large number representing the disruption cost of a disconnected *s*-*t* pair

D_{st} = 1: if node pair *s*-*t* is disconnected due to the disabled node(s) *r*, 0: otherwise

$\Omega^{sq}_{\text{-}D_{st}}$ = network cost at status quo of operable s-t pairs when node(s) *r* disabled, which is an offset constant derived from the results of the SPN at status quo, calculated by equation

$$\Omega^{sq}_{\text{-}D_{st}} = \Omega^{sq} - \sum_s \sum_t \sum_i \sum_{j \in N_i} W_{st} C_{ij} X_{stij} \qquad \forall st \in P_{st}; i \neq j$$

P_{st} = a set of disconnected *s*-*t* pairs due to the disabled node(s) *r*

R = a set of disabled nodes *r*

N_r^R = a set of nodes directly connected to disabled node r, $r \in R$

$|N_r^R|$ = the number of nodes directly connected to disabled node r, $r \in R$

Z_r = 1: if node r is disabled, 0: otherwise

N_i, W_{st}, X_{stij}, Ω^{sq} are as defined in equation 3.4

C_{ij} is as defined in equation 3.1

The objective function 3.8 minimizes the total network disruption cost when specific node(s) r is disabled in the network. The objective function emphasizes the completeness of network structure (depicting by the number of disconnected *s-t* pairs) and the increase of transport cost due to flow reroutes caused by disabled node(s) r. The offset constant $\Omega^{sq}_{-D_{st}}$ reflects network cost at status quo of operable s-t pairs when node(s) r is disabled, which is obtained from the results of status quo. $\Omega^{sq}_{-D_{st}}$ is subtracted from the objective function, in order to purely reflect the total network disruption cost. M is a large number representing the disruption cost of a single disconnected *s-t* pair; the number of M is decided by the total number of disconnected *s-t* pairs. Constraints 3.9 are the modified flow balance constraints when node(s) r is disabled. Constraints 3.10 ensure that no flow can pass through the disabled node r and its incident arcs. Integer restrictions of X_{stij} are set by constraints 3.11.

The AND-SPN chiefly focuses on measuring a network's vulnerability to a targeted nodal disruption scenario by node(s) r, and evaluating the effect of each nodal disruption on the network. From a view of transportation network protection, a node that incurs a greater network cost is considered more critical to the network, and thus should be protected with higher priority than other nodes.

3.3.3 Criticality indicators

As mentioned, however, the criticality of a node can be measured using different network metrics. Supplementary metrics often help network planners assess a scenario and make a decision best suited to specific network protection circumstances. In the AND-SPN or SPN, three metrics —average arc use (AAU), average *s-t* pair cost (APC), and average arc flow (AAF) — are prescribed as supplementary criticality indicators in equations 3.12, 3.13, and 3.14, respectively.

$$AAU = \sum_{s}\sum_{t}\sum_{i}\sum_{j \in N_i} X_{stij} / |A^{av}| \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.12}$$

$$APC = \sum_{s}\sum_{t}\sum_{i}\sum_{j \in N_i} C_{ij} X_{stij} / |P^{av}| \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.13}$$

$$AAF = \sum_{s}\sum_{t}\sum_{i}\sum_{j \in N_i} W_{st} X_{stij} / |A^{av}| \qquad \forall s,t,i,j; s \neq t, i \neq j \tag{3.14}$$

where

$|A^{av}|$ = total number of available arcs

$|P^{av}|$ = total number of available *s*-*t* pairs

X_{stij}, C_{ij}, W_{st} are as defined in equation 3.4

The first criticality indicator, AAU, is the average times an arc is used as members of shortest paths in the network. AAU depicts the average frequency of utilization of operable links. In the context of the AND-SPN, the difference of AAU between disruption scenario *k* and the status quo indicates on average how much additional use of links is necessary to help flows complete their journey. The larger the AAU, the heavier the workload and the greater the burden borne by the network. For example, if all arcs (*i, j*) are used 12 times for the shortest paths of all origin–destination pairs at scenario *k*, and the total number of arcs is 8, the AAU at scenario *k* is 12/8=1.5, meaning that on average an arc is used 1.5 times as members of shortest paths. The second indicator, APC, is the average of travel cost for all origin–destination pairs. Because APC eliminates the effect of the volume of flows, it is associated with network operation cost. In results achieved using the AND-SPN, APC reflects the average travel cost needed for flows to complete their trips, with larger values associated with greater operation cost in the network. For example, if the total arc cost is 114 and 12 node pairs remain operable at scenario *k*, the APC at scenario *k* is 114/12=9.5, meaning that flows need to spend 9.5 on average to complete their trips. The third indicator, AAF, embodies the average flow load on each arc used. Larger values of AAF represent greater congestion, indicating that the network is more likely to suffer from intense congestion.

3.4 Numerical Experiment

Figure 3.1 presents an example directional network with six nodes, with the numbers on arc represent travel distance of each arc. If one passenger is to be delivered between all *s*-*t* pairs, total network cost is 502 calculated by the SPN, which happens to equal to the results obtained from the all-pairs shortest paths problem by summing up the length of all shortest paths (sum of distance is 502). However, when travel demand changes, the SPN distinguishes itself from the all-pairs shortest paths problem. For example, if nine passengers are assigned to be delivered from node #1 to #5 as well as node #5 to #1 and travel demand among all other node pairs remains as one passenger, the total network cost increased to 854 calculated by the SPN. In this case, the all-pairs shortest paths problem is no longer applicable. The SPN is capable of distributing various flow demands via shortest paths across the network.

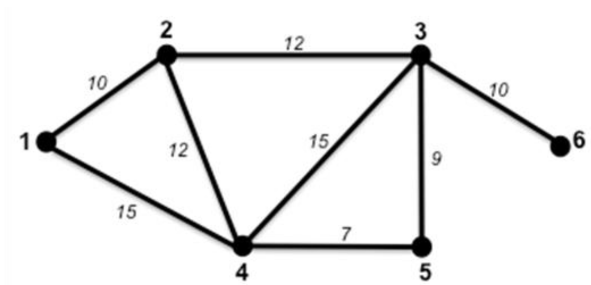


Figure 3.1 Example directional network
Note: number on each arc is travel distance.

Applying the AND-SPN to the sample network in Figure 3.1, the impact of node failures is assessed. Table 3.1(a) summarizes the numerical results of single-node disruption scenarios when one passengers is to be delivered between each *s-t* pair. For example, when node #4 is disrupted, the total disruption cost (i.e., the obj. value of the AND-SPN) is 1022. The nodes are ranked by objective values from the AND-SPN, where the larger the total disruption cost due to node removal, the more critical the node is. The most critical node is obviously node #3, whose malfunction incurs the largest network disruption cost, followed by node #2 and node #4. It is not surprising that nodes #1, #5, and #6 are ranked as the forth place, since they do not serve as intermediate station on shortest path of any node pair.

However, when flow demand changes between *s-t* pairs, the node criticality rankings assessed by the AND-SPN changes as well. Table 3.1(b) summarizes the numerical results of single-node disruption scenarios when nine passengers are assigned to be delivered from node #1 to #5 as well as node #5 to #1 and travel demand among all other node pairs remains as one passenger. The most critical node is found to be node #3 according to the AND-SPN results in Table 3.1(b), followed by node #4 and node #2. Compared with the results in Table 3.1(a), the criticality ranking of node #4 is elevated to the second place, since travel demand between node pairs

Table 3.1 Assessing single-node disruptions for sample network presented in Figure 3.1

(a) Same travel demand among all node pairs

Ranking	Node ID Disrupted	Total Disruption Cost (Obj. values)[a]	Criticality Indicators		
			Average Arc Use (AAU)	Average Pair Cost (APC)	Average Arc Flow (AAF)
1	#3	1800	2.00	14.17	2.00
2	#2	1032	3.20	19.20	3.20
3	#4	1022	4.50	18.80	4.50
4	#5	1000	2.50	17.50	2.50
4	#1	1000	2.33	15.00	2.33
4	#6	1000	1.86	14.30	1.86
-	*status quo*	-	2.88	16.73	2.88

Note: [a] M=100.

(b) Different travel demand

Ranking	Node ID Disrupted	Total Disruption Cost (Obj. values)[a]	Criticality Indicators		
			Average Arc Use (AAU)	Average Pair Cost (APC)	Average Arc Flow (AAF)
1	#3	18000	2.00	14.17	6.00
2	#4	10166	4.50	18.80	10.50
3	#2	10032	3.20	19.20	6.40
4	#5	10000	2.50	17.50	2.50
4	#1	10000	2.33	15.00	2.33
4	#6	10000	1.86	14.30	4.14
-	*status quo*	-	2.88	16.73	4.88

Note: [a] M=1000.

#1&#5 and #5&#1 increases greatly, and node #4 is on the shortest path between node pairs #1&#5 and #5&#1 at status quo.

In the test of the AND-SPN for Table 3.1(a), a value of 100 is assigned to M, which is sufficient to represent the disruption cost, since multiple trials of M at 100,000, 10,000, 1,000, and 100 return the same solution. In the test of the AND-SPN for Table 3.1(b), a value of 1000 is assigned to M, which is sufficient to represent the disruption cost, since multiple trials of M at 100,000, 10,000, and 1,000 return the same solution. It is noteworthy that, the absolute value of M and thus the total disruption cost (i.e., the objective value) is not physically meaningful, only the ranking ordered by the total disruption cost is meaningful.

The three criticality indicators – AAU, APC, and AAF – gauge the criticality of nodes in a more comprehensive manner. For example, in Table 3.1(a), from an AAU standpoint, node #4 is the most critical node: its malfunction increases arc usage most. In contrast, when APC is of concern, reflecting a desire to maintain network performance by minimizing operation cost, node #2 is the most critical node. Similarly, when seeking to minimize network congestion—the AAF approach—node #4 is of highest priority.

3.4.1 Data for experiments

The Amtrak rail system is one of the largest rail systems providing long-haul passenger transport across the United States. We selected this network as a case study because its topology is relatively simple, particularly when compared to other transportation network systems such as subway or highway systems, which feature multiple lanes and nonplanar structure. Because the Amtrak system uses a streamlined network where many stations are connected by a single path (Peterson and Church, 2008), the impact by any disruptions, such as failure of a trestle, outage of stations and links, or presentation of natural hazards, are readily apparent. In short, stations' high level of interdependency presents the potential for a series of consecutive failures in other interdependent stations, manifesting in various outcomes (Murray and Grubesic, 2007; Carreras et al., 2002; Little, 2002). For example, on March 15, 2007, a trestle fire in Sacramento, California, added a 125-mile detour to passenger and freight routes and incurred considerable operation costs. On May 12, 2015, an Amtrak train derailed in the Port Richmond neighborhood of Philadelphia, PA, when bounding from Washington, D.C. to New York City, NY. The mass transit adjustments lasted for a few days aftermath, suspending Amtrak service, New Jersey

Transit service, and SEPTA Regional Rail in the surrounding areas. In a most recent incident, on December 18, 2017, an Amtrak train from Seattle, WA to Portland, OR derailed onto a major highway, leaving Amtrak service south of Seattle stopped after derailment, and road traffic had to detour at least 1.5 hours since the train cars remained dangling from the overpass. Sometimes, even minor incidents (e.g., rockslides, and snow storms) interdict the Amtrak network substantially, such as on October 5, 2015, an Amtrak Train collided with debris fallen on the track after a rock slide near Northfield, VT, causing the locomotive and four coach cars to derail.

In the same way as presented in the small sample network, we can explore which stations are critical from the perspective of a shortest paths network, allowing the preparation of a secondary plan designed to reduce network costs incurred by failures cascading from critical stations. In our test, 62 nodes from the entire Amtrak rail system are investigated based on their behaviors in the proposed models. We selected the 62 nodes using two criteria. One type of station exhibits a high degree of node (≥ 3) and a high volume of total ridership (i.e., the top 25 stations by ridership data in Amtrak Fact Sheets, 2013). The other type includes ending stations in the network to keep the network in a completed form. Figure 3.2 shows the ridership pattern of the Amtrak rail system for the selected 62 nodes. The passenger flow is weighted heavily toward the U.S. West Coast and the U.S. Northeast. Some stations transport significant passenger flow and come with large node degree (e.g., Chicago, IL). Others exhibit a significant nodal degree but carry little passenger flow (e.g., Hammond-Whiting, IN). The ending nodes chiefly claim small to decent passenger flow.

3.4.2 Assessing nodal disruption

Table 3.2 provides the computational results for single-node disruption scenarios of Amtrak stations using the AND-SPN. All problems were solved to optimality within 10 seconds using CPLEX 12 on an Intel i-5 processor paired with 8 GB of RAM. Based on the objective values, a ranking of criticality is obtained. The higher-ranked station is the more critical one to the system's operation, for objective value corresponds to network disruption cost. In reality, when a station is rendered inoperable, the situation implies a direct cost for reconstruction expenses necessary for resumption of operations as well as an increase in total transport cost incurred by the disabled stations. These stations influence the routes of other *s–t* pairs, whose shortest paths are directly associated with a disabled station, forcing them to reroute passengers via alternative shortest paths. In extreme cases, moreover, the disruption of a particular node, if it is a gateway station or hub, disconnects all paths to other nodes completely, causing significant loss of paths from other

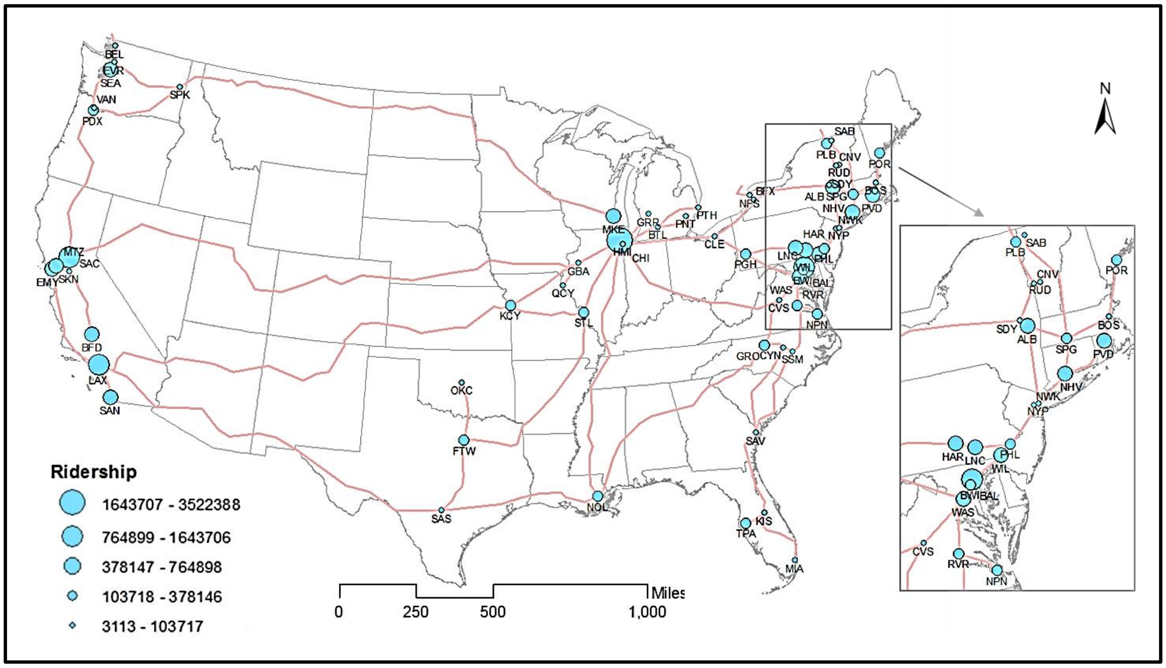


Figure 3.2 Ridership pattern of Amtrak rail system for the 62-node dataset

nodes. For the nodes with the same ranking by total disruption cost, the node ridership (i.e., total flows departure from and arrive at this node) at status quo is used to assist ranking (highlighted with blue color in Table 3.2). With similar influence on the network when disrupted, the node with larger ridership is considered as more critical since it impacts more flows departure from and arrive at this node.

The geographic pattern of nodal criticality presented in Figure 3.3 clearly shows the locations and criticality of stations in the network. To aid visualization, the 62 Amtrak stations are grouped into four classes — most, more, less, and least critical — based on quantile classification. The most critical stations are characterized in two ways. First, stations having high nodal degree and significant passenger flow, such as Chicago, IL (CHI) and Los Angeles, CA (LAX), are highly critical, for they play an important role as hubs in carrying and transferring passengers. Second, several critical stations such as Savannah, GA (SAV), Schenectady, NY (SDY), and Kissimmee, FL (KIS) are the main channel connecting substantial parts of the network. These essentially intermediate nodes allow stations that have a high flow volume to access the rest of the network. In contrast, the least critical stations are all end nodes and chiefly found at the corners or in the outskirts of the network, such as at Bellingham, WA (BEL) in the northwest corner; Tampa, FL (TPA) and Miami, FL (MIA) in the southeast corner; St. Albans, VT (SAB), Plattsburgh, NY

Table 3.2 Results of single-node disruption scenarios for Amtrak rail stations using the AND-SPN.

Rank	Disabled Station Name	Station Code	Total Disruption Cost (Obj. values)[a]	Station Ridership *status quo*	Time (sec.)	Iterations	Criticality Indicators		
							AAU[b]	APC[c]	AAF[d]
1	Hammond-Whiting, IN	HMI	5,850,760,978,800	7,763	9.31	77,540	183	3,089	10,270
2	Savannah, GA	SAV	4,704,313,371,800	71,658	9.02	70,969	140	2,438	6,008
3	Schenectady, NY	SDY	4,700,201,074,900	61,803	8.63	69,962	150	2,559	6,589
4	Battle Creek, MI	BTL	3,600,000,000,000	49,203[e]	9.59	77,122	152	2,525	7,031
5	Kissimmee, FL	KIS	3,600,000,000,000	44,906	9.41	74,144	145	2,443	6,232
6	Castleton, VT	CNV	3,600,000,000,000	4,211	9.13	74,687	146	2,532	6,881
7	Galesburg, IL	GBA	2,422,073,932,400	103,717	9.20	75,465	163	2,578	6,939
8	Los Angeles, CA	LAX	2,420,357,116,500	1,643,706	8.33	66,267	157	2,447	7,285
9	Buffalo, NY	BFX	2,420,284,892,200	6,019	9.36	78,074	173	2,554	6,867
10	Everett, WA	EVR	2,420,239,040,900	43,115	8.53	72,263	153	2,384	6,607
11	Richmond, VA	RVR	2,420,046,901,200	266,237	9.09	74,330	156	2,544	7,252
12	Springfield, MA	SPG	2,420,041,127,900	202,095	9.13	72,117	158	2,520	7,140
13	Fort Worth, TX	FTW	2,420,012,504,500	129,389	9.11	75,552	155	2,470	6,979
14	Boston, MA	BOS	2,420,002,913,100	50,204	8.81	74,856	151	2,511	7,024
15	Stockton, CA	SKN	2,420,000,000,000	39,782	9.03	71,113	154	2,383	6,611
16	Chicago, IL	CHI	1,230,626,974,900	3,522,388	9.23	78,655	217	3,416	12,758
17	Martinez, CA	MTZ	1,225,309,987,400	473,836	8.33	65,328	160	2,568	5,729
18	Emeryville, CA	EMY	1,224,670,499,300	598,859	8.52	68,633	157	2,541	6,433
19	Sacramento, CA	SAC	1,223,378,907,700	1,132,750	8.08	61,324	168	2,613	6,064
20	Portland, OR	PDX	1,222,832,248,500	190,353	8.66	66,011	157	2,563	6,779
21	Vancouver, WA	VAN	1,222,764,746,600	98,473	8.23	63,486	159	2,565	6,659
22	Washington, DC	WAS	1,222,087,402,800	598,859	8.50	70,043	167	2,790	6,891

Table 3.2 Continued. Results of single-node disruption scenarios for Amtrak rail stations using the AND-SPN.

Rank	Disabled Station Name	Station Code	Total Disruption Cost (Obj. values)[a]	Station Ridership *status quo*	Time (sec.)	Iterations	Criticality Indicators		
							AAU[b]	APC[c]	AAF[d]
23	Philadelphia, PA	PHL	1,221,898,154,100	331,045	7.75	63,974	154	2,653	7,638
24	Newark, NJ	NWK	1,221,827,989,100	12,971	8.95	72,547	152	2,637	7,476
25	New York, NY	NYP	1,221,822,188,700	12,749	8.50	69,972	155	2,630	7,531
26	Cleveland, OH	CLE	1,221,821,659,700	50,940	9.36	79,220	193	2,644	9,610
27	Kansas City, MO	KCY	1,221,125,169,000	164,859	9.72	78,538	162	2,517	7,549
28	Milwaukee, WI	MKE	1,221,109,477,600	617,153	9.66	78,369	165	2,639	6,237
29	Pittsburgh, PA	PGH	1,221,063,420,200	135,137	9.17	78,044	165	2,557	8,361
30	BWI Airport, MD	BWI	1,221,040,661,800	152,403	9.05	73,499	158	2,595	7,092
31	Spokane, WA	SPK	1,221,035,659,900	63,975	8.19	67,037	165	2,535	6,038
32	Wilmington, DE	WIL	1,220,795,181,700	738,313	8.67	72,845	157	2,589	7,159
33	Baltimore, MD	BAL	1,220,602,502,600	1,065,576	8.75	72,781	157	2,590	7,124
34	San Antonio, TX	SAS	1,220,444,630,600	68,268	9.38	75,386	161	2,522	7,316
35	Lancaster, PA	LNC	1,220,346,193,500	578,731	9.09	77,160	162	2,529	7,343
36	Harrisburg, PA	HAR	1,220,331,408,500	571,940	9.17	77,545	164	2,525	7,355
37	Albany-Rensselaer, NY	ALB	1,220,272,964,500	764,898	8.80	73,573	167	2,581	7,318
38	Greensboro, NC	GRO	1,220,254,082,800	139,869	9.33	79,246	167	2,548	7,671
39	New Haven, CT	NHV	1,220,171,011,300	745,530	9.02	74,031	159	2,520	7,268
40	St. Louis, MO	STL	1,220,156,875,500	378,146	9.00	75,118	163	2,536	7,554
41	New Orleans, LA	NOL	1,220,104,648,100	212,426	9.55	79,709	162	2,514	7,401
42	Cary, NC	CYN	1,220,103,751,000	88,669	9.17	75,663	165	2,529	7,962
43	Charlottesville, VA	CVS	1,220,099,439,900	29,434	9.33	78,639	167	2,530	7,316
44	Selma-Smithfield, NC	SSM	1,220,058,786,100	13,222	8.75	74,424	162	2,523	7,457

Table 3.2 Continued. Results of single-node disruption scenarios for Amtrak rail stations using the AND-SPN.

