Optimization of Multimodal Evacuation of Large-Scale Transportation Networks

By
Hossam Mohamed Abd El-Hamid Abd El-Gawad

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Graduate Department of Civil Engineering
University of Toronto

© Copyright by Hossam Abd El-Gawad, 2010
Optimization of Multimodal Evacuation of Large-Scale Transportation Networks

Hossam Abd El-Gawad

Doctor of Philosophy

Department of Civil Engineering

University of Toronto

2010

Abstract

The numerous man-made disasters and natural catastrophes that menace major communities accentuate the need for proper planning for emergency evacuation. Transportation networks in cities evolve over long time spans in tandem with population growth and evolution of travel patterns. In emergencies, travel demand and travel patterns drastically change from the usual everyday volumes and patterns. Given that most US and Canadian cities are already congested and operating near capacity during peak periods, network performance can severely deteriorate if drastic changes in Origin-Destination (O-D) demand patterns occur during or after a disaster. Also, loss of capacity due to the disaster and associated incidents can further complicate the matter. Therefore, the primary goal when a disaster or hazardous event occurs is to coordinate, control, and possibly optimize the utilization of the existing transportation network capacity. Emergency operation management centres face multi-faceted challenges in anticipating evacuation flows and providing proactive actions to guide and coordinate the public towards safe shelters.

Numerous studies have contributed to developing and testing strategies that have the potential to mitigate the consequences of emergency situations. They primarily investigate the effect of some proposed strategies that have the potential of improving the performance of the evacuation process with modelling and optimization techniques. However, most of these studies are inherently restricted to evacuating automobile traffic using a certain strategy without considering other modes of transportation. Moreover, little emphasis is given to studying the interaction between the various strategies that could be potentially synergized to expedite the evacuation process.
process. Also, the absence of an accurate representation of the spatial and temporal distribution of the population and the failure to identify the available modes and populations that are captive to certain modes contribute to the absence of multimodal evacuation procedures. Incorporating multiple modes into emergency evacuation has the potential to expedite the evacuation process and is essential to assuring the effective evacuation of transit-captive and special-needs populations.

This dissertation presents a novel multimodal optimization framework that combines vehicular traffic and mass transit for emergency evacuation. A multi-objective approach is used to optimize the multimodal evacuation problem. For automobile evacuees, an Optimal Spatio-Temporal Evacuation (OSTE) framework is presented for generating optimal demand scheduling, destination choices and route choices, simultaneously. OSTE implements Dynamic Traffic Assignment (DTA) techniques coupled with parallel distributed genetic optimization to guarantee a near global optimal solution. For transit evacuees, a Multi-Depots, Time Constrained, Pick-up and Delivery Vehicle Routing Problem (MDTCPD-VRP) framework is presented to model the use of public transit vehicles in evacuation situations. The MDTCPD-VRP implements constraint programming and local search techniques to optimize certain objective functions and satisfy a set of constraints. The OSTE and MDTCPD-VRP platforms are integrated into one framework to replicate the impact of congestion caused by traffic on transit vehicle travel times.

A proof-of-concept prototype has been tested; it investigates the optimization of a multimodal evacuation of a portion of the Toronto Waterfront area. It also assesses the impact of multiple objective functions on emergency evacuation while attempting to achieve an equilibrium state between transit modes and vehicular traffic. Then, a large-scale application, including a demand estimation model from a regional travel survey, is conducted for the evacuation of the entire City of Toronto.

This framework addresses many limitations of existing evacuation planning models by: 1) synergizing multiple evacuation strategies; 2) utilizing robust optimization and solution

---

1 Transit-dependent and vulnerable populations are frequently termed special-needs populations; these include people who may lack access to a private vehicle, may also need assistance in evacuation, and depend primarily on transit for transport.
algorithms that can tackle such multi-dimensional non deterministic problem; 3) estimating the spatial and temporal distribution of evacuation demand; 4) identifying the transit-dependent population; 5) integrating multiple modes in emergency evacuation. The framework presents a significant step forward in emergency evacuation optimization.
ACKNOWLEDGEMENT

Throughout my doctorate journey, I have received tremendous support from many people, probably too many to count. Without them, I would not reach up to this point now while writing these humble words expressing my deep gratitude to them.

I must first and foremost thank Baher Abdulhai, my thesis advisor, for his sincere guidance and support during this journey. Baher: I can’t say enough to deeply thank you for your moral support before your intellectual and technical support, for being a friend before being an advisor. How many times I dropped by your office knocking on your door for advice and you never let me down. Let me just say that I am really proud for our joint accomplishments since 2006.

Special thanks to my family, although away in Egypt. I thank my parents for their love, support and for raising me up to this point. Without your phone calls and your continuous follow-up on my PhD progress, I would never have been able to complete this stage in my life.

Special thanks also to my wife Samah for her long-standing support and love. Thank you Samah for your great and charming spirit, you have always absorbed me when I was stressed, you have always guided me when I sought advice. To my lovely daughter, Nour, for her smiles that have never left her face, a smile that makes me forget all the stress and duties that are left over in my day. Nour: thank you for your understanding (probably not understanding) why Daddy could not spend as much time as he would love to with you.

I would like to extend my gratitude to Matthew Roorda. Matthew: I gained remarkable experience while working closely with you, thank you for your good spirit. When I see your level of enthusiasm, I feel more accountable and work more and more. Thank you for helping me to understand and retrieve some of the TTS data through the DMG program.

I am thankful to Eric Miller and Amer Shalaby for informal discussions about my thesis and for their valuable guidance.

I have known and worked closely with many colleagues at the Urban Transportation Research and Advancement Centre (UTRAC), University of Toronto. I am indebted to all of them for all the moral support and encouragement.
Special thanks to Asmus Georgi, the ITS Lab Manager. He has worked very hard, during
daytime and night-time, during normal weekdays and holidays to set up the departmental
computing facility that I have used in my research. Although always claiming that “Hossam
alone needs a full time IT assistant”, I believe no one else can achieve what Asmus has achieved
given his very tight schedule and his other duties in the Department of Civil Engineering. Special
thanks go also to Mohamed Wahba (aka Medhat) for his continuous support and the countless
informal discussions we had. He has been a model for me as a PhD. student that I have tried to
follow, although it was a challenge. I would like to thank Bahaa Alshalalfah for answering many
questions about ILOG. Also special thanks go to Mohamed Masoud and Osama Hassan, the
programmers and developers who contributed to the development/deployment of GENOTRANS
on Weblogic and WebSphere, respectively.

Finally, I would like to thank God’s divine guidance and divine provision in my life. He
surrounds me with so many blessings and loving people. His grace is abundant throughout my
life.
In this thesis, portions of Five chapters have been reproduced (with modifications) from published material.

These chapters are:

Chapter 2: Literature Review


Chapters 3, 4, 6 and 8: Optimal Spatio-Temporal Evacuation, Optimization of Transit Shuttling during Evacuation, and Prototype Implementation-Toronto Waterfront Application, Large-Scale Application- Evacuation of the City of Toronto


# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>RESEARCH OBJECTIVES</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>THESIS ROADMAP</td>
<td>5</td>
</tr>
<tr>
<td>II</td>
<td>ABSTRACT</td>
<td>II</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>EVACUATION SCHEDULING</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>THE CONCEPT OF ONE DESTINATION</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>TRAFFIC SIGNAL CONTROL</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>TRAFFIC ROUTING AND CONTROL STRATEGIES IN EMERGENCY EVACUATION</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>ROLE OF INTELLIGENT TRANSPORTATION SYSTEMS (ITS) IN EMERGENCY EVACUATION</td>
<td>13</td>
</tr>
<tr>
<td>2.6</td>
<td>CONTRA-FLOW OPERATION IN EMERGENCY EVACUATION</td>
<td>14</td>
</tr>
<tr>
<td>2.7</td>
<td>TRAVEL BEHAVIOUR IN EMERGENCY EVACUATION</td>
<td>17</td>
</tr>
<tr>
<td>2.8</td>
<td>SUMMARY</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>OPTIMIZATION OF AUTOMOBILE EVACUATION</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>SCHEDULING AND DESTINATION CHOICE</td>
<td>20</td>
</tr>
<tr>
<td>3.2</td>
<td>TIME STRUCTURES IN EVACUATION</td>
<td>22</td>
</tr>
<tr>
<td>3.3</td>
<td>TYPICAL OBJECTIVE FUNCTIONS IN EMERGENCY EVACUATION</td>
<td>23</td>
</tr>
<tr>
<td>3.4</td>
<td>OPTIMAL SCHEDULING AND DESTINATION CHOICE PROBLEM FORMULATION</td>
<td>25</td>
</tr>
<tr>
<td>3.5</td>
<td>THE OPTIMIZATION APPROACH</td>
<td>26</td>
</tr>
<tr>
<td>3.5.1</td>
<td>TYPICAL SOLUTION ALGORITHMS VS. EVOLUTIONARY ALGORITHM APPROACH</td>
<td>26</td>
</tr>
<tr>
<td>3.5.2</td>
<td>PARALLEL DISTRIBUTED GENETIC ALGORITHM (PDGA): DESIGN AND IMPLEMENTATION</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>OPTIMIZATION OF TRANSIT SHUTTLING DURING EVACUATION</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>VEHICLE ROUTING PROBLEM AND TRANSIT EVACUATION PROBLEM</td>
<td>37</td>
</tr>
<tr>
<td>4.2</td>
<td>VRP AND EMERGENCY EVACUATION: BACKGROUND</td>
<td>38</td>
</tr>
<tr>
<td>4.3</td>
<td>VEHICLE ROUTING AND MASS TRANSIT EVACUATION: AN ANALOGY</td>
<td>39</td>
</tr>
<tr>
<td>4.4</td>
<td>SOLUTION ALGORITHM: A CONSTRAINT PROGRAMMING APPROACH</td>
<td>45</td>
</tr>
</tbody>
</table>
4.4.1 Building the Evacuation Model ........................................................................................................ 46
4.4.2 Solving the Evacuation Routing and Scheduling Problem .............................................................. 47
  4.4.2.1 Route Construction Heuristic to Find Initial Solution ............................................................... 48
    • Insertion Heuristic .......................................................................................................................... 49
  4.4.2.2 Local Search Methods to Improve Initial Solution ................................................................. 49
5 Multimodal Evacuation Framework ...................................................................................................... 52
  5.1 Role of Transit in Emergency Evacuation: Lessons Learned ............................................................ 52
  5.2 Modes in Evacuation ......................................................................................................................... 54
  5.3 The Overall Framework: Multimodal Evacuation Plan .................................................................... 54
6 Prototype Implementation - Toronto Waterfront Application ........................................................... 57
  6.1 Application Context .......................................................................................................................... 57
  6.2 Design of Experiments ...................................................................................................................... 59
  6.3 Phase A: Evacuation Scheduling and Destination Choice under Various Levels of Demand and Time Pressures ............................................................................................................. 60
    6.3.1 Scenario Design ......................................................................................................................... 60
    6.3.2 Phase A Results and Analysis .................................................................................................. 61
      6.3.2.1 Average Evacuation Time and Trip Distance .................................................................... 64
  6.4 Phase B: Multiple Objective Optimization for Multiple Modes in Emergency Evacuation ................. 66
    6.4.1 Scenario Design ......................................................................................................................... 66
    6.4.2 Phase B Results and Analysis .................................................................................................. 68
      6.4.2.1 Optimal Scheduling, Destination, and Routing ................................................................. 68
      6.4.2.2 Optimal Routing and Scheduling of Transit Vehicles ....................................................... 70
  6.5 Phase C: Impact of Traffic and Transit Mode Split on Evacuation Time ........................................... 76
7 Emergency Evacuation Demand Estimation from Regional Travel Survey ........................................ 79
  7.1 Data Source: The Transportation Tomorrow Survey (TTS) ............................................................. 79
  7.2 Data Management Group (DMG) at the University of Toronto .................................................... 81
  7.3 Demand Estimation Method ............................................................................................................ 81
  7.4 GTA Demand Estimation Results .................................................................................................... 82
  7.5 City of Toronto Demand Estimation Results .................................................................................. 88
8 Large-Scale Application - Evacuation of the City of Toronto .......................................................... 95
  8.1 Application Context and Analysis Scope .......................................................................................... 95
LIST OF FIGURES

