Criticality Analysis of Surface Transportation Infrastructures based on Freight Flow Network Optimization

Abdullah Al Khaled

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
https://scholarsjunction.msstate.edu/td/1223

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.
Criticality analysis of surface transportation infrastructures based on freight flow network optimization

By
Abdullah Al-Khaled

A Dissertation
Submitted to the Faculty of Mississippi State University
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Industrial and Systems Engineering in the Department of Industrial and Systems Engineering

Mississippi State, Mississippi
May 2013
Criticality analysis of surface transportation infrastructures based on freight flow network optimization

By

Abdullah Al-Khaled

Approved:

Burak Ekşioğlu
Associate Professor of Industrial and Systems Engineering
(Major Professor)

Mingzhou Jin
Committee Participant of Industrial and Systems Engineering
(Committee Member)

Sandra D. Ekşioğlu
Associate Professor of Industrial and Systems Engineering
(Committee Member)

Li Zhang
Associate Professor of Civil and Environmental Engineering
(Committee Member)

Kari Babski-Reeves
Associate Professor of Industrial and Systems Engineering
(Graduate Coordinator)

Sarah A. Rajala
Dean of James Worth Bagley College of Engineering
The purpose of this dissertation is to develop models and solution approaches to identify the critical hierarchies of railroad and surface transportation network infrastructures, and to facilitate re-routing options that will be necessary for traffic management decision makings in the event of a disaster. We focus on building mathematical models for routing/re-routing of traffic considering the congestion effects which are obvious in the disrupted networks due to disaster. Based on these models, the critical hierarchies of infrastructures are determined. For railroad, we develop two different models: the first one considers ‘unit’ train re-routing and the other one considers Train Design approach. For intermodal system, the optimization model facilitates optimal re-routing of traffic using three surface transportation modes: highway, railway and waterway, considering the congestion characteristics of each mode.

For the first model of railroad routing, the optimization model optimally routes unit trains based on a minimum cost network flow formulation with nonlinear objective function. The nonlinear objective function is approximated with a piece-wise linear function to make the model computationally tractable. The second model, known as
Train Design optimization, is a highly combinatorial and complex optimization problem. The developed model’s computational complexity suggests us to use heuristic solution procedures. We develop a special heuristic algorithm to route the traffic in the congested network. In this heuristic procedure, we divide the problem into two sub-problems (SPs): SP-1 is termed as Block-to-Train Assignment (BTA) problem, and SP-2 is termed as Train Routing (TR) problem. BTA problem provides a feasible solution that includes the minimum number of required trains with the pick-up and drop-off points of the blocks carried by these trains, and TR problem ensures the optimal routing of these trains.

Similar to railroad, an optimization model is developed for optimal routing/re-routing of traffic using the intermodal network. It is a mixed integer programming (MIP) problem, which is not practical to solve for real-world problem instances within reasonable amount of time. Linear relaxation to this model provides a very good lower bound closer to optimal solution. Therefore, we implemented our case-study for a real-world intermodal transportation system of five U.S. states.
DEDICATION

To my parents, brother, sister, beloved wife, and the most precious daughter

Ilham.
ACKNOWLEDGEMENTS

I would like to express my gratitude to all the people who helped me both directly and indirectly in completion of my Ph.D. at Mississippi State. First of all, I would like to take the opportunity of expressing the heartiest gratitude to my advisor, Dr. Mingzhou Jin, for mentoring me throughout my doctoral studies. His inspiring enthusiasm, guidance, constant support and insightful comments have been extremely helpful throughout my dissertation work.

I would like to specially acknowledge Dr. Burak Ekşioğlu for serving as my major professor of my dissertation when my advisor left the university, and for being so helpful as my major professor. I would like to thank Dr. Sandra Duni Ekşioğlu and Dr. Li Zhang for serving on my dissertation committee, and always being there for me. I also express my sincere gratitude to Dr. Xiaopeng Li for his insightful suggestions in conducting my research. I would like to express my appreciation to all my friends who made Starkville a very special place during the three years of my stay.

Finally, I would like to thank my family. I would like to thank my parents; their love, encouragement, and eternal support have motivated me each day of my life. I would like to express my special thanks to my sister Ruma, brother Shuhel and my four beloved nieces. I am deeply grateful to my wonderful wife, Seratun Jannat. Without her unwavering and unconditional support, I would never have completed my study. Also, special thanks to my little angel, Ilham, for being such a blessing from the Almighty.
TABLE OF CONTENTS

DEDICATION .................................................................................................................... ii

ACKNOWLEDGEMENTS ............................................................................................... iii

LIST OF TABLES ............................................................................................................. vi

LIST OF FIGURES ......................................................................................................... vii

CHAPTER

I. INTRODUCTION ........................................................................................................1

1.1 Motivation ...........................................................................................................5
1.2 Background ........................................................................................................8
1.3 Dissertation objectives and Contributions ......................................................11

II. LITERATURE REVIEW ........................................................................................15

2.1 Railroad transportation system .......................................................................15
  2.1.1 Railroad blocking problem ....................................................................17
  2.1.2 Train routing, timetabling and dispatching problem .........................18
  2.1.3 Block-to-train assignment (BTA) problem .......................................19
  2.1.4 Combined BTA and train routing problem ....................................21
2.2 Intermodal transportation system ..................................................................24

III. DATA AVAILABILITY ........................................................................................29

3.1 Network data ......................................................................................................29
  3.1.1 Railroad network data for criticality analysis using unit train re-routing optimization ..........................................................30
  3.1.2 Railroad network data for criticality analysis using train design optimization ..............................................................30
  3.1.3 Intermodal network data ...................................................................32
3.2 Freight flow data .................................................................................................36
  3.2.1 Railroad freight flow data for criticality analysis using unit train re-routing optimization ......................................................37
  3.2.2 Railroad freight flow data for criticality analysis using train design optimization ..............................................................39
  3.2.3 Intermodal freight flow data ................................................................40
IV. RAILROAD CRITICALITY ANALYSIS USING RE-ROUTING OPTIMIZATION .............................................................................................................42

4.1 Criticality analysis of links using unit train re-routing optimization under heavy congestion .................................................................42
  4.1.1 Problem description ................................................................ 43
  4.1.2 Mathematical model formulation ..............................................44
    4.1.2.1 Parameters .........................................................................44
    4.1.2.2 Variables ..........................................................................45
  4.1.3 Case study ...............................................................................46
  4.1.4 Conclusion .............................................................................49

4.2 Criticality analysis of all infrastructures using train design optimization under heavy congestion .................................................................49
  4.2.1 Problem Statement and model formulation ..................................52
    4.2.1.1 Decision variables .................................................................54
    4.2.1.2 MIP model ..........................................................................54
  4.2.2 Solution approach ....................................................................59
    4.2.2.1 Initialization .......................................................................60
    4.2.2.2 Block-to-train assignment (BTA) sub-problem ..................61
      4.2.2.3.1 Pre-Process .................................................................67
      4.2.2.3.2 Variables ...................................................................68
  4.2.3 Numerical Experiments ............................................................71
  4.2.4 Case Study ...............................................................................72
  4.2.5 Conclusion and future work .....................................................74

V. EVALUATING CRITICALITY OF INTERMODAL NETWORK INFRASTRUCTURES USING FLOW OPTIMIZATION ........................................76

5.1 Problem description .......................................................................77
5.2 Mathematical model formulation ...................................................80
  5.2.1 Parameters .............................................................................81
  5.2.2 Variables .................................................................................82
5.3 Numerical experiments ...................................................................86
5.4 Computational study .....................................................................87
5.5 Conclusion and future work .........................................................90

REFERENCES ..........................................................................................93
# LIST OF TABLES

3.1 Sample data of Class-I railroad network attributes for unit train routing ........31
3.2 Sample data of Class-I railroad link attributes .................................................32
3.3 Sample data of Class-I railroad yard/ station attributes .................................33
3.4 Sample network data for intermodal links .......................................................35
3.5 Sample network data for intermodal transfer terminals .................................36
3.6 Sample unit train flow data ..............................................................................38
3.7 Sample railroad freight flow data for train design optimization ......................39
3.8 Sample intermodal freight flow data .................................................................41
4.1 Sample train travel time components for different links .................................47
4.2 Results of the MIP and heuristic solution approaches for small instances .......71
4.3 Other input parameters .....................................................................................72
5.1 Results of the MIP and LP relaxation solution approaches for small instances .........................................................87
5.2 Various cost components for intermodal transportation ................................88
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Hierarchical structure of decisions in railroad operations [22]</td>
<td>16</td>
</tr>
<tr>
<td>4.1</td>
<td>Piecewise linear approximation of the cost-flow function</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Network showing the criticality of the links</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Framework of the proposed heuristic algorithm</td>
<td>60</td>
</tr>
<tr>
<td>4.4</td>
<td>Steps of initialization algorithm for creating block sub-networks and ‘unit’ trains</td>
<td>61</td>
</tr>
<tr>
<td>4.5</td>
<td>Block-to-train assignment (BTA) algorithm</td>
<td>62</td>
</tr>
<tr>
<td>4.6</td>
<td>Map showing the criticality of Class-I railroad network infrastructures</td>
<td>73</td>
</tr>
<tr>
<td>5.1</td>
<td>Sample intermodal network (a) original network, and (b) equivalent network</td>
<td>80</td>
</tr>
<tr>
<td>5.2</td>
<td>Piece-wise linear approximation of the non-linear cost-flow function for links</td>
<td>84</td>
</tr>
<tr>
<td>5.3</td>
<td>Piece-wise linear approximation of the non-linear cost-flow function for intermodal transfer nodes</td>
<td>85</td>
</tr>
<tr>
<td>5.4</td>
<td>Map showing the criticality of intermodal network infrastructures</td>
<td>89</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Transportation is an essential and critical part of a nation’s economy. Historically, it is obvious that each great nation of civilization had an efficient transportation system to support its economic growth. In particular, freight transportation is acknowledged as a fundamental basis of economic strength of a country. Like every other country in the world, freight transportation is also a vital component of the American economy. It has been playing a significant role in the growth and competitiveness of the U.S. economy since the last couple of centuries. For instance, according to an estimate by the Office of Freight Management and Operations of the U.S. Department of Transportation’s Federal Highway Administration (FHWA), in 2011, the nation’s transportation network served about 117 million households, 7.4 million business establishments, and 89,500 governmental units that represent a huge part of a U.S. economy [1].

The significance of freight transportation and its increased contribution to the U.S. economy can be realized by the following facts. Due to the increased globalization of economic activities, the imports and exports of goods within and across U.S. borders is increasing every year. These increased merchandise activities are escalating the generation of freight movements in the United States proportionately. For instance, the volume of freight movement within U.S. boundary has increased by 60 percent from 1975 to 1997 [2]. In 1998, about 15 billion tons of freight was transported in the U.S.
freight network that valued at more than $9 trillion [3]. In 2002, more than 19 billion tons of freight was moved over the transportation network of United States that accounted for $13.3 trillion [4]. Also, in 2007, about 21 billion tons of freight, with a value of approximately $14.3 trillion, was transported in United States [5]. Based on the current increasing trend of freight transportation and the policies of the Obama Administration, the relative significance of this sector in the overall U.S. economy is expected to continue growing. For example, in 2010, President Obama announced his plan of doubling exports by 2015 through the National Export Initiative (NEI) [6].

Though the United States has the largest freight transportation system in the world, the transportation infrastructure may need to be expanded to meet the increasing demand for freight movement. But infrastructure expansion is a very time consuming and expensive process. Rather, if the freight movements are efficiently managed within and among the individual transportation sectors, this issue may be resolved to a great extent. It is thought to be the critical enabler of future U.S. economic development and competitiveness. Before we suggest for improvement through managing the movement of goods in freight transportation and its infrastructure, it is useful to describe the system and its current state. The U.S. freight-transportation network includes about 4 million miles of public roads and highways, 140 thousand miles of railroad tracks, 25 thousand miles of navigable waterways, 5,200 public-use airports, 1.3 million miles of gas pipeline and 180 thousand miles of oil pipelines [7]. It is worth mentioning that freight movement within U.S. boundary dominates imports and exports both in weight and value [5]. That means we need to focus on the efficient movement of goods within national borders to contribute to the growth and competitiveness of U.S. economy. Particularly, surface
transportation modes consisting of highway railway and waterway play the most important role in the transportation economy. Among these three modes, railway is playing the most important role in terms of total freight ton-miles (a measurement describing the distance one short ton of freight travels within a region or country), though highway is contributing the most in terms of value and weight of freight [5]. Again, the share of intermodal transportation system is increasing each year to keep pace with the increased world-trade and domestic, economic strength. Intermodal transportation is in the rising trend because the demand for goods has grown gradually over the past half century that needs a cost-effective freight transportation to effectively contribute to the national economy, and the underlying principle of intermodalism is to be able to use the most cost-effective modes of transport to ship products from its origin to destination [8]. According to Crainic and Kim, intermodal freight transportation can be broadly defined as the transportation of a shipment from its origin to its destination by a sequence of at least two transportation modes, the transfer from one mode to the next being performed at an intermodal terminal [9]. In intermodal freight transportation, the goods are usually moved in the standardized containers from one location to other locations using the surface transportation modes: highway, railway and waterway. This relatively new concept of transportation has been growing since the enactment of Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 and the following Transportation Equity Act for the 21st Century (TEA-21) of 1998 [10, 11].

The U.S. transportation network and its infrastructure have been susceptible to disruption for many years. Although these interruptions have been caused mostly by natural disasters (e.g., floods, storms, fires, etc.) and man-made disasters (e.g., terrorist
attacks) on these facilities have been taking place with increased rate since the last
decade. These disruptions are crucial to the national economy, and the nation’s way of
life as transportation is highly interlinked with them. Since transportation security and
resilience are critical to the economy and homeland security, the Transportation Systems
Sector-Specific Plan (SSP), a sector that includes all modes of U.S. transportation, has set
the vision: “Our vision is a secure and resilient transportation network, enabling
legitimate travelers and goods to move without undue fear of harm or significant
disruption of commerce and civil liberties [12].”

Research in the field of rail routing/re-routing optimization has not been done
much though it is one of the oldest freight transportation sectors, but it is growing fast
since last two decades. Again, as intermodal transportation system is a relatively newer
concept, research on routing and re-routing optimization in this field has been growing
rapidly since 1991 due to ISTEA. Operation research (OR) applications on efficient
routing/ re-routing of freight would play a pivotal role in these two sectors. It will also
help to identify components/areas of high consequence and vulnerability from which
relevant stakeholders can better develop tactics to protect the critical transportation
infrastructures. The aim of our research is to apply network optimization models which
would help in efficient routing/re-routing of freight movements in the railroad network
and also intermodal network comprising of highway, railway and waterway networks of
USA. These models will also help to determine the critical infrastructures of these two
modes of transportation.
1.1 Motivation

Freight transportation by railroads is an essential part of U.S. economy and its increasing the share of shipments since 1980. In 1980, the rail freight industry moved about 27 percent of the total ton-miles of U.S. freight, and this share increased to over 40 percent by 2007 [13]. While freight movement through railway is increasing every year, some commodities are mostly dependent on it. For example, railroad moves 70 percent of domestically produced automobiles, 70 percent of coal delivered to power plants, and about 35 percent of the U.S. grain harvest each year [13]. Another important trend in the rail freight industry shows that from 1980 to 2006, total tonnages originated by Class I railroads increased 31 percent, ton-miles increased 93 percent and revenues from freight movement increased 91 percent [13]. Again, among all four major freight modes, class I railroads had the highest fuel efficiency of 337 Btu per ton-mile, compared to 514 Btu, 3,357 Btu, and 9,600 Btu per ton-mile for domestic waterborne, heavy trucks, and air freight in 2005 [14]. In terms of air quality, all freight railroads in the US emit less than 1 percent of the greenhouse gas compared to 5.8 percent for trucking, assessed by Environmental Protection Agency (EPA) [15]. In contrast, rail infrastructure has shrunk, and excess capacity has been reduced due to carriers’ consolidated operations since the last two decades. For instance, rail network miles have been decreased by 42 percent from 1980 to 2006. Thus, the railroads have significantly increased their traffic density.

In addition to economy, railroads are also critical to national security as the U.S. armed forces rely on commercial rail carriers to meet many of their landside logistics movement. In case of national crises that necessitate U.S. military response, railroads’ reduced capacity may limit the capability to quickly move troops and equipment to the
desired locations. Since railway freight security and resilience are critical to the U.S. economy and homeland security, the Freight Rail Division has been established under the Transportation Security Administration (TSA)’s Office of Transportation Sector Network Management (TSNM). The vision of this division is to “ensure the secure movement of all cargo on our nation’s freight rail systems, and promote the free flow of commerce by working with our public and private sector partners to maintain a secure, resilient, and sustainable network” [16]. To support this vision, it is necessary to develop models to identify rail areas of high consequence and vulnerability, from which railway stakeholders can better develop tactics to protect the increasingly stressed railway infrastructure.