Rank	Disabled Station Name	Station Code	Total Disruption Cost (Obj. values)[a]	Station Ridership *status quo*	Time (sec.)	Iterations	Criticality Indicators		
							AAU[b]	APC[c]	AAF[d]
45	Seattle, WA	SEA	1,220,052,604,300	640,054	8.89	74,174	157	2,449	6,956
46	Providence, RI	PVD	1,220,002,111,400	660,267	9.26	77,694	156	2,505	7,147
47	San Diego, CA	SAN	1,220,000,000,000	686,953	9.66	78,824	155	2,448	7,069
48	Bakersfield, CA	BFD	1,220,000,000,000	546,439	9.34	78,260	155	2,435	6,783
49	Portland, ME	POR	1,220,000,000,000	190,353	9.48	79,554	154	2,500	7,035
50	Plattsburgh, NY	PLB	1,220,000,000,000	175,299	9.44	79,774	155	2,504	7,061
51	Tampa, FL	TPA	1,220,000,000,000	139,412	9.53	79,773	154	2,480	6,839
52	Newport News, VA	NPN	1,220,000,000,000	128,317	9.58	80,230	155	2,511	7,043
53	Miami, FL	MIA	1,220,000,000,000	84,293	9.28	79,973	154	2,472	6,855
54	Rutland, VT	RUD	1,220,000,000,000	68,463	9.50	79,836	155	2,507	7,061
55	Bellingham, WA	BEL	1,220,000,000,000	68,267	9.48	79,398	155	2,436	6,813
56	Oklahoma City, OK	OKC	1,220,000,000,000	54,952	9.52	80,471	156	2,477	6,961
57	Grand Rapids, MI	GRR	1,220,000,000,000	51,993	9.63	80,737	157	2,510	7,075
58	Quincy, IL	QCY	1,220,000,000,000	50,378	9.69	80,781	156	2,506	7,075
59	Port Huron, MI	PTH	1,220,000,000,000	29,461	9.56	80,347	156	2,501	7,109
60	Pontiac, MI	PNT	1,220,000,000,000	17,077	9.55	80,439	156	2,502	7,109
61	St. Albans, VT	SAB	1,220,000,000,000	3,592	9.89	79,661	155	2,499	7,055
62	Niagara Falls, ON	NFS	1,220,000,000,000	3,113	9.80	80,533	157	2,512	7,053

Note: [a] M=10,000,000,000; [b] AAU at *status quo* is 159; [c] APC at *status quo* is 2,493; [d] AAF at *status quo* is 7,147; [e] values highlighted with blue color indicate where the Node Ridership at *status quo* serves as auxiliary ranking.

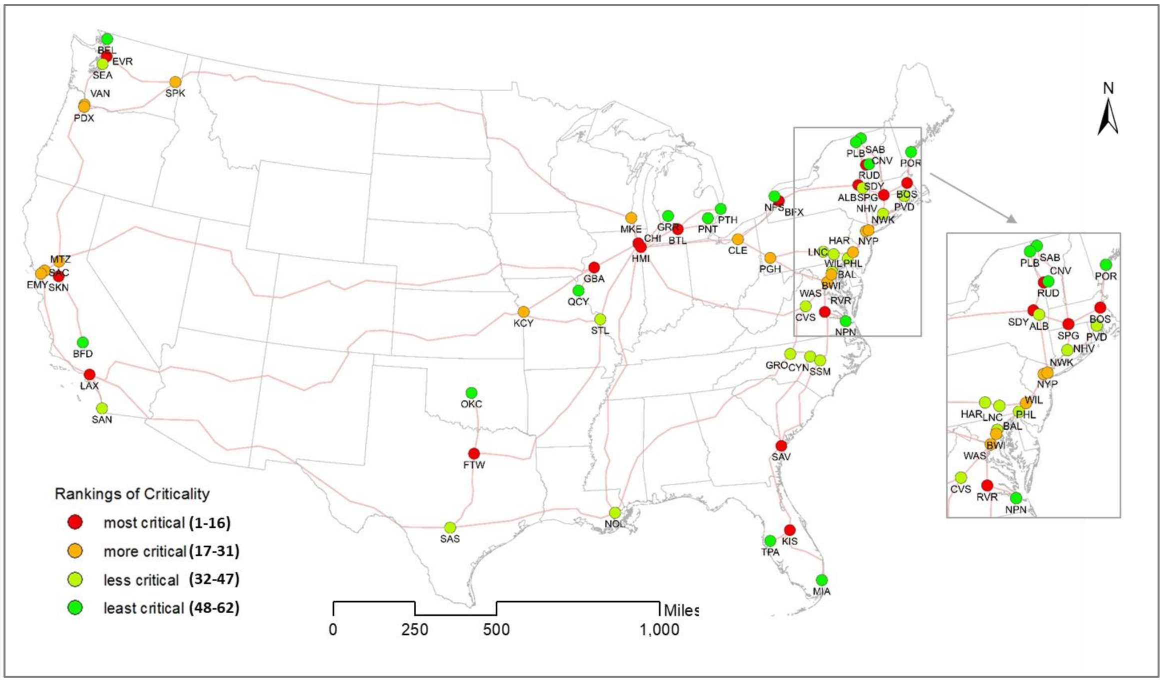
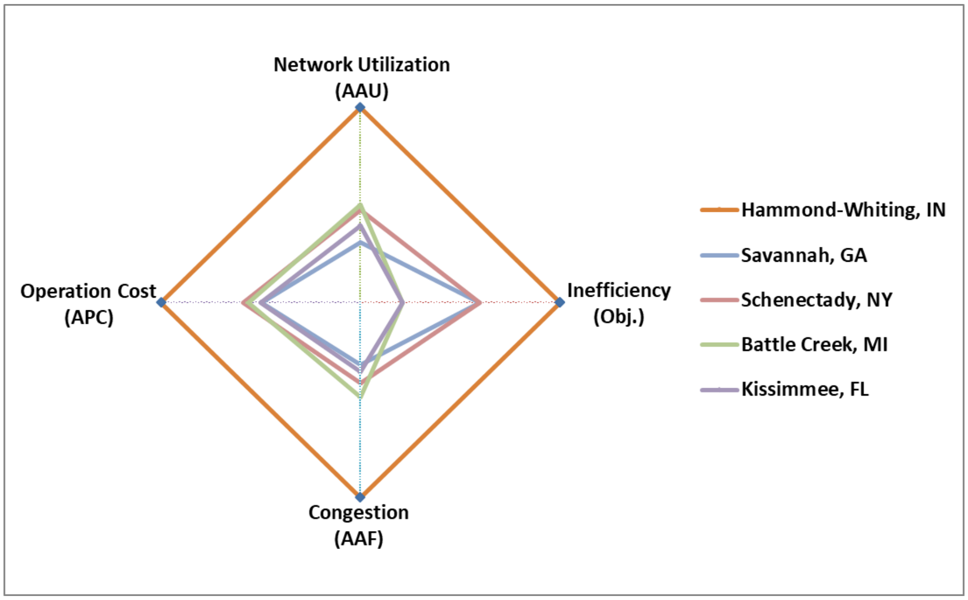


Figure 3.3 Classification of Amtrak stations based on criticality

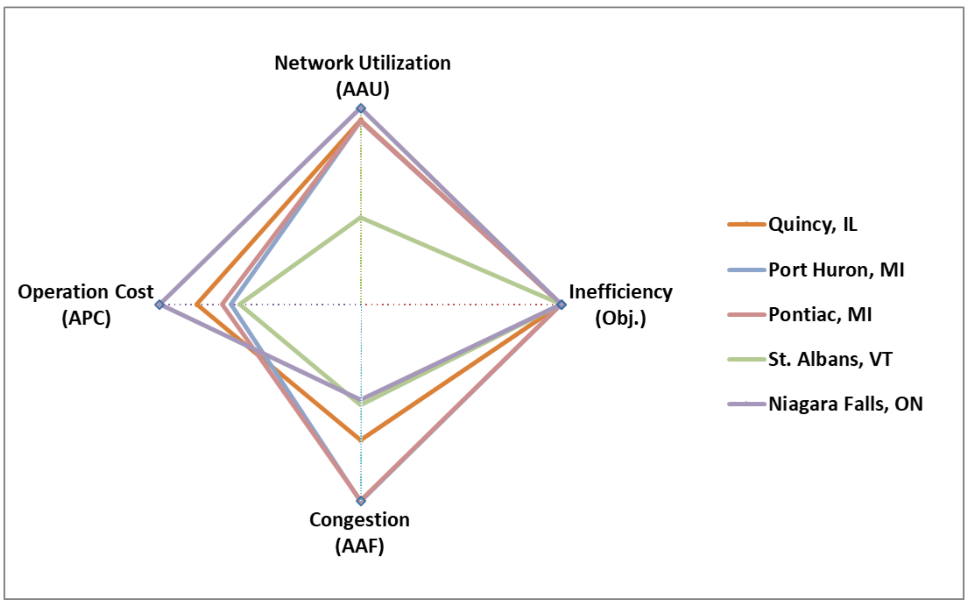
(PLB), and Portland, ME (POR) in the northeast corner; and Port Huron, MI (PTH) and Pontiac, MI (PNT) in the far north. In what seems to be a common characteristic, the highly ranked stations are characterized as having higher nodal degree, higher passenger volume or as being a crucial gateway or hub.

Table 3.2 also provides additional important information about the anticipated network response when using three different criticality indicators: AAU, APC, and AAF. Each indicator reflects a different aspect of nodal criticality, allowing decision makers to assess stations' resilience. As a good example, Schenectady, NY (SDY), the third most critical station as ranked by the objective function, is not significant from an AAU (59th), APC (18th) and AAF (55th) standpoint. Conversely, Chicago, IL (CHI), although only ranked the sixteenth most critical station by the objective function, affects other network performance significantly — AAU (1st), APC (1st), and AAF (1st) — owing to its vital location in connecting east and west, its high node degree, and its significant passenger flow. Similarly, Cleveland, OH (CLE) greatly affects AAU and AAF because of its role in bridging northeastern stations, and Washington, D.C. (WAS) exerts an obvious influence on APC by facilitating passenger flow between southeastern and northeastern areas.

In Figure 3.4, positionality maps demonstrate how stations' criticality can be assessed through the



(a) Five most critical stations in criticality ranking



(b) Five least critical stations in criticality ranking

Figure 3.4 Positionality maps to compare stations' criticalities positionality of nodal criticalities in terms of four criticality indicators among stations, as reported in Table 3.2. A positionality map allows network planners to characterize stations by type of criticality. To see how effectively the positionality map highlights differences of criticality, the five most (Figure 3.4-a) and five least (Figure 3.4-b) critical stations are selected. The position of each indicator is scaled for comparison. When defining each indicator, four indicators—objective value, AAU, APC, and AAF—serve as proxies for network inefficiency, network utilization, operation cost, and congestion, respectively. On the scaled *x*–*y* coordinate plane, the farther the position is from the center, the more critical it is. On the *x*-axis, we find inefficiency and operation cost, two well-accepted network performance measures. When the operation of a station is disrupted, passengers at origins must detour to reach their destinations, increasing total disruption costs and travel distance, which are seen as inefficiency (objective function), and increase of operation cost (APC). On the *y*-axis, network utilization (AAU) and congestion (AAF) are counter-metrics describing links' ability to absorb the burden of rerouted flow resulting from the disruption. A station is considered as more critical if higher levels of network utilization are needed to mitigate the effect of network inoperability by splitting the burden through the arcs available in the network, whereas a lower level of AAF indicates a lower congestion level among arcs and a commensurately lower network delay for scheduled trains. Thus, the larger the rectangle needed to encompass the four charting positions of a station on the *x*–*y* positionality plane, the greater the effects of the disruption.

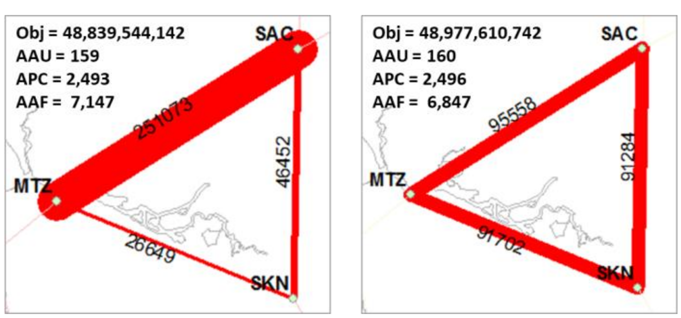
As shown in Figure 3.4(a), the Hammond-Whiting, IN station is clearly the most critical station among the top five stations, with all metrics located at the farthest possible point on both the *x* and *y* axes. Notice that such stations as Battle Creek, MI and Savannah, GA are, however, regarded as critical in different aspects. The disruption of Battle Creek, MI would not entail a greater inefficiency, but it exerts a greater influence on network utilization (AAU) and congestion (AAF) than that of Savannah, GA would. The positionality map also indicates that Savannah, GA and Schenectady, NY, whose rectangles are of similar shape, are thus similarly critical stations. Figure 3.4(b) indicates that Pontiac, MI and Niagara Falls, ON are of very different classes of criticality, with Pontiac, MI stretched on the *y*-axis to represent its relevance for network utilization and congestion. In contrast, Niagara Falls, ON is more associated with operation cost (APC), so its rectangle is wider along the *x*-axis.

3.4.3 Effect of capacity restriction

Restrictions on an arc's capacity for passenger flow are well addressed in transportation network design when excessive concentration of flows on particular arcs is anticipated. Overutilization of network links is usually associated with the potential for traffic congestion, which can cause chronic delays and the degradation of on-time performance in the rail system.

The effects of capacity restriction are well demonstrated in Figure 3.5(a). The busiest arcs of the Amtrak system at *status quo* are California's Sacramento–Martinez (SAC-MTZ) and Martinez–Sacramento (MTZ-SAC) arcs, where the sum of the amount of flows for both directions is 251,073; other connected linkages from MTZ are significantly underused at 26,649 (MTZ-SKN & SKN-MTZ) and 46,452 (SKN-SAC & SAC-SKN). However, if a CAP of 125,000, which is an appropriate level for this demonstration, is imposed for each directional link to the two busiest linkages, SAC-MTZ and MTZ-SAC, through the constraints (3.7) in the SPN, then flows are redistributed over three arcs as shown in Figure 3.5(b). The total travel cost increases, because the excessive flow that initially traveled through the link SAC-MTZ must reroute via SAC-SKN and SKN-MTZ. Accordingly, the level of AAU and APC also increases slightly. However, note that this CAP restriction produces a more balanced flow distribution, sparing each arc from excessive concentration of flow and reducing the gap in linkage use, as indicated by the decrease in AAF.

Capacity restriction also affects network performance and flow patterns in disruption cases. Table 3.3 provides the results of selected multi-nodal disruption scenarios under different CAP conditions using the AND-SPN. The disruption scenarios range from $|R|=1$ to $|R|=3$, and the worst performance scenarios associated with each $|R|$ are selected so as to apply two conditions for capacity restriction on the two busiest linkages (2-CAP) and the ten busiest linkages (10-CAP).



(a) No CAP on each link (b) CAP=125,000 on each link

Figure 3.5 Change of flow distribution with capacity restriction on the links SAC-MTZ and MTZ-SAC in California at *status quo*

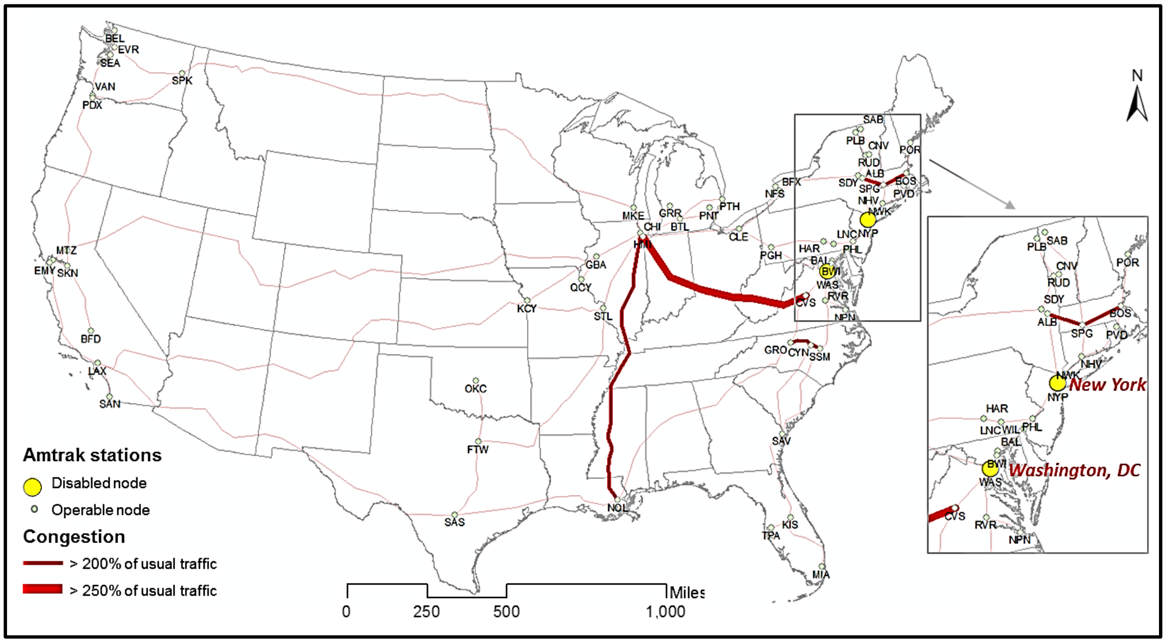
Table 3.3 Results of capacity restriction on linkages to disruption scenarios.

|*R*|*	**Disabled Station Code**	**Number of Arcs Imposing CAP (CAP=100,000)**	**Obj. values**	**AAU**	**APC**	**AAF**
1	HMI	No CAP	5,850,760,978,800	183	3,089	10,270
1	HMI	2-CAP	5,850,966,207,800	186	3,097	9,577
1	HMI	10-CAP	6,396,483,086,100	181	3,279	6,364
2	HMI, SDY	No CAP	10,150,279,588,400	158	3,158	9,475
2	HMI, SDY	2-CAP	10,150,479,807,400	160	3,166	8,808
2	HMI, SDY	10-CAP	10,675,886,740,400	155	3,354	5,870
3	HMI, SDY, SAV	No CAP	14,130,041,237,400	141	3,145	8,398
3	HMI, SDY, SAV	2-CAP	14,130,236,161,400	143	3,153	7,780
3	HMI, SDY, SAV	10-CAP	14,137,807,334,400	144	3,517	4,227

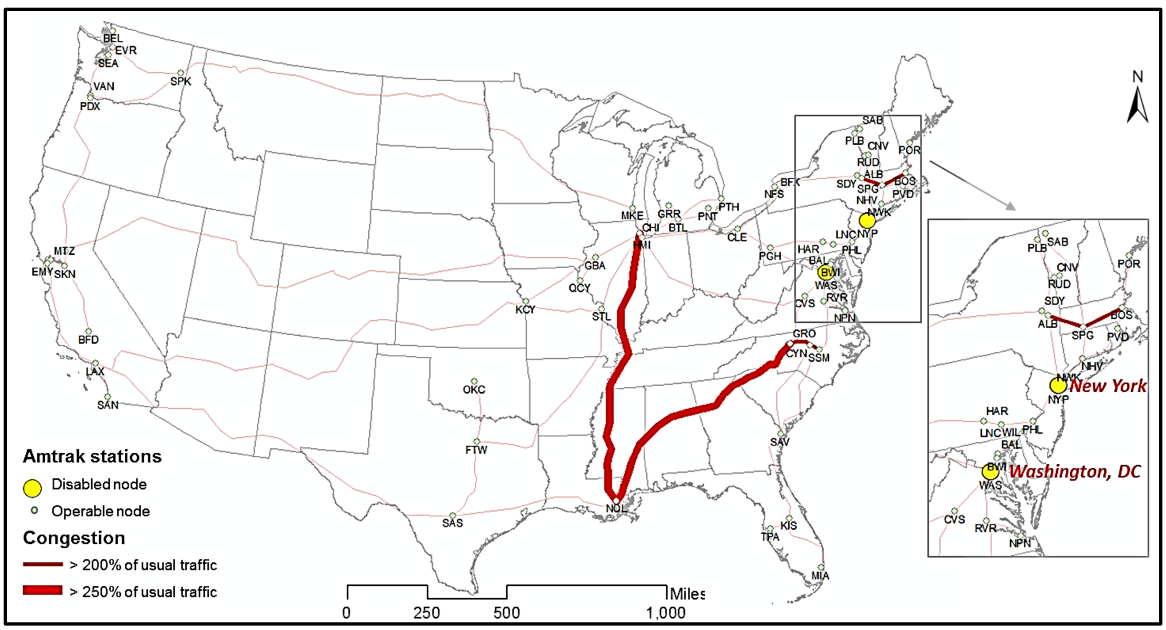
Note: * |*R*| = the number of disabled nodes *r*.