Figure 1.1 Thesis Organization.................................................................5
Figure 2.1 Original Multiple-Destination (nD) and Modified One-Destination (1D) Networks (Yuan et al., 2006) .............................................11
Figure 3.1 Effect of Evacuation Scheduling and Destination Choice on the Loading and Evacuation Curves .............................................21
Figure 3.2 Multiple Objectives in Evacuation ........................................23
Figure 3.3 The Basic Cycle of EAs .........................................................28
Figure 3.4 Development Phases of GENOTRANS ...............................30
Figure 3.5 Parallel and Distributed GA Architecture ...........................31
Figure 3.6 Single and Parallel Distributed Genetic Algorithms ..............34
Figure 4.1 Traditional Vehicle Routing Problem Representation ............40
Figure 4.2 MDTCPD-VRP Representation .............................................42
Figure 4.3 Framework for Integrating Constraint Programming and Local Search Techniques for MDTCPD-VRP ............................48
Figure 4.4 Insertion Heuristic ...............................................................49
Figure 4.5 Local Search Concept ..........................................................50
Figure 5.1 Factors Affecting the Role of Transit Systems in Emergency Evacuation .................................................................53
Figure 5.2 Framework for Optimization of Multimodal Evacuation .......56
Figure 5.3 Framework for Integrating Constraint Programming and Local Search Techniques for MDTCPD-VRP ............................56
Figure 6.1 Transportation Network, Evacuation Zones and Pick-up Points ..................................................................................59
Figure 6.2 Loading and Evacuation Curves for Different Scenarios at Different Risk Levels ......................................................63
Figure 6.3 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 60 min) ..........................................................64
Figure 6.4 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 120 min) ..........................................................64
Figure 6.5 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 180 min) ..........................................................65
Figure 6.6 Relative Effectiveness of Total Trip Distance for OSTE, OTE and SE for ML = 60 min ..................................................66
Figure 6.7 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 60 min) ..........................................................66
Figure 6.8 Example of Routing and Scheduling of Transit Vehicle .......72
Figure 6.9 Effect of Evacuee Waiting Time and Vehicle Cost on In-Vehicle Travel Time .............................................................74
Figure 6.10a Total Travel Time of Transit and Vehicular Evacuation vs. Percentage of People Evacuated by Car ..............................................78
Figure 6.10b Average Travel Time of Transit and Vehicular Evacuation vs. Percentage of People Evacuated by Car ..............................................78
Figure 7.1 GTA Regions for the year 2001 (source: http://www.jpint.utoronto.ca/gta01/GTA.html) ..................................................80
Figure 7.2 Drive, NonDrive and Resident Distributions in the Greater Toronto Area ..........................................................83
Figure 7.3 Total Population Spatial and Temporal Distribution in the Greater Toronto Area ..........................................................84
Figure 7.4 Drive Population Spatial and Temporal Distribution in the Greater Toronto Area ..........................................................85
Figure 7.5 NonDrive Population Spatial and Temporal Distribution in the Greater Toronto Area ..........................................................86
Figure 7.6 Resident Population Spatial and Temporal Distribution in the Greater Toronto Area ..........................................................87
Figure 7.7 Total Population Distribution of Evacuees in the City of Toronto ..........................................................88
Figure 7.8 Drive Population Spatial and Temporal Distribution in the City of Toronto ..........................................................90
Figure 7.9 NonDrive Population Spatial and Temporal Distribution in the City of Toronto ..........................................................91
Figure 7.10 Resident Population Spatial and Temporal Distribution in the City of Toronto ..........................................................92
Figure 7.11 Population Distribution within the City of Toronto Zoning System by Time of Day ..........................................................93
Figure 7.12 Location of the Maximum Number of People in the City of Toronto ..........................................................94
Figure 8.1 The GTA Traffic Analysis Zones .............................................99
Figure 8.2 The Greater Toronto Area Simulation Model .......................100
Figure 8.3 The City of Toronto Simulation Model ................................100
Figure 8.4 The TTC Rapid Transit Lines, Stops, and Fleet Characteristics ..................................................................................102
Figure 8.5 Representation of the Noncompliant Traffic across the GTA 106
Figure 8.6 Geo-Locations of Evacuees within TAZ (partial view of the City of Toronto) ..........................................................107
Figure 8.7 Buffer Zone Representation and Geo-Locations for Evacuation Demand Assigned to the Subway ..................................................108
Figure 8.8 Distribution of Evacuation Demand at Subway Stations ..........109
Figure 8.9 Geo-Locations of Evacuees Assigned to Shuttle Buses ...........110
Figure 8.10 Major Bus Stops and TTC Routes for Bus Shuttle Evacuation ..................................................................................111
Figure 8.11 Distribution of Evacuation Demand at Bus Stations ..........112
Figure 8.12 Spatial Distribution of the TTC Fleet at the Onset of the Evacuation Scenario ..........................................................114
Figure 8.13 Evolution of the Optimization Process using the Genetic Algorithm ..................................................................................117
LIST OF TABLES

Table 2.1 Contra-flow Flow Rates for Four-Lane Freeways ................................................................. 15
Table 3.1 OSTE Formulation .................................................................................................................. 25
Table 4.1 Comparative Literature Survey ............................................................................................. 37
Table 4.2 MDTCPD-VRP Formulation ................................................................................................... 42
Table 6.1 Comparative Analysis of Evacuation Scenarios ................................................................. 73
Table 7.1 Survey Area Zone Numbering for GTA Regions ............................................................... 81
Table 8.1 The GTA and the City of Toronto Model Characteristics ...................................................... 101
Table 8.2 Evacuation Demand in Each Zone in the City of Toronto by Mode at Noon ................... 104
Table 8.3 Genetic Algorithm Methods and Parameter Values .............................................................. 115
Table 8.4 Comparative Analysis of Evacuation Strategies ................................................................. 127
Table 8.5 Measures of Effectiveness of Rapid Transit Lines * ............................................................... 134
1 Introduction

1.1 Background

Canada’s history of man-made disasters and natural threats mimics that of the United States but to a lesser extent. While emergency evacuation and management systems in Canada have not experienced the frequency or severity of events in the United States, emergency management in Canada has drawn significant experience from the events in the United States.

To date, in Canada, flooding has contributed the most in terms of property damage. Spring thaw caused severe flooding; for example, the 1997 Red River flood in Manitoba resulted in huge losses. Transportation accidents have been also known as a significant contributor to property damage and losses in Canada; for example, the Halifax explosion in 1917 led to the launch of the first initiative for academic research in disaster management. It was also considered the largest loss of life (estimated 1960 deaths) in a single event in Canada. Fortunately, Canada has been spared the large-scale acts of terrorism that other countries have been recently experienced. Nevertheless, Canada has experienced some acts of violence such as the 1989 killing of 14 women at the École Polytechnique in Montreal and the Front de libération du Québec (FLQ) kidnapping in 1970 (Lindsay, 2009).

The following paragraphs chart the events that are relevant to planning for emergency evacuation from a transportation perspective and shed some light on the lessons learned from them.

In 1954, Hurricane Hazel blew up from the Caribbean, crossed North Carolina, passed over Washington DC and New York State, and then the storm intensified and hit the Toronto area, bringing 110 kph winds and 285 millimetres of rain. At that time, the management system was limited in its response to this storm and the Humber River flooded, resulting in 30 deaths and 14 homes on Raymore Drive were washed away; this brought the total losses to 81 lives and damage estimated at $1 billion.

---

2 Statistics in this section are from a paper written by John Lindsay, Assistant Professor and Chair, Applied Disaster and Emergency Studies, Brandon University, who discusses the history of Emergency Management in Canada.
In 1979, a huge explosion was caused by a train derailment in Mississauga, Ontario. Most of the train cars were carrying hazardous material that could potentially produce a toxic cloud of chlorine, a threat that led officials of the city to order the largest peacetime evacuation in Canada. Despite the fact that 225,000 people evacuated the city and there were no deaths, this event resulted in renewed interest in proper planning for emergency evacuation.

In 1998, an ice storm struck Ontario causing significant losses in peoples’ lives in the national capital. The death of 28 people, the chaos of 600,000 evacuees and the declaration of over 250 communities a state of emergency made it Canada’s largest disaster.

In 2005, the Atlantic coast of the United States was hit by Hurricanes Katrina and Rita. In hurricane Katrina, out of the 1.4 million inhabitants in the hazard areas, 200,000 to 300,000 people did not have access to reliable personal transportation. In Hurricane Rita, an estimated 3 million people evacuated the Texas coast, creating 100 mile-long traffic jams. After drivers crawled 30 miles in severe congestion, some decided to turn back home (Litman, 2006). Numerous lessons were learned from the consequences of these two hurricanes; most notable were the poorly integrated public transportation services in Katrina and the lack of communication and excessive dependence on automobile evacuation in Rita.

Given the drastic consequences of large-scale emergencies, proper development of emergency evacuation plans is paramount in which the core objective is to move people from potential hazard zones to safe destinations in the quickest and most efficient way. Also, given the typically diverse demographic characteristics of most communities, an efficient evacuation plan should integrate multiple modes to aid transit-dependent people who have no access to automobiles at the time of evacuation or at all.

Numerous notable emergency evacuation planning models have been developed over the past few decades. They propose and investigate the effect of one or more strategies that have the potential to improve the performance of the evacuation process. Approaches in the literature use various modelling and optimization techniques. However, these studies are typically focused on automobile-based evacuation using a certain strategy (e.g. evacuation scheduling) without considering other modes of transportation or attempting to simultaneously synergize several other possible strategies, such as destination choice optimization, route selection optimization,
etc. In addition, the absence of an accurate representation of the spatial and temporal distribution of the population makes estimating the evacuation demand difficult and hence renders the interpretation and conclusions of the methods themselves questionable. Furthermore, the lack of any quantification of mode-specific populations (e.g. transit-dependent) magnifies the vulnerability of these people to threats in the case of large-scale emergencies, a major drawback of existing planning models.

In order to address the limitations of existing evacuation planning models, an integrated multimodal evacuation modelling framework is needed that:

1. synergistically combines all or some of the promising strategies to further improve the efficiency of the evacuation process.
2. optimizes multimodal evacuation using global and efficient optimization techniques while addressing multiple possible objectives.
3. is sensitive to time-dependent and stochastic traffic and transit service characteristics (supply modelling)
4. accurately estimates the spatial and temporal distribution of the population (demand modelling).

1.2 Research Objectives

The objective of this dissertation is to develop a novel framework that optimizes the evacuation of densely populated areas using multiple modes, including vehicular traffic and public transit. It also presents a large-scale application of the proposed modelling framework to the evacuation of the entire City of Toronto in the case of an emergency.