Transportation resilience and security are critical to the nation’s supply chain, national and local economy, and overall community resilience to natural disasters or terrorist attacks. Resilience is usually defined as “the ability to recover from or adjust easily to misfortune or change” [17] and indicates system performance under disruption, recovery speed, and required assistance from outside. During any natural or man-made disaster, it is essential to keep the freight flow by efficiently re-routing the disrupted traffic. To re-route the traffic in the disrupted network, congestion is a logical consequence. So, it is necessary to develop the models that will re-route the traffic in the disrupted network taking into consideration of the congestion effect. Unfortunately, it is rare to find in the literature about rail routing models that considered the congestion effects. In this research, we will develop optimization models and solution algorithms that will consider the congestion effects in routing or re-routing the trains in the disrupted network.
As mentioned earlier, another emerging sector in transportation industry is intermodal transportation, and this sector is increasing the total ton-miles of freight movement every year. Intermodalism pledges to lower the total transportation costs by permitting each mode to be used for the portion of the trip for which it is best suited, and thus reducing the trouble on overstressed network components by changing use to infrastructures with surplus capacities [18]. It is assumed to be the backbone of future freight transportation system in U.S. transportation industry. In 2007, the percentage share of intermodal freight transportation in terms of total ton-miles was about 13 percent of total ton-miles of U.S. freight [14], and Federal Highway Administration (FHWA) estimates this share to increase over 21 percent by 2035 [19]. Container-related transportation activities in global market have been growing remarkably since 1993, and growth rate of container traffic is significant from one year to the next year [9]. While containers are the basic units of intermodal transportation, this increasing trend of container movement clearly indicates the importance of intermodal freight transportation in future. While the volume of operations in this area are increasing, efforts to enhance the overall performance of intermodal transportation system as well as the algorithms necessary to model and optimize operations on them are found very diminutive. In comparison of several other application areas, operations research (OR) models and methods for intermodal freight transportation is still much undeveloped area. To enhance the competitiveness of the economy at all levels as well as meeting the federal requirements, there are pressing needs to research, and develop procedures and methodologies to model freight flows on intermodal freight transportation [20]. In this dissertation, we develop an optimization model that takes into account the congestion
effects which may occur during any disruption in the network, and solve optimally to efficiently flow the freight along the intermodal network consisting of three surface transportation modes.

Once there is a disruption in the transportation infrastructure or any particular region, better traffic management during response and recovery stages can reduce its impact on the regional or national economy and security. The results of our models are expected to enhance surface transportation resilience and security by helping to prioritize protection initiatives during preparedness stages and to better manage freight flow re-routing during the response stages.

1.2 Background

Today, U.S. transportation industry requires an efficient and reliable freight transportation system to remain competitive in global economy, ensure national security and provide better resilience during any disaster. Since our research interest includes rail and intermodal freight transportation, we need to investigate the background of these two areas.

Railroad Industry faces numerous problems and decisions to be made at different levels: strategic, tactical, and operational, which are generally more complex compared to other modes of transportation. The hierarchical structure of decision problems taking place in rail freight shipment at different levels is presented by Assad [21] and Kwon [22]. There are numerous decisions involved that are very complex in nature. Strategic decisions involve resource acquisition like network expansion, increasing line capacities; building, expanding, and closing stations and yards, purchasing new locomotives, railcars, and other pricey equipments; tactical decisions include train routing and make-
up policy, classification policy and maintenance plans; operational decisions cover train dispatching, yard operations, empty railcar movement, locomotive and crew scheduling, etc. [23]. We will focus our review on optimization models used in tactical and operational areas of railroad operations. While the majority of railroad transportation problems can be modeled and solved using mathematical optimization techniques, the related research was not very successful in the past as they failed to incorporate the real-life characteristics (Assad [24] and Haghani [25]). The development of optimization models was stalled for a long time due to the large size of the problems, lack of optimization software and limited computing capabilities of computers. In the recent past, with the ever-increasing computing capacity of computers and due to increased competition among the freight industries, have motivated the researchers to develop and use of optimization models at various levels of railroad transportation.

In the recent years, many optimization models relating a number of aspects in rail operations, like train routing, blocking plan, block-to-train assignment (BTA), locomotive and crew scheduling, etc. are available in OR literature. For instance, some recent articles on railroad blocking problem that decides which blocks are to be classified at each yard, and which railcars are to be added in each block are performed by Newton et al. [26], Barnhart et al. [27], and Ahuja et al. [28]. Models for train scheduling problem that describes the train routing and time-tableing issues are done by Farvolden and Powell [29], Campbell [30], Kraft [31], Brannlund et al. [32] and Jha [33]. The references on BTA problem that includes which blocks to be transported by each train are by Kwon et al. [34] and Jha et al. [35]. Literature found on crew management and relevant issues are by Caprara et al. [36] and Liu [37] and on locomotive scheduling problem are by Ziarati
et al. [38], Ziarati et al. [39] and Liu [37]. There are some integrated models found in literature that concurrently decide train routing and BTA, known as “Train Design”. Relevant models can be found in Crainic and Rosseau [40], Haghani [41], Keaton [42 - 45], Gorman [46], Newman and Yano [47, 48] and Dorfman and Medanic [49].

All the above mentioned models focused on the various issues commonly faced by the railroad companies. But if there is any disaster that becomes a national emergency issue and major transportation infrastructure is disrupted, in that case all the railroad companies may need to work together for the quick resilience. It may result in congestion in the network. In this dissertation, we will present both the unit train routing and train design problem that will consider the congestion issue constraint to other regular issues faced by the railroad companies.

Another goal of this dissertation work is to develop an optimization model and solution procedure that will efficiently route the traffic of different origin-destination (OD) pairs in the intermodal network (consisting of highway, waterway and railway), taking into account the congestion characteristics of individual modes. We review the related works done so far in the literature in order to justify our work. Compared to other individual transportation modes, OR models and solution algorithms for intermodal freight transportation is still an adolescent field. There are few review papers available in literature that provide an overview of planning decisions and proposed solution methods involved in intermodal freight transport. Reviews on these issues may be found in Macharis and Bontekoning [50], Crainic and Kim [9] and Caris et al. [51]. They identify the planning problems in intermodal freight transportation consists of four types of decision makers: drayage, terminal, network and intermodal operators. Our focus covers
the literature related to *intermodal operators*, which select the optimal routing of shipments through the whole intermodal network [51]. Literature related to this field of intermodal transportation is found by Barnhart and Ratliff [52], Boardman et al. [53], Ziliaskopoulos and Wardell [54], Min [55] and Erera et al. [56]. In this dissertation work, we develop an optimization model that determine the optimal routing of shipments for different OD pairs considering the effect of congestion in the overall network, similar to railroad transportation model mentioned above.

**1.3 Dissertation objectives and Contributions**

The focus of this dissertation work is to determine the criticality of the freight transportation infrastructures, and also optimal routing/re-routing of freight traffic in the disrupted network. The experimental studies focus on the Class I freight railroad and intermodal freight transportation system, consisting of aforementioned three surface transportation modes of the United States.

For railroad freight transportation, we present two optimization models. First one considers only unit trains of different OD-specific pairs to be transported in the disrupted network. We have considered only link capacity of the network; the capacity of the yard/station is ignored. As mentioned earlier, due to disruption in the physical network, congestion is desirable due to re-routing of the original demand flow in the disrupted network. So, we have incorporated the congestion effects in our model. As a result, the objective function of the train-only model contains a nonlinear function. We have approximated the function as piece-wise linear to make the model linear. We have solved the approximated optimization model with the commercial solver, CPLEX 9.0. We have determined the criticality of the links, and also re-routing of each individual train in the
disrupted network. Our second model for the railroad optimization is more detailed. We consider both link and yard/station capacities. This model is referred as train design optimization model. We have developed a MIP model, and it is a hard combinatorial optimization problem. We have also considered the congestion effect in our model that makes the model more complex to solve optimally. For a small instance, this model can be solved optimally using CPLEX 12.3, but it cannot solve a moderate sized problem optimally using the solver. Hence we developed a heuristic method to solve a real-life train design problem. Our heuristic solution provides good performance with respect to both time and quality of solution.

For intermodal freight transportation system, we present an optimization model that provides the optimal flow of traffic for all OD specific shipments. It will also be able to re-route all commodities in a disrupted network, and therefore we have incorporated congestion effect in the developed model. Due to the nature of the problem, the optimization model is formulated as a MIP model. We have also solved the model using CPLEX 12.3. As MIP models are computationally complex and time consuming for the large scale problem, we tried to solve it using linear relaxation technique. The linear relaxation to this MIP model provides a very good lower bound, which is close to the optimal solution. Therefore, we implement our real-world case study applying the linear relaxation technique to the original optimization model.

We summarize the contributions made in this dissertation below:

- We formulate two models for railroad criticality analysis with re-routing optimization: one considers link-only capacity based on ‘unit’ train re-
routing, and another one considers all infrastructures’ capacities based on train design approach.

- For the first model of railroad criticality analysis, the objective function is an exponential function that accounts the congestion effect due to disruption in the network, and hence a non-linear programming problem. We approximated the nonlinear objective function as piece wise linear to make the model linear programming (LP) problem and the model becomes computationally tractable for even a large size problem.

- For the second model of railroad criticality analysis, we develop an MIP model that also considers the congestion factor into account. It is a combinatorial optimization problem and solves optimally for small-sized problems using commercial optimization solver CPLEX 12.3. As it is a combinatorial optimization problem it cannot solve a large scale problem optimally, and hence we develop a heuristic solution procedure that can handle the large scale problems with reasonable amount of times, and gives near optimal solutions.

- We develop an optimization model for criticality analysis of the intermodal network infrastructures, and determined optimal re-routing of the traffic in the disrupted network. Our contribution to this research is to incorporate the congestion effect in our model that may occur due to infrastructure damage during any emergency situation. We have implemented a real-world case study for the U.S. intermodal network by
applying linear relaxation technique to the MIP model that provides a very good solution, near to optimality.
CHAPTER II
LITERATURE REVIEW

2.1 Railroad transportation system

Compared to other competitive modes of freight transportation, U.S. railroad industry plays a significant role by providing efficient and cost-effective freight services for the shipment of different commodities. At the same time, it also faces numerous problems which are usually interrelated and complex in nature. Therefore, it needs decisions to be made at different levels of planning operation. These levels are broadly classified into three categories based on planning horizon and the level of decision making: strategic, tactical, and operational. The hierarchical structure of these levels is proposed by Assad [21] and Kwon [22]. Typical decisions in each category and their interrelationships are shown in Figure 2.1. As our dissertation work covers optimization of train design and routing, we will focus on the literature related to OR models and methods applied to tactical and operational areas of railroad industry. Although U.S. freight railroad industry has been in existence for over two centuries, it has not benefitted from OR tools and methodologies for many years because of the high level of complexity in railroad operations and also the size of the real-life problem. Earlier attempts of optimization in this field were not very successful as they failed to incorporate the real-life characteristics (Assad [24], Haghani [25], and Newman et al. [57]). Due to the lack of optimization software and limited computing capabilities of computers, research on the
problem of railroad optimization has developed slowly in the past. However, in the recent years, due to increased profit incentives in railroad, competition among the freight industries and also ever-increasing computing capacity of computers, railroads have started to implement optimization-based decision support systems to solve their real-life problems, and researchers have been motivated to develop optimization models and solution approaches at various levels of railroad transportation.

Figure 2.1 Hierarchical structure of decisions in railroad operations [22]
In the following, we will provide a brief overview of railroad planning and scheduling problems that are available in OR literature.

2.1.1 Railroad blocking problem

This is one of the most important and initial planning problems in railroad operation that involves the grouping of shipments to form blocks. In general, a shipment, consisting of a set of individual cars, passes through several classification yards from its origin to destination. To avoid individual shipment to be reclassified at every yard in its path as they travel over the railroad network, a set of shipments are grouped together to create a block. Once a shipment is attached to a block, it is not reclassified at any yard until it arrives at the destination of that block. In the blocking problem, the aim is to identify the classification plan for all shipments that will minimize the total shipment cost in the network. This is commonly known as the blocking plan. In short, the blocking plan aggregates a vast number of shipments into blocks of shipments as they move from their origins to destinations to minimize the overall intermediate handling costs. The literature related to the blocking problem is provided below.

One of the first models in this area is developed by Bodin et al. [58]. They formulated a nonlinear MIP problem for solving the blocking problem. This is modeled as a multi-commodity flow problem with some additional side constraints. It concurrently determines the optimal blocking plans for all the classification yards in a network. The model is solved using a commercial software and provides an upper bound on the number of cars that can be classified in any given yard. Some of the other earlier papers in this area are done by Van Dyke [59, 60], Keaton [42, 44], Huntley et al. [61] and Gorman [46]. Van Dyke [59, 60] proposed a heuristic approach to improve an existing blocking
plan that solves a series of shortest-path problems, with each arc in the network represents an available block. Keaton [42, 44] provided a mixed integer nonlinear programming (MINLP) formulation to solve the blocking problem; Huntley et al. [61] and Gorman [46] proposed meta-heuristic solution algorithms: simulated annealing and genetic algorithm for blocking and train scheduling, respectively.

Some of the recent articles on this area are by Newton et al. [26], Barnhart et al. [27], and Ahuja et al. [28]. Newton et al. [26] formulated the problem as a MIP, and proposed column generation and branch-and-price algorithms to solve the model. Barnhart et al. [27] proposed the Lagrangian relaxation technique to decompose the problem into two sub-problems to solve the same model developed by Newton et al. [26]. Most recently, Ahuja et al. [28] formulated the model as a MIP problem and developed an algorithm known as very large-scale neighborhood (VLSN) search to solve the model. It solves the problem to near optimality within reasonable amount of computational time of computer.

2.1.2 Train routing, timetabling and dispatching problem

Once the blocking plan is known, the next important step is to design a train schedule or train timetable in order to carry the blocks from its origin to destinations with the objective of minimizing the overall shipping cost. Designing train schedule is to identify train routes, and their timetables in an efficient manner that will minimize the total cost of shipping blocks. Once it is developed, it is normally repeated every period of a given time horizon. After the train timetable is fixed, the train dispatching problem identifies a possible plan of meets and passes that has to satisfy a set of additional constraints on the train operations.
Few recent papers in this area include Farvolden and Powell [29], Campbell [30], Kraft [31] and Brannlund et al. [32]. Farvolden and Powell [29] provide a multi-commodity network flow model for train scheduling problem and develop a local-improvement heuristics based upon sub-gradient methods for solving shipment routing sub-problem. Campbell [30] develops a process for scheduling on trains in an intermodal environment with the objective to improve the reliability of railroad service commitments, increase asset utilization, and broaden the service options available to customers using simulation. Kraft [31] developed a methodology to collectively analyze the service offer and car movements in order to maximize profits using a “bid price” revenue management approach, in which bid prices solve a “Train Segment Pricing” problem on a multi-commodity space time network using a standard sub-gradient step size procedure. Brannlund et al. [32] provide a timetabling procedure on a single rail track, whereas the track is divided into segments of one train-capacity. It gives the train itineraries with the focus on departure time and enroute delay. They developed a Lagrangian heuristic to solve the problem relaxing the segment capacities.

2.1.3 Block-to-train assignment (BTA) problem

Sample After identifying the blocking plan and train schedule, the following step is to determine which trains should carry which blocks to move the shipments of different OD pairs and it’s known as the block-to-train assignment (BTA) problem. The objective is to ship all blocks with minimum cost satisfying all the capacity and operational constraints and it is possible that a block can be carried by a number of trains as it moves from its origin to the its destination.
For many years, most railroad companies handled the BTA problem manually due to the complexity of the problem to be solved optimally. In the literature, efforts to solve exclusively BTA problem using optimization algorithms are very few. Nozick and Morlok [62] developed an integer linear programming model to address a finite horizon, discrete time problem of minimizing the cost of moving cars for a fixed train schedule, while satisfying the due dates. They propose a heuristic procedure that solves linear programming relaxation and rounding few of the resulting fractional values iteratively until a integral feasible solution is found and it provides very good solutions, within 1% of the known optimal solution to the relaxed problem. Kwon et al. [34] address a dynamic freight car routing and scheduling model that improve a given blocking plan and block-to-train assignment problem using a time-space network representation technique. They formulate the problem as a linear multi-commodity flow problem, and use the column generation technique to solve the model.

There are also few models existing in the literature that indirectly consider the BTA problem in their formulations. One of the earlier works in this area is done by Thomet [63]. In this paper, he proposes a cancellation procedure that helps to progressively replace the direct shipments with a series of intermediate train connections so that operational and delay costs are minimized. Crainic et al. [64] developed a nonlinear, mixed integer, multi-commodity flow model that deals the problems of routing freight traffic, scheduling train services and allocating classification work between yards on a rail network. They proposed a heuristic algorithm to solve the model and tested an example at the Canadian National Railroads. Keaton [42] addresses the problem of car blocking, train routing and makeup by developing a MIP model that minimizes the sum
of train costs, car time costs, and classification yard costs, while satisfying the constraints on train size and yard volumes. He proposes a Lagrangian relaxation based heuristics to solve the model. Jha et al. [35] address two formulations for the BTA problem: an arc-based formulation and a path-based formulation; whereas they developed exact and heuristic algorithms for the path-based formulation as this formulation can better handle practical constraints. Their heuristic algorithms include both Lagrangian relaxation-based method and greedy construction method.

2.1.4 Combined BTA and train routing problem

Sample Train design, consisting of BTA and train routing, is one of the most difficult and fundamental problems faced by the freight rail companies. It is also known as train scheduling design problem. The inputs for this problem are the blocking plan and the physical rail network. In the train design optimization, the objective is to identify the optimal routes for the trains, and associated BTA, satisfying different capacity and operational constraints. Due to the nature of this problem, it is highly combinatorial and complex optimization problem. There are several attempts found in the literature to solve special cases of the problem. In the following, we will present those efforts available in the literature.