To enhance the comparison, results without capacity restriction (No CAP) are presented for each scenario. The CAP level is set to 100,000, a figure chosen in response to computational experiments conducted at status quo. For each |*R*|, it is clearly observed that the objective values and APC increase, but that the values of AAF decrease, with the number of CAP-imposed linkages. Interestingly, the values of AAU do not respond consistently to changes in CAP.

Figure 3.6 shows how imposition of a CAP affects the geographic pattern of passenger flows, focusing on changes in congested paths over the network. To examine the changed behavior, two stations—Washington, D.C. (WAS), and New York, NY (NYP)—were selected as a case study for a disruption scenario. In this case, we tailor an adaptive capacity restriction that is 2.5 times the amount of flow to *each* link at status quo instead of applying a uniform capacity to the entire links, because the capacities of each arc are actually different in practice. As seen in Figure 3.6(a), the disruption scenario without CAP shows that the paths between HMI-NOL, ALB-BOS (> 200% of excessive flows at status quo), and HMI-CVS (> 250%) are anticipated to have a high flow volume. Of particular interest, then, is the question of which paths would suffer from the congestion if the congested path HMI-CVS were limited by capacity restriction so as to resolve traffic congestion. Figure 3.6(b) shows the result. When the burden on the path HMI-CVS is released, the flows are split to nearby paths HMI-NOL and CYN-NOL, which then emerge as the new busiest paths, with highly congested flow as their by-product.



(a) Congested linkages without capacity restriction



(b) Congested linkages with capacity restriction

(Note: CAP = 2.5 multiplied by the amount of passenger flow to each link at *status quo*)

Figure 3.6 Congested linkages with capacity restriction

3.5 Concluding Remarks

This research culminated in an alternative method of measuring the level of nodal criticality and network vulnerability as reflected in the shortest paths used in a given network. Specifically, not only the objective function was designed to measure the effect of disruption on stations but also three complementary indicators—AAU, APC, and AAF— were prescribed to provide network planners as well as decision makers with the information they need to help them reduce network vulnerability and reinforce transportation security in multiple dimensions. To facilitate the numerical experiment, two models were constructed: the SPN, an expanded version of the shortest path problem in which a shortest path is identified for a single *s*-*t* pair, minimizes the total network transport cost through finding the shortest paths for every *s*–*t* pair in the network and distribute optimally via the shortest paths; the SPN was further expanded into the AND-SPN, which measures the impact of node failures to network disruption cost.

In our case study of the Amtrak rail system, we evaluated stations' ability to affect the performance of the entire network in terms of network disruption cost and supplementary indicators, emphasizing that the effect of a disruption might not be identically represented by all criticality indicators. By implementing the AND-SPN model using well-estimated passenger flows in a future time, the obtained list of stations ranked by criticality can help decision makers decide which stations are their highest priority during special events or disasters or when enhancing transportation network security.

The selected Amtrak set with 62 nodes was dedicated in demonstrating numerical results using the SPN and the AND-SPN models. However, the dataset is not ideal for practical implementations. A possible direction of future research is to apply the proposed models in multi-modal transportation networks, such as the combination of Amtrak rail and road transport systems, considering that when Amtrak stations are halted from operation, flow reroute through road systems is usually a timelier solution.

Finally, it is worth noting that application of the SPN and the AND-SPN is not limited to planar-type networks, but also applicable to non-planar networks such as air transportation routes and subway systems with scheduled routes. The AND-SPN model is formulated in a way that only accommodates bi-directional arcs, but it can be modified to solve shortest path network problems with one-way arcs as well.

CHAPTER 4

PARTIAL NODE FAILURE

IN SHORTEST PATH NETWORK PROBLEMS

This chapter is a manuscript in preparation for submission to an academic journal.

Abstract

To address partial node failure in network modelling, this paper expands the Shortest Path Network model (SPN) from a previous study to accommodate partial node failures, by introducing link attribute index and distance update constraints to the model. The influence of partial node failures in a flow-based shortest path network is assessed by two indicators, which are flow reroute cost and amount of pass-through flows caused by partially failed nodes. Numerical experiments are carried out on a small sample network, and then applied to the Washington Metropolitan Area Transit Authority network (WMATA). Analysis results detail the change in network flow distribution after a station partially fails, and identify impacted stations where pass-through flows increase. Various disruption scenarios are simulated using different station-to-station flows, providing comprehensive information to evaluate plans for partial node failures such as scheduling maintenance. The proposed model and simulation procedure could be used as a tool to schedule maintenance or plan for temporary closure when needed in a future time if its flows could be well-predicted. It is noteworthy that the rankings of station criticality do not remain the same using different station-to-station flows, which addresses the importance of flows in modelling shortest path network problem in the real-world networks.

Key words: Partial node failure, Shortest Path Network problems for Partial node failure (SPN^{Pr}), link attribute, distance update, flow reroute cost, ViaFlow, the Washington Metropolitan Area Transit Authority network (WMATA)

4.1 Introduction

Aside from complete failure of network facilities in most literature of network vulnerability studies, partial failure has been brought to attention recently with some studies going beyond the conventional topological analysis of complete node or link failures. Partial failures are prone to occur more frequently than complete failures due to administrative or maintenance activities, and thus may result in substantial impacts and significant costs. Nagurney and Qiang (2007) studied the impact of varying degrees of link capacity reductions on the relative decrease of network efficiency. Sullivan et al. (2010) analyzed the influence of different link-based capacity-disruption values on the identification and rankings of critical links in road transport networks, and found that the criticality ranking of links is sensitive to the level of capacity reduction. Burgholzer et al. (2013) researched the impact of partial disruptions for selected links under different scenarios regarding time-of-day, duration and level of capacity reduction in multimodal transport networks. Cats and Jenelius (2016) investigated the relation of a range of link capacity reductions on performance of public transport networks, and showed that reduced line frequency impacted network operation as passenger loads increased.

In Chapter 3 of this dissertation, complete node failures of selected Amtrak stations were simulated. However, disruptions do not always amount to complete node failures but are caused by partial node failures of stations. In addition, partial node failure could be planned ahead for management purpose or scheduled maintenance. In the events of traffic control or facility maintenance at a transfer station, some passengers have to detour via alternative shortest paths where an alternative transfer station is available. Thus, the Shortest Path Network (SPN) model needs to be extended further by adding some constraints so that not only complete node failure but also partial node failure can be simulated in the SPN model to assess nodal criticality of heavy rail stations.

The aim of this chapter is to implement partial node failure into the SPN model. Specially, a case study of this chapter aims to answer the question that, if a Mandatory Transfer (MT) has to be partially closed (i.e., only allow traffic to go through) for one-day period due to scheduled heavy maintenance or major upgrade, which day is the best day of week so that the total additional network cost is minimum.

4.2 Methodology and Formulations

4.2.1 Distance Update constraints and Link Attribute index

Since partial node failure does not simply follow all-or-nothing scenario, the distance matrix needs to be updated when partial node failure happens. Moreover, when the SPN is applied to heavy rail network, the rail lines should be distinguished from each other in order to decide whether the action of transfer is required at a station on the shortest path for an Origin-Destination (OD) pair. To update distance matrix and differentiate rail lines, two elements are introduced in this study, which are distance update constraints 4.1 and link attribute index 4.2.

$$d'_{ij} = \begin{cases} \infty & i = r;\ \text{or}\ j = r; or\ i \neq j \neq r,\ d_{ij} = \infty,\ l_{ij} = 0; \\ d_{ir} + d_{rj} & i \neq j \neq r,\ d_{ij} = \infty,\ l_{ij} = 1; \\ d_{ij} & i \neq j \neq r,\ d_{ij} \neq \infty \end{cases} \quad \forall i, j \ \in N_r \tag{4.1}$$

where

$$l_{ij} = \begin{cases} 1, & if\ L_{ir}\ and\ L_{rj}\ contain\ same\ value \\ 0, & otherwise \end{cases} \tag{4.2}$$

where

i, j = index of operable node

r = index of partially failed node

N_r = a set of nodes directly connected to the partially failed node r

d'_{ij} = distance between nodes i and j when node r partially fails

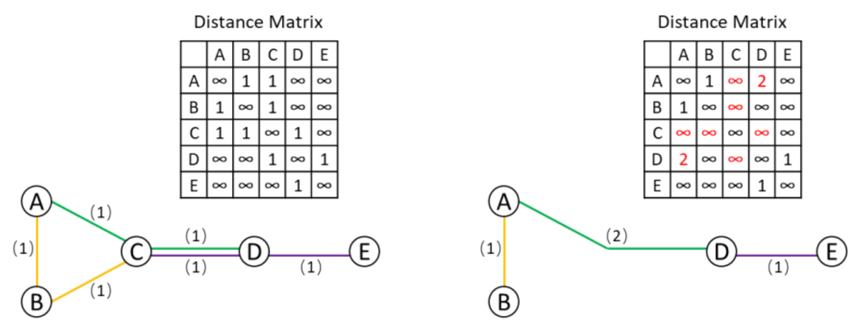
d_{ij} = distance between nodes i and j at *status quo*

l_{ij} = link attribute index

L_{ir} = a set of links between nodes i and r

Note that, $d_{ij} = \infty$ means there is no linkage between nodes i and j (i.e., connectivity = 0). To explain the application of equations 4.1 and 4.2, a sample directional network is presented with five nodes and three lines in Figure 4.1(a). All distance among directly connected node pairs is set to be 1 for simplicity.

In the context of rail network operation, a *line* is serviced by a specific route with trains stopping along the line's stations. The lines are usually distinguished by colors, numbering, names, or a



(a) *status quo* (b) partial node failure of node C

Figure 4.1 The update of distance matrix after partial node failure combination thereof. For example, in Figure 4.1(a), the green line is serviced by a bidirectional route from nodes A to C to D, the yellow line is serviced by a bidirectional route from nodes A to B to C, and the purple line is serviced by a bidirectional route from nodes C to D to E. In contrast, a *link* refers to the line segment connecting two nodes. For example, there is one link between nodes A and C, and two links between nodes C and D.

At status quo, four links exist between node C and the node set N_C: links A-C (green), B-C (yellow), D-C (green), and D-C (purple). When node C partially fails, some of the four links are removed while others are preserved. Specifically, links B-C (yellow) and D-C (purple) are disabled, since the partially failed node C becomes a dead end for these two links (in other words, L_{BC} and L_{DC} do not contain same value); links A-C (green) and D-C (green) remain operable, since the partially failed node C is merely an "intermediate" node connecting the two links on the green line (in other words, L_{AC} and L_{DC} contain a same value of "green"). The network structure after node C partial fails is presented in Figure 4.1(b).

In the partial failure of node C, a four-step procedure is employed to update the distance matrix. First, the set of nodes directly connected to node C (i.e., N_c) is identified as below to decide the distance of which node pairs to be updated.

N_C = {A, B, D}

Second, the sets of links between node C and the three nodes in N_C are as follows.

L_{AC} = {green},
L_{BC} = {yellow},
L_{DC} = {green, purple}.

Third, the node pair among the three nodes contained in N_C are AB, AD, and BD. Link attributes of the three node pairs are as below.

$l_{AB} = 0$ (L_{AC} and L_{BC} do not contain same value),
$l_{AD} = 1$ (L_{AC} and L_{DC} contain a same value of "green"),
$l_{BD} = 0$ (L_{BC} and L_{DC} do not contain same value).

Fourth, according to the distance update constraints, the updated distance is obtained as below.

1. when $i=r$ or $j=r$, $d'_{ij}=\infty$

Therefore, $d'_{AC}=\infty$, $d'_{BC}=\infty$, $d'_{DC}=\infty$, $d'_{CA}=\infty$, $d'_{CB}=\infty$, $d'_{CD}=\infty$.

2. when $i \neq j \neq r$

1) when $d_{ij} \neq \infty$, $d_{ij}=d'_{ij}$

Therefore, $d'_{AB}=d_{AB}=1$, $d'_{BA}=d_{BA}=1$.

2) when $d_{ij}=\infty$

(1) when $l_{ij}=0$, $d'_{ij}=\infty$

Therefore, $d'_{BD}=\infty$, and $d'_{DB}=\infty$.

(2) when $l_{ij}=1$, $d'_{ij}=d_{ir}+d_{rj}$

Therefore, $d'_{AD}=d_{AC}+d_{CD}=2$, and $d'_{DA}=d_{DC}+d_{CA}=2$.

The updated distance matrix of the sample network is provided in Figure 4.1(b). Marked in red color are the differences of distance matrix in Figure 4.1(b) from that in Figure 4.1(a).

4.2.2 The SPN model for Partial node failure (SPN^{Pr})

With distance matrix updated, the SPN is modified to be the SPN^{Pr} to solve each partial node failure scenarios. In equations 4.3 to 4.5, the superscript "P_r" for Ω, C_{ij}, and X_{stij} stands for "Partial" node failure of node r. It is obvious that since distance matrix is modified in partial node failures, some of the shortest paths from origin to destination will be different from *status quo*, and thus change the total network cost.

$$\text{Minimize} \quad \Omega^{P_r} = \sum_s \sum_t \sum_i \sum_{j \in N_i} W_{st} C_{ij}^{P_r} X_{stij}^{P_r} \tag{4.3}$$

subject to

$$\sum_{j \in N_i} X_{stij}^{P_r} - \sum_{j \in N_i} X_{stji}^{P_r} = \begin{cases} 1, \forall i = s \\ 0, \forall i \neq s,t \\ -1, \forall i = t \end{cases} \qquad \forall s,t,i,j; s \neq t \neq r, i \neq j \neq r \tag{4.4}$$

$$X_{stij}^{P_r} \in \{0,1\} \qquad \forall s,t,i,j; s \neq t \neq r, i \neq j \neq r \tag{4.5}$$

where

Ω^{Pr} = objective function when node r partially fails

r = index of disabled node(s)

s = index of origin

t = index of destination

i, j, k = index of operable node

N_i = set of nodes connected directly to node i by an arc (i, j)

W_{st} = amount of flows (passenger or freight) to be delivered from s to t

C_{ij}^{Pr} = travel cost of arc (i, j) when node r partially fails ($i \neq j \neq r$)

X_{stij}^{Pr} = 1: if arc (i, j) is the member of the shortest path from s to t, 0: otherwise when node r partially fails

4.2.3 Flow Reroute Cost

Nodal criticality by partial node failure is assessed by the additional network cost caused by flow reroutes. If node r partially fails, the flows that initially transfer at node r will have to reroute through other available paths and they travel longer distance than at *status quo*, unless there are multiple optimal paths between a node pair. The amount of additional network cost caused by flow reroutes is considered as an important indicator for the impact of partial node failure. It is noteworthy that, in this study, flows departing from or arriving at the partially failed node are not of concern in analyzing flow reroute for two reasons. For one thing, the amount of flows departing from and arriving at a partially failed node could be easily identified by looking at its ridership at *status quo*; for another reason, the re-assignment of these flows at the partially failed node could be complicated (e.g., it is uncertain that how passengers or cargos would access the network by connecting other transportation modes such as bus, taxi, bike, or even walk).

Equation 4.6 illustrates the calculation of flow reroute cost denoted by $Cost_{reroute}^{P_r}$, which is obtained by subtracting total network cost when node r partially fails (Ω^{Pr}) by total network cost at *status quo* excluding the network cost related with node r (Ω^{sq-r}).

$$Cost_{reroute}^{P_r} = \Omega^{P_r} - \Omega^{sq-r} \tag{4.6}$$

where

$Cost^{P_r}_{reroute}$ = flow reroute cost when node *r* partially fails

Ω^{sq-r} = total network cost at *status quo* excluding the network cost related with node *r*, calculated by equation

$$\Omega^{sq-r} = \Omega^{sq} - (\sum_{t(t \neq r)} \sum_{i} \sum_{j \in N_i} W_{rt} C_{ij} X_{rtij} + \sum_{s(s \neq r)} \sum_{i} \sum_{j \in N_i} W_{sr} C_{ij} X_{srij}) \ \forall s,t,i,j; i \neq j \ ,$$

where Ω^{sq} = objective function at *status quo*; C_{ij} = travel cost of arc (*i*, *j*) at *status quo*; X_{stij} = 1: if arc (*i, j*) is the member of the shortest path from *s* to *t*, 0: otherwise at *status quo*

s, t, i, j, r, N_i, Ω^{Pr}, and W_{st} are as defined in equation 4.3.

4.2.4 ViaFlow and NodeFlow

In this study, the term *ViaFlow* is defined as the amount of flows pass through a specific node. Due to the flow reroutes caused by a partial node failure, some nodes may transport more passengers or goods than usual, since these nodes are now on the shortest path for some node pairs that initially not pass through them. Equations 4.8 and 4.9 calculate the ViaFlow at node *k* when node *r* partially fails ($ViaFlow_k^{Pr}$) and at *status quo* ($ViaFlow_k^{sq}$), respectively. Equation 4.7 calculates their difference through subtracting $ViaFlow_k^{Pr}$ by $ViaFlow_k^{sq}$.

The term *NodeFlow* is defined as the sum of all flows at a node, including all flows departing from and arriving at the node as well as all inbound and outbound pass-through flows. Equations 4.10 and 4.11 detail the calculation of NodeFlow at node *k* when node *r* partially fails ($NodeFlow_k^{Pr}$) and at *status quo* ($NodeFlow_k^{sq}$), respectively. Equation 4.12 calculates the ridership at node *k* when node *r* partially fails, by subtracting flow between node *k* and node *r* from its ridership at *status quo*.

$$DiffViaFlow_k^{P_r - sq} = ViaFlow_k^{P_r} - ViaFlow_k^{sq} \tag{4.7}$$

where

$$ViaFlow_k^{P_r} = \frac{NodeFlow_k^{P_r} - Ridership_k^{P_r}}{2} \qquad \forall k \neq r \tag{4.8}$$

$$ViaFlow_k^{sq} = \frac{NodeFlow_k^{sq} - Ridership_k^{sq}}{2} \qquad \forall k \neq r \tag{4.9}$$

$$NodeFlow_k^{P_r} = \sum_s \sum_{t(t\neq s)} \sum_{j\in N_k} W_{st} X_{stkj}^{P_r} + \sum_s \sum_{t(t\neq s)} \sum_{i\in N_k} W_{st} X_{s\,t\,i}^{P_r}$$
$$\forall s,t,i,j;\, s\neq t\neq r, i\neq k\neq r,\, j\neq k\neq r \tag{4.10}$$

$$NodeFlow_k^{sq} = \sum_s \sum_{t(t\neq s)} \sum_{j\in N_k} W_{st} X_{stkj} + \sum_s \sum_{t(t\neq s)} \sum_{i\in N_k} W_{st} X_{s\,t\,i}$$
$$\forall s,t,i,j;\, s\neq t\neq r, i\neq k\neq r,\, j\neq k\neq r \tag{4.11}$$

$$Ridership_k^{P_r} = Ridership_k^{sq} - (\sum_k \sum_r W_{kr} + \sum_r \sum_k W_{rk}) \qquad \forall k \neq r \tag{4.12}$$

Where

k = index of operable node

$DiffViaFlow_k^{Pr\text{-}sq}$ = difference of $ViaFlow_k^{Pr}$ from $ViaFlow_k^{sq}$

$ViaFlow_k^{Pr}$ = number of passengers or cargos pass through node k when node r partially fails

$ViaFlow_k^{sq}$ = number of passengers or cargos pass through node k at *status quo*

$NodeFlow_k^{sq}$ = sum of inbound and outbound flow at node k at *status quo*

$NodeFlow_k^{Pr}$ = sum of inbound and outbound flow at node k when node r partially fails

$Ridership_k^{sq}$ = number of passengers or cargos whose origin or destination is node k at *status quo*

$Ridership_k^{Pr}$ = number of passengers or cargos whose origin or destination is node k when node r partially fails

$s, t, i, j, r, N_k, W_{st}, X_{stij}$, and X_{stij}^{Pr} are as defined in equations 4.3 and 4.6.

4.3 Numerical Experiments

The proposed methodology in equations 4.3 to 4.12 is demonstrated using the same small sample network presented in Figure 4.1(a) at status quo and Figure 4.1(b) for partial node failure of MT_C. For simplicity, one passenger is assigned to travel between each OD pair; the unit of distance is assigned to be mile, and thus the unit of network cost is passenger*miles.

Table 4.1 summarizes the key numerical results of the small sample network. At *status quo*, using the SPN model from Chapter 3, the optimal solution is found to be 34 passenger*miles (Ω^{sq}), meaning that if every passenger takes the shortest path to reach his or her destination, the total network cost is 34 passenger*miles at *status quo*. When MT_C partially fails, the optimal solution is found to be 28 passenger*miles (Ω^{P_c}). Ω^{P_c} does not take into account the unrealized trips departing from or arriving at MT_C. To make a fair comparison, the network cost related with MT_C at *status quo* (which is 10 passenger*miles) should be deducted from the total network cost at *status quo* (Ω^{sq-c} = 34 - 10 = 24). Therefore, the flow reroute cost when MT_C partially fails ($Cost^{P_c}_{reroute}$) is 4 passenger*miles calculated by equation 4.6.