This investigation focuses on the issues concerning the optimization of multimodal emergency evacuation, in particular the optimization of two modules: vehicular traffic and mass transit evacuation. The study considers multiple dimensions of the vehicular traffic optimization problem: evacuation scheduling, destination choice and route choice. Such dimensions were either simplified or isolated in existing approaches in the literature. It also considers multiple variants of the Vehicle Routing Problem (VRP) to model the mass transit evacuation problem: Multiple Depots, Time Constrains, and Pick-up and Delivery. Both modules are integrated in one
framework that is capable of modelling the dynamics of the evacuation problem, as well as best utilizing the existing infrastructure (i.e. transportation network) in the case of emergencies.

The proposed framework integrates two modules, namely the Optimal Spatio-Temporal Evacuation–OSTE and the Multiple Depot Time Constrained Pick-up and Delivery VRP–MDTCPD-VRP modules. OSTE is achieved by optimizing the evacuation scheduling problem and the destination choice problem simultaneously in a Dynamic Traffic Assignment (DTA) environment while solving the optimization problem using Evolutionary Algorithms (EAs). The MDTCPD-VRP is formulated with additional supply, demand and time constraints to better mimic the evacuation situation; therefore, constraint programming (CP) and neighbourhood search techniques are simultaneously used to solve a constraint satisfaction problem (CSP) and an optimization problem. Both optimization platforms are further augmented with an accurate representation of the transportation supply using simulation-based Dynamic Traffic Assignment, and an accurate estimation of the evacuation demand using regional travel survey data.

In summary, the following items represent the research objectives:

- Accurate assessment and representation of the transportation infrastructure, specifically the roadway and public transit networks (transportation supply).
- Estimation of the spatial and temporal distribution of population (transportation demand).
- Identification of available modes and captive populations of certain modes.
- Integrated framework that accounts for various evacuation strategies such as evacuation scheduling, route choice and destination choice simultaneously.
- Multimodal evacuation strategies that synergize the effects of multiple modes.
- Accounting for background traffic or for noncompliant evacuees (the percentage of travellers not following the evacuation plan).

While the aforementioned elements constitute separate modelling, analysis, optimization and operational tasks, they are closely interrelated. Each is indispensable for the design and implementation of an effective emergency evacuation plan. This dissertation is geared towards amalgamating all these elements in one framework that is designed to optimize plans for realistic large-scale multimodal emergency evacuations.
1.3 Thesis Roadmap

Based on the functions and objectives of the evacuation planning framework, the primary research sections of this dissertation have been organized into ten chapters. The thesis roadmap and the interrelations among these chapters are illustrated in Figure 1.1.

Figure 1.1 Thesis Organization

The introduction chapter of the dissertation starts with a brief description of Canada’s history of disasters and the challenges that faced the early emergency management systems. It also
addresses the major limitations of existing approaches and outlines the motivation and research objectives. The chapter also provides a high-level description of the proposed framework.

Chapter 2 provides an overview of the emergency evacuation concepts and approaches in the literature. It reviews and addresses the challenges in emergency evacuation planning and modelling. Section 2.8 summarizes the major challenges and gaps in the existing literature.

Chapter 3 discusses the conceptual development of the optimal spatio-temporal evacuation (OSTE) approach. It provides a high-level description of the framework components. It then introduces the time structures incorporated in emergency evacuation and their role in achieving certain objectives in emergency situations. Then it provides a mathematical formulation for OSTE. Discussion of the solution algorithm is presented while emphasizing the multi-dimensional and non-deterministic nature of the problem. Performance details of the approach are presented through a prototype implementation in Chapter 6 and a large-scale application in Chapter 8.

Chapter 4 discusses the conceptual development of the transit shuttling approach during evacuation. It provides a general description of the framework components and the connection to the typical Vehicle Routing Problem (VRP). It then introduces the analogy between the VRP and the mass transit evacuation problem. It also provides a mathematical formulation of the MDTCPD-VRP approach. Discussion of the solution algorithm is presented while emphasizing the constrained nature of the problem. Performance details of the approach are presented through a prototype implementation in Chapter 6 and a large-scale application in Chapter 8.

Chapter 5 presents the main lessons and findings of relevance to this research effort, attempting to build on these lessons while integrating multiple modes in the overall framework. It then highlights the modes that can be potentially incorporated in emergency evacuation plans. It ends with the description of a multimodal evacuation framework that combines the two optimization platforms discussed in Chapters 3 and 4.

Chapter 6 demonstrates the applicability and feasibility of the approaches presented in Chapters 3 and 4 in a prototype implementation. This prototype shows the main principles of the evacuation planning model; it does not, however, represent a large-scale implementation of the
model. A demand estimation model for emergency evacuation is paramount when producing realistic and meaningful results in the case of large-scale evacuation. Chapter 7 proposes a demand estimation model, in which not only the value of the evacuation demand per Traffic Analysis Zones (TAZs) is determined but also the spatio-temporal distribution of demand is examined.

The application of OSTE and MDTCPD-VRP to a large-scale case study that necessitates the evacuation of the entire City of Toronto is presented in Chapter 8. The large-scale application demonstrates the essence of the proposed approach and builds on the lessons learned from the prototype implementation described in Chapter 6.

Chapter 9 provides a new vision for a comprehensive evacuation planning model. It begins by contrasting the typical four stage planning model to the current state-of-the-practice emergency evacuation planning models. It ends with a description of the presented planning model and how it could be extended to a closed-loop evacuation control system.

Chapter 10 summarizes the conclusions and contributions of this thesis with a discussion of a future research agenda.
2 Literature Review

This chapter provides an overview of emergency evacuation concepts and approaches. Numerous approaches have been the focus of many researchers to model/solve the emergency evacuation problem; therefore, this chapter categorizes the main approaches into subsections and, in each subsection, the previous studies are reviewed, the gaps/limitations are highlighted, and potential angles for new contributions are presented.

Emergency evacuation is the collective movement of people using multiple modes of transport from a hazard area (emergency protection zone or EPZ) to safe destinations (shelters) via specific routes. Emergency evacuation plans are necessary both in the case of man-made disasters (e.g. nuclear reactor failures or leaks or terrorist attacks) and natural disasters (e.g. hurricanes, floods, tsunamis, earthquakes, or tornados).

Transportation networks in cities evolve over long spans of time in tandem with population growth and evolution of travel patterns. In emergency cases, travel demand and travel patterns drastically change from normal everyday volumes and patterns. Given that most US and Canadian cities are already congested and operating near capacity during peak periods, network performance can severely deteriorate if such drastic changes in Origin-Destination (O-D) demand patterns occur during or after a disaster (Tuydes and Ziliaskopoulos, 2004). Loss of capacity due to the disaster and associated incidents can further complicate the matter. The primary goal when a disaster or hazardous event occurs is to coordinate, control and possibly optimize the utilization of the existing transportation network capacity. Emergency operation management centres face multi-faceted challenges in anticipating evacuation flows and taking proactive actions to guide and coordinate the public towards safe shelters (Chiu et al., 2006).

Designing a transportation network for evacuation demand patterns is financially infeasible. One option is to better utilize the available network capacity by reallocating it more efficiently, possibly in the form of “contra-flow” operation (i.e. reversing the capacity in the opposite direction into the direction of evacuation). Another alternative is to stage evacuation demand
such that the population is advised to evacuate according to an announced schedule rather than simultaneously. In general, the evacuation problem can be modelled as an optimization problem to achieve certain objectives (e.g. minimize the evacuation time, minimize the network clearance time). The control variables can combine capacity allocation, evacuation scheduling, destination (shelter) choice, traffic routing, and traffic control, to name a few.

Emergency evacuation has been extensively investigated in the last two decades due to frequent natural and man-made disasters. Numerous studies have explored the emergency evacuation problem; some are robust and have quantitatively addressed pressing issues; others have presented qualitative measures of emergency evacuation. The following sections review and highlight the challenges in emergency evacuation planning and modelling.

2.1 Evacuation Scheduling

Scheduled/staged evacuation is a widely used control strategy to guide evacuation flows. Evacuation scheduling aims to better distribute/manage the evacuation demand over an evacuation horizon. In simultaneous evacuations, evacuees are advised to evacuate immediately to their destination; whereas in staged evacuation, evacuees are advised when to evacuate so as to achieve certain objectives (e.g. minimize network clearance time) (Sbayti and Mahmassani, 2006). By managing the evacuation surge by holding some evacuees at their origins, scheduled evacuation can effectively reduce overall network congestion and, more importantly, mitigate potential casualties, stress levels, and chaos caused by evacuees being blocked in hazard areas (Liu, 2007).

In staged evacuation, the most critical decision is the time to issue the evacuation orders for the evacuation zones. Once an evacuation order is announced, the evacuees’ responses will determine the demand generation, which requires continuous monitoring to track network conditions and the evolution of the evacuation process. Determining these starting times during a staged evacuation is the subject of many studies with varying degrees of comprehensiveness. Chen and Zhan (2006) investigated the effectiveness of simultaneous (concurrent) and staged evacuation strategies in three road network structures using Paramics as a microscopic simulator. The study concluded that staging the evacuation process is essential in communities where the street networks have a “Manhattan structure” and the population density is high.
Using a small hypothetical network, Mitchell and Radwan (2006) proposed a heuristic prioritization of emergency evacuation to reduce the network clearance time. Zonal parameters that might affect the staging decision are defined, these include: population density, road exit locations/capacity and major evacuation routes. Chiu et al. (2006) presented a system optimal dynamic traffic modelling technique for solving the evacuation destination-route-flow-staging problem for no-notice events. The algorithm was based on the Cell Transmission Model (CTM) (Daganzo, 1994) and the LP formulation proposed by Ziliaskopoulos (2000). The framework was applied to a simple hypothetical evacuation event in which cell flows at each time interval were reported. The study concluded that the optimal solution of the presented LP depicts the optimal joint evacuation destination-flow-staging decision in an effective manner.

Sbayti and Mahmassani (2006) introduced an optimal evacuation scheduling approach with two assumptions to relax destination selection and evacuee compliance, in which they assumed that evacuees would adhere to evacuation guidance information and not switch to different departure times, destinations, or paths. It was also assumed that a controller would provide pre-trip, variable message signs and en-route information for non-evacuees who are on their way to the impacted zone for necessary detours and route changes. The final output of the study was an optimal loading curve that minimized the total network clearance time. Abdelgawad and Abdulhai (2009) proposed an optimal spatio-temporal evacuation (OSTE) strategy that optimizes the scheduling and destination choice problems simultaneously. The OSTE platform is built on the interaction between DTA and Evolutionary Algorithms (EAs). The output of OSTE is guidance to evacuees as to when to leave their origins, where to go (optimal destination) and how to get there (by which route) in the quickest possible way.

### 2.2 The Concept of One Destination

In conventional evacuation planning models, evacuees are assigned to pre-determined destinations that are based primarily on the geographical context and their daily activities. The traditional method of assigning evacuees to a pre-specified fixed OD table might result in sub-optimal system performance due to congestion, road blockage, chaos, incidents, hazards associated with the emergency situation and limited destination/shelter capacity. One promising concept to address this problem is to relax the constraint of assigning evacuees to fixed destinations. In other words, instead of assigning the demand to fixed destinations, evacuees are
directed to the destinations that they can reach in minimal time. From a modelling perspective, this can be achieved by adding one augmented dummy destination beyond all destinations as shown in Figure 2.1.