Morlok and Peterson [65] propose one of the earliest models for combined routing and scheduling decisions. They develop timetables: departure time, routing, intermediate stops and assigning of freight to each potential train. The model is developed as a multi-commodity flow problem and solved using branch and bound algorithm for a small instance. Assad [66] proposes a hierarchical viewpoint of the modeling issues and develops a class of models for car routing and train makeup problem using the concept of
network flows and combinatorial optimization. He suggests Bender’s decomposition or Lagrangian relaxation as the possible techniques to solve the models but he doesn’t solve it. Crainic [67] proposes to different methods: a deterministic simulation method and a network optimization method for train routing, classification and make-up policies in yard at the tactical planning level. He develops a computer-assisted network analysis tool, named ‘CANAT’, for the simulation of tactical level of train operations and tested for a commercial railroad company. It consists of three modules that interactively create trains to meet the demands for traffic shipments by checking the yard workloads. He also formulates a mixed integer, nonlinear, multi-flow, multimode model that considered the delay issue at the yard and link travel during classification and routing of trains. He solved the model for a small instance using decomposition method. Haghani [41] presents a mixed-integer programming problem with nonlinear objective function and linear constraints for a combined train routing and makeup, loaded and empty car distribution problem. He developed a heuristic procedure based on decomposition technique that divides the problem into smaller sub-problems to solve the model, and tested on a small instance problem.

More recently, Keaton [42] formulates an integer programming optimization model for determining train connections, its frequencies, blocking and routing plans for general commodity shipments. He considered train travel cost, car travel cost and sorting cost into the objective and train size, yard capacity and maximum transit time for each OD pair as restrictions. He implemented the Lagrangian relaxation technique using a dual adjustment procedure to solve the problem. Marín and Salmerón [68] formulate a non-convex, aggregate freight planning model that concurrently decide the optimal
assignment of the trains to the service network including their stops, frequencies and number of cars using each service. They develop different heuristics to solve the model. It includes Simulated Annealing, Tabu Search and a special heuristic, named as 'Descending method'. Among the heuristics, Simulated Annealing provides the best solution but requires more computational time, whereas descending method shows better efficiency compared to Tabu search method for large networks. Newman and Yano [47, 48] formulate an integer programming model that simultaneously determines the schedule for both direct and indirect trains and assigns intermodal containers to these trains. They develop a novel decomposition procedure to solve the problem that provides near optimal solutions faster (less than one-third) than the optimization software. They also develop a method based on valid inequalities to get the tight lower bounds to the original model. Jha [33] develops a three phase decomposition approach to solve a zero-based train design problem, where decisions are made at the planning level about the trains to be made, their frequency, timings and their routes in the physical network. In Phase I, II and III, the author solves the network design problem, network flow problem and scheduling problem, respectively. Most recently, Yue et al. [69] provide an integer programming formulation for combined blocking and train routing problem that minimizes the operating costs of shipping all traffic on the railway network while satisfying all resource and capacity constraints and the priority constraints for shipments. They propose an improved Ant Colony (AC) algorithm for solving this model and tested on a real life problem of coal heavy haul rail network in China.

To summarize, we can state that a good number of OR models and heuristic solution procedures have been developed recently in the literature of railroad planning.
and scheduling problems considering several aspects: blocking, train routing, BTA and combined BTA & train routing. These papers have developed various solution approaches, for example, relaxation, decomposition and heuristics procedures to solve the complex problems of real-life railroad planning and scheduling operations. There are also some other issues available in OR literature of railroad operations – crew scheduling, locomotive scheduling, empty car distribution, etc. that we didn’t overview in the literature review as those are not relevant to our dissertation work. In our dissertation work, we have developed two models: one considers only for train routing and another one is an integrated model that considers combined BTA and train routing, also referred to as Train Design. The uniqueness to our models is incorporation of congestion issue that is a growing concern in railroad freight transportation in recent years. Besides, as the ultimate goal of the dissertation work is to determine critical infrastructure of the railroad network based on re-routing optimization, we need to consider the disruptions to the existing network and this will obviously affect the network with increased congestion. In our survey to the literature related to train routing and train design optimization, we hardly find the inclusion of congestion issues in their models. This factor makes our model more complex than the existing models in the literature.

2.2 Intermodal transportation system

As stated earlier, intermodal transportation is a very young but an emerging field in transportation research. Since traditional freight transportation system has been facing several problems (e.g., congestion, environmental concerns and traffic safety) with the increase in overall freight flow and also the growing concern of speed, agility and cost efficiency in supply chain, intermodal freight transportation has received increased
attention in recent years. There are couple review papers available in the literature that addresses the domain of problems and related operations research models existing in intermodal freight transportation. The reviews may be found in Crainic et al. [9] and Caris et al. [51]. According to these reviews, the planning problems in intermodal freight transportation system can be grouped into four types of decision makers or operators, based on the four main activities in this transportation system. They are *drayage*, *terminal*, *network* and *intermodal operators*. *Drayage operators* are responsible for the planning and scheduling of trucks between terminals and shippers and receivers; *terminal operators* are accountable for managing the transshipment operations from road to rail or barge, or from rail to rail or barge to barge; *network operators* are liable for the infrastructure planning and organization of rail or barge transport and *intermodal operators* are users of the intermodal infrastructure and services and identify the suitable routes for shipping the goods along the whole intermodal network. Since we are interested in optimal routing of shipments through the whole intermodal network, we will concentrate our literature related to intermodal operators that select the appropriate path for the shipments in the intermodal network. Interested readers for the other three operators (drayage, terminal and network) of intermodal transportation are referred to [9, 51].

Guélat et al. [70] formulate a multi-mode multi-product network assignment model to determine the optimal routes for the products in the intermodal network between the given origins and destinations. The modeling framework is made up of road and rail modes, nodes, links, and intermodal transfers, where the multiple commodities are to be shipped from their origins to their destinations using this multimodal network.
They develop a Gauss-Seidel-Linear Approximation Algorithm to solve the model and tested on a real-life multi-modal network. Barnhart and Ratliff [52] formulate a model that identifies the minimum cost intermodal routings in order to optimize total transportation costs on the truck and/or rail combination of intermodal network. They developed an algorithm to identify two types of decisions based on who owns the equipment and who is providing the service. Boardman et al. [53] develop a decision support system (DSS) that uses a robust analytical method to determine the optimal routing of commodities to help shippers in selecting the best combinations of intermodal transportation modes based on cost, service level, and the nature of the commodity being shipped. The model used to develop the DSS is based on the K-shortest path algorithm that minimizes both cost and time as opposed to distance between the origins and destinations. The DSS has an interface with the commercial geographic information system (GIS) software package to help users in visualizing the result. Ziliaskopoulos and Wardell [54] develop a time-dependent intermodal least-time path (TDILTP) algorithm on a multidimensional network that computes optimum routes from all origins to all destinations with time-dependent arc travel times and switching delays and without explicitly expanding the network. They design a preprocessor that constructs the required inputs from common transit timetables and a simple data structure that significantly improves the efficiency of the algorithm. They tested the algorithm on a realistic size network. The complexity of the algorithm doesn’t depend on the number of modes and computational time is almost linearly related to the number of nodes in the network and the number of time intervals. Min [55] proposes a chance-constrained goal programming model that identifies the most effective intermodal routes with the objective of
minimizing cost and risk by satisfying the on-time service requirements. The transportation modes are distinguished based on costs, market coverage, equipment capacity, speed, reliability and risk. Arnold et al. [71] present an integer linear programming (ILP) formulation to optimally locate the intermodal transfer terminals for an integrated railroad and highway transportation system. Their formulation is based on the multicommodity fixed-charge network design problems, which is an alternative to the general hub location formulation. They develop a heuristic procedure, referred as ITLSS (Intermodal Terminals Location Simulation System) to solve the ILP formulation. More recently, Erera et al. [56] formulate an integrated model for shipping loaded tank containers and repositioning empty tank containers on an intermodal network. The deterministic model is formulated based on time-space network. Their study indicates that integrated container management can significantly reduce empty container repositioning costs.

In summary, we can claim that optimization models related to intermodal routing of commodities are not rich in literature compared to other traditional individual modes of transportation. This is due to the fact that the concept of intermodal transportation is a very young domain. Though the OR models in this area are not affluent, we can see some excellent works in optimal routing of commodities that cover several issues of intermodal transportation in the literature. They have developed several heuristic and decomposition methods to solve the model near to optimality. In our dissertation work, we propose an optimization model that ships all commodities from their origins to their destinations using the intermodal network. In the proposed model, we have incorporated the congestion effects for links of individual modes, intermodal transfers and also locks of
waterway system based on their characteristics that may happen due to the heavy traffic and also re-routing of traffic due to disruption (both natural or man-made) in the network. The model is expected to re-route the traffic optimally in the disrupted network. We have implemented a real-life case study of U.S. intermodal network to determine the critical infrastructures of the network using the optimization model.
CHAPTER III
DATA AVAILABILITY

In order to test and validate the developed models and solution approaches, and also to implement the case studies, we basically need two major types of data: network data and freight flow data. As stated earlier, in order to evaluate criticality of network infrastructures, we propose two models for railroad and one model for intermodal transportation system. For all of these models, we use three different sets of network and freight flow data, all of which belong to U.S. freight transportation system. In the following sections, we briefly discuss about these two categories of data used for each case study, and the sources of these data from where they are collected.

3.1 Network data

Network data includes physical attributes and operating characteristics of particular freight transportation network that determines the capacity of that network. Physical attributes comprise distance between nodes (e.g., stations, yards, ports, locks, intermodal transfer terminals), number of tracks/lanes, geographical locations, and operating characteristics include signal types, traffic systems, etc. Based on these attributes and characteristics, the capacities of the network components of each individual mode are determined. For each individual model and case study, sample network data and their sources are described in the following sub-sections.
3.1.1 Railroad network data for criticality analysis using unit train re-routing optimization

The physical network, used to determine the criticality of links using unit train routing optimization, is relatively a small network. It includes Class-I railroad network of Mississippi and its adjacent four states (Alabama, Tennessee, Arkansas and Louisiana) partially. We consider only the major stations/yards as the nodes and the capacities of these nodes (stations/yards) are assumed to be unlimited (i.e., uncapacitated) in this study. Only link (track) capacities are considered to implement the study. The network contains approximately 50 major rail stations/yards (nodes) and 70 links. The link attributes (length and number of tracks) and track operating characteristic (e.g., signal types) are collected using the software North American Railroad Map, Version 3.11, which reflects the North American railroad network in 2010 [76]. Using these link features, we estimated the capacity in the number of trains per day based on [77]. Table 3.1 provides sample network data used in this study.

3.1.2 Railroad network data for criticality analysis using train design optimization

This study uses the network of one of the major Class-I railroad industries of USA. The network has approximately 400 nodes and 500 links. The network attributes and operating characteristics are collected from the same source as mentioned in the previous sub-section (Sub-section 3.1.2). Using on these characteristics, link capacities are estimated similarly in terms of trains per day, based on [77]. Besides, other parameters (e.g., length and tonnage restrictions allowed for each travelled train) of each link are considered in this study. Information available at the railroad’s website aided this
process. A sample network data for links, used in this case study is presented in Table 3.2.

Table 3.1 Sample data of Class-I railroad network attributes for unit train routing

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Length (miles)</th>
<th>Capacity (trains/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexandria</td>
<td>Baton Rouge</td>
<td>105</td>
<td>16</td>
</tr>
<tr>
<td>Amory</td>
<td>Birmingham</td>
<td>110</td>
<td>30</td>
</tr>
<tr>
<td>Amory</td>
<td>Columbus</td>
<td>37</td>
<td>16</td>
</tr>
<tr>
<td>Bald Knob</td>
<td>Little Rock</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Baton Rouge</td>
<td>New Orleans</td>
<td>78</td>
<td>36</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Tuscaloosa</td>
<td>48</td>
<td>30</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Montgomery</td>
<td>118</td>
<td>46</td>
</tr>
<tr>
<td>Brinkley</td>
<td>Pine Bluff</td>
<td>77</td>
<td>30</td>
</tr>
<tr>
<td>Columbus</td>
<td>Demopolis</td>
<td>84</td>
<td>16</td>
</tr>
<tr>
<td>Hattiesburg</td>
<td>Mobile</td>
<td>92</td>
<td>16</td>
</tr>
<tr>
<td>Corinth</td>
<td>Tupelo</td>
<td>42</td>
<td>30</td>
</tr>
<tr>
<td>Camden</td>
<td>Texarkana</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>Chattanooga</td>
<td>Birmingham</td>
<td>130</td>
<td>18</td>
</tr>
<tr>
<td>Memphis</td>
<td>Tupelo</td>
<td>115</td>
<td>30</td>
</tr>
</tbody>
</table>

We also include the station/yard capacities (in terms of number of cars and trains) in this study. These are derived from the railroad’s websites and correspondence with
railroad personnel. A sample network data for nodes (stations/yards), used in this case study is presented in Table 3.3.

Table 3.2 Sample data of Class-I railroad link attributes

<table>
<thead>
<tr>
<th>Origin ID</th>
<th>Destination ID</th>
<th>Distance (miles)</th>
<th>Capacity (feet)*</th>
<th>Capacity (tonnage)*</th>
<th>Capacity (trains/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>40.1</td>
<td>4905</td>
<td>12870</td>
<td>30</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>62.4</td>
<td>5355</td>
<td>14220</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>58.7</td>
<td>4905</td>
<td>12870</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>67.4</td>
<td>3825</td>
<td>11880</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>69.8</td>
<td>4905</td>
<td>12870</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>31.0</td>
<td>4905</td>
<td>12870</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>25.9</td>
<td>4905</td>
<td>12870</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>48.0</td>
<td>4905</td>
<td>12870</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>21.0</td>
<td>5355</td>
<td>14220</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>62.4</td>
<td>5355</td>
<td>14220</td>
<td>30</td>
</tr>
</tbody>
</table>

*Capacity for each individual train traveling through the link

3.1.3 Intermodal network data

Intermodal network data, consisting of surface transportation modes: highway, railroad, and waterway, along with intermodal terminals, are collected from the Center for Transportation Analysis (CTA) at Oak Ridge National Laboratory (ORNL) [84]. The CTA at ORNL maintains a comprehensive database of North American transportation infrastructure of all modes. For each mode of the intermodal network, we collect the link
attributes and operating characteristics from CTA database [84]. For each link, for instance, we collect the length, number lanes, and types for highway; the length, number of tracks and signal types for railway; lock characteristics, type of inland waterway for waterway. Using on these attributes and characteristics, we estimate the link capacities of each mode, in terms of TEUs (Twenty-foot equivalent unit) per day, of the links for each mode based on [77, 85, 86]. The capacities of the intermodal terminals, measured in TEUS per day, are derived partially from the intermodal sections of the railroad’s website and also correspondence with relevant authority.

<table>
<thead>
<tr>
<th>Node ID</th>
<th>Capacity (cars/day)</th>
<th>Capacity (trains/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2000</td>
<td>48</td>
</tr>
<tr>
<td>1</td>
<td>1600</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>2200</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>1500</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>1200</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>2600</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>1500</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>1200</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>1800</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>1500</td>
<td>32</td>
</tr>
</tbody>
</table>
In this study, we use the intermodal network comprising five states of USA: Alabama, Arkansas, Louisiana, Mississippi, and Tennessee. The network consists of approximately 5500 nodes and over 6800 links. Among these nodes, there are over 200 intermodal transfer terminals. A sample intermodal network data for links used in this case study is presented in Table 3.4. Also, a sample network data for intermodal transfer terminals, used in this case study, is presented in Table 3.5
<table>
<thead>
<tr>
<th>Origin ID</th>
<th>Destination ID</th>
<th>Direction</th>
<th>Capacity (TEUs/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3618</td>
<td>3647</td>
<td>2</td>
<td>4000</td>
</tr>
<tr>
<td>3646</td>
<td>3558</td>
<td>2</td>
<td>5000</td>
</tr>
<tr>
<td>3612</td>
<td>3886</td>
<td>2</td>
<td>9600</td>
</tr>
<tr>
<td>66</td>
<td>389</td>
<td>2</td>
<td>15023</td>
</tr>
<tr>
<td>3722</td>
<td>3586</td>
<td>2</td>
<td>16000</td>
</tr>
<tr>
<td>3605</td>
<td>3881</td>
<td>2</td>
<td>20000</td>
</tr>
<tr>
<td>348</td>
<td>352</td>
<td>2</td>
<td>30046</td>
</tr>
<tr>
<td>4894</td>
<td>4895</td>
<td>2</td>
<td>78000</td>
</tr>
<tr>
<td>4926</td>
<td>4920</td>
<td>2</td>
<td>62000</td>
</tr>
<tr>
<td>5687</td>
<td>3246</td>
<td>1</td>
<td>52399</td>
</tr>
<tr>
<td>547</td>
<td>559</td>
<td>2</td>
<td>30046</td>
</tr>
<tr>
<td>380</td>
<td>951</td>
<td>2</td>
<td>15023</td>
</tr>
<tr>
<td>5180</td>
<td>5567</td>
<td>1</td>
<td>62000</td>
</tr>
<tr>
<td>5567</td>
<td>4132</td>
<td>1</td>
<td>52399</td>
</tr>
<tr>
<td>4014</td>
<td>5572</td>
<td>1</td>
<td>62399</td>
</tr>
<tr>
<td>5572</td>
<td>1116</td>
<td>1</td>
<td>72399</td>
</tr>
<tr>
<td>4449</td>
<td>4147</td>
<td>2</td>
<td>5000</td>
</tr>
<tr>
<td>4149</td>
<td>4150</td>
<td>2</td>
<td>9600</td>
</tr>
</tbody>
</table>
Table 3.5 Sample network data for intermodal transfer terminals

<table>
<thead>
<tr>
<th>Intermodal transfer node ID</th>
<th>Capacity (TEUs/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5258</td>
<td>2400</td>
</tr>
<tr>
<td>5268</td>
<td>2450</td>
</tr>
<tr>
<td>5288</td>
<td>2550</td>
</tr>
<tr>
<td>5309</td>
<td>2200</td>
</tr>
<tr>
<td>5325</td>
<td>2500</td>
</tr>
<tr>
<td>5366</td>
<td>2600</td>
</tr>
<tr>
<td>5386</td>
<td>2700</td>
</tr>
<tr>
<td>5467</td>
<td>2600</td>
</tr>
<tr>
<td>5479</td>
<td>2800</td>
</tr>
<tr>
<td>5516</td>
<td>2800</td>
</tr>
<tr>
<td>5687</td>
<td>2800</td>
</tr>
<tr>
<td>5688</td>
<td>2800</td>
</tr>
<tr>
<td>5689</td>
<td>2800</td>
</tr>
<tr>
<td>5690</td>
<td>2300</td>
</tr>
</tbody>
</table>

3.2 Freight flow data

In this section, we discuss various OD specific freight flow data that are used in implementing the models for case studies along with their sources. All freight flow datasets are collected from Freight Analysis Framework (FAF) database. This data source is maintained by the Office of Freight Management and Operations of Federal Highway
Administration (FHWA) which is a branch of US department of transportation (USDOT) and funded by federal government for its overall management. So, data from this source are publicly available for free. It combines data from several sources to create a comprehensive image of freight movement among states and major zones by all transportation modes. FAF version 3 provides the most recent freight flow data of tonnage and value, by commodity type, mode, and OD, and we collect data from this version for all our case studies. In the following sub-sections, a brief description of freight flow data for each case study is presented.