With respect to ViaFlow, when MT_C partially fails, all nodes experience flow decreases except node A, since passengers have to reroute through node A when MT_C partially fails, and node A would expect 4 more passengers pass through it than usual (at *status quo*, no passenger has to pass through node A to reach his or her destination).

In detail, at the partial node failure of MT_C, the passenger from node B to D has to reroute

Table 4.1 Summary of key numerical results of the sample network at *status quo* and when MT_C partially fails

key elements	**value**
Ω^{P_c}	28
Ω^{sq-c}	34 - 10 = 24
$Cost^{P_c}_{reroute}$	28 - 24 = 4
$NodeFlow_A^{sq}$	8
$Ridership_A^{sq}$	8
$NodeFlow_A^{P_c}$	14
$Ridership_A^{P_c}$	6
$ViaFlow_A^{sq}$	(8 - 8) / 2 = 0
$ViaFlow_A^{P_c}$	(14 - 6) / 2 = 4
$DiffViaFlow_A^{P_c\text{-}sq}$	4 – 0 = +4

through the path B-A-D (previously B-C-D at *status quo*) and thus travels 1 mile more than *status quo*; the passenger from node D to B also has to travel 1 more mile by detouring through node A; the passenger from node B to E has to reroute via the path B-A-D-E (previously B-C-D-E at *status quo*) and thus travels 1 mile more than *status quo*; the passenger from node E to B has to travel 1 more mile by detour through node A as well. Therefore, altogether 4 passenger*miles increase in total network cost and 4 passengers increase in ViaFlow at node A, confirming the results obtained from model testing.

4.3.1 The dataset – the WMATA network

When partial node failure happens in short-haul heavy rail networks, passengers are likely to reroute within the network rather than using other modes of transportation such as bus, because most of the stations are not far away from each other and thus the reroute will not be too lengthy and time-consuming by taking alternative routes. The selected network for the case study is the Washington Metropolitan Area Transit Authority (WMATA) network with 9 MTs out of 91 stations. The WMATA network possesses moderate complexity that is appropriate for the case study compared with the other U.S. heavy rail networks, which are either too simple with star- or tree-shaped network structure such as the Metropolitan Atlanta Rapid Transit Authority (MARTA) and the Greater Cleveland Regional Transit Authority (GCRTA) networks, or too large with several hundred nodes such as the New York City Transit (NYCT) network and too complicated with inner loops such as the Miami-Dade Transit (MDT) network.

Figure 4.2 shows the complete structure of the WMATA network with the three types of transfer stations (including 9 MTs, 31 non-MTs, and 3 ETs) as well as end stations and intermediate stations in the year of 2015. The lengths of each line segment are proportional to real physical distance. The station-to-station distance is obtained by using Washington Metropolitan Area Transit Authority API on developer.wmata.com. The distance is in miles, and the Origin-Destination distance ranges from 0.29 to 5.65 miles with an average of 1.25 miles.

Table 4.2 lists station names for all the 91 stations by node ID, and describes their station type and link attribute. In addition to the complete network structure showing the connectivity among stations and rail line attributes and station-to-station distance (C_{ij}), one more dataset is needed to implement the SPN^{Pr} model, which is the station-to-station passenger flow (W_{st}).

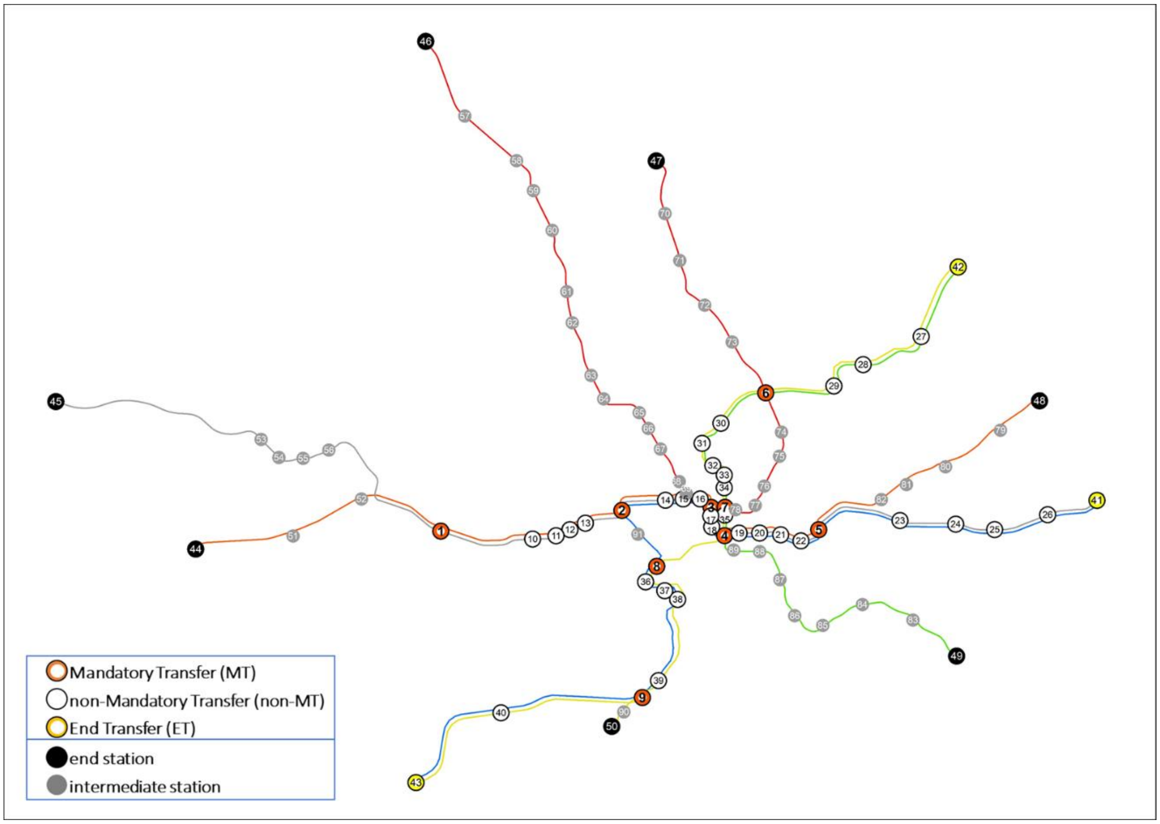


Figure 4.2 The WMATA network (proportional to real physical distance), 2015

Note: The network was modified and redrawn by author based on the WMATA network (Source: opendata.dc.gov/datasets/metro-stations-regional and opendata.dc.gov/datasets/metro-lines-regional); the GIS layers are projected to USA_Contiguous_Equidistant_Conic coordinate system to preserve distance in map visualization.

The station-to-station passenger flow (i.e., OD flow) of the WMATA network is downloaded from PlanItMetro by Washington Metropolitan Area Transit Authority[1]. The most recent publicly available OD dataset is for October 2015. The dataset presents OD flow in one month by each day of week (Monday through Sunday). In other words, for example, the passenger flow on Monday is the average daily trips of all Mondays in October 2015. Shown in Figure 4.3 is the daily average ridership of the nine MTs in the WMATA network from Monday to Sunday in October 2015.

Time periods within a day are not differentiated although the original OD dataset comes with more detailed passenger ridership during morning, midday, afternoon, evening, and late night,

[1] https://planitmetro.com/?s=Metrorail+Ridership+Data+Download%2C+October+2015

Table 4.2 Station ID, name, type, and link attribute of the WMATA network, 2015

Node ID	Station Name	Type	Link Attribute
1	East Falls Church	MT	Orange, Silver
2	Rosslyn	MT	Blue, Orange, Silver
3	Metro Center	MT	Blue, Orange, Red, Silver
4	L'Enfant Plaza	MT	Blue, Green, Orange, Silver, Yellow
5	Stadium-Armory	MT	Blue, Orange, Silver
6	Fort Totten	MT	Green, Red, Yellow
7	Gallery Place/Chinatown	MT	Green, Red, Yellow
8	Pentagon	MT	Blue, Yellow
9	King St-Old Town	MT	Blue, Yellow
10	Ballston-MU	non-MT	Orange, Silver
11	Virginia Square-GMU	non-MT	Orange, Silver
12	Clarendon	non-MT	Orange, Silver
13	Court House	non-MT	Orange, Silver
14	Foggy Bottom-GWU	non-MT	Blue, Orange, Silver
15	Farragut West	non-MT	Blue, Orange, Silver
16	McPherson Square	non-MT	Blue, Orange, Silver
17	Federal Triangle	non-MT	Blue, Orange, Silver
18	Smithsonian	non-MT	Blue, Orange, Silver
19	Federal Center SW	non-MT	Blue, Orange, Silver
20	Capitol South	non-MT	Blue, Orange, Silver
21	Eastern Market	non-MT	Blue, Orange, Silver
22	Potomac Ave	non-MT	Blue, Orange, Silver
23	Benning Rd	non-MT	Blue, Silver
24	Capitol Heights	non-MT	Blue, Silver
25	Addison Rd/Seat Pleasant	non-MT	Blue, Silver
26	Morgan Blvd	non-MT	Blue, Silver
27	College Park-U of Md	non-MT	Green, Yellow
28	Prince George's Plaza	non-MT	Green, Yellow
29	West Hyattsville	non-MT	Green, Yellow
30	Georgia Ave-Petworth	non-MT	Green, Yellow
31	Columbia Heights	non-MT	Green, Yellow
32	U St/African-Amer Civil War Memorial/Cardozo	non-MT	Green, Yellow
33	Shaw-Howard Univ	non-MT	Green, Yellow
34	Mt Vernon Sq/7th St-Convention Center	non-MT	Green, Yellow
35	Archives/Navy Memorial-Penn Quarter	non-MT	Green, Yellow
36	Pentagon City	non-MT	Blue, Yellow

Table 4.2 Continued. Station ID, name, type, and link attribute of the WMATA network, 2015

Node ID	Station Name	Type	Link Attribute
37	Crystal City	non-MT	Blue, Yellow
38	Ronald Reagan Washington National Airport	non-MT	Blue, Yellow
39	Braddock Rd	non-MT	Blue, Yellow
40	Van Dorn St	non-MT	Blue, Yellow
41	Largo Town Center	ET	Blue, Silver
42	Greenbelt	ET	Green, Yellow
43	Franconia-Springfield	ET	Blue, Yellow
44	Vienna/Fairfax-GMU	end station	Orange
45	Wiehle-Reston East	end station	Silver
46	Shady Grove	end station	Red
47	Glenmont	end station	Red
48	New Carrollton	end station	Orange
49	Branch Ave	end station	Green
50	Huntington	end station	Yellow
51	Dunn Loring/Merrifield	intermediate station	Orange
52	West Falls Church/VT/UVA	intermediate station	Orange
53	Spring Hill	intermediate station	Silver
54	Greensboro	intermediate station	Silver
55	Tysons Corner	intermediate station	Silver
56	McLean	intermediate station	Silver
57	Rockville	intermediate station	Red
58	Twinbrook	intermediate station	Red
59	White Flint	intermediate station	Red
60	Grosvenor-Strathmore	intermediate station	Red
61	Medical Center	intermediate station	Red
62	Bethesda	intermediate station	Red
63	Friendship Heights	intermediate station	Red
64	Tenleytown-AU	intermediate station	Red
65	Van Ness-UDC	intermediate station	Red
66	Cleveland Park	intermediate station	Red
67	Woodley Park/Zoo/Adams Morgan	intermediate station	Red
68	Dupont Circle	intermediate station	Red
69	Farragut North	intermediate station	Red
70	Wheaton	intermediate station	Red
71	Forest Glen	intermediate station	Red
72	Silver Spring	intermediate station	Red
73	Takoma	intermediate station	Red
74	Brookland-CUA	intermediate station	Red

Table 4.2 Continued. Station ID, name, type, and link attribute of the WMATA network, 2015

Node ID	Station Name	Type	Link Attribute
75	Rhode Island Ave/Brentwood	intermediate station	Red
76	NoMa-Gallaudet U	intermediate station	Red
77	Union Station	intermediate station	Red
78	Judiciary Square	intermediate station	Red
79	Landover	intermediate station	Orange
80	Cheverly	intermediate station	Orange
81	Deanwood	intermediate station	Orange
82	Minnesota Ave	intermediate station	Orange
83	Suitland	intermediate station	Green
84	Naylor Rd	intermediate station	Green
85	Southern Ave	intermediate station	Green
86	Congress Heights	intermediate station	Green
87	Anacostia	intermediate station	Green
88	Navy Yard-Ballpark	intermediate station	Green
89	Waterfront	intermediate station	Green
90	Eisenhower Ave	intermediate station	Yellow
91	Arlington Cemetery	intermediate station	Blue

because the goal of the case study is to find out the impact of partial node failures for at least 24 hours when heavy maintenance could be carried out. Otherwise, if the maintenance is not heavy and could be done within a few hours, it is most likely that it would take place overnight especially on weekdays when metro rail does not operate for a few hours. In 2015, WMATA was operated from 5am to midnight from Monday to Thursday (five hours out of operation), 5am – 3am on Friday (two hours out of operation), 7am – 3am on Saturday (four hours out of operation), and 7am to midnight on Sunday (seven hours out of operation).

In Figure 4.3, most of the MTs have the smallest ridership on Sunday and Saturday and the largest ridership on Monday, which is in accordance with ridership of the entire network for the seven days of week. Therefore, it might be expected in the simplest situation when ridership is the top concern, most MT maintenance (which only causes partial node failure at MTs) could be carried out on Sunday and Saturday and avoid Monday, in order to minimize their influence on passengers departing from and arriving at the MT. However, in this study, network flow distribution is the top concern, rather than station ridership. The following section will present analysis results where passenger reroute is the key point to be addressed when minimizing the influence of partial node failure on network flow distribution.

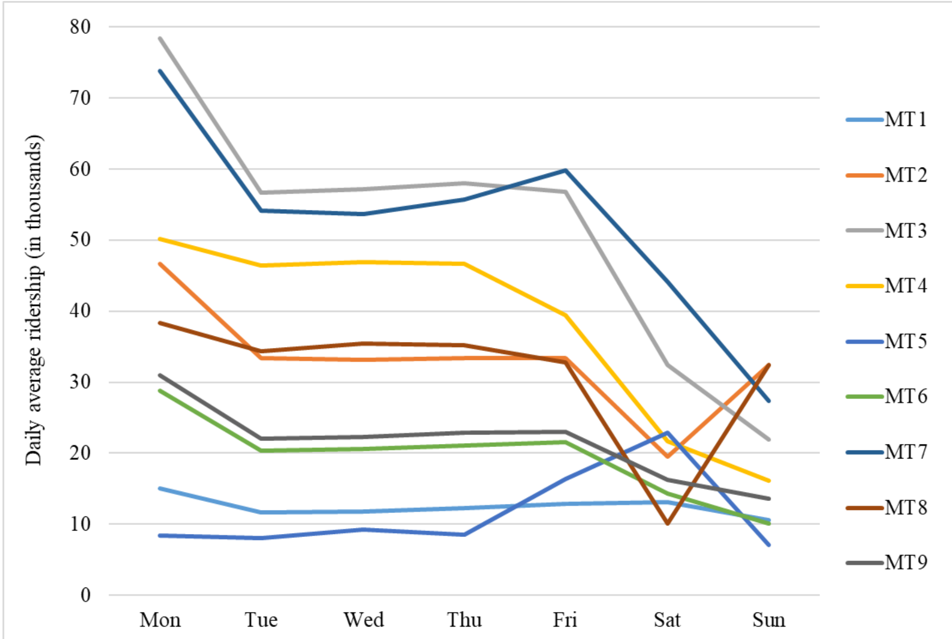


Figure 4.3 Daily average ridership of the nine MTs in October 2015

4.3.2 The impact of partial node failure of MTs in the WMATA network

In the case study of the WMATA network, the analysis focuses on MTs because of their unique roles in transferring passengers within the network. With the SPN^{Pr} model, nodal disruption scenarios of the nine MTs in the WMATA network are conducted. Table 4.3(a) presents the computational results for partial node failure scenarios for the WMATA network. Partial node failure of "None" represents the situation at *status quo*. The Obj. values are total passenger*miles of the entire network assuming all passengers take the shortest distance path from origin to destination. All problems were solved to optimality within 12 seconds using CPLEX 12 on an Intel i-7 processor paired with 16 GB of RAM. The number of iterations and nodes in each solution are reported in Table 4.3(a) as well. Based on the $Cost_{reroute}^{P_r}$ values, the smallest and second smallest values could be identified among partial node failure scenarios of each MT.

For each MT, during the seven days of week, a day is considered as the best day to schedule MT maintenance (which only triggers partial node failure of the MT) if $Cost_{reroute}^{P_r}$ is the smallest; likewise, the alternative day to schedule MT maintenance is the day when $Cost_{reroute}^{P_r}$ is the

Table 4.3 Results and summary of partial node failure scenarios for the WMATA network during the seven days of week in October 2015

(a) Results of partial node failure scenarios

Partial Node Failure	Day	Total Passenger*Miles (Obj. values)	Solution Time (sec.)	Iterations	Nodes	$Cost_{reroute}^{P_r}$ (Passenger *Miles)
None	Monday	8,938,458	9.91	50,742	0	N/A
None	Tuesday	7,011,337	9.95	50,929	0	N/A
None	Wednesday	7,064,503	9.92	50,858	0	N/A
None	Thursday	7,353,972	11.17	51,002	0	N/A
None	Friday	7,408,573	10.13	50,742	0	N/A
None	Saturday	4,816,424	11.36	50,609	0	N/A
None	Sunday	3,526,568	11.16	49,939	0	N/A
MT_1	Monday	8,823,267	9.25	49,767	0	6,974
MT_1	Tuesday	6,918,687	9.41	49,549	0	5,540
MT_1	Wednesday	6,971,922	9.52	49,695	0	5,647
MT_1	Thursday	7,255,287	9.38	49,775	0	5,992
MT_1	Friday	7,304,426	9.58	49,789	0	6,702
MT_1	Saturday	4,703,423	9.25	49,399	0	3,878
MT_1	Sunday	3,436,934	9.58	48,607	0	2,051
MT_2	Monday	8,725,023	9.00	44,382	0	42,425
MT_2	Tuesday	6,856,648	9.08	44,289	0	34,840
MT_2	Wednesday	6,910,940	9.16	44,192	0	35,595
MT_2	Thursday	7,203,264	9.03	44,159	0	38,893
MT_2	Friday	7,262,983	9.30	44,400	0	40,325
MT_2	Saturday	4,730,392	9.06	43,911	0	25,003
MT_2	Sunday	3,356,169	9.27	43,394	0	32,273
MT_3	Monday	8,609,317	9.00	43,811	0	180,933
MT_3	Tuesday	6,785,886	9.00	43,759	0	146,572
MT_3	Wednesday	6,841,368	8.97	44,161	0	149,835
MT_3	Thursday	7,136,343	8.91	44,019	0	160,362
MT_3	Friday	7,218,614	9.78	43,665	0	171,308
MT_3	Saturday	4,729,130	8.89	43,350	0	121,566
MT_3	Sunday	3,472,257	9.13	43,013	0	82,516
MT_4	Monday	8,739,334	8.38	32,916	0	114,487
MT_4	Tuesday	6,815,577	8.45	32,954	0	96,767
MT_4	Wednesday	6,869,544	8.48	32,867	0	99,314
MT_4	Thursday	7,168,453	8.42	33,004	0	107,243
MT_4	Friday	7,286,959	9.16	33,049	0	116,125
MT_4	Saturday	4,778,639	8.41	32,736	0	90,138
MT_4	Sunday	3,491,884	8.53	32,329	0	60,286

Table 4.3(a) Continued. Results of partial node failure scenarios

Partial Node Failure	Day	Total Passenger*Miles (Obj. values)	Solution Time (sec.)	Iterations	Nodes	$Cost_{reroute}^{P_r}$ (Passenger *Miles)
MT_5	Monday	8,890,791	9.66	49,911	0	1,891
MT_5	Tuesday	6,963,518	9.34	49,716	0	1,550
MT_5	Wednesday	7,005,761	9.72	49,863	0	1,466
MT_5	Thursday	7,302,765	9.84	49,736	0	1,771
MT_5	Friday	7,287,289	10.22	49,767	0	1,960
MT_5	Saturday	4,662,281	9.64	49,322	0	1,583
MT_5	Sunday	3,473,958	9.45	48,970	0	1,244
MT_6	Monday	8,848,170	6.39	27,903	0	94,301
MT_6	Tuesday	6,944,764	6.42	27,894	0	70,124
MT_6	Wednesday	6,995,616	6.41	27,863	0	70,229
MT_6	Thursday	7,282,473	6.59	27,949	0	72,870
MT_6	Friday	7,339,107	6.95	27,960	0	78,751
MT_6	Saturday	4,776,994	6.34	27,636	0	60,194
MT_6	Sunday	3,495,971	6.48	27,255	0	39,606
MT_7	Monday	8,607,383	8.77	47,265	0	91,978
MT_7	Tuesday	6,760,527	9.00	47,257	0	73,410
MT_7	Wednesday	6,818,075	9.38	47,220	0	74,345
MT_7	Thursday	7,100,146	9.34	47,179	0	77,862
MT_7	Friday	7,149,937	10.25	47,342	0	90,140
MT_7	Saturday	4,613,313	9.36	46,796	0	65,947
MT_7	Sunday	3,406,132	8.88	46,421	0	38,637
MT_8	Monday	8,724,544	9.61	48,469	0	1,206
MT_8	Tuesday	6,813,712	9.44	48,547	0	805
MT_8	Wednesday	6,862,968	9.37	48,444	0	757
MT_8	Thursday	7,151,361	9.36	48,610	0	1,057
MT_8	Friday	7,219,607	11.25	48,662	0	982
MT_8	Saturday	4,763,017	9.61	48,255	0	1,320
MT_8	Sunday	3,309,214	9.59	47,575	0	1,628
MT_9	Monday	8,662,722	9.38	49,706	0	715
MT_9	Tuesday	6,804,826	10.03	49,959	0	594
MT_9	Wednesday	6,854,973	9.52	49,627	0	611
MT_9	Thursday	7,136,922	9.72	49,809	0	612
MT_9	Friday	7,190,445	10.17	49,785	0	671
MT_9	Saturday	4,663,032	9.66	49,589	0	413
MT_9	Sunday	3,408,351	9.50	48,999	0	287

Table 4.3(b) Summary of partial node failure scenarios

	MT	Day	$Cost_{reroute}^{P_r}$	
Best Day for MT Maintenance	MT_3	Sunday	82,516	
	MT_4	Sunday	60,286	
	MT_6	Sunday	39,606	
	MT_7	Sunday	38,637	
	MT_2	Saturday	25,003	
	MT_1	Sunday	2,051	
	MT_5	Sunday	1,244	
	MT_8	Wednesday	757	
	MT_9	Sunday	287	
				Percentage Increase compared with Best Day
Alternative Day for MT Maintenance	MT_3	Saturday	121,566	47.3%
	MT_4	Saturday	90,138	49.5%
	MT_7	Saturday	65,947	70.7%
	MT_6	Saturday	60,194	52.0%
	MT_2	Sunday	32,273	29.1%
	MT_1	Saturday	3,878	89.1%
	MT_5	Wednesday	1,466	17.9%
	MT_8	Tuesday	805	6.4%
	MT_9	Saturday	413	43.9%

second smallest. This assumption of "best day for MT maintenance" concerns about passengers whose origin or destination are not the partially failed MT, they would complete their trips via reroute path in the enclosed WMATA network without need of other transportation modes. Table 4.3(b) summarizes, on which day of the week, a partial node failure of a MT causes the smallest and second smallest flow reroute cost in October 2015. It is found that, Sunday is the best day for maintenance of most MTs except MT_2 and MT_8. For MT_2, Saturday is the best day for maintenance, and the alternative day is Sunday. For MT_8, Wednesday is the best day for maintenance, and the alternative day is Tuesday. For all the other MTs, the alternative day is mostly Saturday, except Wednesday for MT_5.