Chiu et al. (2006) and Yuan et al. (2006) proposed the concept of a One-Destination (1D) evacuation model in which the traditional road network with \(m\) origins to \(n\) destinations (as shown in Figure 2.1-a) has been transformed to a network with \(m\) origins to one destination (as shown in Figure 2.1-b). The modified network is augmented with dummy edges that link each real world destination to the one common dummy destination (D*). The added dummy links are assumed to have unlimited capacity. Yuan et al. (2006) reported that a reduction of approx. 60% in the overall evacuation time can be achieved when testing the case of a regional evacuation due to a nuclear power plant mishap, and a reduction of 80% in the overall evacuation time when modelling traffic routing and en-route information accompanied by the 1D framework.

![Figure 2.1](image)

Figure 2.1 Original Multiple-Destination (nD) and Modified One-Destination (1D) Networks (Yuan et al., 2006)

### 2.3 Traffic Signal Control

As an efficient control strategy, traffic signal control has been widely accepted to improve arterial road capacity and reduce congestion during daily traffic and, more importantly, during emergency evacuation.

Sisiopiku et al. (2004) utilized SYNCHRO (software for optimizing traffic signal timing) to set up the optimal signal-timing plans for a small area in Birmingham, Alabama. CORSIM, is then
used to test different evacuation scenarios and evaluate the impacts of signal-timing optimization on selected measures of effectiveness. The study concluded that signal optimization during evacuation could significantly reduce average vehicle delays and improve network clearance time. Chen et al. (2007) applied CORSIM to two evacuation corridors in Washington, D.C. and examined four different signal-timing plans: Red Flash, Yellow Flash, Minimal Green, and Ordinary Peak Hour Plan. Although this study offered some insights into the effect of various timing plans, its analysis of plan selection under various evacuation scenarios is mostly qualitative.

Another area of interest in emergency signal control is emergency vehicle pre-emption (EVP), in which the emergency vehicles override all other traffic movements and, thus, may affect the evacuation traffic. An impact analysis of EVP using CORSIM was conducted by McHale and Collura (2003). Among a series of studies to evaluate EVP impacts on traffic conditions, Louisell et al. (2004) proposed a conflict point analysis approach to evaluate the potential safety benefits of EVP. Furthermore, they developed a worksheet method to assess the crash reduction benefits of EVP on a given intersection during a pre-emption signal phase. However, to the author’s knowledge and based on the literature, no published research has been reported in the area of emergency vehicle pre-emption during evacuation scenarios.

2.4 Traffic Routing and Control Strategies in Emergency Evacuation

Traffic routing, as one of the main traffic control strategies, aims to identify the best set of routing decisions so as to fully utilize the available capacity of a transportation network and assign traffic to those routes accordingly. User equilibrium (UE) assignment means that drivers follow time-dependent least travel time paths, while system optimum (SO) assignment results from drivers following time-dependent least marginal travel time paths. Under UE, all used routes between a given O-D pair have the same travel time and hence the network reaches a stable equilibrium state. Under SO on the other hand, some users may be assigned to routes that are longer while others are assigned to shorter routes. SO traffic patterns are therefore neither stable nor equitable. Simulation-based approaches have become flexible enough to perform UE, SO, or multiple user class assignment, although the equilibrium conditions are only heuristically approximated (Peeta and Ziliaskopoulos, 2001).
From this perspective, there are two schools of thought in DTA during evacuation. One group argues that in emergency evacuations, the major concern of planners is the overall system performance; therefore, it is more plausible to use system-optimized traffic assignment models. Among these studies are the works conducted by Sbayti and Mahmassani (2006) on evacuation scheduling and by Han and Yuan (2005) on destination optimization for emergency evacuation assignment using SO DTA. Conversely, the other group argues that aiming for SO in actual evacuation operations is neither practical nor desirable from an equity perspective because the SO traffic assignment assumes full adherence of evacuees to the evacuation guidance information, departure times, destinations and paths. On the other hand, it is implausible to expect that routes would have equal and minimum travel times, as in fact many evacuees will most likely choose suboptimal routes. The evacuation conditions therefore do not resemble the basic definition of Wardrop’s user equilibrium or any other known assignment technique (Brown et al., 2009), a concern that motivated the work done by Chiu and Mirchandani (2008) on dynamic traffic management for emergency evacuation. They quantified system performance in the case of emergency evacuation in a small hypothetical network. The study compared the results from the route choice decisions made by evacuees and the SO pre-trip route guidance scheme. A controller was designed to influence system performance towards an optimum level. The results showed suboptimal system performance due to driver deviation from optimal paths.

2.5 Role of Intelligent Transportation Systems (ITS) in Emergency Evacuation

Real-time traffic information is crucial during emergency evacuation. Several states/cities plan to incorporate ITS technologies into the evacuation process. ITS allow traffic information to be collected remotely from the field and disseminated through a traffic management centre. This information can help emergency managers monitor the status of contra-flow operations, evacuation scheduling, congested traffic routes, incidents; and thus proactively deal with the evacuation situation. A wide matrix of tools can be utilized in the case of emergency evacuation. Closed Circuit Television (CCTV) cameras can provide detailed information to transportation officials and emergency managers. CCTV cameras can provide real-time traffic data for incident detection and verification that an incident has been cleared. This information can then be disseminated to evacuees or used by the Traffic Management Centre (TMC) to divert evacuees to less congested routes (Urbina, 2001).
In addition, Variable Message Signs (VMSs) and Highway Advisory Radio (HAR) (FHWA, 2010) have been shown to be useful tools for disseminating evacuation information and controlling traffic routing. VMSs can be fixed or mobile and are typically installed before bifurcation points to reroute traffic during emergency evacuation. Under routine conditions, HAR is used for traffic information broadcasting over small geographical areas. During emergency evacuation, evacuees can use these information technology systems to obtain alternative evacuation route information, congestion locations, incident information and shelter locations. In summary, numerous ITS technologies and communication techniques should be considered to improve emergency evacuation and potentially reduce evacuation time.

2.6 Contra-flow Operation in Emergency Evacuation

Contra-flow operations refer to flows in the reverse direction on one or more lanes of a road segment, thereby increasing the flow capacity in the heavier flow direction without constructing additional lanes (Urbina, 2001). The contra-flow concept has been considered in many cities in the United States as a way to improve the roadway capacity during routine rush traffic flow hours. Contra-flow has also been implemented during special events (e.g. concerts, football games) to accommodate the outbound traffic at the end of the event (FHWA, 2010).

The concept of contra-flow was first proposed in the 1980s by the Federal Emergency Management Agency (FEMA) for use as a last resort during emergencies and was originally planned for use during potential nuclear missile attacks. Considering the increased frequency of hurricanes in heavily populated areas, contra-flow operations have become a valuable option for moving people out of threatened areas (Wolshon, 2001).

Although contra-flow evacuation has been widely implemented, there is little comprehensive research on the costs and benefits of its use. FEMA (Post, 2000) investigated the cost/benefit of capacity improvements resulting from the implementation of contra-flow operations. In their study, traffic volumes collected by Florida, Georgia, and South Carolina during Hurricane Floyd were considered as a basis for computing planning-level roadway capacities for evacuating traffic under different conditions. According to this report, depending on the configuration of contra-flow lanes, a 30 to 70 per cent increase in capacity over conventional operations can be achieved by the use of contra-flow lanes. Table 2.1 shows an example of such values.
Table 2.1 Contra-flow Flow Rates for Four-Lane Freeways

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Estimated Average Total Outbound Capacity (veh/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Two Way Operations</td>
<td>3,000</td>
</tr>
<tr>
<td>Three Lane (one contra-flow lane)</td>
<td>3,900</td>
</tr>
<tr>
<td>Three Lane (using outside shoulder)</td>
<td>4,200</td>
</tr>
<tr>
<td>All lanes reversed for evacuation (no shoulder lanes)</td>
<td>5,000</td>
</tr>
</tbody>
</table>

Source: (Post, 2000)

Most of the research conducted on the contra-flow problem is based on the Cell Transmission Model (CTM). The CTM is based on fundamental traffic flow theory diagrams (e.g. trapezium shape) (Daganzo, 1994; Muñoz et al., 2006). The CTM discretizes the time period of interest into small intervals and then divides every link of the transportation network into homogenous segments, called cells, such that the length of each cell is travelled in one time interval while moving at free-flow speed. For each cell, the hydrodynamic flow equations result in two sets of equations, cell mass conservation and flow propagation. However, this approach uses a continuous variable to represent reversible capacity, which may not be practically implementable (Bell and Iida, 1997). Furthermore, the absence of dynamics incorporated in the traffic assignment process is a major limitation of such models especially in the case of emergency evacuation. Alternatively, microscopic simulation models (e.g. CORSIM, PARAMICS) were assessed and used to model contra-flow operations (Theodoulou and Wolshon, 2004; Lim and Wolshon, 2005; Meng et al., 2008). However, CORSIM and PARAMICS, like most traffic simulators, do not explicitly support the creation of reversible flow freeway segments or the behavioural characteristics of evacuation drivers. Therefore, special attention should be given to reasonably modelling the contra-flow operation.

Termination points of contra-flow segments are the points where traffic is diverted from the contra-flowing lane to normal lanes. Although critical to the operation of contra-flow, these points are not adequately investigated in the literature. Lim and Wolshon (2005) assessed and compared the operational characteristics of contra-flow evacuation termination points in an attempt to fill this gap. CORSIM is used to model contra-flowing freeway traffic under evacuation conditions. The microscopic model was used to comparatively rank the pre-defined termination designs and to investigate the effectiveness of various models, including the effect of reducing traffic volumes before termination. The study concluded that it is better to maintain all
lanes through the termination point. In addition, it is preferred to reduce the volume entering the termination point by maintaining exit points along the route.

In an attempt to solve the contra-flow problem, Meng et al. (2008) addressed the optimal contra-flow lane configuration. They formulated the problem as a bi-level programming model in which the upper level problem is a binary-string formulation that aims to minimize the total travel time, while the lower level problem is a microscopic traffic simulation model (PARAMICS) that dynamically simulates the reaction of drivers resulting from a contra-flow lane configuration. Due to the complexity and nonlinearity of the problem, a genetic algorithm (GA) is used to search for the optimal lane configuration at the upper level. In this study, a feasible set of links are defined as candidate links prior to the optimization process. One downside of the proposed study is that extra lanes (shadow lanes) need to be created in the opposite direction of each candidate link in the simulation model so that the necessary direction-reversing operation can be implemented through PARAMICS. The GA searches in the vicinity of the candidate links for the optimal set of links and the lanes to be reversed.

Recently, there has been strong interest among researchers and planners in adapting contra-flow during emergency evacuation to better optimize the use of the existing infrastructure capacity. Although most of these studies, (Theodoulou and Wolshon, 2004; Tuydes and Ziliaskopoulos, 2004; Lim and Wolshon, 2005; Kim et al., 2008; Meng et al., 2008) show operational improvements, safety and practical operation continue to be issues, especially for unfamiliar drivers with the contra-flow operation. Brian Wolshon (2001) pointed out that reverse flow scenarios are not without significant problems. These problems include the safety risks associated with reverse flow on interstate freeways; the fact that traffic control devices and safety devices are not designed to accommodate contra-flow; and potential problematic operation near access and termination points. Despite the wide acceptance of contra-flow operations in practice, limited research has been published regarding how to choose the links or lanes to be reversed for contra-flow operations in an optimal manner that maximizes effectiveness under resource limitations.
Travel Behaviour in Emergency Evacuation

Travel behaviour and the compliance of evacuees with an evacuation order are important factors that should be modelled or realistically taken into account during evacuation operation. Travel behaviour under emergency conditions is expected to differ considerably from day-to-day travel patterns. Fu and Wilmot (2004) proposed a sequential logit model to simulate evacuee behaviour in the case of hurricane evacuation. However, little research has been conducted in the area of driver stress and aggression that would certainly increase under such conditions. Zhi et al. (2010) investigated the driver perception-reaction times (PRT) under emergency evacuation situations. Using a driving simulator to model emergency situations and a survey to validate the driving simulator environment, the study concluded that the value of PRT in normal situations is greater than that under emergency situations. It has also been hypothesized that higher levels of confusion might result from unfamiliarity of driving in contra-flow conditions. Moreover, it is anticipated that more incidents will occur under emergency evacuations; however, incidents during evacuation are rarely modelled in the literature apart from the sensitivity analysis performed by Chen et al. (2007). They randomly assumed three incident levels (minor, medium, major) in different locations on a small tested network to capture the effect of incidents.