3.2.1 Railroad freight flow data for criticality analysis using unit train re-routing optimization

The flow data used in this case study covers the area of Mississippi and adjacent four states: Alabama, Tennessee, Arkansas, and Louisiana, partially. We also consider the through traffic for these five states in order to have a realistic essence of the load on the test network. The database provides data based on FAF zones to zones OD flow in thousands of tons. This study uses 2009 rail freight data based on FAF zones to zones OD flow by tonnage from [72]. Then the data is converted into node (station/yard) to node (station/yard) OD flow in the units of trains per day based on [73]. A sample of unit trains flow data is presented in Table 3.6.
Table 3.6  Sample unit train flow data

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>KTON 2009</th>
<th>Demand (trains/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Orleans</td>
<td>Birmingham</td>
<td>5922.9006</td>
<td>4</td>
</tr>
<tr>
<td>New Orleans</td>
<td>Nashville</td>
<td>3770.8709</td>
<td>3</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Shreveport</td>
<td>2710.7119</td>
<td>2</td>
</tr>
<tr>
<td>Montgomery</td>
<td>New Orleans</td>
<td>1919.89245</td>
<td>2</td>
</tr>
<tr>
<td>New Orleans</td>
<td>Memphis</td>
<td>1998.918</td>
<td>2</td>
</tr>
<tr>
<td>Shreveport</td>
<td>Birmingham</td>
<td>2103.8464</td>
<td>2</td>
</tr>
<tr>
<td>Baton Rouge</td>
<td>Birmingham</td>
<td>2024.2187</td>
<td>2</td>
</tr>
<tr>
<td>Memphis</td>
<td>West Point</td>
<td>1865.11717</td>
<td>2</td>
</tr>
<tr>
<td>Memphis</td>
<td>Hattiesburg</td>
<td>1988.93804</td>
<td>2</td>
</tr>
<tr>
<td>Memphis</td>
<td>Jackson</td>
<td>2066.46912</td>
<td>2</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Little Rock</td>
<td>866.7976</td>
<td>1</td>
</tr>
<tr>
<td>Montgomery</td>
<td>Baton Rouge</td>
<td>127.3409</td>
<td>1</td>
</tr>
<tr>
<td>Mobile</td>
<td>Shreveport</td>
<td>68.79995</td>
<td>1</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Memphis</td>
<td>51.0329</td>
<td>1</td>
</tr>
<tr>
<td>Memphis</td>
<td>New Orleans</td>
<td>178.07145</td>
<td>1</td>
</tr>
<tr>
<td>Memphis</td>
<td>Baton Rouge</td>
<td>178.07145</td>
<td>1</td>
</tr>
</tbody>
</table>

*KTON-Kilo (thousand) Ton
3.2.2 Railroad freight flow data for criticality analysis using train design optimization

The flow data used in this study includes a major Class-I railroad industry spanning 28 U.S. states, which covers middle to western part of USA. This study also uses 2009 rail freight data for all OD pairs by tonnage from [72]. We convert the annual tonnage flow into daily number of railcars (i.e., blocks) moving from each origin to each destination based on [74]. Each block has its unique attributes such as origin, destination, number of cars, length and tonnage and these attributes don't change throughout the journey when it travels with different trains. A sample flow data is given in Table 3.7.

Table 3.7 Sample railroad freight flow data for train design optimization

<table>
<thead>
<tr>
<th>Block ID</th>
<th>Origin ID</th>
<th>Destination ID</th>
<th>No. of Cars</th>
<th>Length (feet)</th>
<th>Weight (tonnage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>25</td>
<td>1362</td>
<td>3575</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>545</td>
<td>1430</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>7</td>
<td>16</td>
<td>873</td>
<td>2288</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>6</td>
<td>25</td>
<td>1362</td>
<td>3575</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>655</td>
<td>1716</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>6</td>
<td>19</td>
<td>1035</td>
<td>2717</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>15</td>
<td>83</td>
<td>3528</td>
<td>10956</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>9</td>
<td>78</td>
<td>3315</td>
<td>10452</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>13</td>
<td>50</td>
<td>2725</td>
<td>7150</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>12</td>
<td>30</td>
<td>1440</td>
<td>4020</td>
</tr>
</tbody>
</table>
3.2.3 Intermodal freight flow data

The freight flow data used in this study spans five U.S. states: Alabama, Arkansas, Louisiana, Mississippi, and Tennessee. In addition to different OD flow data within this study area, we also include the through traffic over the study area in order to have realistic picture of the traffic flow. In this research, we assume that all the freights moving within the U.S. boundary can be transported by intermodal network (i.e. using all three surface transportation modes simultaneously). That is why, we combine all the freights moving individually using these three modes, and this intermodal freight flow data is prepared by combining total annual tonnage of freight flow of these three surface transportation modes. This study uses 2011 freight data for all OD pairs from [72]. We convert the annual tonnage movement into total number of TEUs flowing daily from their origins to destinations using the intermodal network. Total number of TEUs shipping daily over the intermodal network is calculated based on [75]. These containers are usually transported by barges, semi-trailer truck, and freight trains as they travel from their origins to destinations without any handling of the freight itself when changing modes; they are transferred between modes by container cranes at intermodal terminals. Table 3.8 provides a sample intermodal freight flow data.
Table 3.8  Sample intermodal freight flow data

<table>
<thead>
<tr>
<th>Shipment ID</th>
<th>Origin ID</th>
<th>Destination ID</th>
<th>Demand (TEUs/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>293</td>
<td>114</td>
<td>935</td>
</tr>
<tr>
<td>1</td>
<td>3594</td>
<td>3748</td>
<td>300</td>
</tr>
<tr>
<td>2</td>
<td>3748</td>
<td>3946</td>
<td>370</td>
</tr>
<tr>
<td>3</td>
<td>270</td>
<td>398</td>
<td>558</td>
</tr>
<tr>
<td>4</td>
<td>878</td>
<td>177</td>
<td>500</td>
</tr>
<tr>
<td>5</td>
<td>269</td>
<td>565</td>
<td>318</td>
</tr>
<tr>
<td>6</td>
<td>3766</td>
<td>4089</td>
<td>64</td>
</tr>
<tr>
<td>7</td>
<td>115</td>
<td>826</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>267</td>
<td>501</td>
<td>219</td>
</tr>
<tr>
<td>9</td>
<td>861</td>
<td>102</td>
<td>152</td>
</tr>
<tr>
<td>10</td>
<td>270</td>
<td>2511</td>
<td>186</td>
</tr>
<tr>
<td>11</td>
<td>369</td>
<td>748</td>
<td>347</td>
</tr>
</tbody>
</table>
CHAPTER IV

RAILROAD CRITICALITY ANALYSIS USING RE-ROUTING OPTIMIZATION

In this chapter, we evaluate the criticality of railroad network infrastructures using freight flow network optimization. The basic idea underlying critical infrastructure identification is to develop optimization models and solution algorithms to efficiently re-route the overall traffic through the disrupted network. In fact, the models and solution algorithms will determine the optimal alternative paths that will minimize total travel time/cost for all traffic in the residual network after disruption. We have developed two different models for railroad re-routing optimization. The first model considers only link (track) capacities of the physical network at the train level. In the second model, we consider the capacity of both links (tracks) and nodes (stations/yards) of the railroad network and blocks with different OD pairs are basis to design and route the trains. We have developed a special heuristic procedure to solve the real-life problem for the second model. In the following sections we present these two models along with the case study results.

4.1 Criticality analysis of links using unit train re-routing optimization under heavy congestion

As an energy-efficient transportation mode, railways play a vital role in U.S. freight transportation [14]. During any natural or man-made disasters, it is essential to keep the freight flow by efficiently re-routing the disrupted traffic. Since railway freight
security and resilience are critical to the U.S. economy and homeland security, it is necessary to develop models to identify rail areas of high consequence and vulnerability, from which railroad stakeholders can better develop tactics to protect the increasingly stressed railway infrastructure. This research work considers the routing of trains to minimize the total travel time for the whole network. The routing problem is formulated as a minimum-cost network flow problem that has a nonlinear objective function of minimizing the total travel time on all links and considers Origin-Destination (OD) specific demand. To make the model computationally tractable, the nonlinear travel time function at each link is approximated with a piece-wise linear function. The criticality of a railway link is evaluated by the increased delay when the link is disrupted. A case study is conducted for the Class I railroad network of the State of Mississippi and its adjacent states (Alabama, Tennessee, Arkansas, and Louisiana) partially.

4.1.1 Problem description

The problem of evaluating the criticality of links for a railway network can be stated as follows. For a given railroad network, we know the capacities of the links and the volume of trains flowing within and through the test network for different OD pairs for a given time period (as described in Chapter 3). The model considers the routing of trains as link-based routing in the network for different OD pairs. In the model, the objective function is to minimize the total routing time of the trains in the network. Each individual link is considered to be disrupted once at a time and we optimize the total routing time. It is worth mentioning that every time the ‘disrupted link’ is removed from the network; meaning that the capacity of that ‘disrupted link’ is regarded as zero. The difference between the objective functions of undisrupted and disrupted situation is the
amount of delay (additional cost) occurred by that disrupted link. We define the most
critical link as the one whose disruption results in maximum delay. The objective
function of this model is considered to be exponential based on delay function developed
by [79]. We simplified the objective function by decomposing it into piece-wise linear
functions. For further simplification and ease of computation, we consider two pieces of
linear functions of the original exponential function. However, it is suggested to make
more pieces of linear functions for exact approximation of the exponential objective
function.

4.1.2 Mathematical model formulation

Consider a graph $G = (V, E)$, where $V$ is the vertex set and $E$ is the bi-directed
link set. In other words, $(k, l) \in E$ implies $(l, k) \in E$. Other notations are defined below.

4.1.2.1 Parameters

- $D^j_i$: demand for an OD pair $(i, j)$, where $i \in V, j \in V$, and $D^j_i = 0$;
- $C^1_{k,l}$: uncongested (normal) marginal time required for unit flow on link $(k, l) \in E$;
- $C^2_{k,l}$: congested marginal time required for unit flow on link $(k, l) \in E$;
- $u_{k,l}$: capacity of a link $(k, l) \in E$;
- $A(k)$: the adjacent node list of node $k$, where $A(k) = \{l \in V : (k, l) \in E\}$;
- $a^k_i$: \begin{align*}
    a^k_i = \begin{cases}
    1, & i = k \\
    0, & i \neq k
    \end{cases}
\end{align*}
    for $\forall i \in V, \forall k \in V$.

$A_0$ & $B_0$: parametric traffic, operating coefficient and constant, respectively.
4.1.2.2 Variables

- $x_{k,l}^{i,j}$ flow on link $(k, l)$ for demand characterized by OD pair $(i, j)$;
- $x_{k,l}$ total flow on link $(k, l)$; where, $x_{k,l} = y_{k,l}^1 + y_{k,l}^2$;
- $y_{k,l}^1$ part of flow on link $(k, l)$ with marginal travel time of $C_{k,l}^1$, which is up to $u_{k,l}$;
- $y_{k,l}^2$ additional flow on link $(k, l)$ with marginal travel time of $C_{k,l}^2$ beyond $u_{k,l}$.

The model can be formulated as follows:

Minimize $\sum_{(k,l) \in E} A_0 e^{B_0x} dx$ \hspace{1cm} (4.1)

Subject to:

\[ a_i^k D_l^j + \sum_{l \in A(k)} x_{l,k}^{i,j} = \sum_{l \in A(k)} x_{k,l}^{i,j} + a_j^k D_l^j \quad i \in V, j \in V, k \in V, i \neq j \] \hspace{1cm} (4.2)

\[ y_{k,l}^1 + y_{k,l}^2 = \sum_{l \in V} \sum_{j \in V, j \neq i} x_{k,l}^{i,j} \quad \forall (k, l) \in E \] \hspace{1cm} (4.3)

\[ y_{k,l}^1 + y_{i,k}^1 \leq u_{k,l} \quad \forall (k, l) \in E \] \hspace{1cm} (4.4)

\[ x_{k,l}^{i,j}, y_{k,l}^1, y_{k,l}^2 \geq 0 \] \hspace{1cm} (4.5)

As mentioned earlier, the nonlinear objective function (4.1) is approximated with the piece-wise linear function so that the total model becomes linear and computation of linear model is relatively easier and faster. After approximation, this nonlinear objective function is converted to the form, Minimize $\sum_{(k,l) \in E}(C_{k,l}^1 y_{k,l}^1 + C_{k,l}^2 y_{k,l}^2)$. This function minimizes the total travel time for the whole system/network. For each link, we consider two components of travel time: normal (uncongested) travel time $C_{k,l}^1$ and congested travel time $C_{k,l}^2$, for normal flow $y_{k,l}^1$ and congested flow $y_{k,l}^2$ respectively. Constraint set
(4.2) enforces the usual flow conservation requirements. Constraint set (4.3) states that the sum of two different piece-wise linear flows in a link should be equal to the total flow in that link. Constraint set (4.4) makes sure the flow volume on link \((k, l)\) with marginal unit travel time \(C_{k,l}^1\) cannot exceed its capacity \(u_{k,l}\). Since \(C_{k,l}^1 \leq C_{k,l}^2\), \(y_{k,l}^2\) will be greater than zero only when \(y_{k,l}^1 = u_{k,l}\). All variables are non-negative as shown by (4.5). The relationship among \(C_{k,l}^1\), \(C_{k,l}^2\), and \(u_{k,l}\) can be shown in Figure 4.1, in which \(TC_{k,l}\) is the total travel time of all traffic on link \((k, l)\), and \(y_{k,l}\) is the total flow on link \((k, l)\), which is equal to \(y_{k,l}^1 + y_{k,l}^2\).

![Figure 4.1  Piecewise linear approximation of the cost-flow function](image)

**4.1.3 Case study**

In this section, we present the case study using the proposed model. As mentioned earlier, the study area includes the Class I railroad network of Mississippi and its adjacent states (Alabama, Arkansas, Louisiana, and Tennessee) partially. It uses the data described
in Chapter 3. In addition, we calculate the data for travel time for each link, and uncongested travel time is calculated based on its distance and normal freight train speed [78]. Congested travel time for each link is calculated by considering arbitrarily as twice the uncongested travel time for that link. A sample of the travel time components for different links, used in this study, is given in Table 4.1.

Table 4.1 Sample train travel time components for different links

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Time¹ (hrs.)</th>
<th>Time² (hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amory</td>
<td>Columbus</td>
<td>3.7</td>
<td>7.4</td>
</tr>
<tr>
<td>Alexandria</td>
<td>Baton Rouge</td>
<td>10.5</td>
<td>21</td>
</tr>
<tr>
<td>Bald Knob</td>
<td>Little Rock</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Baton Rouge</td>
<td>New Orleans</td>
<td>7.8</td>
<td>15.6</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Tuscaloosa</td>
<td>4.8</td>
<td>9.6</td>
</tr>
<tr>
<td>Birmingham</td>
<td>Montgomery</td>
<td>11.8</td>
<td>23.6</td>
</tr>
<tr>
<td>Brinkley</td>
<td>Pine Bluff</td>
<td>7.7</td>
<td>15.4</td>
</tr>
<tr>
<td>Brookhaven</td>
<td>McComb</td>
<td>2.8</td>
<td>5.6</td>
</tr>
<tr>
<td>Brookhaven</td>
<td>Bogalusa</td>
<td>8.5</td>
<td>17</td>
</tr>
<tr>
<td>Jackson</td>
<td>Hattiesburg</td>
<td>9.5</td>
<td>19</td>
</tr>
</tbody>
</table>

Using these data, we have implemented the approximated LP model using C++ on a PC with 2.40GHz CPU and 2 GB RAM, and we call ILOG CPLEX 9.0 to optimize it. First, the existing network containing all links is considered to determine the total optimal travel time for the whole network. This optimal total travel time is regarded as the “base”
travel time for the network. Next, all the links are considered to be disrupted separately, one at a time, and the total travel times are optimized for those disrupted situations. The most critical link is the one that causes maximum delay when that particular link is disrupted in the network.