The MTs in Table 4.3(b) are arranged by descending order of $Cost_{reroute}^{P_r}$, and thus reflect criticality ranking of MTs. Rankings in both the best day and alternative day for MT maintenance agree that, MT_3 is the most critical MT in the WMATA network since its partial failure triggers the largest flow reroute cost, followed by MT_4. Most of the rankings in the best day are in accordance with those in alternative days, except MT_6 and MT_7 although their flow reroute cost is very close to each other. However, when comparing the percentage increase of flow reroute cost on the best day and alternative day, the percentage increase ranges widely from 6.1% to 89.1%, meaning the difference varies greatly of conducting MT maintenance on the best day or alternative day. For MT_8, maintenance can be conducted either on Wednesday or Tuesday since the difference in flow reroute cost is minor. From another perspective, if MT_8 has to close for two consecutive days for maintenance, Tuesday and Wednesday are the best choices to minimize flow reroute cost. For MT_1, Sunday is substantially better than Saturday for maintenance since flow reroute cost is almost doubled on Saturday compared with Sunday.

It is interesting to take a closer look at partial node failure of MT_8, which is a good example to show that ridership does not always seem to dominate the results of $Cost_{reroute}^{P_r}$. It is noteworthy that, although ridership of the entire network excluding MT_8 is the smallest on Sunday and the second smallest on Saturday, partial node failure of MT_8 yields an additional network cost of 1,628 and 1,320 passenger*miles on Sunday and Saturday respectively, ranking the largest and second largest $Cost_{reroute}^{P_8}$ among the seven days of week. To explain the 'unusual' result of MT_8 in more details, it is mainly caused by larger passenger flow on weekends than weekdays between an intermediate station #91 and the stations to the east of MT_4 on orange, silver, and blue lines, where MT_8 is on the shortest path. After MT_8 partially fails, these passengers would have to reroute through the path between MT_4 and MT_2 along orange, silver, and blue lines, increasing their travel distance by 1.17 miles.

The next section illustrates and visualizes passengers reroute when MT maintenance takes place on the "best day" where $Cost_{reroute}^{P_r}$ is the smallest. Presented in each "best day" scenario are stations that would expect more passengers and the number of passengers they would expect.

4.3.3 Impacted stations with increased ViaFlow

The information of increased ViaFlow is important in examining the redistribution of network flow and evaluating the level of traffic load on each station due to flow reroutes caused by a partially failed MT. With this information, decision makers and management staff can prepare for the flow increase at specific station(s). In each partial node failure scenario, the difference of ViaFlow ($DiffViaFlow_k^{Pr\text{-}sq}$) at each node could be obtained by subtracting $ViaFlow_k^{sq}$ from

$ViaFlow_k^{Pr}$. If the resulted difference is a positive value, it means node *k* has more ViaFlow (i.e., increased number of passengers pass through node *k*) when MT_r partially fails than at *status quo*.

Among the nine MT partial failure scenarios on their best day for maintenance, stations are only impacted with increased ViaFlow in four scenarios – the partial node failure of MTs #3, #4, #6, and #7. For the other five scenarios, all stations experience decreased or unchanged ViaFlow. It is noteworthy that, decreased or unchanged ViaFlow at node *k* does not mean no passengers reroute via node *k*; instead, it means maybe some passengers reroute via node *k*, but the number of these reroute passenger is offset by the decrease of ViaFlow at node *k* caused by a partially failed MT.

Table 4.4 presents the stations with increased ViaFlow. For example, on Sunday, 153,684 passengers pass through node #4 at status quo, and the number of ViaFlow is raised to 195,284 when MT_3 partially fails. Therefore, decision makers and management staff could prepare for ViaFlow increase at station #4, knowing that 41,600 more passengers would pass through station #4 than usual due to the partial failure of MT_3. By comparing the values in each row of Table 4.4, it is found in some scenarios that, although the increase of ViaFlow is not substantial, a node could still experience a high volume of ViaFlow. For example, on Sunday, due to the partial failure of MT_6, only 2,132 more passengers pass through station #7; however, station #7 has the largest ViaFlow either at *status quo* or partial node failure scenario compared with stations #74 to #78.

Figure 4.4 visualizes the impacted stations with increased ViaFlow, presented as the red nodes with white labels in the maps. The figures clearly show the locations of impacted stations, and how much they are impacted. When a MT partially fails, not only some of the directly connected stations experience ViaFlow increase, but also some of the indirectly connected stations (even some far away stations). For example, in Figure 4.4(b), when MT_4 partially fails, its directly connected non-MT_{35} hosts 35,866 more ViaFlow than usual, some nearby indirectly connected stations (i.e., MT_3, MT_7, non-MT_{17}) hosts 5,300 to 47,405 more ViaFlow, even a faraway station MT_2 experiences 32 more passengers although the increase is minor compared to others.

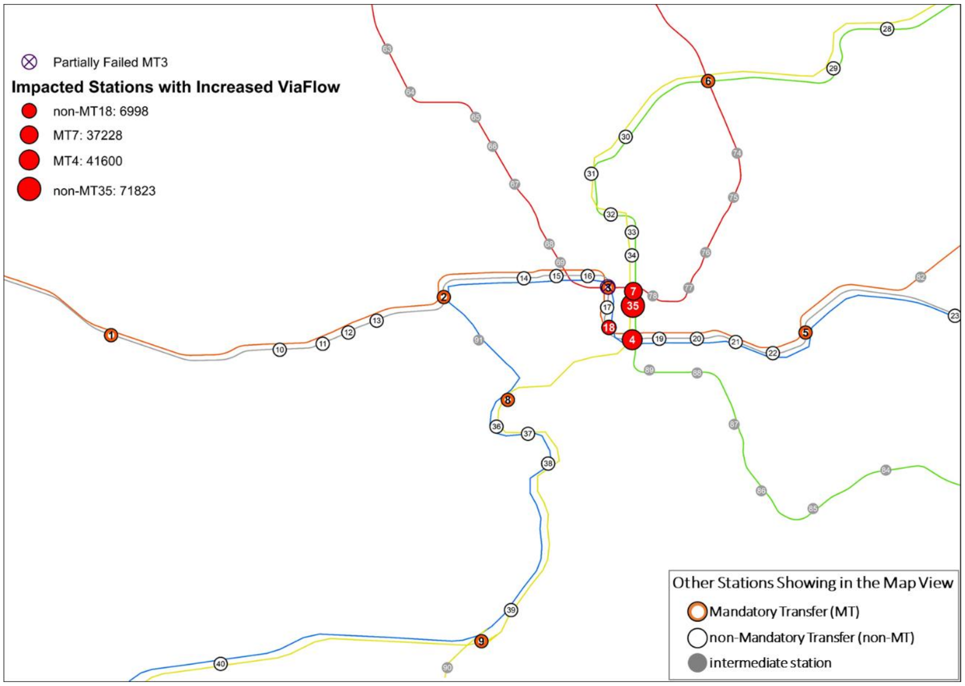
4.3.4 Variation of criticality rankings using flows in different temporal scales

Without flows, a network is nothing but a static topology with fixed nodes and links. Freight and passengers bring a network to life with flow exchange among fixed node locations through links in between. Majority of this study has focused on analysis using OD flow on different days. In

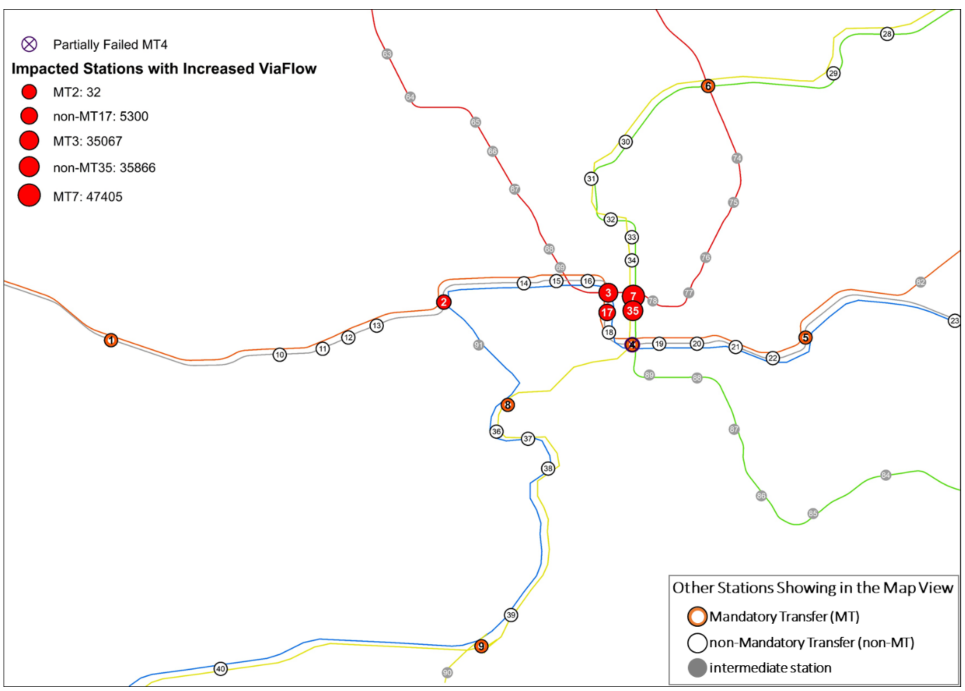
Table 4.4 MTs with increased ViaFlow on their best day for maintenance of the WMATA network

	Node ID	$ViaFlow_k^{sq}$	$ViaFlow_k^{Pr}$	$DiffViaFlow_k^{Pr\text{-}sq}$ **(with Increased ViaFlow)**
MT_3	4	153,684	195,284	+ 41,600
Partially	7	117,004	154,232	+ 37,228
Fails	18	65,233	72,231	+ 6,998
(Sunday)	35	60,854	132,677	+ 71,823
MT_4 Partially Fails (Sunday)	2	77,620	77,652	+ 32
	3	148,726	183,793	+ 35,067
	7	117,004	164,409	+ 47,405
	17	77,820	83,120	+ 5,300
	35	60,854	96,720	+ 35,866
MT_6 Partially Fails (Sunday)	7	117,004	119,136	+ 2,132
	74	6,823	29,862	+ 23,039
	75	12,246	37,155	+ 24,909
	76	15,921	42,084	+ 26,163
	77	18,979	46,736	+ 27,757
	78	35,899	64,652	+ 28,753
MT_7 Partially Fails (Sunday)	3	148,726	151,928	+ 3,202
	4	153,684	162,970	+ 9,286
	17	77,820	104,825	+ 27,005
	18	65,233	93,401	+ 28,168
	74	6,823	27,978	+ 21,155
	75	12,246	32,512	+ 20,266
	76	15,921	35,413	+ 19,492
	77	18,979	37,535	+ 18,556
	78	35,899	53,478	+ 17,579

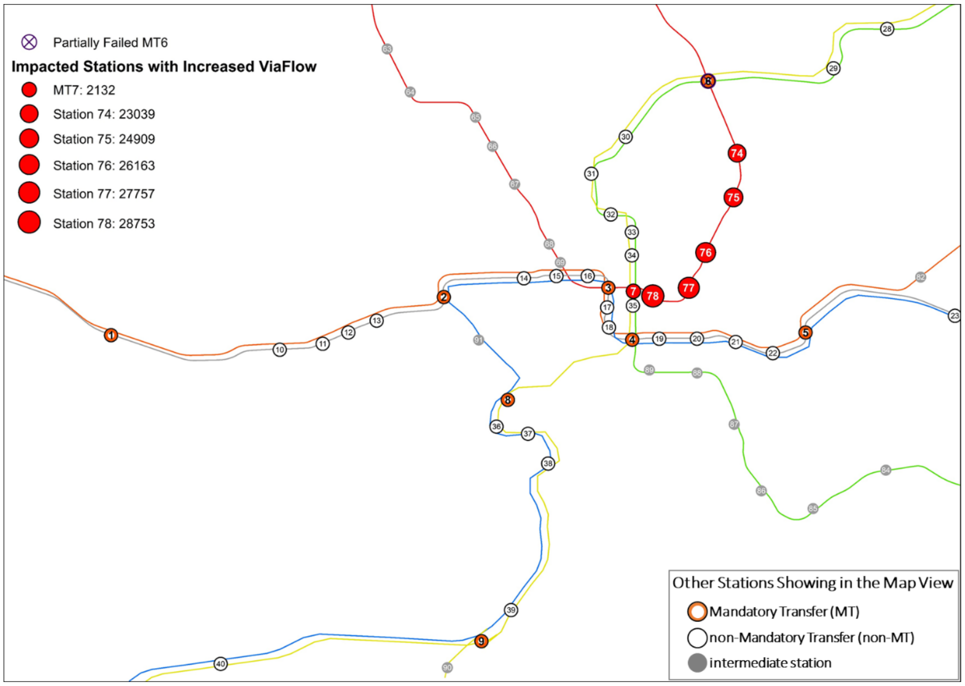
Figure 4.4 Impacted stations with increased ViaFlow on best day for maintenance of MT_3, MT_4, MT_6, and MT_7 of the WMATA network in October 2015



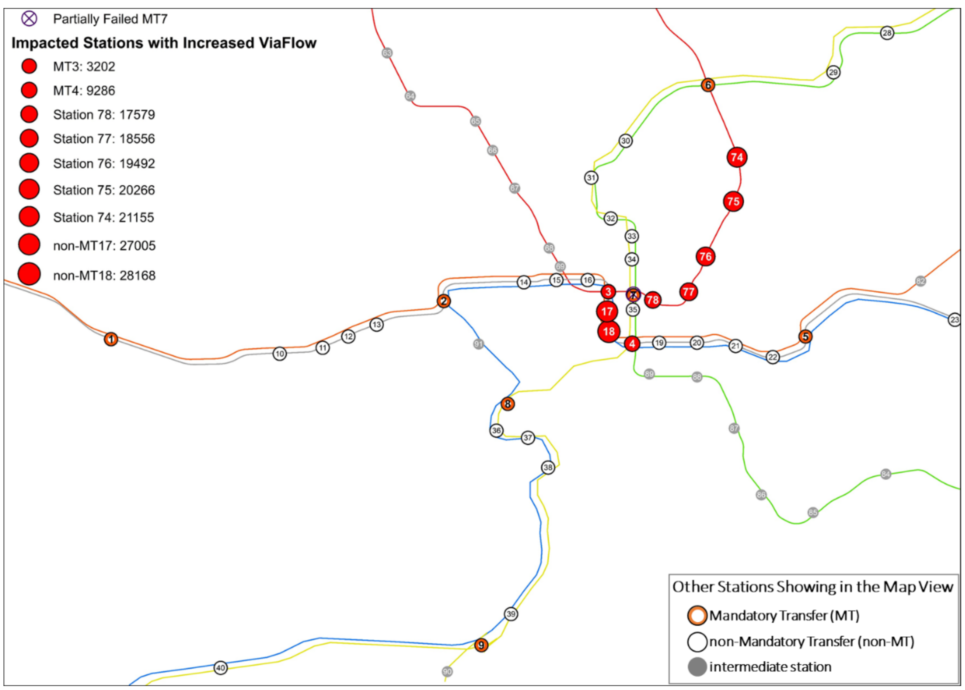
(a) partial failure of MT_3 (Sunday)



(b) partial failure of MT_4 (Sunday)



(c) partial failure of MT_6 (Sunday)



(d) partial failure of MT_7 (Sunday)

Figure 4.4 Continued. Impacted stations with increased ViaFlow on best day for maintenance of MT_3, MT_4, MT_6, and MT_7 of the WMATA network in October 2015 this section, the OD flow changes to monthly total flows of October 2015, in order to further examine the impact of different flows on network disruption cost and station criticality.

Table 4.5(a) presents criticality rankings of MTs in the WMATA by flow reroute cost. The MTs are arranged in an order based on station criticality ranked by the results of partial MT failures using monthly flow. Criticality rankings of MTs calculated by daily flows are provided as well for comparison. For MT_3, MT_4, MT_2, MT_1, and MT_9, their criticality ranking stays the same across various temporal scales. For the other four MTs (the rows highlighted in grey color in Table 4.5-a and -b), criticality rankings are not consistent across different time periods. The values in red color highlight the discrepancy. MT_7 is mostly ranked as the 3^{rd} place, but on Monday and Sunday, its ranking changes to the 4^{th} place; MT_6 is mostly ranked as the 4^{th} place, but on Monday and Sunday, its ranking changes to the 3^{rd} place; MT_5 and MT_8 are mostly ranked as the 7^{th} and 8^{th} place, respectively, but on Sunday, their ranking switches with each other. Table 4.5(b) provides the corresponding flow reroute cost in passenger*miles caused by partially failed MTs in the WMATA across different temporal scales. On Monday and Sunday, flow reroute cost of MT_6 becomes 2.5% greater than that of MT_7. On Sunday, flow reroute cost of MT_8 exceeds that of MT_5 by 30.9%.

4.4 Concluding Remarks

This chapter implemented partial node failure in the SPN model. The model and indicator equations were firstly tested using a small sample network with five nodes and three different lines to demonstrate the analysis procedure of the SPN model for partial node failure (SPN^{Pr}).

Then the model and indicators were applied into a real network – the WMATA network – to show their capability in solving real-world network problems. The case study focused on partial node failure scenarios of the nine MTs in the WMATA network. The distance matrix was the station-to-station distance in miles. The flow matrix was the station-to-station passenger flow in October 2015, which provided the average daily trips of the seven days in a week from Monday to Sunday, and total monthly trips. Altogether 80 scenarios were carried out to optimize network flow distribution based on shortest paths, including 8 scenarios at *status quo* and 72 scenarios for partial node failure of the nine MTs on Monday to Sunday and monthly sum. The influence of partial node failure on the entire network was assessed by the additional network cost due to flow

Table 4.5 Criticality rankings of MTs in the WMATA in different time periods

(a) Criticality rankings of MTs by flow reroute cost

MT	Month	Mon.	Tues.	Weds.	Thurs.	Fri.	Sat.	Sun.
MT_3	1	1	1	1	1	1	1	1
MT_4	2	2	2	2	2	2	2	2
MT_7	3	4	3	3	3	3	3	4
MT_6	4	3	4	4	4	4	4	3
MT_2	5	5	5	5	5	5	5	5
MT_1	6	6	6	6	6	6	6	6
MT_5	7	7	7	7	7	7	7	8
MT_8	8	8	8	8	8	8	8	7
MT_9	9	9	9	9	9	9	9	9

Note: Rows in grey color indicate where criticality rankings are not consistent across different time periods; values in red color highlight the discrepancy.