Alsnih and Stopher (2004) acknowledged that there is a major gap in modelling driver behaviour. It is a challenging issue to resolve, since it involves identifying how evacuees would perceive an evacuation order, as mandatory or merely recommended. Alsnih and Stopher (2004) also concluded that households may not follow an evacuation order for numerous reasons, such as preferring to stay to protect property, not seeing neighbours evacuate, and obviously the inconvenience associated with the evacuation process. How to model and capture the above behaviour is still a major concern needing further investigation. Southworth and Chin (1987) posed the question of whether evacuees seek the safest, nearest, or farthest destination from the hazard location. The study showed that evacuees’ choice of exit (destination) can be taken from one of the following possibilities:

- Exit to nearest shelter;
- Exit depending on location of friends and relatives and travel speed of approaching hazard;
- Exit towards pre-specified destinations, depending on evacuation plan in operation;
Evacuees departing from the area according to the underlying traffic conditions of the network at the time of evacuation (allows for myopic evacuee behaviour).

2.8 Summary

Despite the numerous approaches that have significantly contributed to improving evacuation strategies, an integrated optimal evacuation strategy still needs further research. More effort is needed to synergistically combine all or some of the promising strategies to further improve the efficiency of the evacuation process. Traffic routing, evacuation staging, destination optimization can be possibly combined into a comprehensive portfolio of solutions. However, optimizing the evacuation problem necessitates extensive modelling, design, and analyses that should capture the dynamic interaction between the analytical part of the evacuation process, the operational side of the transportation network and the behaviour of evacuees. Some analytical challenges faced in the development of such integration are:

- The lack of a hazard prediction method that could potentially define the evacuation area and duration and risk level associated with the hazard.
- The lack of a travel demand estimation method or model for predicting the population to be evacuated within the hazard area and their flow patterns. This demand model should incorporate, identify, and predict demand by mode (transit, driving, etc.) under atypical evacuation circumstances.
- The lack of understanding of how best to disseminate evacuation orders so as to produce the desired scheduling or staging of the evacuation demand.
- The lack of estimation and modelling of destination selection and shelter capacity which can affect the success or failure of an evacuation plan.
- The lack of methods that are capable of dynamically assigning traffic to the transportation network while capturing the spatio-temporal characteristics of travel in such chaotic situations.

Alongside the analytical challenges, there are many pressing operational issues, such as:
The lack of traffic management tools that propose optimal control strategies while capturing the operational constraints embedded in each control strategy.

The lack of ubiquitous traveller information dissemination systems to update the system state and inform travellers of any disruptions during the evacuation process.

The lack of understanding of how to integrate the various control strategies in the field, such as contra-flow, staging, traffic routing, signal control and the use of ITS.

While the aforementioned issues constitute separate modelling, analysis or operational tasks, they are closely interrelated. Each is indispensable for the design and implementation of an effective emergency evacuation plan.
3 Optimization of Automobile Evacuation

This chapter discusses the conceptualization and development of the optimal spatio-temporal evacuation (OSTE) approach to the evacuation problem. It provides a high-level description of the framework components. It then introduces the time structures incorporated in emergency evacuation and their role in achieving certain objectives in emergency situations. It also provides a mathematical formulation of the approach. Discussion of the solution algorithm is presented while emphasizing the multi-dimensional and non-deterministic nature of the problem. Details on the performance of the approach are presented in later chapters with a prototype implementation in Chapter 6 and a large-scale application in Chapter 8.

3.1 Scheduling and Destination Choice

In emergency situations that require population evacuation, particularly in dense urban areas, evacuees may react chaotically causing severe congestion, gridlock and excessive delays. Such uncontrolled evacuation may also expose evacuees to further harm, especially in cases with a high time pressure to evacuate. Therefore, the challenge of prompt evacuation of dense urban areas has made evacuation demand management an essential priority in planning for emergency situations. In addition, given the fact that the distribution of evacuation demand has unique spatio-temporal characteristics that are different from common everyday demand patterns used in conventional transportation planning, studying the demand distribution in both time and space is crucial. Concurrent temporal and spatial management of evacuation demand is central for successful emergency evacuation planning. Optimal Spatio-Temporal Evacuation (OSTE) is achieved by optimizing the evacuation scheduling and destination choice problems simultaneously. OSTE, a comprehensive and extendible tool that is capable of generating optimal demand management plans for realistic-size networks, is founded on the interaction between DTA, evacuation demand scheduling and destination choice.

Conceptually, the evacuation process follows two stages: mobilizing evacuees into the transportation network and evacuating the hazard area towards safe shelters. The two stages are
represented by two cumulative flow curves as shown in Figure 3.1. The loading (mobilization) curve, \(L(t)\), represents the demand entering the system and the evacuation curve, \(E(t)\), represents the arrival to safety, i.e. exit flows. In the first stage, evacuee loading can be managed by temporally staging the demand in an optimal manner that fulfils a predefined objective function. Objectives can include, but are not limited to, maximizing the number of evacuees reaching safe destinations over a pre-defined evacuation horizon or minimizing the number of evacuees en route or minimizing the waiting time.

![Figure 3.1 Effect of Evacuation Scheduling and Destination Choice on the Loading and Evacuation Curves](image)

Destination choice refers to guiding evacuees to the nearest safe refuge and not necessarily their homes or routine daily destinations. Choice of destinations where evacuees seek refuge can significantly improve the efficiency of the evacuation process by reducing unnecessary longer trips (Chiu et al., 2006; Yuan et al., 2006). Unlike the fixed, predefined destinations in conventional transportation planning models, optimal destination choice offers more flexibility in both destination and route choice. This may prevent overloading critical routes given the fact that some routes to destinations may be blocked or damaged by the disastrous event. Therefore, it may be better to direct traffic to the nearest safe zone, i.e. optimize destination choice to fulfil the desired objective. This can be achieved by amalgamating all destinations into one hypothetical super-zone or safe destination. Evacuees would seek this destination in the fastest possible manner, i.e. reaching a safe zone quickly which is not necessarily their home.
For illustration purposes, three sets of curves are sketched conceptually in Figure 3.1, in which each set characterizes a certain level of demand control; simultaneous evacuation (SE), Optimal Temporal Evacuation (OTE) and Optimal Spatio-Temporal Evacuation (OSTE). The case of SE mimics an evacuation scenario where evacuees are advised to leave the area instantaneously without prior information about the destination, i.e. no control. In the case of OTE, evacuees receive guidance on when to start the evacuation but without specifying a destination. In the case of OSTE, evacuees receive guidance on when to evacuate (departure time) and where to go (optimal destination) as well as how to get there (optimal route). As depicted in the figure, the horizontal distance between the loading and evacuation curves is the travel time experienced by evacuees, and the vertical distance is the percentage of evacuees queued in the system, i.e. en route, Q(t). Therefore the ultimate goal is to push the loading and evacuation curves upwards and to the left as shown in the figure while minimizing the area in between them, hence minimizing both the travel time and number of vehicles en-route. It is clear that the optimal evacuation curve and the optimal loading curve are interdependent.

The concept of evacuees seeking a safe super-zone destination in the case of emergency evacuation can be modelled using any plausible simulation tool that has the capability of modelling complex network structures. In a modified network representation with hypothetical zero travel time links added from all destination zones to the super sink zone, drivers seek and reach the super-zone destination in the simulation model in the shortest or fastest way. Therefore, the simulation model produces the optimal “exit” point for each vehicle as will be illustrated further in the case study. To deploy the evacuation plan, as in our case, evacuees are instructed to head for the corresponding exit points/destinations obtained from the analysis. Communicating such information to evacuees is possible through numerous emerging ITS technologies and pervasive mobile communication and computing platforms.

3.2 Time Structures in Evacuation

The evacuation problem comprises three interconnected time structures: warning time, mobilization time and evacuation time (Sorensen and Sorensen, 2006). The warning time is announced based on the nature of the event; in no-notice evacuation the warning time is zero, while in expected evacuation scenarios such as hurricanes, jurisdictions typically announce a warning to the public according to a predefined plan. The mobilization time/loading time (how
groups of people are evacuated over time) is highly variable depending on the nature of the event and the level of urgency. The mobilization or loading pattern does significantly affect the overall evacuation process. Following the mobilization time is the evacuation time, which is the time evacuees spend travelling through the transportation network seeking safe destinations. Figure 3.2 illustrates the temporal patterns of mobilization and evacuation, i.e. the mobilization/loading curve, L(t), and evacuation curve, E(t). The time at which all evacuees reach a safe destination is defined as the network clearance time, T. It is important to note that both curves are dynamically interacting in emergency evacuation. In other words, the loading curve can be optimized rather than being assumed fixed as in most conventional hurricane evacuation studies. Also, the evacuation curve can be optimized by shifting it as close as possible to the loading curve in order to minimize the number of evacuees trapped in congestion.

![Figure 3.2 Multiple Objectives in Evacuation](image)

### 3.3 Typical Objective Functions in Emergency Evacuation

The evacuation problem has been typically solved as an optimization problem that minimizes/maximizes a certain objective function subject to supply, demand and time constraints. Numerous objective functions have been formulated in the literature with the goal of expediting the evacuation process. The most common objective functions addressed in the emergency evacuation literature can be summarized according to the following taxonomy (see Figure 3.2):

- Minimize the Evacuation Travel Time (area A₁ in Figure 3.2) (Hobeika and Kim 1998; Tuydes and Ziliaskopoulos 2004; Chiu, Villalobos et al. 2006; Tuydes and Ziliaskopoulos 2006; Yuan, Han et al. 2006; Liu, Ban et al. 2007)
- Minimize Network Clearance Time (point T in Figure 3.2) (Sattayhatewa and Ran, 2000; Sbayti and Mahmassani, 2006)
- Maximize Throughput for a specified clearance time (Liu, 2007; Abdelgawad and Abdulhai, 2009)
- Minimize Total Travel Time and Total Waiting (mobilization) Time (Area $A_1 + A_2$ in Figure 3.2) (Sbayti and Mahmassani, 2006; Abdelgawad et al., 2010)

As presented in the above taxonomy, various objectives have been investigated in the literature. Although some objective functions seem to be similar, some key differences reside in how the objective function is formulated and solved. For instance, some studies show that the minimization of total system travel time is equivalent to providing system optimal information to evacuees from a traffic management centre with 100% compliance. System optimal traffic assignment is an appealing system-wide target; however, it implies that some evacuees will experience travel times longer than their best attainable individual travel time. It seems unrealistic to expect evacuees to offer such sacrifice during evacuation in order for the system to be optimized, not to mention the tort liability of the system operator. In addition, minimizing travel (en-route) times ignores the waiting time, i.e. how long evacuees have to be held at the origin before being allowed to commence evacuation. The network clearance time is certainly essential for evacuation planning; however, considering solely the network clearance time (end point T in Figure 3.2) might be misleading. This is due to the fact that two evacuation curves might result in the same network end/clearance time, but the network conditions at intermediate times and the system stability for both cases might differ significantly (an upward or downward curve at intermediate points) (Han et al., 2007). Alternatively, another measure for the effectiveness of the evacuation process is the time at which multiple percentages of evacuees have reached a safe destination; for example, Yuan et al. (2006) measured the time at which 25, 50, 75, 95, and 100% of the population have evacuated the hazard area.