The criticality of each individual link is shown in Figure 4.2.
The widths of the links represent their level of criticality. The most critical link is the one that has maximum width. This means that disruption to that particular link will result in maximum delay on the total travel time for the whole network. The links with the minimum width are the least critical links of the test rail network. Disruptions to these links will have the least impact on the overall delay. For example, for the rail network of the state of Mississippi and its adjacent states, the top five critical links are Jackson-Monroe, Birmingham-Tuscaloosa, Meridian-Tuscaloosa, Jackson-Newton and New Orleans-Gulfport, sequentially, as shown in Figure 4.2.

4.1.4 Conclusion

This research presents a piece-wise linear approximation of the non-linear programming problem for vehicle routing through the railway network to determine the criticality of the railway links. Our simple translation of the non-linear programming problem into a piece-wise linear program is very effective. Its performance is satisfactory. It is obvious that determining criticality of the railway links is of great importance because it will help the authority to take necessary steps for the security and also resilience of the critical links in the network. It is expected that the proposed model will help the railway management to enhance the preparedness capability relevant to transportation security.

4.2 Criticality analysis of all infrastructures using train design optimization under heavy congestion

In the previous section, we have presented a train routing problem that has some limitations. First, we have considered only ‘unit trains’ for routing that doesn’t represent the real picture of the railroad industry. In practice, it represents a certain fraction of the
overall traffic in railroad industry. Also, stations/yards capacities are ignored but this is a very important component in defining the capacity of the network. Finally, link (track) capacities are restricted only to the number of trains that can travel per unit time. In reality, in addition to number of trains, each link is also restricted to length and weight limitations of each train travel through it.

This section proposes a system-wide optimization model and a solution approach considering the congestion effects and capacity restrictions to route traffic to evaluate the effect of a disruption, and to determine the railroad network infrastructure’s criticality. Similar to previous section, an infrastructure element’s criticality is evaluated by estimating the increased cost (delay) when this element is disrupted. After a disruption, traffic needs to be re-routed over the residual network. The disruption may cause huge congestions, directly or indirectly over a big part of the network and make the flows at many links and yards close to their capacities.

As stated in Chapter2, railroad operations are highly complex, and real-world problems are large in size [24]. Early optimization attempts had limited success and could not incorporate real-life characteristics [21, 25, 57]. In recent years, railroads have started to implement optimization-based decision support systems to address operational issues, such as blocking planning, block-to-train assignment (BTA), train scheduling, crew scheduling, and locomotive scheduling. Modern optimization models and solution approaches are founded on more sophisticated problem formulations, the ever-increasing computing capacity, and much improved data [27, 28, 32, 35]. Freight rail operations often start by developing a blocking plan. It aggregates a vast number of shipments into blocks as they move together to reduce reclassification at yards and overall intermediate
handling costs [27, 35, 26, 69]. Based upon a blocking plan, train routing (TR) or scheduling identifies train routes, timetables, and frequencies [29, 32]. BTA is then conducted to determine which trains should carry which blocks [34, 35]. The combination of the two tasks of BTA and TR is called train design. Our research focuses on the train design problem in re-routing freight traffic to determine rail infrastructure criticality. Train design is a highly combinatorial and complex optimization problem. A number of efforts have been made to solve this problem using different cost terms and business constraints [25, 41, 42, 44, 46-49, 66, 68]. None of these papers considered the congestion factor that may occur during any catastrophes. Most of these researches [25, 41, 42, 44, 68] divided the overall problem into two separate problems of train design and block routing and solved them iteratively. Two papers [46, 49] used the integrated approach to solve the whole problem. More recently, a few works addressed the large-scale train design problems, but without considering congestion issues in their models [23, 80-82]. North American railroads widely use the Multi-Modal system to create the sequence of blocks and train routes with both optimization and simulation techniques [83]. Multi-Modal creates block sequences by minimizing the number of switches and total car-miles under various constraints through shortest-path algorithms. The system is used for railway planning under normal conditions without considering heavy congestion. To evaluate the effect of disruptions at individual infrastructure element this research develops a train design approach that incorporates capacity constraints at links and yards, operational constraints of trains, and congestion effects at links. The purpose of the proposed approach is not for normal operational planning.
4.2.1 Problem Statement and model formulation

Different from train forming and routing under normal condition, a disruptive event can cause the rail network operate under high congestion and close to its capacity at both links and yards. An optimization model is necessary to incorporate the following congestion concerns.

- At each link, the travel time depends on the number of trains traveling on the link. When more trains are assigned to it, the speed on the link may be reduced. The decreasing speed-volume relationship depends on link features, such as the number of tracks and its signal system. Furthermore, the model also has an upper bound on the daily number of trains that can travel on each link.

- At each classification yard, increased volume can increase the waiting time for an arriving train to be humped but may decrease the waiting time to form an departing train so that the cost (or time) for a railcar to be classified at one specific yard is assumed to be independent from the volume. However, a yard can only handle a limited number of trains and/or cars. If the volume exceeds the capacity, the traffic may overflow and block the traffic on the main tracks. This capacity concern is not important under the current normal practice.

The following train forming and routing model with congestion concerns considers a directed railway network $G = (N, A)$, where $N$ is the set of nodes, and $A = \{(i,j): i,j \in N\}$ is the set of directed links. $(i,j) \in A$ implies $(j,i) \in A$. $N_c$ is the set of nodes for classification and $N_c \subseteq N$. At each classification node (also known as yard)
\( i \in N_c \), a capacity is defined by the number of trains and the number of cars that the yard can handle every day, \( U^T_i \) and \( U^C_i \) respectively. The cost (or time) for a railcar to be classified at yard \( i \in N_c \) is assumed at \( C^C_i \), which is independent from the classification volume. The distance of each link \( (i, j) \in A \) is \( d_{ij} \). The travel cost (related to travel time) per railcar depends on the number of trains traveling on the link. If \( l \) trains travel on link \( (i, j) \) per day, including both directions, the travel cost per railcar is \( C^T_{ij} \). At most \( U^M_{ij} \) trains can be assigned to link \( (i, j) \), where \( i < j \). Furthermore, the train length and weight is limited by \( U^L_{ij} \) and \( U^W_{ij} \) on each link \( (i, j) \in A \). In addition, each train incurs a start cost of \( C^S \) and travel cost \( C^T_{ij} \) for passing link \( (i, j) \). \( C^T_{ij} \) is the product of the train travel cost \( C^T \) per mile and the link distance \( d_{ij} \). Both \( C^S \) and \( C^T_{ij} \) are independent from the travel time and therefore from the volume over links. All the four cost items are normalized into a momentary unit though several of them are originally based on ton-hours. A set of blocks \( B \) needs to be shipped from their origins to destinations, where \( b \) is its index. For each block \( b \), we know the following information: number of railcars \( r_b \), origin and destination \( o_b \) and \( d_b \), length \( L_b \), and weight \( w_b \). Whenever a train stops to either pick up or drop-off blocks, it is called a work event. Each train cannot exceed its maximum allowable work events \( R^{we} \) in its route. Each train is restricted to carry a maximum number of blocks \( R^{bt} \). Without the loss of generality, this paper further assumes that each train can only depart a node once; a train can start and end at the same node; and a train can only pick up and drop off one block by at most once. A mixed integer program (MIP) is developed to represent the described problem with the following decision variable definitions.
4.2.1.1 Decision variables

- $z_{ij}^t$ whether train travels along link $(i, j), t \in T, (i, j) \in A$;
- $\gamma_{ij}^t$ whether there are $l$ trains travels along link $(i, j)$ on both directions, $i < j, l \in \{0, \ldots, U_{ij}^M\}$;
- $x_i^t$ whether train has a working event at node $i, t \in T, i \in N_c$;
- $u_i^t$ whether train $t$ starts at node $i, b \in B, i \in N$;
- $v_i^t$ whether train $t$ ends at node $i, b \in B, i \in N$;
- $a_i^t$ whether train $t$ starts and ends at node $i, b \in B, i \in N$;
- $p_i^{b,t}$ whether train $t$ picks up block $b$ at node $i, b \in B, t \in T, i \in N_c$;
- $q_i^{b,t}$ whether train $t$ drops block $b$ at node $i, b \in B, t \in T, i \in N_c$;
- $g_i^t$ length of train $t$ after visiting node $i$;
- $h_i^t$ tonnage of train $t$ has after visiting node $i$;
- $\theta_i^t$ number of railcars carried by train $t$ after visiting node $i$;
- $\gamma_i^t$ number of railcars carried by train $t$ after conducting classification at yard $i$;
- $\varphi_{ij}^t$ number of railcars that train $t$ carries along link $(i, j), t \in T, (i, j) \in A$;
- $\tau_{ij}^t$ total car travel cost over link $(i, j), (i, j) \in A, i < j$, for both direction;
- $k_i^t$ an artificial variable to eliminate sub-tours, $t \in T, i \in N$.

4.2.1.2 MIP model

\[
\text{Min} \quad \sum_{i \in T} \sum_{(i, j) \in A} C_{ij}^t z_{ij}^t + C_s^t \sum_{i \in T} \sum_{i \in N} u_i^t + \sum_{i \in N_c} \sum_{i \in T} C_i^c \gamma_i^t + \sum_{(i, j) \in A, i < j} \tau_{ij}^t \tag{4.6}
\]
S. T.

\[
\sum_{\forall j: (i,j) \in A} z_{ij}^t - \sum_{\forall j: (j,i) \in A} z_{ji}^t = u_i^t - v_i^t \quad \forall t \in T, i \in N \tag{4.7}
\]

\[
\sum_{\forall j: (i,j) \in A} z_{ij}^t \leq 1 \quad \forall t \in T, i \in N \tag{4.8}
\]

\[
\sum_{\forall j: (i,j) \in A} z_{ji}^t \leq 1 \quad \forall t \in T, i \in N \tag{4.9}
\]

\[
\sum_{t \in T} z_{ij}^t + \sum_{t \in T} z_{ji}^t = \sum_{l \in \{0, \ldots, u_{ij}^M\}} l y_{ij}^l \quad \forall (i,j) \in A, (i < j) \tag{4.10}
\]

\[
\sum_{l \in \{0, \ldots, u_{ij}^M\}} y_{ij}^l = 1 \quad \forall (i,j) \in A, (i < j) \tag{4.11}
\]

\[
\sum_{i \in N} u_i^t \leq 1 \quad \forall t \in T \tag{4.12}
\]

\[
2a_i^t \leq v_i^t + u_i^t \quad \forall t \in T, i \in N \tag{4.13}
\]

\[
k_i^t \geq k_i^t + 1 - Ma_j^t - (1 - z_{ij}^t)M \quad \forall t \in T, (i,j) \in A \tag{4.14}
\]

\[
\left\{ \begin{array}{l}
\sum_{t \in T} p_{i}^{b,t} = 1 - \sum_{t \in T} q_{i}^{b,t} = 1 \quad \text{if } i = o_b \\
\sum_{t \in T} q_{i}^{b,t} = 1 - \sum_{t \in T} q_{i}^{b,t} = 1 \quad \text{if } i = d_b \\
\sum_{t \in T} p_{i}^{b,t} = \sum_{t \in T} q_{i}^{b,t} \quad i \in N_c - o_b - d_b
\end{array} \right. \quad \forall b \in B, i \in N_c \tag{4.15}
\]

\[
x_i^t \geq p_{i}^{b,t} \quad \forall t \in T, i \in N, b \in B \tag{4.16}
\]

\[
x_i^t \geq q_{i}^{b,t} - v_i^t \quad \forall t \in T, i \in N, b \in B \tag{4.17}
\]
\[ x^t_i \leq \sum_{j : (i,j) \in A} z^t_{ij} \quad \forall t \in T, i \in N_c \quad (4.18) \]

\[ \sum_{b \in B} p^b_{i,t} = \sum_{b \in B} q^b_{i,t} \quad \forall b \in B, t \in T \quad (4.19) \]

\[ k^t_j \geq k^t_i + 1 + M(p^b_{i,t} + q^b_{j,t} - v^t - 2) \quad \forall t \in T, i, j \in N, b \in B \quad (4.20) \]

\[ g^t_j \geq g^t_j + \sum_{b \in B} p^b_{j,t} L_b - (1 - z^t_{ij})M - \sum_{b \in B} q^b_{j,t} L_b - Mv^t \quad \forall t \in T, (i,j) \in A \quad (4.21) \]

\[ (u^t_i - 1)M + \sum_{b \in B} p^b_{i,t} L_b \leq g^t_i \leq \sum_{j : (i,j) \in A} U^t_{ij} z^t_{ij} \quad \forall t \in T, i \in N \quad (4.22) \]

\[ h^t_j \geq h^t_i + \sum_{b \in B} p^b_{i,t} w_b - (1 - z^t_{ij})M - \sum_{b \in B} q^b_{j,t} w_b - Mv^t \quad \forall t \in T, (i,j) \in A \quad (4.23) \]

\[ (u^t_i - 1)M + \sum_{b \in B} p^b_{i,t} w_b \leq h^t_i \leq \sum_{j : (i,j) \in A} U^t_{ij} z^t_{ij} \quad \forall t \in T, i \in N \quad (4.24) \]

\[ \theta^t_j \geq \theta^t_i + \sum_{b \in B} p^b_{j,t} r_b - (1 - z^t_{ij})M - \sum_{b \in B} q^b_{j,t} r_b - Mv^t \quad \forall t \in T, (i,j) \in A \quad (4.25) \]

\[ \theta^t_i \geq (u^t_i - 1)M + \sum_{b \in B} p^b_{i,t} r_b \quad \forall t \in T, i \in N \quad (4.26) \]

\[ \phi^t_{ij} \geq \theta^t_i - M(1 - z^t_{ij}) \quad \forall t \in T, (i,j) \in A \quad (4.27) \]

\[ \sum_{t \in T} x^t_i \leq U^T_i \quad \forall i \in N_c \quad (4.28) \]

\[ \gamma^t_i \geq \theta^t_i - M(1 - x^t_i) \quad \forall t \in T, i \in N_c \quad (4.29) \]
The objective function (4.6) minimizes the total cost, including train travel cost, train start cost, railcar classification costs, and railcar travel cost along links, which depends on traffic volume. Constraint set (4.7) keeps the train flow conservation. Constraint sets (4.8) and (4.9) make sure that one train can leave each node (link) at most once. Constraint sets (4.10) and (4.11) are used to obtain the number of trains traveling over link \((i, j)\), represented by \(y_{ij}^t\). They also make sure that the number of trains on both directions over link \((i, j)\) will not exceed its capacity of \(U_{ij}^M\). Constraint set (4.12) guarantees that each used train only starts once while constraint set (4.13) permits a train to start and end at the same node. Sub-tours are eliminated by constraint set (4.14). When node \(j\) is visited by train \(t\) immediately after node \(i\), \(k_{ij}^t\) will be greater than \(k_{ji}^t\) by at least 1. Here, \(M\) is a big number. Constraint set (4.15) guarantees a block is picked up (dropped off) by a train at either its origin (destination) or in the classification yard where
some train drops (picks) it and a block is not dropped (picked) at its origin (destination).
Constraint sets (4.16) and (4.17) together make \( x^t_i = 1 \) when train \( t \) picks or drops
(except its terminating yard) some blocks at yard \( i \). Constraint set (4.18) guarantees that a
train definitely passes the node where the train has a work event (classification). When a
train picks up a block, the train must drop it, as indicated by (4.19). A train must pick up
a block before it can drop that block, guaranteed by constraint set (4.20) using the
ordering variable of \( k^t_i \) obtained from (4.14). Please note that (4.20) does not matter when
the dropping point is the termination node of the train, which is necessary when a train
starts and terminates at the same node. The length of train \( t \) after visiting a node is
obtained by (4.21) and (4.22) and satisfies the maximum train length restriction on each
link. Similarly, the weight restriction of trains over each link is met by (4.23) and (4.24).
The number of railcars a train has after visiting a node is obtained by (4.25) and (4.26).
The number of railcars that a train carries over a link is obtained by (4.27). Constraint set
(4.28) restricts the number of trains that can be classified at each yard \( i \) while (4.29) and
(4.30) restrict the number of railcars each classification yard can handle. (4.31) restricts
the maximum number of work events a train can have. The number of blocks one train
can carry cannot exceed the maximum number of blocks that a train is allowed to
transport, restricted by constraint set (4.32). (4.33) determines the railcar travel cost
traveling both directions over each link, which is used in the objective to calculate the
total railcar travel cost for all links. Finally, (4.34) is used to reduce the symmetry of
solutions to improve the computational efficiency. This MIP model is a highly
combinatorial optimization problem that cannot be solved for a large network using exact
optimization techniques, though small instances can be solved optimally using

58
commercial software (e.g., IBM CPLEX). Hence a heuristic algorithm is proposed in the next section to obtain quality solutions.