(b) Flow reroute cost (passenger*miles) caused by partially failed MTs

MT	Month	Mon.	Tues.	Weds.	Thurs.	Fri.	Sat.	Sun.
MT_3	4,505,603	180,933	146,572	149,835	160,362	171,308	121,566	82,516
MT_4	3,050,952	114,487	96,767	99,314	107,243	116,125	90,138	60,286
MT_7	2,283,222	91,978	73,410	74,345	77,862	90,140	65,947	38,637
MT_6	2,156,123	94,301	70,124	70,229	72,870	78,751	60,194	39,606
MT_2	1,101,630	42,425	34,840	35,595	38,893	40,325	25,003	32,273
MT_1	163,714	6,974	5,540	5,647	5,992	6,702	3,878	2,051
MT_5	51,178	1,891	1,550	1,466	1,771	1,960	1,583	1,244
MT_8	34,383	1,206	805	757	1,057	982	1,320	1,628
MT_9	17,312	715	594	611	612	671	413	287

Note: Rows in grey color indicate where criticality rankings are not consistent across different time periods; values in red color highlight the discrepancy.

reroute, regardless of the unrealized flow on the partially failed node. By comparing and summarizing computational results from the case study, the best day and alternative day for maintenance were identified for each of the nine MTs in October 2015, where the total additional network cost caused by passengers reroutes was the minimum or second smallest during the seven days of week. Four best-day situations with increased ViaFlow were mapped out to visualize which stations would expect higher pass-through passenger volume when a specific MT partially failed, and how many more passengers they could expect exactly. The visualizations also helped in understanding how a partial node failure impacted passenger flow on its directly and indirectly connected nodes. Last but not least, station criticality was ranked in eight different temporal units, showing that different flow demand in different time period impact the criticality of the nine MTs.

Major contributions of this chapter are threefold. First, the study extends the SPN model by adding two elements for partial node failure: distance update constraints (d'_{ij}) and link attribute index (l_{ij}). To implement partial node failure in the SPN model, the distance matrix needs to be updated for each scenario; and link attribute index is necessary for computing distance update. Second, two indicators are constructed to assess the impact of partial node failures on network flow distribution based on shortest paths – the additional network cost caused by flow reroute ($Cost^{P_r}_{reroute}$) and the change in flow passing through each node ($DiffViaFlow_k^{Pr\text{-}sq}$). Third, from the practical perspective, the assessment procedure in this study demonstrates a way to identify proper day for scheduling heavy maintenance or major upgrade of a rail station in terms of least additional network cost caused by passengers reroutes, and to predict the change in pass-through traffic at every station, if OD flow can be well estimated in a future time period.

CHAPTER 5

CONCLUSIONS AND DISCUSSIONS

5.1 Summary of Dissertation Research

Network vulnerability and nodal criticality has been of research focus in network modelling and analysis. The malfunction of network components would impact the operation of network to various extent. In network modelling, all-or-nothing scenarios are frequently simulated to assess how the removal of nodes or links would impact network operation. However, in the real world, partial failures usually take place more often than complete node failures, since partial failures could happen occasionally in daily operation to serve administrative or maintenance purposes. Although the influence of partial failures might not be as severe as complete failures, the overall impact of partial failures could cause significant loss from time to time, and thus the study of partial failures in network modelling should be paid the same attention to if not more than rare extreme cases such as terrorist attack or natural disasters.

This dissertation addresses the network vulnerability and nodal criticality from the perspectives of nodal connectivity, partial node failure, and shortest path network problems. Case studies are carried out using major U.S. heavy rail systems. The dissertation starts from the improvement of nodal degree, which is a traditional and frequently used measure in network connectivity. A comprehensive Degree of Nodal Connection (DNC) index is developed to depict and compare station criticality in major U.S. heavy rail networks. Four local and global indexes are constructed to evaluate the impact of nodal disruption. Most importantly, network disruption is distinguished into two categories: partial node failure, and complete node failure. Results are analyzed and compared between the two types of node failure. Then a third network component – network flow – is considered in addition to nodes and links, which is embedded into shortest path problem to evaluate network vulnerability and nodal criticality. Several models and constraints are developed to assess network vulnerability and nodal criticality in flow-based networks using shortest paths in both complete and partial node failure scenarios.

Specifically, the research re-classifies transfer stations in rail networks into three types – Mandatory Transfer (MT), non-Mandatory Transfer (non-MT), and End Transfer (ET); proposes two new measures of Link Degree (LD) and Total Link Degree (TLD), and a new index of DNC index to comprehensively assess nodal criticality in rail networks especially when multi-lines present. The LD and TLD, along with the traditional nodal degree and Derrible's multiple edges, are applicable only to undirected networks. And thus, the DNC index is applicable only to undirected networks. The measures and index can be used in both planar and non-planar networks.

LD, TLD, and DNC are scalable across different networks. The ranges of DNC values can be adjusted by modifying the scaling factor in the DNC equation.

A major theme of this dissertation presents the concept and definition for partial node failure, in contrast of complete node failure. Numerical results are provided to show the impact of partial node failures for major U.S. heavy rail networks; results of complete node failures are analyzed as well in order to make comparison with partial node failures. Network vulnerability of the eleven U.S. heavy rail networks is measured by four indexed – two DNC-loss indexes at local level and two DNC-resilience indexes at global level, where worst-case damage and average damage are both evaluated for nodal disruption scenarios. Various results regarding partial node failures are specifically useful for policy making and operation management, since partial node failures are more realistic in real world situations despite all-or-nothing scenarios.

The Shortest Path Network (SPN) model is developed and formulated in the form of mixed-integer programming to identify the best flow distribution across the entire network based on shortest paths. The Assessing Nodal Disruption in SPN (AND-SPN) model is constructed to evaluate the impact of complete node failure on the entire network, and to rank nodal criticality accordingly. The SPN and the AND-SPN are applicable to both directed and undirected networks, planar and non-planar networks. However, negative arc is not allowed in solving the SPN and the AND-SPN. The two models can be expanded with fixed or variable flow capacity constraints when arc capacity needs to be restricted. In the SPN and the AND-SPN, along with objective functions, three supplementary criticality metrics are provided to present more comprehensive numerical results, which are average arc use (AAU), average arc cost (AAC), and average arc flow (AAF). Computational experiments by mixed-integer programming via CPLEX are carried out on a small toy network first, and then applied to the Amtrak rail system with selected 62 nodes set. The selected 62 nodes are ranked according to the obtained objective values from the AND-SPN. The numerical experiment also examines the effect of capacity restriction and visualizes its influence on flow distribution.

The SPN model is further expanded to the SPN^{Pr} to accommodate partial node failure, through incorporating distance update constraints and link attribute index. Flow reroute cost ($Cost_{reroute}^{P_r}$) and pass-through flows (ViaFlow) are constructed to evaluate the impact of partial node failures on network operation, which are then applied to the WMATA network to assess the impact of partial node failures of the nine MTs in the system. The numerical experiment identifies the impacted stations with increased ViaFlow caused by partially failed MTs, and shows the influence of different flow demand on the variation in criticality rankings of the nine MTs.

Above all, the dissertation [1] improves existing nodal degree indexes and develops easy-to-use index-based metrics to evaluate comprehensive nodal criticality and network vulnerability; [2] distinguishes partial node failure instead of all-or-nothing scenarios, and addresses the importance and usefulness of disruption results in partial node failures in decision making and operation management; [3] develops the SPN model to find the best flow distribution in a network with no negative arcs, and expands it to the AND-SPN model for assessing network vulnerability in the SPN; [4] develops the SPN^{Pr} model by involving distance update constraints and link attribute index, so that not only complete node failure, but also partial node failure can be modeled in shortest path networks. The index-based metrics are not only applicable to rail networks, but also applicable to all other undirected networks. The optimization-based models are not only applicable to transportation networks, but also other networks where shortest paths are desired as long as no negative arc presents.

5.2 Discussions and Future Research Directions

Some limitations of this dissertation rest in three aspects. The first aspect is data collecting and preprocessing. For the eleven heavy rail networks, we are not able to find the exact physical lane information, and thus the topology of the eleven heavy rail networks are solely based on their operating line graph. Nevertheless, the impact is double-sided. On one hand, since the topology of the elven networks do not represent physical rail lanes, the usefulness of the numerical results for station criticality and network vulnerability are uncertain. On the other hand, the uncertainty protects rail transport security, since the numerical results will be not suitable to release to public if the topology of operating rail lines exactly match that of the physical rail lanes. With respect to the Amtrak dataset, we do not have the actual OD flow among stations, and passenger flows between origin to destination are estimated through OD matrix given the ridership at each station. How much the estimation resembles the reality could influence optimization results of the SPN and the AND-SPN. In the field of transportation engineering, flow estimation is an important topic and could be sophisticated, which is not of concern for the dissertation though.

The second aspect of limitation rest in the index-based metrics. The LD, TLD, and DNC only work on undirected networks, therefore the eleven U.S. heavy rail networks have to be generalized and treated as undirected networks. This is not an issue for some of the networks where all arcs are bi-directional. For some other networks with one-way arcs, the index-based metrics are not fully suitable for application. Especially when one-way arcs account for a significant proportion of the entire network, the usefulness of the metrics degrades to a large extent. Future research could focus on solving this problem and developing more delicate metrics for directed networks.

The third limitation is in the mixed-integer programming used in optimizing the SPN and the AND-SPN. To ensure optimality, we used mixed-integer programming to solve the SPN and the AND-SPN with CPLEX software. However, there is usually a trade-off between optimality and solution time. When the node set exceed about 100 nodes, mixed-integer programming in CPLEX cannot guarantee feasible solution within feasible time period. This is another main reason that we have to select a dataset consisting of 62 important nodes from over 500 Amtrak stations instead of using the entire node set. Future research could focus on developing algorithms to solve the SPN and the AND-SPN for larger dataset, such as the entire Amtrak dataset with over 500 stations. One possible direction is heuristic algorithm. Although the optimality could not be guaranteed in heuristic methods, the near-optimal solution could still be achieved in a short amount of time so that a good solution could still be feasible for large networks.

Model construction and numerical analysis are improved in Chapter 4 compared with Chapter 3. On one hand, in terms of the models and indicators, all the values obtained from optimization and mathematical equations have true physical meanings in Chapter 4. For example, the value of objective function (Ω^{Pr}) from the SPN model for partial node failure represents total network cost, and its unit is passenger*miles (if given the unit of travel cost is miles and the unit of flow is number of passengers); the value of $Cost_{reroute}^{P_r}$ represents additional network cost caused by flow reroutes when node r partially fails, with the same unit of passenger*miles; the value of $DiffViaFlow_k^{Pr\text{-}sq}$ represents the difference of ViaFlow at node k when node r partially fails compared with that at *status quo*, and its unit is passengers. In contrast, among the optimization models and indicator equations from Chapter 3, the value of objective function (Ω^r) of the AND-SPN model does not represent true value with physical meaning, due to the fact that M (the large number representing the disruption cost) is added as penalty in the model. As a result, the rankings of the obtained objective values (Ω^r) are viable, rather than the numerical values that do not come with physical unit. On the other hand, in terms of the dataset used in case study, the accuracy of the obtained numerical results is improved substantially for assessing partial node failures in the WMATA network, because real OD flow dataset is adopted in Chapter 4. The real OD flow dataset of the WMATA network contains the number of passengers from every station to every other station in October 2015. In contrast, the OD flow of Amtrak network in Chapter 3 is estimated from ridership at each station.

For future research, the study in Chapter 4 can be extended and further developed in several directions. A first possible direction is to analyze the re-assignment of unrealized flows at the partially failed node, which is ignored intentionally in Chapter 4. In other words, the question is how to allocate the unrealized travel demand at a partially failed station, assuming passengers would take the metro rail to reach their destinations anyway. Take the partial node failure of a MT in the WMATA network for example, if the partially failed MT is not their origin or destination, passengers can complete their trip – at most via longer reroute path – since none of the partially failed MTs would separate the WMATA network into sub-networks. However, if the partially failed MT is their origin, passengers might walk or bike or take bus to a nearby alternative station to get on the train; if the partially failed MT is their destination, passengers might choose to get off the train at a nearby alternative station and walk or bike or take a bus to their destination. Here, the alternative stations are defined as the stations directly connect with the partially failed MT. Refer to Figure 4.2 for the complete structure of the WMATA network in 2015, if MT_1 partially fails, passengers have three alternative stations, which are stations #10, #52, and #56. Then we need to know (or simulate) which station passengers will choose from the three alternatives, and how much additional cost will be triggered for the passengers due to the unavailability of the partially failed MT. Even further, a second possible direction is to analyze the re-assignment of unrealized flows by incorporating networks of other transportation modes such as bus routes, and to find if there is any better alternative plan to reroute passengers so that the additional network cost is minimum.

LIST OF REFERENCES

Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: *Network flows: theory, algorithms, and applications*. Prentice Hall, New Jersey (1993)

Akgun, I., Kandakoglu, A., Ozok, A.F.: Fuzzy integrated vulnerability assessment model for critical facilities in combating the terrorism. *Expert Systems with Applications.* 37, 3561-3573 (2010)

Albert, R., Jeong, H., Barabasi, A.L.: Error and attack tolerance of complex networks. *Nature*. 406, 378-382 (2000)

Amtrak Fact Sheets 2013 https://www.amtrak.com/state-fact-sheets

Angeloudis, P., Fisk, D.: Large subway systems as complex networks. *Physica A: Statistical Mechanics and its Applications*. 367, 553-558 (2006)

Arulselvan, A., Commander, C.W., Elefteriadou, L., Pardalos, P.M.: Detecting critical nodes in sparse graphs. *Comput. Oper. Res.* 36(7), 2193-2200 (2009)

Bellman, R., Kalaba, R.: On kth best policies. *J. Soc. Indust. Appl. Math*. 8(4), 582-588 (1960)

Bellman, R.: On a routing problem. *Quarterly of Applied Mathematics.* 16, 87-90 (1958)

Bock, F., Kantner, H., Haynes, J.: *An algorithm (the rh best path algorithm) for finding and ranking paths through a network*. Research Report, Armour Research Foundation, Chicago, Illinois (1957)

Bruun, E.C.: *Better public transit systems: analyzing investments and performance (second edition)*. Routledge, London (2014)

Burgholzer, W., Bauer, G., Posset, M., Jammernegg, W.: Analysing the impact of disruptions in intermodal transport networks: a micro simulation-based model. *Decision Support Systems.* 54, 1580-1586 (2013)

Carreras, B.A., Lynch, V.E., Dobson, I., Newman, D.E.: Critical points and transitions in an electric power transmission model for cascading failure blackouts. *Chaos.* 12(4), 985-994 (2002)

Caschili, S., Reggiani, A., Medda, F.: Resilience and vulnerability of spatial economic networks. *Networks and Spatial Economics.* 15, 205-210 (2015)

Cats, O., Jenelius, E.: Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability. *Transportmetrica B: Transport Dynamics*. DOI: 10.1080/21680566.2016.1267596 (2016)

Chiou, S.W.: Vulnerability analysis of a signal-controlled road network for equilibrium flow. In: Márquez, F.P.G., Lev, B. (eds.) *Advanced Business Analytics*, pp. 109-142. Springer International Publishing, Switzerland (2015)

Church, R.L., Scaparra, M.P., Middleton, R.S.: Identifying critical infrastructure: the median and covering facility interdiction problems. *Annals of the Association of American Geographers.* 94(3), 491-502 (2004)

Clarke, S., Krikorian, A., Rausan, J.: Computing the n best loopless paths in a network. *J. of SIAM*. 11(4),1096-1102 (1963)

Corley, H.W., Sha, D.Y.: Most vital links and nodes in weighted networks. *Operations Research Letters.* 1, 157-160 (1982)

Cormican, K.J., Morton, D.P., Wood, R.K.: Stochastic network interdiction. *Operations Research.* 46(2), 184-197 (1998)

D'Este, G.M., Taylor, M.A.P.: Network vulnerability: an approach to reliability analysis at the level of national strategic transport networks. In: Iida, Y., Bell, M.G.H. (eds.) *The Network Reliability of Transport*, pp. 23-44. Pergamon, New York (2003)

D'lima, M., Medda, F.: A new measure of resilience: an application to the London Underground. *Transportation Research Part A*. 81, 35-46 (2015)

Dantzig, G.B.: On the shortest route through a network. *Management Science*. 6(2), 187-190 (1960)

Demsar, U., Spatenkova, O., Virrantaus, K.: Identifying critical locations in a spatial network with graph theory. *Transactions in GIS*. 12, 61-82 (2008)

Derrible, S., Kennedy, C.: Characterizing metro networks: state, form, and structure. *Transportation*. 37(2), 275-297 (2010a)

Derrible, S., Kennedy, C.: The complexity and robustness of metro networks. *Physica A: Statistical Mechanics and Its Applications*. 389(17), 3678-91 (2010b)

Dijkstra, E.W.: A note on two problems in connexion with graphs. *Numerische Mathematik.* 1(1), 269-271 (1959)

El-Rashidy, R.A., Grant-Muller, S.M.: An assessment method for highway network vulnerability. *Journal of Transport Geography*. 34, 34-43 (2014)

Floyd, R.W.: Algorithm 97: Shortest path. *Communications of the ACM*. 5(6), 345 (1962)

Ford, Jr., Lester, R.: *Network flow theory*. Paper P-923. RAND Corporation, Santa Monica, California (1956)

Fulkerson, D.R., Harding, G.C.: Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming.* 13, 116-118 (1977)

Galindo, G., Batta, R.: *Prepositioning of supply points in preparation for foreseeable disasters*. INFORMS, Austin, TX (2010)

Grubesic, T.H., O'Kelly, M.E., Murray, A.T.: A geographic perspective on commercial Internet survivability. *Telematics and Informatics*. 20(1), 51-69 (2003)

Holmgren, A.J.: A framework for vulnerability assessment of electric power systems. In: Murray, A.T., Grubesic, T.H. (eds.) *Critical Infrastructure*, pp. 31-55. Springer, Heidelberg (2007)

Israeli, E., Wood, R.K.: Shortest-path network interdiction. *Networks.* 40, 97-111 (2002)

Johnson, D.B.: Efficient algorithms for shortest paths in sparse networks. *Journal of the ACM*. 24(1), 1-13 (1977)

Kim, H., Kim, C., Chun, Y.: Network reliability and resilience of rapid transit systems. *The Professional Geographer.* 68(1), 53-65 (2016)

Kim, H., Ryerson, M.S.: Ad-hoc port assignment due to a port outage. *Proceedings of the 2ndINFORMS Transportation Science and Logistics (TSL) Society Workshop: Maritime Transportation and Port Logistics*. 83-89 (2013)

Kim, H., Ryerson, M.S.: The q-Ad hoc hub location problem for multi-modal networks. *Networks and Spatial Economics.* 17(3), 1015-1041 (2017)

Kim, H., Song, Y.: Examining accessibility and reliability in the evolution of subway systems. *Journal of Public Transportation*. 18(3), 89-106 (2015)

Kim, H.: Geographical analysis on network reliability of public transportation systems: a case study of subway network system in Seoul. *Journal of the Korean Geographical Society*. 44(2), 187-205 (2009)

Kim, H.: *p*-Hub protection models for survivable hub network design. *Journal of Geographical Systems.* 14(4), 437-61 (2012)

Kuby, M., Xu, Z., Xie, X.: A minimax method for finding the k best "differentiated" paths. *Geographical Analysis.* 29(4), 298-313 (1997)

Kurauchi, F., Sumalee, A., Seto, Y.: Network evaluation based on connectivity vulnerability. In: Lam, W.H.K., Wong, S.C., Lo, H.K. (eds.) *Transportation and Traffic Theory*, pp. 637-76. Routledge, New York (2009)

Laporte, G., Marin, A., Mesa, J.A., Perea, F.: Designing robust rapid transit networks with alternative routes. *Journal of Advanced Transportation*. 45(1), 54-65 (2011)

Li, Y., Kim, H.: Assessing survivability of the Beijing subway system. *International Journal of Geospatial and Environmental Research*. 1(1), 1-22 (2014)

Lim, C., Smith, J.C.: Algorithms for discrete and continuous multicommodity flow network disruption problems. *IIE Transactions.* 39, 15-26 (2007)

Little, R.G.: Controlling cascading failure: understanding the vulnerabilities of interconnected infrastructure. *Journal of Urban Technology.* 9(1), 109-123 (2002)

Matisziw, T.C., Murray, A.T., Grubesic, T.H.: Exploring the vulnerability of network infrastructure to interdiction. *The Annals of Regional Science*. 43(2), 307-21 (2009)

Matisziw, T.C., Murray, A.T., Grubesic, T.H.: Strategic network restoration. *Networks and Spatial Economics.* 10(3), 345-361 (2010)

Matisziw, T.C., Murray, A.T.: Modeling s–t path availability to support disaster vulnerability assessment of network infrastructure. *Computers & Operations Research.* 36, 16-26 (2009)

Murray, A.T., Grubesic, T.: *Reliability and vulnerability in critical infrastructure: a quantitative geographic perspective*. Springer, Heidelberg (2007)

Murray, A.T., Grubesic, T.H.: *Critical infrastructure: reliability and vulnerability*, Chapter 1: Overview of reliability and vulnerability in critical infrastructure. Springer-Verlag, Berlin (2007)

Murray, A.T., Matisziw, T.C., Grubesic, T.H.: A methodological overview of network vulnerability analysis. *Growth and Change.* 39(4), 573-592 (2008)

Murray, A.T., Matisziw, T.C., Grubesic, T.H.: Critical network infrastructure analysis: Interdiction and system flow. *Journal of Geographical Systems.* 9, 103-117 (2007)

Murray, A.T.: An overview of network vulnerability modeling approaches. *GeoJournal*, 78(2), 209-221 (2013)