Very few studies have explicitly considered minimizing the waiting time of evacuees at origins as well as the travel time of evacuees in the transportation network. A good evacuation model is one that minimizes the total system evacuation time including both waiting time and travel time; however, both objectives may be in conflict and the trade-off between the two is challenging. As shown in Figure 3.2, minimizing the waiting time for evacuees implies evacuating all the
population instantly, i.e. simultaneous evacuation which may lead to longer travel times in the system and longer total evacuation time. Simultaneous evacuation typically results in early gridlock and under-utilization of the available infrastructure. Also, minimizing the travel time implies delaying the evacuees at the origin (increasing the waiting time) so that the network conditions remain stable and evacuees can reach their destination in the least possible travel time. Although Sbayti and Mahmassani (2006) considered both the waiting time and travel time of evacuees in the optimal scheduling problem, the waiting time was not explicitly optimized. One way to achieve this compromise is defined by Abdelgawad et al. (2010) in which the problem is solved in order to minimize the waiting time of evacuees at origins as well as the en-route travel time of evacuees.

3.4 Optimal Scheduling and Destination Choice Problem Formulation

The decision variables in the formulation below are the staging percentages \( \mu = (\ldots, \mu_t, \ldots) \) for all evacuation zones with a modified super-destination network representation to account for the optimal destination. The problem is formulated as a multi-objective optimization problem in which waiting time and travel time are minimized, as opposed to the typical minimization of vehicle travel time. The formulation is presented in Table 3.1.

Table 3.1 OSTE Formulation

| Sets | | | |
|------|---|---|
| \( T \) | Evacuation planning horizon divided into equal evacuation (departure) intervals of time \( t \), where \( t = 1, \ldots, T \) |
| \( I \) | Set of evacuation zones (origins); \( i \in I \) |
| \( J \) | Set of safe shelters (destinations); \( j \in J \) |

| Parameters | | | |
|------------|---|---|
| \( D_{ij} \) | Evacuation demand from origin \( i \) to destination \( j \) |

| Decision Variables | | |
|-------------------|---|
| \( \mu \) | Demand scheduling/staging vector for evacuation: \( \mu = (\ldots, \mu_t, \ldots) \) where \( \mu_t \) is the percentage of demand released at time interval \( t \). The destination choice component is optimized through the simulation model. |

| Objective Function | | |
|-------------------|---|
• **Min Vehicle Travel Time**

\[
\text{Min } \sum_{i \in I} \sum_{j \in J} \sum_{t \in T} Y_{ijt} TT_{ijt}
\]

where \( TT_{ijt} \) is the evacuation travel time from origin \( i \) to destination \( j \) at time \( t \). It should be noted that travel times are dynamically changing according to the scheduling vector and the DTA model.

• **Min Waiting Time**

\[
\text{Min } \sum_{i \in I} \sum_{j \in J} \sum_{t \in T} Y_{ijt} WT_{ijt}
\]

where \( WT_{ijt} \) is the waiting time for evacuees travelling from origin \( i \) to destination \( j \) at time \( t \). It should be noted that waiting times are also dynamically changing according to the scheduling vector.

• **Min Waiting Time and Vehicle Travel Time**

\[
\text{Min } \sum_{i \in I} \sum_{j \in J} \sum_{t \in T} Y_{ijt} (TT_{ijt} + WT_{ijt})
\]

where:

\( Y_{ijt} \) = number of evacuees leaving evacuation zone \( i \) to go to destination \( j \) in time interval \( t \),

\( Y_{ijt} = \mu_i D_{ij} \)

<table>
<thead>
<tr>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow Conservation Constraint</td>
</tr>
<tr>
<td>Scheduling Range</td>
</tr>
</tbody>
</table>

### 3.5 The Optimization Approach

#### 3.5.1 Typical Solution Algorithms vs. Evolutionary Algorithm Approach

The evacuation planning problem has been typically formulated as an iterative bi-level optimization problem where, in the upper level, the objective function is optimized while keeping the traffic assignment parameters fixed (route flow patterns), whereas in the lower level, the traffic assignment problem is solved while keeping the upper level optimization parameters fixed (Chen and Zhan, 2006). The conventional solution algorithms for such bi-level optimization problems iterate between the two optimization levels (Sbayti and Mahmassani, 2006; Meng *et al.*, 2008; Xie *et al.*, 2009). To the best of the author’s knowledge and based on the literature, the optimization problem has been solved at the upper level with deterministic approaches such as: hill-climbing search method, the simplex method, the Frank-Wolfe algorithm, and gradient descent approach. These traditional optimization methods have been found problematic when applied to non-linear or high-dimensional problems in which the objective function cannot be analytically represented in closed form (Kruchten, 2003).
Furthermore, these approaches inherently search in the vicinity of the starting point and thus create a high dependency on the initial solution which can be an issue in very large search spaces. Moreover, traditional deterministic approaches follow a single path in the search space and may get stuck in local minima. This is to be contrasted to methods that rely on the evolution of multiple solutions. In addition, large-scale applications such as the evacuation of a large city may require ample computer processing power or even parallel processing. Gradient-based algorithms are not amenable to parallelization (Bethke, 1976). Therefore, the existing bi-level deterministic iterative optimization approaches may not be best adapted to tackling large-scale evacuation problems, particularly when a significant difference exists between the optimum solution and the initial starting point and when the search process may get stuck in local minima.

Global optimization offers a myriad matrix of potential heuristic approaches that can be utilized to optimize the evacuation problem. Among the most promising global optimization approaches are Evolutionary Algorithms (EAs). EAs are methods of searching in multi-dimensional space while satisfying certain criterion (Bäck, 1996). Recently, EAs have emerged as one of the leading methodologies for powerful searches and optimizations of problems in high dimensional and non-differentiable search spaces. Examples of EAs are Genetic Algorithms (GAs) and Evolution Strategies (ES). OSTE utilizes GAs instead of the more traditional approaches in the literature. GAs sidestep all of the problems associated with traditional deterministic optimization methods; they start the search from a population of solutions; therefore, the odds of finding the global optima without getting stuck in local minima are higher than in most conventional approaches; and they don’t require differentiation of the objective function. Moreover, GAs are inherently parallelizable, allowing the power of several computers or CPUs to be harnessed and the future use of High Performance Computing (HPC) clusters (see sections 3.5.2 and 8.5.1.2). The number of processors in a typical HPC can be in the hundreds, i.e. extensible computing power can be availed as the size of the evacuation problem grows, thereby reducing the computation time almost linearly with the number of available processors (Kruchten et al., 2004).

In a GA, a population of artificial solutions (chromosomes) is created. A chromosome is one feasible point or a candidate solution which carries the encoded values of the decision variables in the form of a string of genes. Each candidate solution from the population is then evaluated to
give some measure of its “fitness” (Garey et al., 1979). Evaluation can be as simple as substituting variable values in a mathematical function or it can be a full simulation experiment such as in our case. After evaluating an initial population by calculating the fitness of each candidate solution, selection and a series of genetic operators work on the population to reproduce a sequence of populations (children), by methods of crossover and mutation, with increasingly enhanced solutions. This cycle of evaluation, fitness assignment, selection, and reproduction continues for a number of generations (iterations) until a certain stopping criterion is met (see Figure 3.3). It is clear to see that GAs are inspired by the process of natural selection and evolution.

![Figure 3.3 The Basic Cycle of EAs](image)

An important property of EAs is that they can be parallelized (Bäck, 1996). In general, three types of EA exist; panmictic EAs, diffusion-style parallel EAs, and island model parallel EAs. In a panmictic EA, reproduction can be conducted between any two chromosomes in the population; whereas in the diffusion-style parallel EA, the chromosomes are spatially distributed (e.g. two-dimensional grid) and only neighbouring chromosomes can be recombined. In the island model parallel EA, semi-independent subpopulations, demes, evolve independently with periodic exchange of some chromosomes through a migration process (Tomassini, 1999). This type of EA exhibits even more correspondence to the evolution theory of species where
thousands of subpopulations (or demes) exist and co-evolve in parallel in the same continuous geography.

It is important to note that parallel in parallel EA is not to be confused with the parallel in parallel computing. Parallel computing distributes the computation across multiple processors simultaneously, in which chromosomes are farmed out to multiple processors for evaluation (i.e. in parallel). On the other hand, the parallel in parallel EA refers to the spatial structure of the population. Therefore, a parallel EA can be executed either sequentially on a single processor (in which sub-populations or demes evolve independently and migration takes places across multiple demes) or across multiple processors. In this thesis, to avoid confusion, the term distributed is used to refer to computation instead of parallel, and the term parallel will only refer to the population structure of EAs (see Figure 3.5).

3.5.2 Parallel Distributed Genetic Algorithm (PDGA): Design and Implementation

A non-iterative formulation for the optimization of the spatio-temporal evacuation pattern is developed. A non-iterative approach means that the simulated flows and scheduling strategy are computed simultaneously rather than iteratively for a given scheduling vector. Therefore, OSTE works in a non-iterative fashion invoking a simulation model only to evaluate a given scheduling vector in a super-zone destination representation. OSTE uses and expands on GENOTRANS (Generic Parallel Genetic Algorithms Framework for Optimizing Intelligent Transportation Systems), developed at the University of Toronto (Mohamed, 2007). GENOTRANS is a genetic algorithm platform in which the objective function is evaluated and constraints are satisfied through a simulation model. The simulation model is used to replicate the transportation network and perform dynamic traffic assignment as well as representing the super-destination approach. Vehicles are loaded onto the super-destination network according to their optimized schedules, and navigate through the network towards their destination(s) while dynamically updating their travel paths until they reach a safe destination.

OSTE expands on GENOTRANS in the following dimensions:

- Incorporate constraints in the design of GAs.
- Fine tune the parallelization of GAs in GENOTRANS.

---

3 Figure 3.4 provides the development phases of GENOTRANS.
- Fine tune the distributed GAs in GENOTRANS by deploying them in the High Performance Cluster facility.
- Remotely evaluate the simulation model.
- Concurrent execution of the simulation model on cluster nodes (CPUs).
- Design of parameters for independent and remotely configured GAs.
- Design of parameters for an independent and remote configuration for the evacuation problem.
- Notification of the evolution of the GA upon its start and termination.
- Design of a generic and comprehensive output format for the GA.

Figure 3.4 Development Phases of GENOTRANS

In addition to the parallelized nature of GAs, OSTE and GENOTRANS also utilize Distributed Genetic Algorithms (DGAs). This distribution is very efficient when dealing with computationally demanding problems (either large network size or need for very long computation time). In this thesis, a PDGA is designed to investigate the possibility of providing better performance in terms of solution quality and convergence speed (Tomassini, 1999) compared to the Single GA (SGA). A Multi-Deme Distributed GA is designed to simultaneously optimize the evacuation schedule and the destination choice for large-scale evacuation problems.
As clearly shown in Figure 3.5, in OSTE and GENOTRANS the core GA engine is both distributed and parallelized.