4.2.2 Solution approach

An iterative heuristic algorithm, illustrated by Figure 4.3, is developed with two major sub-problems of BTA and TR. The inputs include network data, block data, and all the cost parameters and restrictions. The initialization stage creates a sub-network for each block and assigns the blocks with reasonably large number of railcars to unit trains, which travel from their origins to destinations without any classifications. The BTA algorithm assigns the remaining blocks one by one to trains and provides a list of trains that are formed to serve all the blocks. At the TR stage, each formed train is routed optimally. BTA and TR stages are repeated until the percent cost gap between two consecutive TR runs falls below a predefined tolerance limit.
4.2.2.1 Initialization

In the Initialization step, detailed in Figure 4.4, a sub-network for each individual block from the whole network is built to reduce computational burden.
Initialization:

1. Create a shortest path for each pair of two nodes \(i, j \in N\) and have the shortest distance \(D_{ij}\),

2. For each block \(b\) in \(B\):
   - Include all nodes in the shortest path \(P_b\) from \(o_b\) to \(d_b\) into set \(N_b\);
   - Add all nodes in \(N \setminus N_b\) whose distance from any node in \(P_b\) is less than or equal to a given real number \(D\) into \(N_b\) (in this case, \(D = 100\) miles);
   - Let \(A_b\) the set of links with both ends in \(N_b\),

3. Let \(B^*\) be the set of blocks with \(r_b \geq R\) (in this case study, \(R = 80\) railcars),

4. For each block \(b\) in \(B^*\):
   - Assign ‘unit’ trains for \(b\), each following the shortest path \(P_b\) and satisfying link weight and length restriction along \(P_b\),
   - Update the available capacity of the network \((N, A)\), and
   - Update \(r_b, L_b\) and \(w_b\) by deducting railcars carried by ‘unit’ trains.

Figure 4.4  Steps of initialization algorithm for creating block sub-networks and ‘unit’ trains

4.2.2.2  Block-to-train assignment (BTA) sub-problem

A set of trains \(T\) serving blocks \(b \in B \setminus B^*\) is assumed to be obtained from the previous iteration. The given information at the beginning of the BTA is as follows.

For each train \(t \in T\):

- The route from start node \(o^t\) to the end node \(d^t\), the set of visited nodes \(N^t\), and the set of visited links \(A^t\),
• The number of carried railcars \( \theta_i^t \), length \( g_i^t \), and weight \( h_i^t \) after visiting node \( i \in N^t \), and
• The number of work events \( X^t \), the set of carried blocks \( B^t \), and their pickup and drop-off points.

For each node \( i \in N_c \):
• The number of trains classified \( \lambda_i \), and
• The number of railcars classified \( \gamma_i \).

For each link \( (i, j) \in A \):
• The total number of railcars carried by all trains \( \varphi_{ij} \), and
• The total number of travelled trains \( l_{ij} \).

---

**BTA:**

1. Randomize the blocks in \( B \setminus B^* \);
2. For each block \( b \in B \setminus B^* \) following the above order,
   • Create a set of candidate trains \( T^b \) following the train selection criteria.
   • For each train \( t \in T^b \)
     o Build the sets of nodes \( N_b^t \) and links \( A_b^t \) that be used for block \( b \) and train \( t \).
     o Calculate the cost of using train \( t \) to carry block \( b \) from node \( i \) to node \( j \), both belonging to \( N_b^t \).
   • Extend the network of \( (N_b^t, A_b^t) \) to \( o_b \) (or \( d_b \)) if \( o_b \) (or \( d_b \)) is not in \( N_b^t \) and \( d^t \neq o^t \).
   • Build a network by combining all \( (N_b^t, A_b^t) \), \( \forall t \in T^b \), plus a new train from \( o_b \) to \( d_b \).
   • Find the shortest cost path from \( o_b \) to \( d_b \) in the combined network.

---

**Figure 4.5** Block-to-train assignment (BTA) algorithm

62
The BTA algorithm, explained in Figure 4.5, assigns blocks in \( B \setminus B^* \), one by one following a random order, to existing trains or a new created train with the least cost. The random order is used to avoid local optimal solutions over iterations. Train information is updated along with the network capacity consumptions after assigning a block. The set of candidate trains \( T^b \) that can carry block \( b \in B \setminus B^* \) is formed by deleting the trains that have carried \( R^{bt} \) blocks belonging to \( B - \{b\} \), have carried only block \( b \), are unit trains, or satisfy \( A^T \cap A_b = \emptyset \) (called the train selection criteria in Figure 4.5). A new variable of \( \rho_t^i \) is introduced to indicate whether train \( t \) is classified at node \( i \in N^t \).

\[
\rho_t^i = \begin{cases} 
0 & \text{If train } t \text{ pickups/drops any blocks in } B - \{b\} \text{ at node } i \in N^t \\
1 & \text{Otherwise.} 
\end{cases}
\]

A new directed network for each train \( t \in T^b \) is built for block \( b \). If \( o_b \in N^T \) and/or \( d_b \in N^T \), we delete all nodes visited after \( d_b \) from \( N^T \) to have \( N_b^t \) and delete all links visited before \( o_b \) or after \( d_b \) from \( A^T \) to have \( A_b^t \). If \( A_b^t = \emptyset \), we do not build a network and skip all future steps for the train. We then delete any link that will exceed its length and weight maximum if train \( t \) carries block \( b \) from \( A_b^t \). For each pair of nodes \( m, n \in N_b^t \), when there is a path from \( m \) to \( n \) in \( (N_b^t, A_b^t) \) and \( \rho_{m}^t + \rho_{n}^t + X^t \leq R^{we} \), \( \lambda_m + \rho_{m}^t \leq U_m^T \), \( \lambda_n + \rho_{n}^t \leq U_n^T \), \( \gamma_m + \rho_{m}^t \theta_m^t \leq U_m^C \) and \( \gamma_n + \rho_{n}^t \theta_n^t \leq U_n^C \), a new link directly from \( m \) to \( n \) called \((m, n)\) is added into \( A_b^t \) with the cost \( DC_{mn}^t \) based on the on the shortest path cost from \( m \) to \( n \) in the network of \( (N_b^t, A_b^t) \). Define a path \( P_{m \rightarrow n} \) that contains the sequence of nodes from \( m \) to \( n \) in \( N_b^t \). The total cost of using train to carry block from \( m \) to \( n \) is \( CN_{mn}^t \) and can be calculated by (4.35). It includes the car travel cost and classification costs.
\[ CN_{mn}^t = DC_{mn}^t r_b + (\rho_m^t \theta_m^t + r_b) C_m^t + \rho_n^t \theta_n^t C_n^t + \sum_{k \in p_{m-n} - (m-n)} (1 - \rho_k^t) r_b C_k^t \]  

(4.35)

The shortest path cost \( DC_{mn}^t \) is calculated by running the shortest cost algorithm in \( (N_b^t, A_b^t) \) and the cost to travel on each link \((i, j) \in A_b^t\) is \( C_{ij}^{l ij} \). Here, \( C_{ij}^{l ij} \) is the travel cost of each railcar on link \((i, j) \in A_b^t\) when \( l_{ij} \) trains have already traveled on that link.

If \( o_b \) (or \( o_b \)) is NOT in \( N_b^t \) and \( d^t \neq o^t \), and \( \sum_{i \in N_t^t} (1 - \rho_i^t) + 1 \leq R_w^t \), we extend the network of \( (N_b^t, A_b^t) \) to \( o_b \) (or \( o_b \)). We add \( o_b \) into \( N_b^t \) with a link from \( o_b \) to each node \( n \in N_b^t \) into \( A_b^t \) if \( 1 + \rho_n^t + X^t \leq R_w^t \), \( \lambda_{ob} + \rho_{ob}^t \leq U_{ob}^t \) and \( \lambda_n + \rho_n^t \leq U_n^t \), \( \gamma_n + \rho_n^t \theta_n^t \leq U_n^c \) with the cost of \( CN_{obn}^t \) as (4.36).

\[ CN_{obn}^t = DC_{obn}^{b,t} + DC_{ob}^{c,t} r_b + \rho_n^t \theta_n^t C_n^c + r_b C_{ob}^c + \sum_{k \in p_{ob-n} - (ob-n)} (1 - \rho_k^t) r_b C_k^c \]  

(4.36)

\( DC_{obn}^{b,t} \) is the shortest path cost from \( o^t \) to \( n \) and is similar to \( DC_{mn}^t \) in (4.35). The shortest path cost \( DC_{ob}^{b,t} \) is calculated by,

- Updating the network of \( (N_b, A_b) \) after removing the links with \( U_{ij}^M \) trains and calculating the travel cost in each remaining link \((i, j) \in A_b\) as

\[ C_{ij} = C_{ij}^{l ij+1} r_b + (C_{ij}^{l ij+1} - C_{ij}^{l ij}) \varphi_{ij} + C_{ij}^t. \]

- Running the shortest cost algorithm from \( o_b \) to \( o^t \) in \( (N_b, A_b) \) to obtain \( DC_{ob}^{b,t} \).

By the above steps, \( o_b \) is into \( N_b^t \). Similarly, the destination of the block \( d_b \) is added into \( N_b^t \) if the following conditions are met.

1. \( d_b \notin N_b^t \); and
2. \( \sum_{i \in N^t} (1 - \rho^t_i) + 1 \leq R^{we} \) if \( o_b \) was in \( N^t_b \) or \( \sum_{i \in N^t} (1 - \rho^t_i) + 2 \leq R^{we} \) if \( o_b \) was not in \( N^t_b \).

We add \( d_b \) into \( N^t_b \) with a link from each node \( m \in N^t_b - \{o_b\} \) to \( d_b \) into \( A^t_b \), if
\[
1 + \rho^t_m + X^t \leq R^{we}, \lambda^t_m + \rho^t_m \leq \lambda^T_{d_b}, \gamma^t_m + \rho^t_m \theta^t_m \leq \gamma^C_{d_b} \quad \text{and} \quad \rho^t_{d_b} \theta^t_{d_b} \leq U^C_{d_b},
\]
with the cost of \( CN^t_{md_{d_b}} \) as (4.37).

\[
CN^t_{md_{d_b}} = DC^t_{md_{d_b}} + DC^{b,t}_{d_{d_b}} + \rho^t_m (\theta^t_{d_b} + r_b) C^C_m + C^C_{d_b} r_b + \sum_{k \in p_{m - d_b} - (m) - (d_b)} (1 - \rho^t_k) r_b C_k^C \quad (4.37)
\]

\( DC^t_{md_{d_b}} \) is calculated similar to \( DC^t_{mn} \) in (4.35), and also \( DC^{b,t}_{d_{d_b}} \) is calculated similar to \( DC^{b,t}_{o_b} \) in (4.36). If both \( o_b \) and \( d_b \) are newly added nodes and \( 2 + X^t \leq R^{we} \) and also \( \lambda^t_{o_b} + 1 \leq \lambda^T_{o_b}, \lambda^t_{d_b} + 1 \leq \lambda^T_{d_b}, \gamma^t_{o_b} + \rho^t_{o_b} \theta^t_{o_b} \leq \gamma^C_{o_b} \) and \( \gamma^t_{d_b} + \rho^t_{d_b} \theta^t_{d_b} \leq U^C_{d_b} \), we add a direct link from \( o_b \) to \( d_b \) into \( A^t_b \) with the cost of \( CN^t_{o_b d_b} \), shown in (4.38).

\[
CN^t_{o_b d_b} = DC^{b,t}_{o_b o_b} + DC^{b,t}_{o_b o_b} + DC^{b,t}_{o_b d_{d_b}} + DC^{b,t}_{o_b d_{d_b}} + \rho^t_{o_b} \theta^t_{o_b} \sum_{k \in p_{o_b - d_b} - (o_b) - (d_b)} (1 - \rho^t_k) r_b C_k^C \quad (4.38)
\]

If \( o_b \) is not a newly added node, and \( 1 + X^t \leq R^{we}, \lambda^t_{o_b} + 1 \leq \lambda^T_{o_b}, \lambda^t_{d_b} + 1 \leq \lambda^T_{d_b}, \gamma^t_{o_b} + \rho^t_{o_b} \theta^t_{o_b} \leq U^C_{o_b} \) and \( \gamma^t_{d_b} + \rho^t_{d_b} \theta^t_{d_b} \leq U^C_{d_b} \), a link from \( o_b \) to \( d_b \) is added into \( A^t_b \) with the cost given in (4.39).

\[
CN^t_{o_b d_b} = DC^{b,t}_{o_b d_{d_b}} + DC^{b,t}_{o_b d_{d_b}} + \rho^t_{o_b} \theta^t_{o_b} \sum_{k \in p_{o_b - d_b} - (o_b) - (d_b)} (1 - \rho^t_k) r_b C_k^C \quad (4.39)
\]

All networks \( (N^t_b, A^t_b) \) for \( t \in T^b \) are combined together into one network called \( (N^C_b, A^C_b) \). For a link traveled by more than one train, the one with the least cost is selected. Nodes and links are deleted from \( (N^C_b, A^C_b) \) if they do not belong to \( (N^b, A^b) \).

We create each link from \( o_b \) to \( n \in N^C_b - \{d_b\} \) as a new train and added it into \( A^C_b \) with
the cost of $CN_{ob,n}^{\text{New}}$ based on (4.40), if $\lambda_{ob} + \rho_{ob}^t \leq U_{ob}^T$ and $\lambda_n + \rho_n^t \leq U_n^T$ and $CN_{ob,n}^{\text{New}}$ is less than the current cost of link $(o_b,n)$ in $(N_b^C, A_b^C)$.

$$CN_{ob,n}^{\text{New}} = DC_{obn}^b + C^S + C_{ob}^C r_b$$ (4.40)

Similarly, we create each link from $n \in N_b^C - \{o_b\}$ to $d_b$ as a new train and add it into $A_b^C$ with the cost of $CN_{n,d_b}^{\text{New}}$ as (4.41), if $\lambda_n + \rho_n^t \leq U_n^T$ and $\lambda_{db} + \rho_{db}^t \leq U_{db}^T$ and $CN_{n,d_b}^{\text{New}}$ is less than the current cost of link $(n,d_b)$ in $(N_b^C, A_b^C)$.

$$CN_{n,d_b}^{\text{New}} = DC_{ndb}^b + C^S + C_n^C r_b$$ (4.41)

Finally, a link from $o_b$ to $d_b$ is created as a new train and added into $A_b^C$ with the cost of $CN_{ob,d_b}^{\text{New}}$ as (4.42) if $\lambda_{ob} + \rho_{ob}^t \leq U_{ob}^T$ and $\lambda_{db} + \rho_{db}^t \leq U_{db}^T$ in $(N_b^C, A_b^C)$.

$$CN_{ob,d_b}^{\text{New}} = DC_{obd_b}^b + C^S + C_{ob}^C r_b$$ (4.42)

The next step of BTA algorithm is to find the shortest cost path from $o_b$ to $d_b$ in the network of $(N_b^C, A_b^C)$. Each link in the shortest cost route represents a different train. Train information and network capacity information is updated after the assignment of a block.

Please note that though the shortest-path algorithm is used frequently in BTA, the definition and the costs associated with those newly defined links in $A_b^C$ are not the same as the physical links in the original network and the mileages of those physical links. Each newly-defined link $(i,j)$ in $(N_b^C, A_b^C)$ indicates that block $b$ is classified at yard $j$ right after yard $i$. The congestion cost at links and classification cost at yards are considered in $CN^t_{ij}$ if block is carried by train $t$ from yard $i$ to yard $j$. These calculations
are very different from the shortest-path algorithms used by Multimodal [83] on the original physical network.

4.2.2.3 Train routing (TR)

Each train, made up in BTA, is re-routed through an optimization model at the TR step. For each train, we know the start, termination, intermediate work event nodes, pick-up and drop-off points of the served blocks from BTA and will be regarded as parameters in TR model. Similar to block sub-networks, we also create a train sub-network for each train. Before routing a train, Pre-Process steps are used to remove routing information of the train under consideration so that congestion effects are not double counted.

4.2.2.3.1 Pre-Process

Each train \( t \in T \) carries a set of blocks \( b \in B^t \) with known pickup and drop-off points \((P^b_t, Q^b_t)\). The path of each train is known along with its \( X^t \) work event nodes and let \( N^t_p \) be the set of nodes in \( N^t \) with a working event (i.e., \( N^t_p = \bigcup_{b \in B^t} \{o^b_t, d^b_t\} \)). A train sub-network is created to reduce problem size by,

- Adding node \( j \) into \( N^t \) if there is a pair of nodes \((i, j)\) such that \( i \in N^t, j \in N \setminus N^t \) and \( D_{i,j} \leq D \), where \( D \) is a predefined number (e.g., 100 miles);
- Adding any links that have both ends in \( N^t \) and are not currently in \( A^t \) into \( A^t \);
- Deleting the links in \( A^t \) that have already reached the maximum train capacity restriction; and
- Letting \( l_{i,j} \) and \( q_{i,j} \) be the numbers of trains and railcars, excluding train \( t \), along link \((i, j) \in A^t \), where \( l_{i,j} = l_{j,i} \) and \( q_{i,j} = q_{j,i} \).
The TR optimization model is provided below with the following decision variable definitions.

### 4.2.2.3.2 Variables

- $z_{ij}$: whether the train travels along link $(i, j), (i, j) \in A^t$;
- $u_i$: whether the train starts at node $i, i \in N^t$;
- $v_i$: whether the train ends at node $i, i \in N^t$;
- $a_i$: whether the train starts and ends at node $i, i \in N^t$;
- $g_i$: length of the train after visiting node $i, i \in N^t$;
- $h_i$: tonnage of the train has after visiting node $i, i \in N^t$;
- $\theta_i$: number of railcars carried by the train after visiting node $i, i \in N^t$;
- $\tau_{ij}$: total additional car travel cost over link $(i, j) \in A^t$; and
- $k_i$: an artificial variable to eliminate sub-tours, $i \in N^t$.