Myung, Y.S., Kim, H.J.: A cutting plane algorithm for computing k-edge survivability of a network. *European Journal of Operational Research*. 156, 579-589 (2004)

Nagurney, A., Qiang, Q.: Robustness of transportation networks subject to degradable links. *Europhysics Letters (EPL).* 80(6), 68001 (2007)

Nakanishi, Y.J.: Securing America's mass transit systems. *Journal of Transportation Security.* 2(1), 9-11 (2009)

Nardelli, E., Proietti, G., Widmayer, P.: Finding the most vital node of a shortest path. *Theoretical Computer Science.* 296, 167-177 (2003)

O'Kelly, M.E., Kim, H.: Survivability of commercial backbones with peering: a case study of Korean networks critical infrastructure. In: Murray, A.T., Grubesic, T.H. (eds.) *Reliability and Vulnerability*, pp. 107-28. Springer Verlag, Berlin (2007)

O'Kelly, M.E.: Network hub structure and resilience. *Networks and Spatial Economics*. 15(2), 235-251 (2014)

Peterson, S.K., Church, R.L.: A framework for modeling rail transport vulnerability. *Growth and Change*. 39(4), 617-641 (2008)

Rashidi, E.: *Optimization models and algorithms for vulnerability analysis and mitigation planning of pyro-terrorism*. ProQuest Dissertations Publishing, Mississippi State University (2016)

Rawls, C.G., Turnquist, M.A.: Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B.* 44, 521-534 (2010)

Rodrigue, J.P., Comtois, C., Slack, B.: The geography of transport systems, 4th edn. Routledge, New York (2017)

Rodríguez-Núñez, E., García-Palomares, J.C.: Measuring the vulnerability of public transport networks. *Journal of Transport Geography*. 35, 50-63 (2014)

Scaparra, M.P., Church, R.L.: An exact solution approach for the disruption median problem with fortification. *European Journal of Operational Research*. 189(1), 76-92 (2008)

Schoon, M.: A short historical overview of the concepts of resilience, vulnerability, and adaptation. *Workshop in Political Theory and Policy Analysis*, Indiana University Working Paper W05-4 (2005)

Snyder, L.V., Scaparra, M.P., Daskin, M.S., Church, R.L.: Planning for Disruptions in Supply Chain Networks. In *INFORMS Tutorials in Operations Research*. https://doi.org/10.1287/educ.1063.0025 (2006)

Strandberg, V.J.: Rail bound traffic – a prime target for contemporary terrorist attacks? *Journal of Transportation Security*. 6(3), 271-286 (2013)

Sullivan, J. L., Novak, D. C., Aultman-Hall, L., Scott, D. M.: Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: a link-based capacity-reduction approach. *Transportation Research Part A*. 44, 323-336 (2010)

Sun, D., Guan, S.: Measuring vulnerability of urban metro network from line operation perspective. *Transportation Research Part A*. 94, 348-359 (2016)

Sun, D., Zhao, Y., Lu, Q.C.: Vulnerability analysis of urban rail transit networks: a case study of Shanghai, China. *Sustainability*. 7(6), 6919-6936 (2015)

Taylor, M.A.P., Sekhar, S.V.C., D'Este, G.M.: Application of accessibility based methods for vulnerability analysis of strategic road networks. *Networks & Spatial Economics*. 6(3-4): 267-291 (2006)

The BART (San Francisco, CA) network https://www.bart.gov/stations

The CTA (Chicago, IL) network http://www.transitchicago.com/assets/1/maps/ctatrainmap.png

The GCRTA (Cleveland, OH) network http://www.riderta.com/sites/default/files/pdf/maps/System_Map_Rapid_Connect.pdf

The LACMTA (Los Angeles, CA) network https://media.metro.net/documents/8f0fe43e-da3b-4a10-bd8e-4cfd54e30eb3.pdf

The MARTA (Atlanta, GA) network http://www.itsmarta.com/train-stations-and-schedules.aspx

The MBTA (Boston, MA) network http://www.mbta.com/schedules_and_maps/subway/

The MDT (Miami, FL) network http://www.miamidade.gov/transit/metrorail-stations.asp and https://www.miami-dade.gov/transit/metromover-stations.asp

The NYCT (New York, NY) network http://web.mta.info/maps/submap.html

The PATH (New York, NY) network http://www.panynj.gov/path/maps.html

The SEPTA (Philadelphia, PA) network http://septa.org/maps/system/

The WMATA (Washington, DC) network
https://wmata.com/schedules/maps/upload/system_map_color.pdf

Ukkusuri, S.V., Yushimito, W.F.: A methodology to assess the criticality of highway transportation networks. *Journal of Transportation Security.* 2(1), 29-46 (2009)

Wan, N., Zhan, F., Cai, Z.: A spatially weighted degree model for network vulnerability analysis. *Geo-spatial Information Science*. 14(4), 274-281 (2011)

Wang, J., Li, Y., Liu, J., He, K., Wang, P.: Vulnerability analysis and passenger source prediction in urban rail transit networks. *PLoS ONE*. 8(11), e80178 (2013)

Whiting, P.D., Hillier, J.A.: A method for finding the shortest route through a road network. *Journal of the Operational Research Society*. 11(1), 37-40 (1960)

Wood, R.K.: Deterministic network interdiction. *Mathematical and Computer Modeling*. 17, 1-18 (1993)

Yates, J., Sanjeevi, S.: Assessing the impact of vulnerability modeling in the protection of critical infrastructure. *Journal of Geographical Systems*. 4(4), 415-435 (2012)

Zenklusen, R.: Network flow interdiction on planar graphs. *Discrete Applied Mathematics.* 158, 1441-1455 (2010)

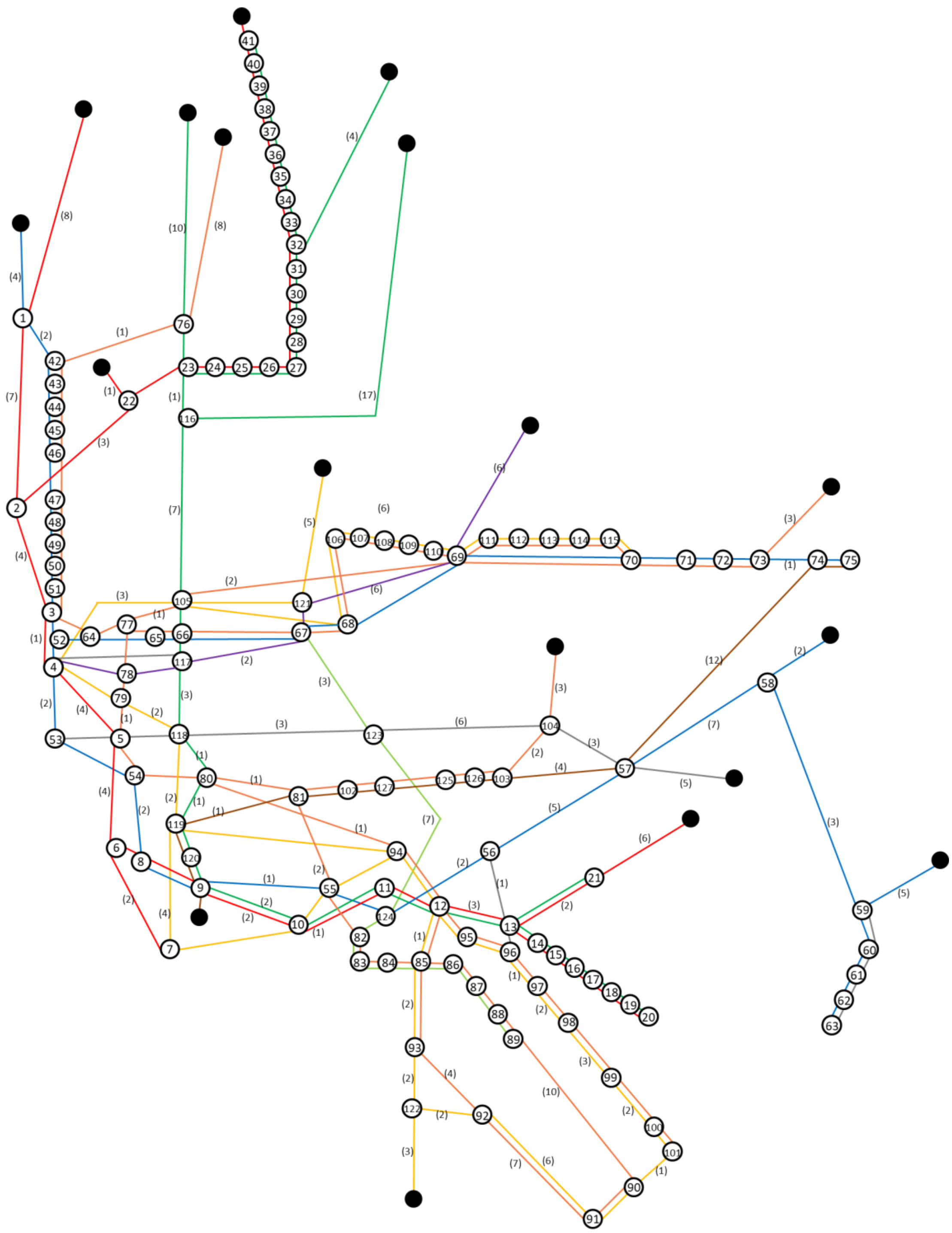
Zhang, X., Miller-Hooks, E., Denny, K.: Assessing the role of network topology in transportation network resilience. *Journal of Transport Geography*. 46, 35-46 (2015)

APPENDIX

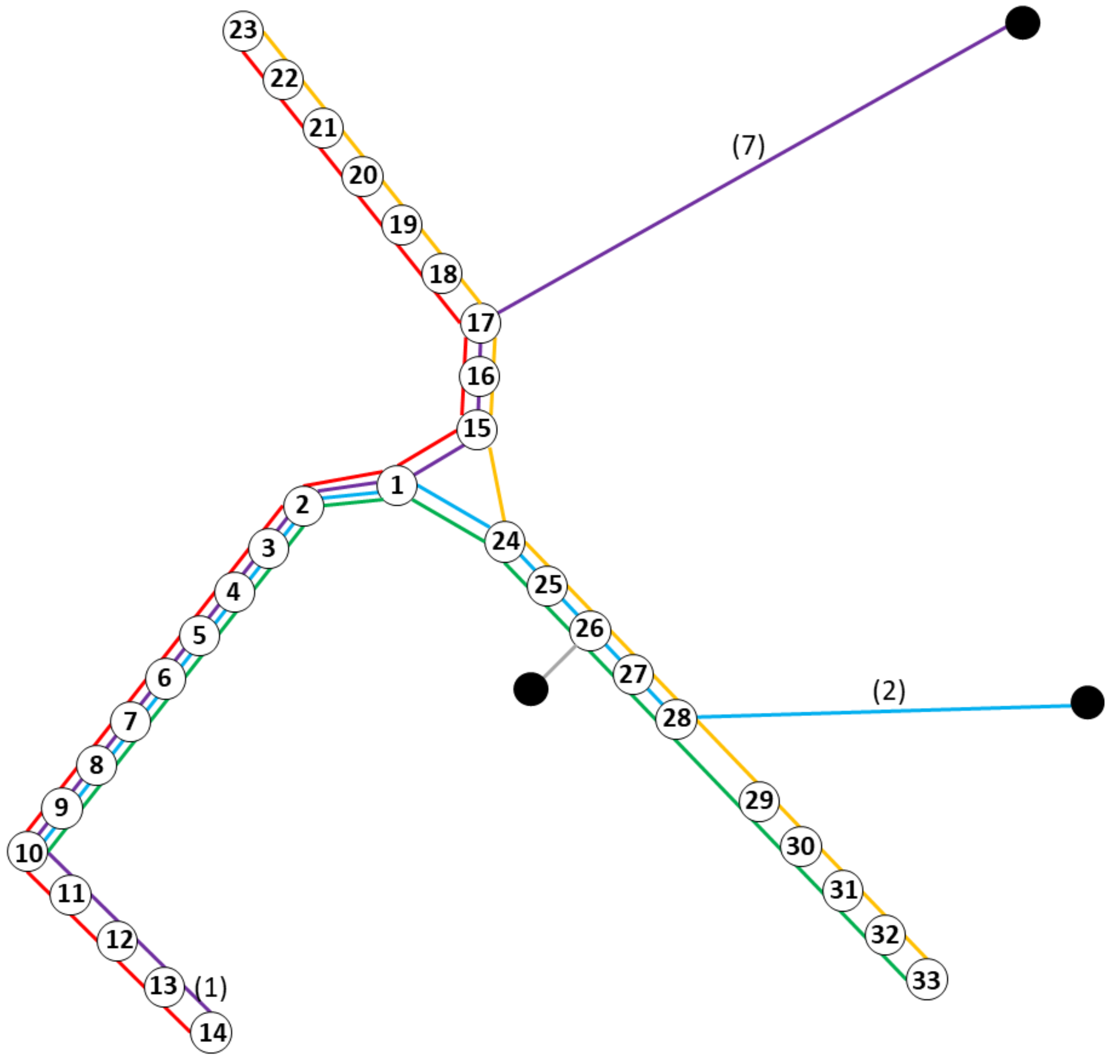
The eleven U.S. heavy rail networks

Note: Only transfer nodes and end nodes are shown in the networks. The value in parentheses indicates the number of intermediate nodes. The graphs are dedicated to present network connectivity, not proportional to physical distance. The order of network figures corresponds the order in Table 2.3.

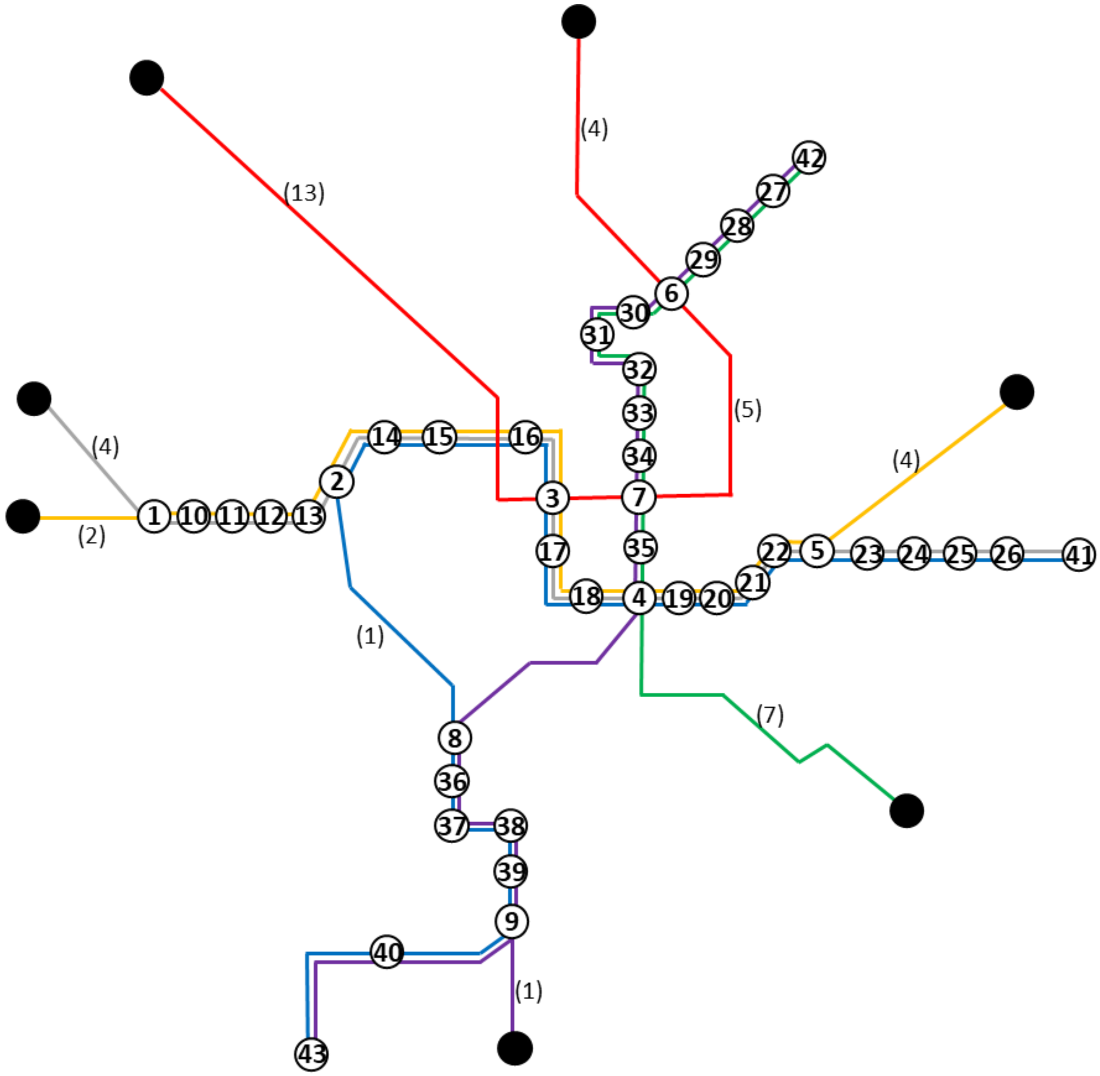
1. MTA New York City Transit (NYCT). New York, NY.



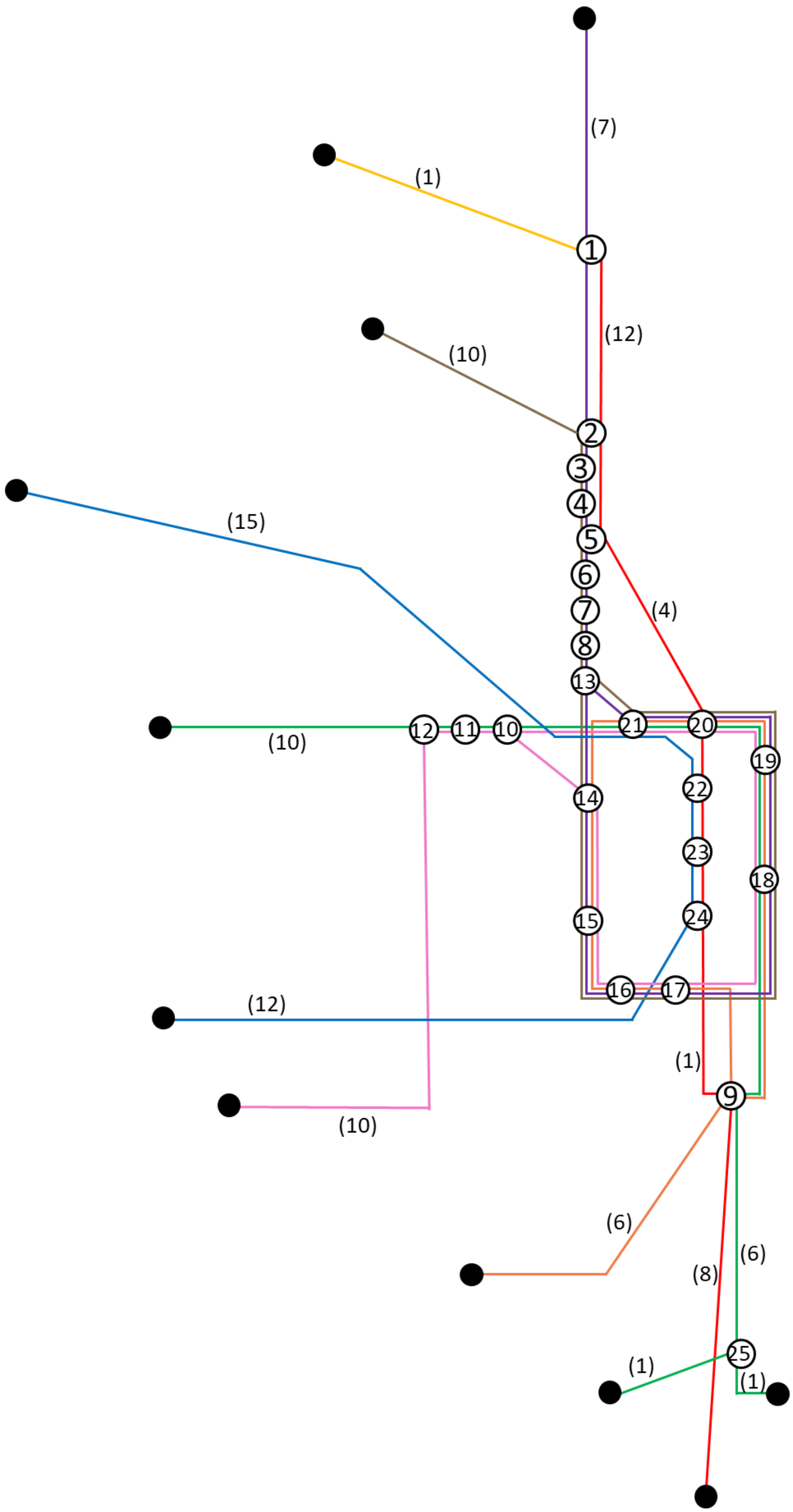
2. San Francisco Bay Area Rapid Transit District (BART). San Francisco, CA.