Figure 3.5 Parallel and Distributed GA Architecture.

This research utilizes a recently acquired High Performance Computing Cluster (HPC) at the Department of Civil Engineering of the University of Toronto. The cluster has 64 Processing Nodes, 44 with 4 GB of memory and 20 with 8 GB of memory, all with two processors, XEON 5150 2.66GHz dual core Woodcrest for a total of 4 processing cores per node (i.e. 265 processors), 36GB 15k rpm SAS hard disk, Dual Gig Ethernet, one public port and one dedicated cluster port. In addition to a master node, 20 nodes/slaves (i.e. 80 processors) were made available to conduct the optimization of OSTE to solve large-scale evacuation problems. It is to
be noted that the cluster is managed through the Compute Cluster Job Manager user interface provided by Microsoft® Windows® Compute Cluster Server (2003). The cluster manager provides an integrated application platform for running, managing, maintaining, monitoring, and developing parallel computing applications⁴.

The new version of GENOTRANS is designed using IBM WebSphere⁵ technology. IBM granted the University of Toronto the opportunity to explore and use WebSphere products through the academic initiative licensing program, which presents a great opportunity to explore the rich functionality of stable and latest software components from IBM. The IBM WebSphere Network Deployment (WebSphere, ND) V7.0 is used to deploy and manage the Websphere cluster nodes centrally using the Websphere Administrative Console. The WebSphere ND has built-in Work Load Management (WLM) and High Availability (HA) features (IBM, 2010) that are utilized in the design and configuration of the distributed GA on the HPC. The core GENOTRANS development is conducted using the Rational Application Developer (RAD) V7.5 (IBM, 2010), which maintains the resources and properties of the developed GENOTRANS Java classes/libraries.

In large-scale optimization problems that require considerable reading and writing of associated network and parameter files, conventional disk drive systems are not efficiently amenable to handling such massive information transfer across HPC nodes. Therefore, RAID (Redundant Array of Independent Disks), a new and efficient technique of improving data availability and transferability using arrays of disks and various data-stripping methodologies, is utilized. RAID is found to be very efficient compared to typical disk drive systems and the communication time (time required to write/read/send the files through the server) in the case of large-scale evacuation optimization was reduced to half.

In OSTE, the chromosome is encoded as the row vector \( \mu \) encoding the evacuation percentages over discrete time intervals. Each gene in the decision variable vector is encoded as a real number with a range between 0 and 1. The chromosome size (number of genes) represents the

---


⁵ “WebSphere is IBM’s integration software platform. It includes the entire middleware infrastructure — such as servers, services, and tools — needed to write, run, and monitor 24x7 industrial-strength, on demand Web applications and cross-platform, cross-product solutions. WebSphere provides reliable, flexible, and robust integration software”. Source (http://www.ibm.com/developerworks/websphere/newto/).
number of departure time intervals. The GA fitness function is the same as the objective function in Table 3.1 which requires the output of the traffic simulator each time an evaluation is performed. Figure 3.6 describes the general structure of the solution algorithm. Figure 3.6-a and Figure 3.6-b depict the SGA and PDGA structures, respectively, in which GENOTRANS interacts with a simulation model of the transportation network to be evacuated.

As shown in Figure 3.6-b, a master process generates the initial population, divides it into the specified number of demes (islands) and manages each deme’s evolution and migration procedure. Each chromosome is evaluated via a complete simulation run covering the entire evacuation period. Based on the literature and pilot experimentation, the multi-deme structure is selected as follows:

1) Deme topology: fully connected multiple deme topology, where every deme exchanges individuals with all the others.

2) Number of demes (d) and deme/subpopulation size: scenario dependent\(^6\),

3) Migration policy: good migrants replace bad individuals,

4) Migration rate (the number of individuals to migrate): 15 % of the population.

---

\(^6\) The relation between the population size and the number of demes is governed by having a divisor resulting from dividing the population size / subpopulation. For example, a population of 80 chromosomes could only be divided to 2 demes, or 4 demes, or 8 demes, or 10 demes, etc.
The design of a GA for a particular application involves the choice of methods and parameters for the population size, initial population, selection, crossover, mutation, stopping criterion. Significant testing is undertaken to find a combination of parameters to refine the choice of the GA parameters. The above values and the following GA parameters are used to illustrate the application of the algorithm itself rather than its sensitivity to parameters. The solution algorithm iterates through the following steps until termination:

- **Step 0. Initialization**
  
  - Set the user-specific GA design parameters.
  
  - Set the generation counter $t = 0$.

- **Step 1. Chromosome Representation**
o Code the decision variable $\mu$ to form a chromosome $C$ of real numbers.

o Chromosome Repairing Procedure: Adjust the chromosome using a scaling algorithm to sum up the staging vector to 100% of the demand being released at each generation (refer to Figure 3.6).

♦ Step 2. Evaluate Objective Function

o Using the evacuation schedule encoded in each chromosome, a simulation-based traffic assignment is performed to evaluate the chromosome’s fitness. In the simulation model, the traffic seeks an amalgamated super-zone and hence the optimal exit points (actual destinations) are ultimately produced. The process is repeated for all the chromosomes in population $P(t)$. Based on the output of each simulation run, the objective function is evaluated.

♦ Step 3. Apply Genetic Operators

o Based on the fitness of each solution, the selection process, which resembles the “survival of the fittest” principle, picks a number of candidate solutions for “breeding” to produce the next generation. Better individuals have higher probabilities of being selected to produce the next generation. Breeding is performed via crossover and mutation.

o The crossover step, also known in the GA literature as recombination, is the process of combining genes from parents in order to produce better offspring.

o The mutation step is the process responsible for exploring new areas in the search space by randomly altering the genetic structure of some chromosomes. The offspring is similar to the parent except for a few changes in the parental genetic information.

♦ Step 4. Chromosome Repairing

o After the genetic operators have been executed and if the offspring is an infeasible solution; repair it by scaling the evacuation percentages to add up to 100%.
Step 5. Fitness Assignment

- Use the objective function values to determine the fitness values of the new population.

Step 6. Apply the migration policy in case of PGA.

Step 7. Test Termination Criterion

If the termination criterion is satisfied, then terminate the GA and output the best solutions from the last iteration. Otherwise, go back to step 3. Termination is based on a pre-specified convergence threshold, or when the pre-set number of generations is exhausted.
4 Optimization of Transit Shuttling During Evacuation

This chapter discusses the conceptualization and development of the transit shuttling approach during evacuation. It provides a general description of the framework components and their connection to the typical Vehicle Routing Problem (VRP). It then introduces an analogy between the VRP and the mass transit evacuation problem. It also provides a mathematical formulation of the approach. A discussion of the proposed solution algorithm is given, while emphasizing the constrained nature of the problem. Performance details of the approach are presented in later chapters with a prototype implementation in Chapter 6 and a large-scale application in Chapter 8.

4.1 Vehicle Routing Problem and Transit Evacuation Problem

The VRP is extensively investigated in the literature; numerous techniques are offered with the common goal of modelling and solving the VRP. However, the solution efficiency differs significantly according to the exactness of the solution and the problem size. Table 4.1 provides a comparative survey of the literature on the VRP and its extensions. The survey is organized to highlight the most relevant studies according to the following categories: exact, approximate, and meta-heuristic approaches. The rows represent the method/approach used to solve the problem and the columns represent the criteria for each study in terms of their relevance to the transit evacuation problem.

Table 4.1 Comparative Literature Survey

<table>
<thead>
<tr>
<th>Method</th>
<th>Solution Procedure</th>
<th>Objective</th>
<th>Time Window</th>
<th>Multiple Depot</th>
<th>Pick-up and Delivery Location</th>
<th>Problem Size and Computation Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>(Fisher, 1994) Solve degree-constraint k-tree problem</td>
<td>Min Total Routing Cost</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>50–199 customers, CPU Time (6–680 min)</td>
</tr>
<tr>
<td>Approximate</td>
<td>(Clarke and Wright, 1964) Assign a vehicle to each customer and then improve the routing using saving algorithm</td>
<td>Min Tour Length</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Small size problems</td>
</tr>
<tr>
<td></td>
<td>(Fisher and Jaikumar, 1981) Generalized assignment procedure with</td>
<td>Min Total Routing Cost using</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>50–199 customers, CPU Time (9–25 min)</td>
</tr>
<tr>
<td>Meta-Heuristic</td>
<td>Heuristic</td>
<td>Min Total Cost of Travel</td>
<td>Min Total Tour Length</td>
<td>Min Total Distance Travelled by Vehicles</td>
<td>Min Total Travel Cost</td>
<td>Min Total Routing Length</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-----------------------</td>
<td>----------------------------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>(Kindervater and Savelbergh, 1997)</td>
<td>K-exchange algorithm to improve and extend TSP solution</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>10–30 customers, CPU Time (0.2–0.5 sec)</td>
<td></td>
</tr>
<tr>
<td>(Rochat and Taillard, 1995)</td>
<td>Probabilistic Diversification and Intensification to overcome local minima and improve computation time</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>50–385 customers, CPU Time (0.2–3000 min) to get 1% improvement over the best known problems</td>
<td></td>
</tr>
<tr>
<td>(Shaw, 1998)</td>
<td>Constraint Programming using Large Neighbourhood Search</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>Performs better than benchmark problems (C, R, RC)</td>
<td></td>
</tr>
<tr>
<td>(Xu and Kelly, 1996)</td>
<td>Tabu Search to Solve Network Flow Model</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>50–199 customers, CPU (4.89–207.8 min)</td>
<td></td>
</tr>
<tr>
<td>(Toth and Vigo, 2003)</td>
<td>Granular Tabu Search</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>50–199 customers, CPU (0.8–3.18 min)</td>
<td></td>
</tr>
<tr>
<td>(Gambardella et al., 1999)</td>
<td>Multiple Ant Colony Optimization</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>50–199 customers, results are reported after stopping criterion (100,200,…, 1800 sec)</td>
<td></td>
</tr>
</tbody>
</table>

4.2 VRP and Emergency Evacuation: Background

Sayyady (2007) investigated the use of public transit systems in no-notice evacuation of urban areas. The author formulated the public transit routing plan (P Trojan) problem as a mixed integer linear program in which a network flow problem is solved with the objective of evacuating as many people as possible from a set of source nodes (transit stations) to a set of exit nodes (shelters) in a given timeframe without violating capacity constraints. However, the bus schedules were predefined with a limited number of tasks to be performed. The evacuation demand was assumed to follow a normal distribution that forms the demand at each visit. The study provided a sensitivity analysis of available fleet size and of the percentage of captive people to transit. The computation times of two solution algorithms, CPLEX and Tabu Search, were evaluated; a bottleneck was found. CPLEX runs out of memory after running for three days.
without reporting the optimal solution. The run time to evacuate 300 to 2000 people ranges from 166 to 167 min using CPLEX, and 4 to 16 min using Tabu Search, respectively.

Murray-Tuite and Mahmassani (2003) provided two linear integer programs to express household behaviour under evacuation conditions. The first formulation determines the meeting location for household members, while the second considers a modified version of the VRP in which the sequence of family member pick-up is determined. The fleet of vehicles available to households is assumed to be heterogeneous depending on the available car capacity. In addition, vehicles are distributed according to the location of their drivers at the onset of the evacuation.