Minimize

\[
\sum_{(i,j) \in A^t} C_{ij} z_{ij} + \sum_{(i,j) \in A^t} \tau_{ij}
\]  

(4.43)

Subject to:

\[
\sum_{\forall j: (i,j) \in A^t} z_{ij} - \sum_{\forall j: (j,i) \in A^t} z_{ji} = u_i - v_i \quad \forall i \in N^t
\]

(4.44)

\[
\sum_{\forall j: (i,j) \in A^t} z_{ij} \leq 1 \quad \forall i \in N^t
\]

(4.45)

\[
\sum_{\forall j: (i,j) \in A^t} z_{ji} \leq 1 \quad \forall i \in N^t
\]

(4.46)
\[
\sum_{i \in N_t} u_i = 1 \quad (4.47)
\]

\[
2a_i \leq u_i + v_i \quad \forall i \in N^t \quad (4.48)
\]

\[
k_j \geq k_i + 1 - M a_j - (1 - z_{ij})M \quad \forall (i, j) \in A^t \quad (4.49)
\]

\[
\sum_{\forall j: (i,j) \in A^t} z_{ij} + v_i \geq 1 \quad \forall i \in N_p^t \quad (4.50)
\]

\[
k_j \geq k_i + 1 + M(p_i^b + q_j^b - v_j - 2) \quad b \in B^t, i, j \in N^t \quad (4.51)
\]

\[
g_j \geq g_i + \sum_{b \in B^t} p_i^b L_b - (1 - z_{ij})M - \sum_{b \in B^t} Q_j^b L_b - M v_j \quad \forall (i, j) \in A^t \quad (4.52)
\]

\[
g_i \geq (u_i - 1)M + \sum_{b \in B^t} p_i^b L_b \quad \forall i \in N^t \quad (4.53)
\]

\[
g_i \leq \sum_{\forall j: (i,j) \in A^t} U_{ij}^t z_{ij} \quad \forall i \in N^t \quad (4.54)
\]

\[
h_j \geq h_i + \sum_{b \in B^t} p_i^b w_b - (1 - z_{ij})M - \sum_{b \in B^t} Q_j^b w_b - M v_j \quad \forall (i, j) \in A^t \quad (4.55)
\]

\[
h_i \geq (u_i - 1)M + \sum_{b \in B^t} p_i^b w_b \quad \forall i \in N^t \quad (4.56)
\]

\[
h_i \leq \sum_{\forall j: (i,j) \in A^t} U_{ij}^w z_{ij} \quad \forall i \in N^t \quad (4.57)
\]

\[
\theta_j \geq \theta_i + \sum_{b \in B^t} p_j^b r_b - (1 - z_{ij})M - \sum_{b \in B^t} Q_j^b r_b - M v_j \quad \forall (i, j) \in A^t \quad (4.58)
\]
\[ \theta_i \geq (u_{i} - 1)M + \sum_{b \in B_t} P^b_ir_b \quad \forall i \in N_t \quad (4.59) \]

\[ \tau_{ij} \geq \left( C^{t+1}_{ij} - C^{t}_{ij} \right) \varphi_{ij} + C^{t}_{ij} \theta_i - M(1 - z_{ij}) \quad \forall (i, j) \in A_t \quad (4.60) \]

\[ z_{ij}, u_i, a_i \in \{0, 1\}; g_i, h_i, \theta_i, v_i, \varphi_{ij}, \tau_{ij}, k_i \geq 0 \]

The objective function (4.43) minimizes the total cost, including train travel cost and railcar travel cost that depends on link volume, which is \( C^{t}_{ij} \). Constraint set (5.44) keeps the flow conversation. Constraint sets (5.45) and (5.46) make sure that train can visit each node (link) at most once. Constraint set (5.47) ensures that the train can start only once while constraint set (5.48) permits the train to start and end at the same node. Sub-tours are eliminated by constraint set (5.49). Constraint set (5.50) guarantees that the train definitely leaves the node where the train has a work event, except the termination node. The train must pick up a block before dropping the block, guaranteed by constraint set (5.51) that uses the variables ordering nodes visited by the train from constraint set (5.49). Please note that constraint set (5.51) is always loose when the dropping point is the termination node. Length restriction of the train over each link is ensured by constraint sets (5.52) to (5.54). Similarly, the weight restriction over each link is met by constraint sets (5.55) to (5.57). The number of railcars the train has after visiting a node is obtained by constraint sets (5.58) and (5.59). Finally, the additional railcar traveling cost caused by routing this train over a link is obtained by constraint set (5.60). Though the model (5.43-5.60) seems complicated, it is much smaller than the original model (5.6-5.34) because it has a much smaller network and is only for one train. The model can be solved to its optimum with commercial solvers, such as IBM CPLEX.
4.2.3 Numerical Experiments

The iterative heuristic algorithm is implemented using C++ and calling IBM CPLEX 12.3 for the TR sub-problem on a PC with 2.40GHz CPU and 2 GB RAM. The MIP model (4.6 - 4.34) of the overall problem is also solved to optimality for small instances by CPLEX 12.3 as a benchmark. Please note the optimization model (4.6 - 4.34) cannot be solved for large-scale instances. Table 4.2 summarizes the results for five randomly generated small-sized instances.

Table 4.2 Results of the MIP and heuristic solution approaches for small instances

<table>
<thead>
<tr>
<th>Instance</th>
<th>Nodes-Arcs-Blocks</th>
<th>CPLEX Value ($)</th>
<th>CPLEX CPU Time (sec.)</th>
<th>Heuristic Solution Value ($)</th>
<th>Heuristic Solution CPU Time (sec.)</th>
<th>Difference Value (%)</th>
<th>Difference CPU Time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4-5-4</td>
<td>15387.5</td>
<td>334.0</td>
<td>15387.5</td>
<td>0.33</td>
<td>0.00</td>
<td>99.901</td>
</tr>
<tr>
<td>2</td>
<td>5-7-5</td>
<td>17377.5‡</td>
<td>10000.0†</td>
<td>17377.5</td>
<td>0.50</td>
<td>0.000</td>
<td>99.995</td>
</tr>
<tr>
<td>3</td>
<td>6-9-6</td>
<td>23737.5‡</td>
<td>10000.0†</td>
<td>23637.5</td>
<td>0.67</td>
<td>-0.421</td>
<td>99.993</td>
</tr>
<tr>
<td>4</td>
<td>7-11-7</td>
<td>29087.5‡</td>
<td>10000.0†</td>
<td>29328.8</td>
<td>0.79</td>
<td>0.830</td>
<td>99.992</td>
</tr>
<tr>
<td>5</td>
<td>8-13-8</td>
<td>42562.5‡</td>
<td>10000.0†</td>
<td>40112.5</td>
<td>0.98</td>
<td>-5.756</td>
<td>99.990</td>
</tr>
</tbody>
</table>

‡ Time limit of 10,000 seconds is reached.
† Best solution when time limit of 10,000 seconds is reached

The second column in Table 4.2 reports the problem size in terms of the numbers of nodes, links, and blocks. The optimization model (4.6 - 4.34) could not be solved to its optimality by CPLEX within the 10,000-second time limit from the second instance. The heuristic algorithm provides quality solutions with much higher speed (less than 1
second). For the first small instance, the heuristic algorithm reaches the same optimal solution from CPLEX. It (Table 4.2) justifies the superiority of the proposed heuristic algorithm regarding both solution quality and computational time.

### 4.2.4 Case Study

The proposed solution approach is applied to the Class-I railroad industry’s network of USA to evaluate the criticality of its infrastructures, aforementioned in Chapter 3. This study uses the two types of data: network data and freight flow data, which are described in Chapter 3. The network consists of approximately 400 nodes and 500 links. Similar to previous section (Section 4.1), it also evaluates the criticality of each element, a node or a link, by comparing the total travel cost (time) before and after the disruption to (removal of) this element. Besides, the study utilizes some other additional data, which are related to cost parameters and train attributes, given in Table 4.3.

**Table 4.3 Other input parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train starting cost ($)</td>
<td>400</td>
</tr>
<tr>
<td>Train travel cost ($/mile)</td>
<td>10</td>
</tr>
<tr>
<td>Car classification cost ($)</td>
<td>40</td>
</tr>
<tr>
<td>Car travel cost without congestion ($/mile)</td>
<td>0.5</td>
</tr>
<tr>
<td>Car travel cost at moderate congestion ($/mile)</td>
<td>0.75</td>
</tr>
<tr>
<td>Car travel cost at high congestion ($/mile)</td>
<td>1.5</td>
</tr>
<tr>
<td>Maximum blocks per train</td>
<td>8</td>
</tr>
<tr>
<td>Maximum number of work events per train</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 4.3 shows various cost parameters and train attributes collected from [23, 77]. Three levels of congestion: normal (no congestion), moderate, and high are considered at links. The proposed heuristic solution approach, outlined in Figure 4.3, is applied to the whole network to obtain the total cost, which is considered the ‘standard’ transportation cost. Subsequently, each network element, link or node, is considered to be disrupted individually (removed from the network) and all traffic is re-routed in the residual network with the heuristic approach. The criticality of each element is estimated by the increased transportation cost caused by the disruption to that element compared to the standard transportation cost.

Figure 4.6 demonstrates the criticality levels of all elements for this case, which is obtained by running the heuristic approach for about 900 times.

Figure 4.6 Map showing the criticality of Class-I railroad network infrastructures
Each run takes about 15 minutes. The relative criticality level of a link or yard is represented by its width or size. Wider links or bigger yards are more critical. A disruption to those elements will result in more transportation cost (delay) for the whole rail freight flows. The stakeholders of rail network should pay more attention to protect or add redundancy to those elements in order to enhance the whole network’s resilience.

4.2.5 Conclusion and future work

This research evaluates the criticality of railroad network components by optimally creating and routing of trains through a disrupted and highly congested network. Because of the expected congestion and capacity constraints, existing train forming and routing models and algorithms in the literature, which were developed for normal operations, could not be used under an event of disruption. This research proposes an optimization model considering the capacity at both links and yards and the speed-volume relationship at links. An iterative heuristic solution method is proposed to solve the large-scale instances. The numerical study shows that the heuristic solution approach is capable of obtaining high quality solutions with high speed. Both the model and the solution method may not be appropriate for normal operational planning because the capacity and congestion at links and nodes are not so important under current normal freight operations. In fact, the normal train forming and routing may have to consider other factors, such as crew scheduling and routing, track rights, locomotive availability, timetables, etc. Criticality of each individual link or yard is measured by the increased transportation cost (delay) after re-routing the trains through the residual network in the absence of the element.
A real-life case study for a Class-I railroad company based on publicly obtained data is implemented to demonstrate the application of the proposed method to determine the critical hierarchy of the rail network infrastructures. The criticality of each infrastructure element, a link or a yard, is calculated by comparing the performance under the base case (the whole network) and under the residual network after removing the element. This critical hierarchy can help the railroad stakeholders to identify areas of high consequence so that they can better develop tactics to protect or add redundancy to the increasingly stressed railroad network. Thus, it is expected to enhance the preparedness capability relevant to transportation security and resiliency. In addition to enhance preparation, the proposed heuristic algorithm can also improve the response capability of homeland security agencies and railroads by optimally re-routing traffic in a disrupted network which could be highly congested and hurt the national security and congestion.

The method could also be used to estimate the effect of a certain disruption scenario that may involve multiple rail infrastructure elements. However, the research does not demonstrate this application because of the large number of possible scenarios and therefore large computational burden. Our methodology for evaluating the criticality might be improved by developing a decision support system (DSS) based on this heuristic algorithm. It should have the feature to arbitrarily disrupt multiple network components and run the model to evaluate their criticality. The DSS may have more practical value if interfaced with the commercial geographic information system (GIS) software.
CHAPTER V
EVALUATING CRITICALITY OF INTERMODAL NETWORK INFRASTRUCTURES USING FLOW OPTIMIZATION

In Chapter 4, we present models and solution procedures for determining criticality of railroad network infrastructures. In this chapter, we extend our research to identify the critical infrastructures of an integrated intermodal network. We present an optimization model that facilitates flow of all OD-specific traffic with minimum cost, utilizing the integrated intermodal network. As stated earlier, the intermodal network consists of all surface transportation modes: highway, railway, and waterway, and the model takes into account the congestion effects of each individual mode, as well as intermodal transfer terminals, which may happen under heavy traffic congestion. We implement a real-world case study to show the critical infrastructures for a given intermodal network of USA. Again, an integrated way to utilize all surface transportation modes is expected to significantly improve the overall freight flow. Making quick and correct traffic management decisions is a big challenge in a damaged or highly-congested transportation network. In addition to criticality analysis, the model will help traffic management decision makings in the event of a disaster by re-routing the traffic in the residual network. The proposed model provides the optimal routing and re-routing paths for the overall freight flow during and in the aftermath of a disaster in order to enhance surface transportation resilience and security.
5.1 Problem description

Intermodal freight transportation is increasingly gaining importance due to technological and equipment innovations, building of partnerships between major rail and truck carriers, and concern about time and cost-effective service by all the providers [52]. To justify, annual intermodal shipments moved in trailers and containers doubled from 3.1 million to 6.2 million by 1980 and 1990, and it increased from about 6.2 million to 11.7 million by 1990 and 2005 [87]. It is thought to be the future of the freight transportation system that is flawless, competent, secure, flexible, and environmentally sound, and meets the requirements of the customers and industries. In Chapter 2, it is observed that there are some optimization models and solution procedures that handle the intermodal transportation issues. Among them, one paper [71] solves the hub location problem for intermodal transfer terminals in a rail/road multimodal environment. Two others [52, 56] develop routing models for the combination of railroad and trucking transportation system. Researches for a combination of more than two modes are found in [53-55]. Most of these works did not implement for the real-world large scale problems except [70, 71]. Again, these researches are conducted for normal operational conditions without considering heavy congestion. In this research, we propose an optimization model and solution approach to determine the criticality of intermodal network infrastructures. Similar to the concept of railroad criticality analysis in Chapter 4, an intermodal infrastructure’s criticality is evaluated by estimating the increased cost (delay) when this element is disrupted. A system-wide optimization model considering the congestion effects at each individual mode and intermodal transfer terminals, and
capacity restrictions of infrastructures are necessary to route traffic to evaluate the effect of a disruption.

To describe briefly about the intermodal network considered in the proposed model, it consists of three surface transportation modes: highway, railway, and waterway, as stated earlier. These three surface transportation modes are connected by the intermodal transfer terminals. Not all transfer terminals are able to connect all three modes at their respective locations. Some transfer terminals are able to connect two modes, which can be either rail/road or rail/water, or road/water combination; whereas some are able to transfer among all three modes. Again, some intermodal transfer terminals are directed, whereas some are undirected. For instance, for a directed rail/road transfer terminal, traffic can move from railroad to truck only; it cannot move from truck to railroad. In contrast, for an undirected rail/road transfer terminal, traffic can move in both directions, railroad to truck and truck to railroad directions. For each mode, links capacities are accounted in the model, whereas nodes are assumed as uncapacitated. For highway, some links are one-way (directed), whereas remaining links are bi-directional (undirected). For railway, all links (tracks) are undirected. Railroad link capacities are limited by the number of tracks and types of signals used. Waterway network consists of ports/terminals, locks and reaches (i.e., links). Ports are the nodes in the waterway network where the barges can load/unload. All traffic movements originate and terminate at ports in the waterway network. Locks are treated as the bottlenecks of the waterway system, where the traffic faces congestion, in general. In our optimization model, locks are represented as links with limited capacities, where the traffic faces congestion effects as it approaches to their capacities.
In the event of a disaster, a disruptive event can cause the surface transportation infrastructures to be unavailable, or the capacities of its elements may be reduced. The traffic on those network elements has to be re-routed to other arcs or nodes of the residual intermodal network. It causes the traffic to operate under high congestion and close to the capacity of the network at both links and intermodal transfer terminals. The optimization model incorporates the following congestion concerns:

- At each link of the surface transportation modes, the travel time (cost) depends on the volume of traffic traveling on the link. When more traffic is assigned to it, the speed on the link may be reduced. The decreasing speed-volume relationship depends on link features, such as physical and operating characteristics. Furthermore, the model also has an upper bound on the daily volume of traffic that can travel on each link.

- At each intermodal transfer terminal and lock of the waterway system, increased volume can increase the waiting time for an arriving traffic to be transferred and locked, respectively, due to their capacity limitations.

Due to congestion effects, the routing time (cost) through the links and transferring time (cost) through the intermodal hubs will not be linear; rather they will follow different non-linear relationships based on the individual nature of the different transportation modes. In the proposed model, these non-linear relationships are approximated to linear functions by converting them into piece-wise linear functions. Finally, the objective is to minimize the overall costs (or, delays) for re-routed traffic. The infrastructure elements that will cause the most delays for the re-routed traffic due to their disruptions are the most critical nodes and arcs in the network. Considering all the
constraints and congestion effects of individual modes, the model will provide the optimal routes for all traffic OD pairs using the integrated intermodal network.

5.2 Mathematical model formulation

Consider a directed graph $G = (N, A)$, where $N$ is the set of nodes and $A$ is the set of directed links of the intermodal network containing highway, waterway and railway modes. For the purpose of model formulation, each intermodal transfer node $i \in N_t$ is further decomposed into a number of dummy nodes and dummy arcs, which is shown in Figure 5.1 for a sample intermodal network.