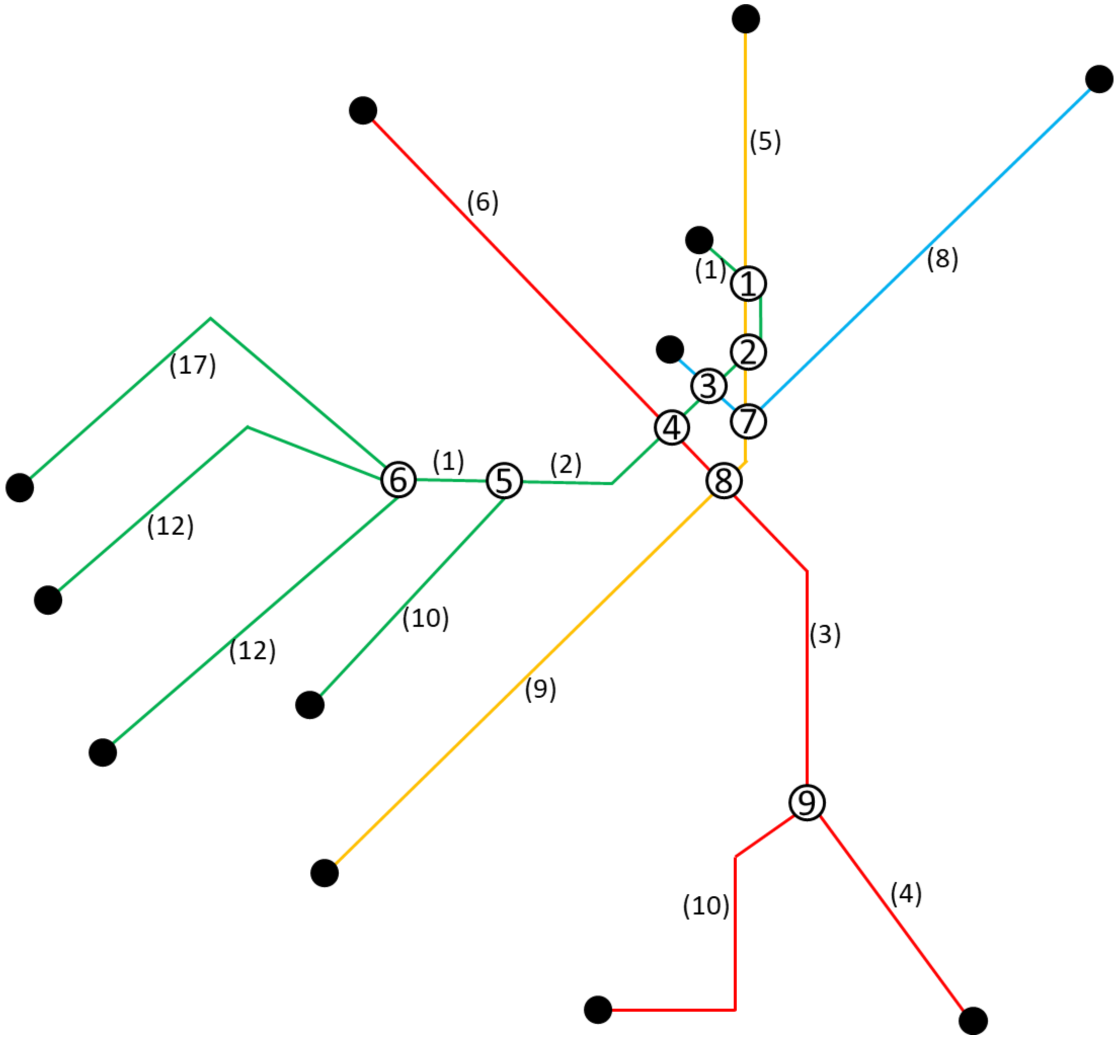
3. Washington Metropolitan Area Transit Authority (WMATA). Washington, DC.



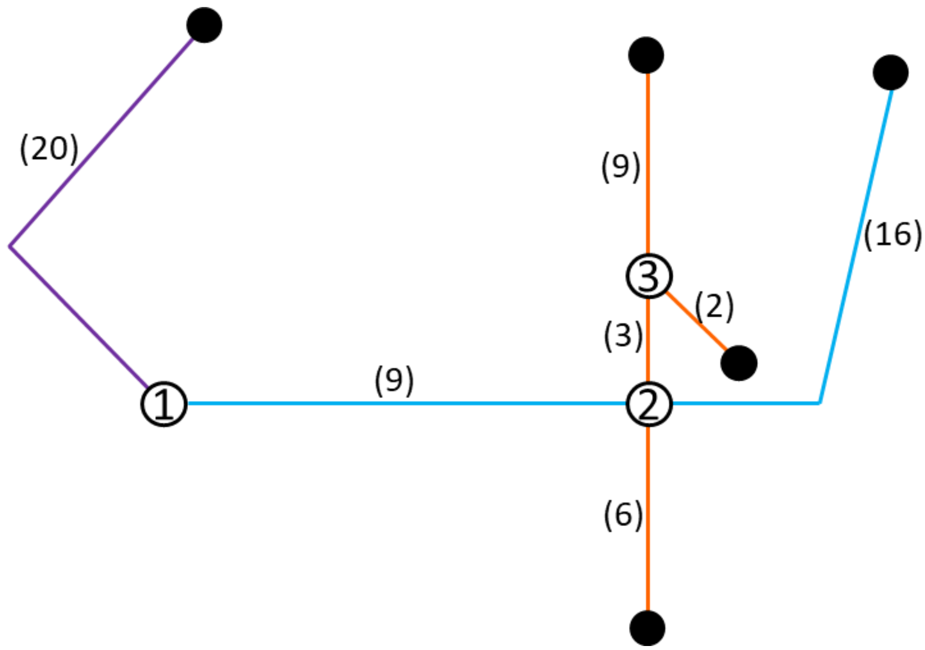
4. Chicago Transit Authority (CTA). Chicago, IL.



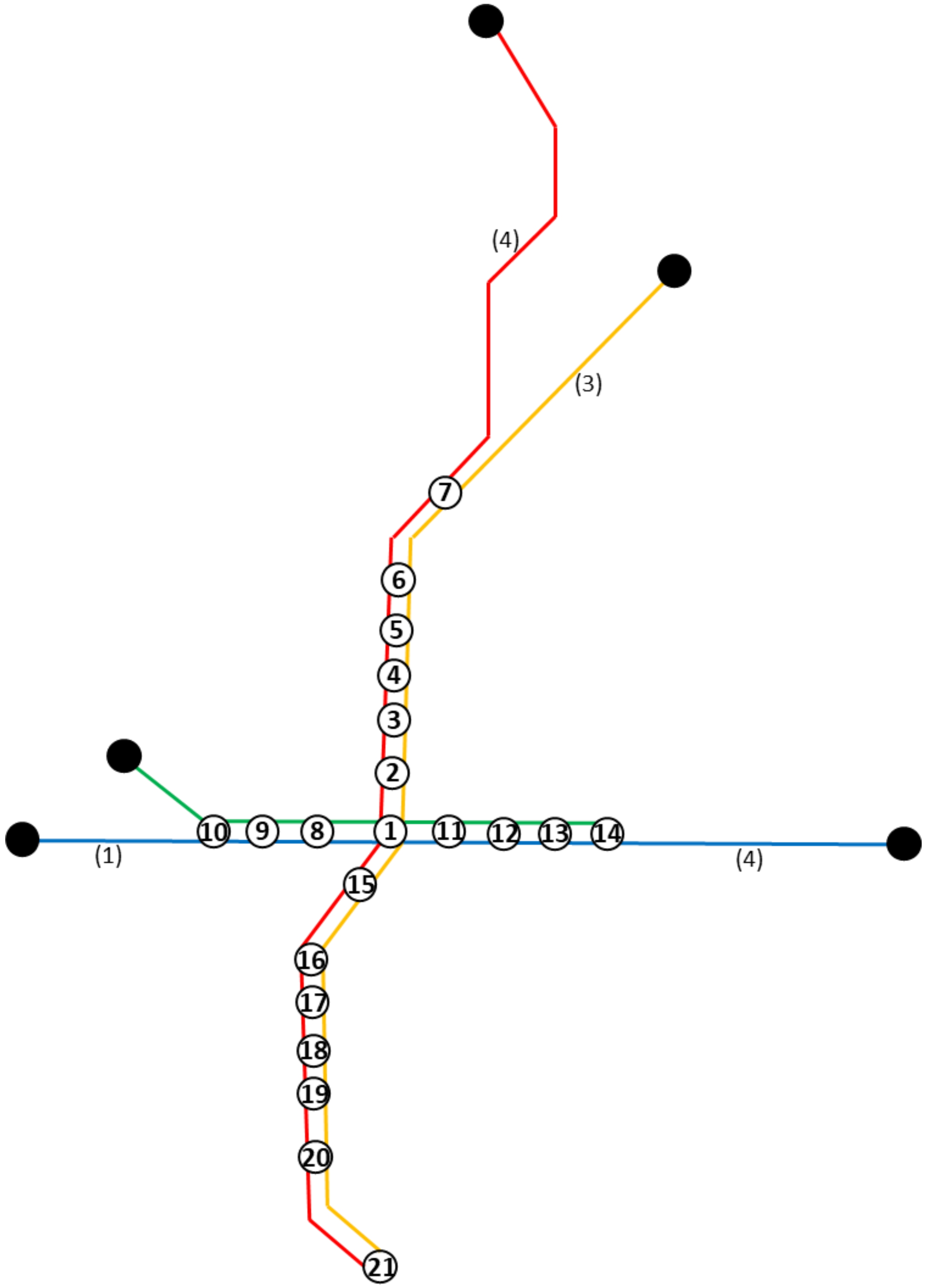
5. Massachusetts Bay Transportation Authority (MBTA). Boston, MA.



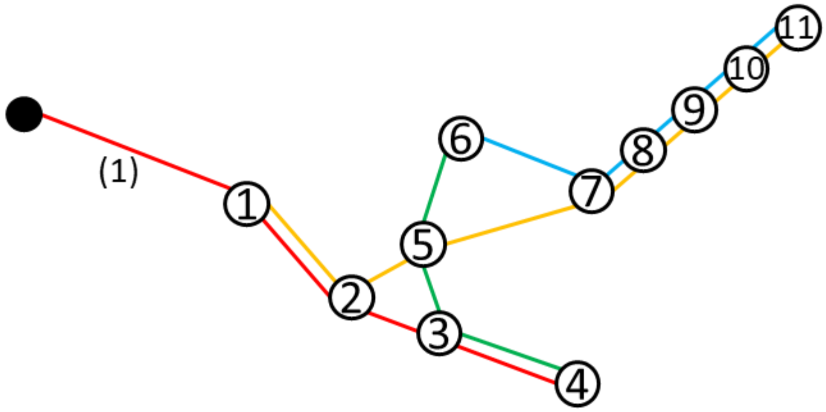
6. Southeastern Pennsylvania Transp. Authority (SEPTA). Philadelphia, PA.



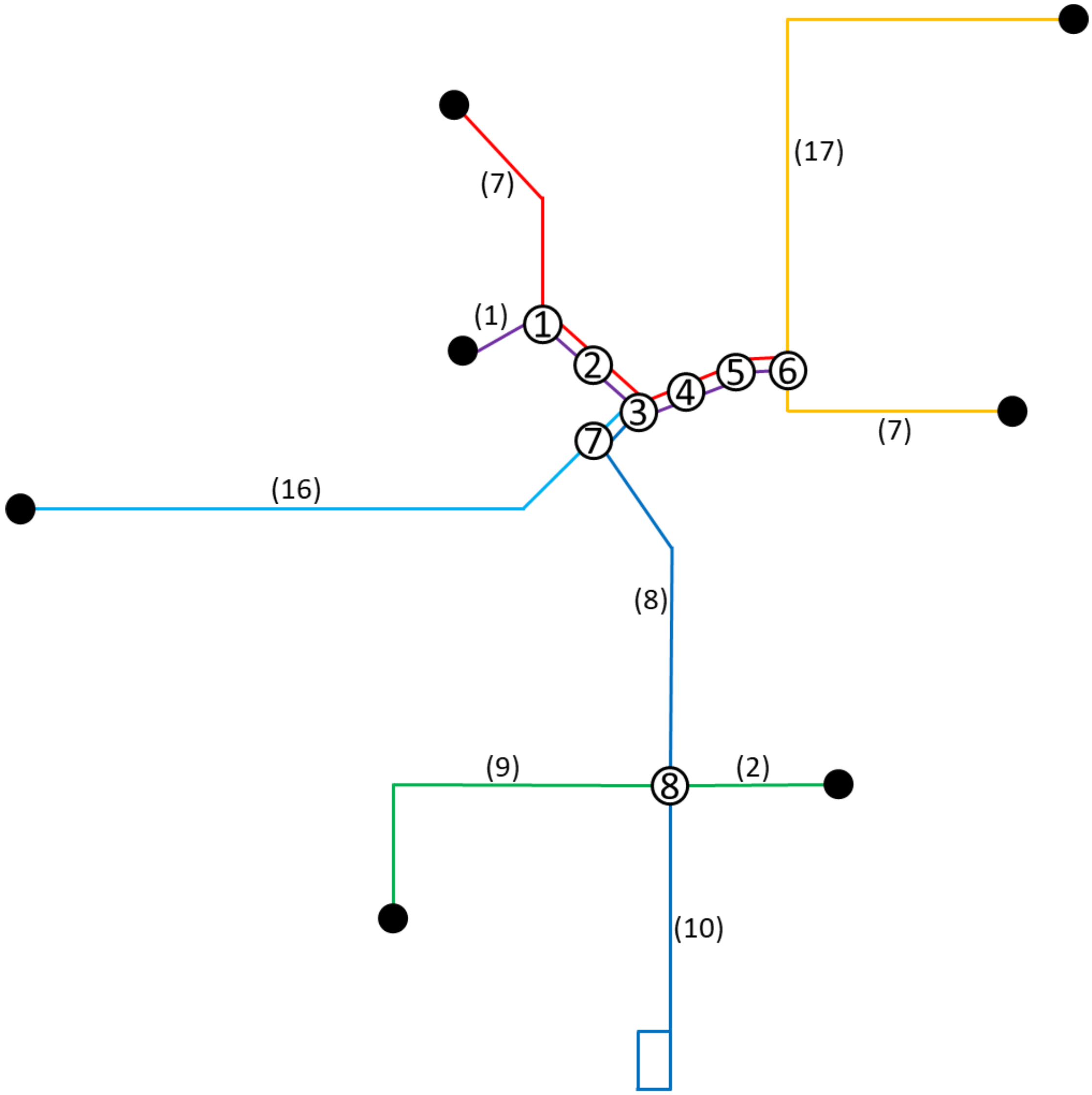
7. Metropolitan Atlanta Rapid Transit Authority (MARTA). Atlanta, GA.



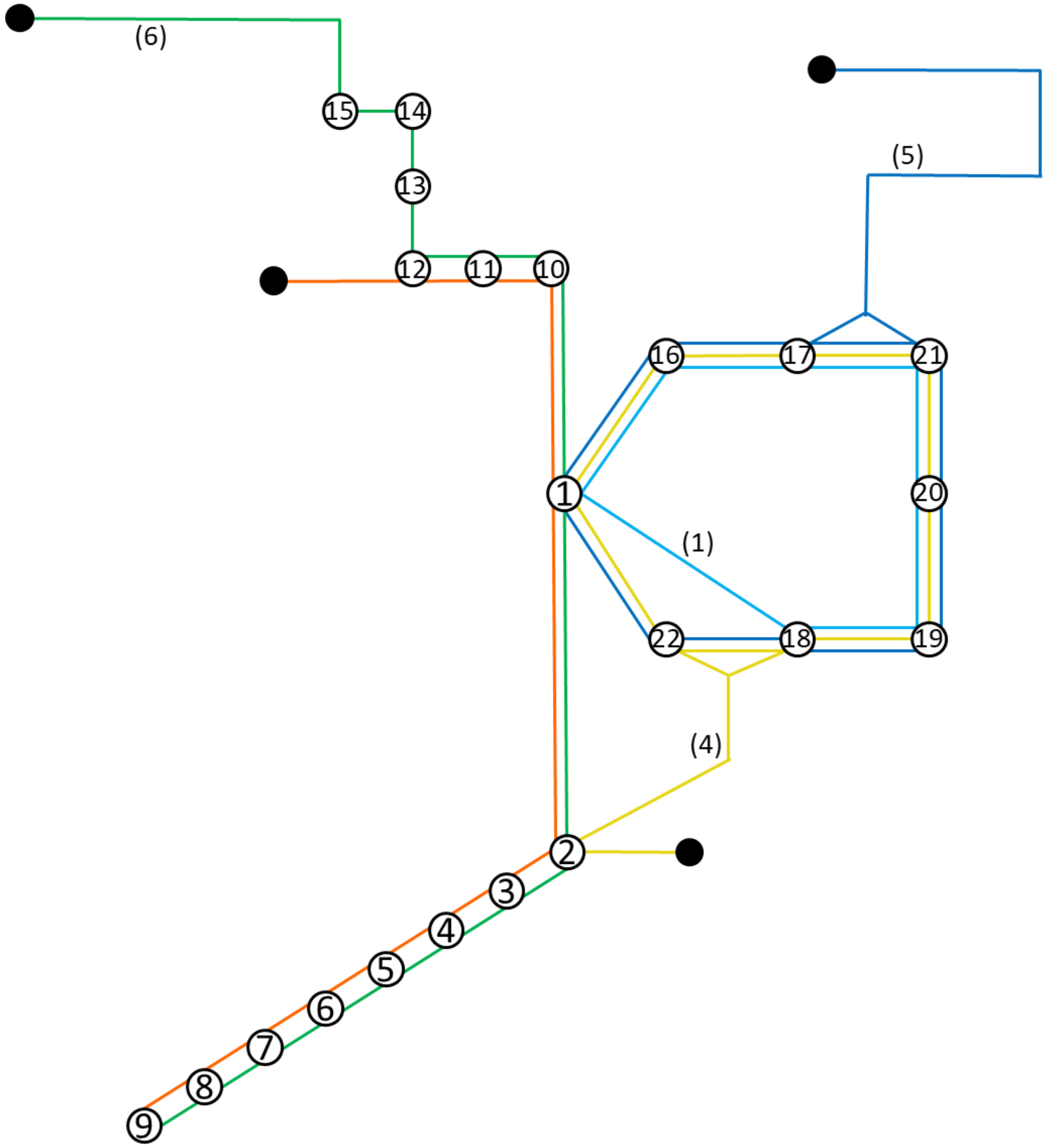
8. Port Authority Trans-Hudson Corporation (PATH). New York, NY.



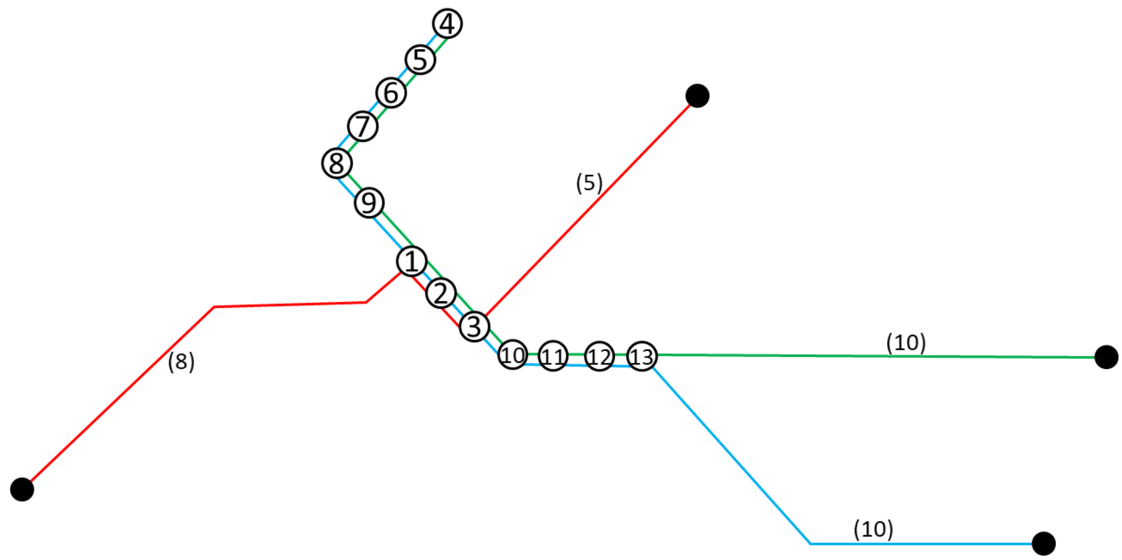
9. Los Angeles County Metropolitan Transp. Auth. (LACMTA). Los Angeles, CA.



10. Miami-Dade Transit (MDT). Miami, FL.



11. The Greater Cleveland Regional Transit Authority (GCRTA). Cleveland, OH.



VITA

Qian Ye was born and raised in Tangshan, Hebei province of China. She received her Bachelor's degree from Beijing Forestry University in 2009, and Master's degree from Beijing Normal University in 2012. She was an exchange graduate student at the University of North Carolina at Greensboro during Fall 2010 and Spring 2011. She came to the University of Tennessee, Knoxville in Fall 2014 to pursue her Doctor of Philosophy degree in Geography. Her research focuses on Quantitative Methods, Network Optimization and Modeling, Geographic Information Science (GIS), and Geography of Transportation. She entered the Intercollegiate Graduate Statistics Program (IGSP) in 2015 to pursue her double degree in Master of Science in Statistics, where she gained a comprehensive knowledge in statistical Data Mining and Machine Learning. She was also involved in the Interdisciplinary Graduate Minor in Computational Science (IGMCS) program to facilitate her research in quantitative fields and mathematical modelling.