![Sample intermodal network](image)

Figure 5.1 Sample intermodal network (a) original network, and (b) equivalent network
Figure 5.1(a), shows the sample intermodal network has one intermodal transfer terminal, which is connected to all three surface transportation modes: highway, railway, and waterway. Highway links have both unidirectional way and bi-directional links, and links of other modes are bi-directed. In addition, there are two locks in the waterway network. The equivalent network, in Figure 5.1(b), decomposes the intermodal transfer node into a number of dummy nodes and arcs based on the number of modes connected to it, and also the permitted directions of transfers. It (Figure 5.1(b)) shows the six dummy intermodal nodes and twelve dummy arcs, and three additional dummy nodes for each mode connected to it. The dummy arcs are directed, implying the directions of allowable transfers. The cost of flow through these dummy arcs is zero. Thus the equivalent network consists of set of dummy nodes $N_d$, dummy arcs $A_d$ in addition to the existing nodes and arcs, which are referred as functional nodes $N_f$ and functional arcs $A_f$, respectively. Thus the overall node set is $N \in \{N_f \cup N_d \cup N_t\}$, and overall arc set is $A \in \{A_f \cup A_d\}$. Again, $A_f \in \{A^b \cup A^u\}$, where $A^b$ is the bidirectional link set and $A^u$ is the unidirectional link set. For the bi-directed link set $A^b$, $(i, j) \in A^b$ implies $(j, i) \in A^b$.

Other notations are defined as follows:

5.2.1 Parameters

$S$ set of source nodes for the commodity, indexed by $k$;

$H_k$ set of destination nodes for the commodity originated at node $k$, $\forall k \in S$

$D^l_k$ demand for an OD pair $(k, l)$, where $k \in S, l \in H_k$ and $D^l_k = 0$;

$\alpha^c_{i,j}$ slope (unit flow cost) of the section of piece-wise linear function of link $(i, j) \in A_f$ for the piece $c$;
\( \beta_{i,j}^c \) intercept of the section of piece-wise linear function of link \((i,j) \in A_f\) for the piece \(c\);

\( C_i \) fixed cost required to flow through intermodal transfer node \(i \in N_t\);

\( \alpha_i^c \) slope (unit flow cost) of the section of piece-wise linear function of intermodal transfer node \(i \in N_t\) for the segment \(c\);

\( \beta_i^c \) intercept of the section of piece-wise linear function for intermodal transfer node \(i \in N_t\) for the segment \(c\);

\( U_{i,j} \) capacity of a link \((i,j) \in A\);

\( V_i \) capacity of an intermodal transfer node \(i \in N_t\);

\( \xi_{i,k} \) \(\begin{cases} 1, & k = i \\ 0, & k \neq i \end{cases}\) for \(\forall k \in N, \forall i \in N\).

### 5.2.2 Variables

\( x_{i,j}^{k,l} \) flow on link \((i,j)\) for demand characterized by OD pair \((k,l)\);

\( y_{i,j} \) total volume of flow on link \((i,j) \in A\);

\( w_{i,j} \) total cost of flow along link \((i,j) \in A_f\);

\( q_i \) \(\begin{cases} 1, & \text{if there is a flow through intermodal transfer node } i \in N_t, \\ 0, & \text{otherwise} \end{cases}\)

\( s_i \) total transfer cost through intermodal transfer node \(i \in N_t\).

The model can be formulated as follows:

Minimize

\[
\sum_{(i,j) \in A_f} w_{i,j} + \sum_{i \in N_t} s_i + \sum_{i \in N_t} C_i q_i
\]  

(5.1)
Subject to:

\[
\xi_k^l D_k^l + \sum_{\forall (i,j) \in A} x_{i,j}^{k,l} = \sum_{\forall (i,j) \in A} x_{i,j}^{k,l} + \xi_l^l D_k^l \quad \forall k \in \mathcal{S}, l \in \mathcal{H}_k, i \in \mathcal{N} 
\]

(5.2)

\[
y_{i,j} = \sum_{k \in \mathcal{S}} \sum_{l \in \mathcal{H}_k} x_{i,j}^{k,l} \quad \forall (i,j) \in \mathcal{A}
\]

(5.3)

\[
y_{i,j} \leq U_{i,j} \quad \forall (i,j) \in \mathcal{A}^u
\]

(5.4)

\[
y_{i,j} + y_{j,i} \leq U_{i,j} \quad \forall (i,j) \in \mathcal{A}^b; i < j
\]

(5.5)

\[
w_{i,j} \geq \alpha_{i,j}^c y_{i,j} + \beta_{i,j}^c \quad \forall (i,j) \in \mathcal{A}^u; c = 1,2, \ldots, Q_{i,j}
\]

(5.6)

\[
w_{i,j} \geq \alpha_{i,j}^c (y_{i,j} + y_{j,i}) + \beta_{i,j}^c \quad \forall (i,j) \in \mathcal{A}^b; i < j; c = 1,2, \ldots, Q_{i,j}
\]

(5.7)

\[
\sum_{\forall (i,j) \in \mathcal{A}_d} y_{i,j} \leq V_i q_i \quad \forall i \in \mathcal{N}_t
\]

(5.8)

\[
s_i \geq \alpha_i^c \sum_{\forall (i,j) \in \mathcal{A}_d} y_{i,j} + \beta_i^c \quad \forall i \in \mathcal{N}_t; c = 1,2, \ldots, Q_i
\]

(5.9)

\[
q_i \in \{0,1\} \quad \forall i \in \mathcal{N}_t
\]

(5.10)

\[
x_{i,j}^{k,l} \in \mathbb{Z}^+ \quad \forall k \in \mathcal{S}, l \in \mathcal{H}_k, (i,j) \in \mathcal{A}
\]

(5.11)

\[
y_{i,j} \geq 0 \quad \forall (i,j) \in \mathcal{A}
\]

(5.12)

\[
w_{i,j} \geq 0 \quad \forall (i,j) \in \mathcal{A}_f
\]

(5.13)

\[
s_i \geq 0 \quad \forall i \in \mathcal{N}_t
\]

(5.14)

The objective function (5.1) minimizes the total cost, including freight flow cost along links, transfer costs through intermodal transfer terminals, both of which depend on traffic volume, fixed cost of freight transfer through intermodal transfer terminals, and fixed cost of freight origination at their origins. Constraint set (5.2) enforces the flow conservation requirements at each node. Constraint set (5.3) ensures that the sum all OD-
specific flows along a link must be equal to the total flow in that link. Constraint sets (5.4) and (5.5) make sure that maximum volume of flow on link \((i,j)\) cannot exceed its capacity \(U_{i,j}\) for unidirectional and bidirectional links, respectively. Constraint sets (5.6) and (5.7) calculate the total variable cost of flow along link \((i,j)\) for unidirectional and bidirectional links, respectively. The non-linear relationship between the total variable flow cost and the total flow on link \((i,j)\) \(\in A_f\) is approximated with a piece-wise linear function, which is shown in Figure 5.2.

![Diagram](image)

**Figure 5.2** Piece-wise linear approximation of the non-linear cost-flow function for links

Constraint set (5.8) ensures total flow through the intermodal transfer node \(i \in N_t\) cannot exceed its capacity \(V_i\). Constraint set (5.9) calculates total variable cost of flow through the intermodal transfer node \(i \in N_T\). Similar to the relationship between total flow cost and total flow on the links, the non-linear relationship between total transfer cost and total amount of transfers through the intermodal transfer terminals \(i \in N_t\) is
approximated with the piece-wise linear function, which is shown in Figure 5.3. Here (Figure 5.3), $z_i$ is the total flow originated from the virtual source node $i \in N_0$. Constraint set (5.10) is binary constraint and Constraint set (5.11) is standard integrality constraint. Finally, constraint sets (5.12) to (5.14) are standard non-negativity constraints.

![Figure 5.3 Piece-wise linear approximation of the non-linear cost-flow function for intermodal transfer nodes](image)

Similar to train design optimization model in Chapter 4, this MIP model is also a combinatorial optimization problem that cannot be solved for a real world large-scale network using exact optimization techniques, though small to moderate size instances can be solved optimally using commercial software (e.g., IBM CPLEX). In order to solve this model for the real world large-scale problem, we tried to implement some relaxation techniques to the MIP model. The effectiveness of the relaxation techniques depend on the quality of the solutions, which is measured by the relative gap between the relaxed and optimum solutions, and also the solution times. The most commonly used relaxation technique is the linear programming (LP) relaxation, which is achieved by dropping the
integrality constraints from the MIP formulation. Compared to original MIP, it is usually much easier to solve using the existing optimization solver. Sometimes the relative gap between the relaxed and optimum solutions is too large to be effective for fulfilling the purpose. In contrast, if the relaxed solution turns out all variables to be integer for any MIP, in that case relaxed LP is said to be *naturally integer*, and very effective for that formulation. For our research, we intend to implement the LP relaxation technique to the original MIP formulation. In order to test the performance of the intended LP relaxation technique to this MIP model based on solution quality and solution time, we implement numerical experiments for some randomly generated problem instances. In the following section, we implement these numerical experiments to judge the acceptability of this technique to apply for the real world case study.

5.3 Numerical experiments

Both MIP and its LP relaxation are implemented using C++ with Concert Technology and solved using IBM ILOG CPLEX 12.3 on a notebook computer with 2.30GHz CPU and 8 GB RAM. Both MIP model (5.1-5.16) and its LP relaxation are solved to optimality by CPLEX for some randomly generated problem instances to check the performance of the relaxed model. Table 5.1 summarizes the results for five randomly generated small and medium sized instances. The second column reports the problem size in terms of the numbers of nodes, links, and shipments of different ODs. For these instances, the relaxed model provides very good lower bounds, which are close to the optimum solutions of original MIP, using CPLEX. Table 5.1 justifies the applicability of the relaxed model for the real-life case study regarding solution quality, and obviously its computational time should be much faster than original MIP.
Table 5.1 Results of the MIP and LP relaxation solution approaches for small instances

<table>
<thead>
<tr>
<th>Instance</th>
<th>No. of Nodes-Links-Ships</th>
<th>MIP Solution</th>
<th>LP Solution</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Value (K$)</td>
<td>Value (K$)</td>
<td>Value (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CPU Time (sec.)</td>
<td>CPU Time (sec.)</td>
<td>CPU Time (%)</td>
</tr>
<tr>
<td>1</td>
<td>38-70-25</td>
<td>12323.1</td>
<td>12322.1</td>
<td>0.0081</td>
</tr>
<tr>
<td>2</td>
<td>50-98-40</td>
<td>12924.8</td>
<td>12923.8</td>
<td>0.0077</td>
</tr>
<tr>
<td>3</td>
<td>66-144-50</td>
<td>13342.2</td>
<td>13340.9</td>
<td>0.0097</td>
</tr>
<tr>
<td>4</td>
<td>90-170-60</td>
<td>15087.1</td>
<td>15072.2</td>
<td>0.0099</td>
</tr>
<tr>
<td>5</td>
<td>100-210-70</td>
<td>16639.5</td>
<td>16634.9</td>
<td>0.0276</td>
</tr>
</tbody>
</table>

K$ - Thousands of dollars

5.4 Computational study

In this section, we present the result of our computational study. The proposed LP relaxation of the original MIP approach is applied to the intermodal network of five U.S. states to evaluate the criticality of its infrastructures, stated in Chapter 3. This study uses the two types of data: network data and freight flow data, belonging to the intermodal transportation study of Chapter 3. The network consists of approximately 5500 nodes and over 6800 links, including over 200 intermodal transfer terminals. Besides, this study uses some other input data, which are related to various cost components for intermodal transportation. These cost parameters are shown in Table 5.2 and collected from [88-90]. Three levels of congestion: normal (no congestion), moderate, and high are considered at links. Again, two levels of congestion: normal (no congestion), and high are considered at links of each individual mode and variable intermodal transfer nodes.

The LP relaxation to the original MIP model is applied to the whole intermodal network to obtain the total cost, which is considered the ‘standard’ transportation cost.
Similar to chapter 4, network elements (links and nodes) are considered to be disrupted (removed from the network) separately, and all traffic is re-routed in the residual network.

Table 5.2 Various cost components for intermodal transportation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit (uncongested) barge transportation cost per container-mile ($/TEU/mile)</td>
<td>0.12</td>
</tr>
<tr>
<td>Unit (uncongested) rail transportation cost per container-mile ($/TEU/mile)</td>
<td>0.35</td>
</tr>
<tr>
<td>Unit (uncongested) truck transportation cost per container-mile ($/TEU/mile)</td>
<td>1.8</td>
</tr>
<tr>
<td>Fixed (transfer) cost per shipment at intermodal terminal for barge ($/shipment)</td>
<td>200</td>
</tr>
<tr>
<td>Fixed (transfer) cost per shipment at intermodal terminal for rail ($/shipment)</td>
<td>150</td>
</tr>
<tr>
<td>Fixed (transfer) cost per shipment at intermodal terminal for truck ($/shipment)</td>
<td>0</td>
</tr>
<tr>
<td>Unit (uncongested) cost per container at intermodal terminal for barge ($/TEU)</td>
<td>44</td>
</tr>
<tr>
<td>Unit (uncongested) cost per container at intermodal terminal for rail ($/TEU)</td>
<td>50</td>
</tr>
<tr>
<td>Unit (uncongested) cost per container at intermodal terminal for truck ($/TEU)</td>
<td>0</td>
</tr>
</tbody>
</table>
There are some links which are very small in length. For those small links which are adjacent to each other and no other links are incident on them, we disrupted multiple of them together, and re-routed the traffic in residual network to save the overall experimentation time. This is a practical approach since no other links are incident on them. The element or group of elements which causes the utmost additional cost due to disruption to that element compared to the standard transportation cost is the most critical infrastructure in the intermodal network. Figure 5.4 demonstrates the criticality levels of all elements for this intermodal network.

![Figure 5.4 Map showing the criticality of intermodal network infrastructures](image)

Each run takes approximately 13-16 minutes. The relative criticality levels of links are represented by its color and width. As shown in Figure 5.4, in terms of color, the
order of criticality for links is red, orange and black, respectively. Again, for a particular color, the wider link is the more critical than others. For nodes, the comparative criticality levels of intermodal transfer nodes are represented by its shape and size. The order of criticality for intermodal nodes is also shown in Figure 5.4, star and circle being the most and the least critical, respectively. A disruption to those elements will result in more transportation cost (delay) for the whole intermodal freight flows. In order to enhance the whole network’s resilience, relevant authorities of the intermodal network should pay more attention to protect or add redundancy to those elements.

5.5 Conclusion and future work

In this research, we evaluate the criticality of intermodal network infrastructures, consisting of three surface transportation modes, by optimally routing the traffic through a disrupted and congested network. During normal operation, highway and waterway locks are subject to operate under congestion effects that are established in the literature [91-96]. Since we re-route all the traffic through the disrupted network, it is very natural to expect heavy congestion in each element (e.g., links, transfer terminals) of the individual modes: highway, railway, and waterway of the intermodal network. Therefore, we consider congestion effect in each element of the intermodal network that has capacity limitations. As intermodal freight transportation is relatively young in transportation research domain, there are very few freight flow models found in the literature. Most of these models are developed for rail/road intermodal transportation under normal operations. Very few existing models that consider more than two modes in intermodal network, did not consider the congestion issues that is obvious during emergency situations, and therefore, these cannot be used for routing under an event of
disruption. In this research, we propose an optimization model considering the capacity limitations at all links (e.g., highway link, railroad track, waterway reach and lock) and intermodal transfer terminals. Speed-volume relationships are different for different surface transportation modes, and hence we consider different cost vs. volume functions (rates) for different types of links and transfer terminals. One of the most commonly relaxation techniques, LP relaxation, is employed to solve the real-life problem. Numerical experiments show that the relaxed optimization model is capable of obtaining high quality solutions with high speed. This model may not be appropriate for normal operational planning because the congestions at railroad and waterway links are not so important under current normal freight operations. Criticality of each individual link or transfer terminal is measured by the increased transportation cost (delay) after re-routing the traffic through the residual network in the absence of that particular element.

A real-world case study for the intermodal network based on publicly obtained data is implemented to demonstrate the application of the proposed model to determine the critical hierarchy of the network infrastructures. The criticality of each infrastructure element, a link or a transfer terminal, is determined by comparing the performance under the base case (the whole network) and under the residual network after removing the element. Similar to railroad in Chapter 4, this critical hierarchy can also help the relevant authorities to identify areas of high importance so that they can better develop plans to protect or add redundancy to the critical infrastructures of the surface transportation network including intermodal terminals. Thus it is expected to improve the preparedness capability, which is significant to intermodal transportation security and resiliency. In addition, the proposed model can also improve the response capability of homeland
security agencies and individual modes by optimally re-routing traffic in a disrupted network which could be highly congested and impair the national security and congestion.

Though the LP relaxation technique, used in this research, provides a good lower bound to the original formulation for small-sized instances, the gap between the solutions would increase for large-scale problems. Therefore, we suggest and hope to implement some other relaxation (e.g., Lagrangian Relaxation) or decomposition (e.g., Benders Decomposition) or any heuristic procedures in our future research as an extension to the current research. The method could also be used to estimate the effect of a certain disruption scenario that may involve multiple intermodal infrastructure elements. Though this research combines multiple links disruption for small length adjacent links, it does not demonstrate this application for other cases because of the large number of possible scenarios, and therefore, large computational burden. This opportunity might be implemented by developing a decision support system (DSS) based on the proposed model. The DSS should have the ability to arbitrarily disrupt multiple network components, and run the proposed model to evaluate the criticality of combined multiple network components disruptions. Commercial geographic information system (GIS) software (e.g., ArcGIS) can be interfaced with the DSS to make it more practical and visually attractive to all users.
REFERENCES


93