Pagès et al. (2006) introduced the mass transport vehicle routing problem (MTVRP) in which a fleet of vehicles (of given capacity) is routed to pick up and deliver passengers. The problem is solved iteratively between two levels; the Transit Problem (TP) and Passenger Problem (PP). The TP is solved once to generate an initial solution and then the PP works on improving the first solution by assigning passengers to routes. The authors compared the computation time of CPLEX when solving the PP for different network sizes (ranging from 5 to 56 links). The benefits of using the proposed algorithm are found to be greater in the case of large networks. However, the network size and problem dimensions are relatively small compared to real-life evacuation scenarios.

Song et al. (2009) proposed an optimal transit routing for emergency situations. The problem is formulated as a location routing problem in which the decision variable is the routing of vehicles and choice of shelter points so as to minimize the vehicle routing times. The problem is solved using a hybrid GA in which 15 constraints have to be satisfied in the initialization and reproduction processes. The problem has a special network structure that is different from the actual transportation network, which necessitates building a set of different link categories for streets, intersection, U-turn links. The proposed algorithm was tested on a small network with hypothetical evacuation parameters to test the feasibility of the proposed hybrid GA compared to a traditional GA and the former turned out to perform better.

### 4.3 Vehicle Routing and Mass Transit Evacuation: An Analogy

The VRP is a generic class of problem in which sets of customers are visited by vehicles. The goal is to solve the routing problem by assigning vehicles to traverse the transportation network
so as to visit customers, satisfy a given set of constraints, and optimize a certain objective function (see Figure 4.1). The VRP comprises several interacting elements, which are summarized as follows with the key attributes of each element:

- **Customers**: demand, time constraints, pick-up and delivery locations, priority
- **Vehicles**: capacity, cost, time window of vehicle availability
- **Depot**: number, location, capacity
- **Network**: time, distance, geographical representation

![Figure 4.1 Traditional Vehicle Routing Problem Representation](image)

To effectively utilize the transit fleet in emergency evacuation situations, the traditional VRP can be extended to include (i) **Multiple Depots** to account for the dispersed transit vehicles in the transportation network, (ii) **Time Constraints** to account for the desired evacuation time window, and (iii) **Pick-up and Delivery** locations for evacuees to allow for picking up evacuees from dispersed stops to avoid excessive walk distances. The Multi-Depot Time Constrained Pick-up and Delivery VRP (MDTCPD-VRP) is known to be an NP-hard problem (Garey *et al.*, 1979). The following points highlight the correspondence between the traditional VRP and the proposed emergency evacuation MDTCPD-VRP (see Figure 4.2), where customers become evacuees; pick-up points are hazard areas to be evacuated; delivery points are safe shelters and vehicles are transit shuttle buses:
• Pick-up and Delivery Problem (PDP) where evacuees (customers) are picked up from hazard zones and delivered to shelters (safe destinations). The transit fleet (shuttle buses) head back empty to pick up more evacuees until all the evacuees are delivered to shelters. Shuttling is not necessarily confined to two fixed ends (pick-up and delivery points) but rather is optimized.

• Multiple Depots VRP (MDVRP) where the transit fleet is stored at multiple locations as opposed to the traditional, fixed, one depot location. This is particularly convenient for cases where buses are initially scattered at different hubs, more so in the case of multiple bus companies such as public transit buses, commuter service buses, school buses and so on. In cases of emergency, buses can be assembled from different providers in accordance with prior agreements and globally managed from a unified emergency command centre.

• Time Constrained Problem (known as VRP with Time Windows, VRPTW). Time window problems occur in many business sectors. The greater the time constraints introduced into the problem, the more challenging the routing plan is. The severity of the emergency situation may dictate certain evacuation time horizons to minimize evacuees’ exposure to risk. Also, if evacuees wait for service for excessively long times, they may become impatient and seek other (uncalculated) alternatives. Therefore, the following time windows are modelled in the evacuation problem:
  o Evacuees have to be picked up before a time threshold ($T_{\text{Min}}$)
  o Evacuees have to be delivered to a safe destination before a time threshold ($T_{\text{Max}}$)
The MDTCPD-VRP is formulated by defining the sets, parameters, decision variables, objective function and constraints as shown in Table 4.2.

Table 4.2 MDTCPD-VRP Formulation

<table>
<thead>
<tr>
<th>Sets</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>All nodes that construct the network (graph)</td>
</tr>
<tr>
<td>(E)</td>
<td>Evacuee pick-up points, (E \subseteq N)</td>
</tr>
<tr>
<td>(S)</td>
<td>Shelter delivery points</td>
</tr>
<tr>
<td>(A)</td>
<td>All arcs ((i,j)) linking connected nodes (i, j \in N)</td>
</tr>
<tr>
<td>(V)</td>
<td>Available fleet of vehicles</td>
</tr>
<tr>
<td>(D)</td>
<td>Set of depot nodes where vehicles (V) are located</td>
</tr>
<tr>
<td>(K)</td>
<td>({K^D \cup K^S}), where (K^D) is the set of routes from depots to shelters and (K^S) is the set of routes from shelter to shelter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(CAP_v)</td>
<td>Capacity of vehicle (v)</td>
</tr>
<tr>
<td>(COST_v)</td>
<td>Cost of using vehicle (v)</td>
</tr>
<tr>
<td>(DEM_i)</td>
<td>The demand of evacuees located at node (i), (i \in E)</td>
</tr>
<tr>
<td>(C_{ij})</td>
<td>The routing cost (travel time) on arc ((i,j)), ((i,j) \in A)</td>
</tr>
<tr>
<td>(T_{MAX})</td>
<td>Time global threshold before which evacuees have to be delivered to shelters</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_{vij}^k)</td>
<td>Indicates whether or not vehicle (v) traverses arc ((i,j)) in route (k)</td>
</tr>
<tr>
<td>(y_{vi}^k)</td>
<td>Indicates whether or not pickup point (i) is served by vehicle (v) in route (k)</td>
</tr>
<tr>
<td>(Z_v)</td>
<td>Indicates whether or not vehicle (v) is dispatched</td>
</tr>
</tbody>
</table>

---

7 - Travel times used in the MDTCPD-VRP is exported from the traffic simulation model (refer to section 5.3) to replicate traffic congestion.
- In the MDTCPD-VRP, depots represent the dispersed locations of transit vehicles at the onset of the evacuation.
- Refer to section 6.4.1 for discussion on the choice of the multi-objective function weights.
<table>
<thead>
<tr>
<th>$t_{vi}^k$</th>
<th>Indicates the time, measured from the onset of the evacuation, until vehicle $v$ reaches pick point $i$ in route $k$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_{vi}^k$</td>
<td>Indicates the number of evacuees picked up by vehicle $v$ from node $i$ in route $k$.</td>
</tr>
</tbody>
</table>

**Objective Function**

- **Min Routing Cost**
  
  $$\text{Min } \sum_{k \in K} \sum_{v \in V} \sum_{(i,j) \in A} X_{vij}^k C_{ij}$$

- **Min Routing Cost and Waiting Time**
  
  $$\text{Min } \left[ W_{RC} \sum_{k \in K} \sum_{v \in V} \sum_{(i,j) \in N} X_{vij}^k C_{ij} + W_{WT} \sum_{k \in K} \sum_{v \in V} \sum_{i \in N} Y_{vi}^k t_{vi}^k q_{vi}^k \right]$$

- **Min Routing Cost and Vehicle Cost**
  
  $$\text{Min } \left[ W_{RC} \sum_{k \in K} \sum_{v \in V} \sum_{(i,j) \in N} X_{vij}^k C_{ij} + W_{VC} \sum_{v \in V} Z_v \text{ COST}_v \right]$$

- **Min Routing Cost, Vehicle Cost, and Waiting Time**
  
  $$\text{Min } \left[ W_{RC} \sum_{k \in K} \sum_{v \in V} \sum_{(i,j) \in N} X_{vij}^k C_{ij} + W_{WT} \sum_{k \in K} \sum_{v \in V} \sum_{i \in N} Y_{vi}^k t_{vi}^k q_{vi}^k + W_{VC} \sum_{v \in V} Z_v \text{ COST}_v \right]$$

**Constraints**

- $X_{vij}^k \in \{0,1\}$ \quad $\forall (i, j) \in A, v \in V, k \in K$
- $Y_{vi}^k \in \{0,1\}$ \quad $\forall i \in N, v \in V, k \in K$
- $Z_v \in \{0,1\}$ \quad $\forall v \in V$
- $t_{vi}^k \geq 0$ \quad $\forall i \in N, v \in V, k \in K$
- $q_{vi}^k \geq 0$ \quad $\forall i \in N, v \in V, k \in K$
- $\sum_{k \in K} \sum_{j \in N} \sum_{i \in D} X_{vij}^k = Z_v$ \quad $\forall v \in V$
- $\sum_{j \in N} X_{vij}^k = Y_{vi}^k$ \quad $\forall v \in V, i \in N, k \in K$

**Flow and Route Conservation Constraints**

- **At Intermediate Nodes**
  
  $$\sum_{j \in N} X_{vih}^k - \sum_{j \in N} X_{vij}^k = 0 \quad h \in N - \{D \cup S\}, \forall v \in V, k \in K$$

- **At Shelter**
  
  $$\sum_{k \in K} \sum_{i \in N} X_{vij}^k = \sum_{k \in K} \sum_{j \in N} X_{vij}^k + Z_v \quad \forall i \in S, v \in V$$

- **At Depot**
  
  $$\sum_{k \in K} \sum_{j \in N} X_{vij}^k \leq 1 \quad \forall v \in V, i \in D$$

**Capacity and Vehicle Constraints**

- **Vehicle Capacity**
  
  $$\sum_{i \in D} Y_{vi}^k q_{vi}^k \leq \text{CAP}_v \quad \forall v \in V, k \in K$$

- **Evacuees Pickup**
  
  $$\sum_{v \in V} \sum_{i \in D} Y_{vi}^k q_{vi}^k \leq \text{DEM}_i \quad \forall i \in E$$

- **Fleet Capacity**
  
  $$\sum_{v \in V} Z_v \leq V$$
Dwell time consists of the time lost prior to opening and after closing the transit vehicle doors, and the time required for boarding/alighting of passengers at most heavily used doors. Factors affecting the calculation of dwell times include: vehicle floor height and platform height, number of boarding/alighting channels (doors), and fare type and fare collection. The dwell time value is calculated based on the following equations\(^8\) (4.1) and (4.2): (Vuchic, 2005):

\[
t_s = t_o + b' \cdot \tau_b + a' \cdot \tau_a
\]

(4.1)

\[
b' = (b' / n') \cdot \zeta_b
\]

(4.2)

where:

- \(t_o\) is the time lost before/after doors are opened
- \(b'\) is the number of boarding riders through the most heavily used boarding door
- \(a'\) is the number of alighting riders through the most heavily used alighting door
- \(\tau_b\) and \(\tau_a\) are, respectively, boarding and alighting times per person
- \(n'\) the number of channels per vehicle = number of channels/door x number of doors
- \(\zeta_b\) is the ratio of maximum to average number of boarding passengers per door = coefficient of distribution among doors.

It should be noted that in emergency evacuation cases, fares are not collected; thus the time spent for boarding and alighting per person should be lower than the standard values. However, due to the chaos coupled with emergency evacuation \(\tau_b\) and \(\tau_a\) are assumed to be, at least, equal to the typical values.

\(^8\) It is worth noting that in case of emergency evacuation, passengers would be either boarding at evacuees’ pickup points or alighting at safe destinations.
4.4 Solution Algorithm: A Constraint Programming Approach

Constraint programming (CP) aims to simultaneously solve a constraint satisfaction problem (CSP) and an optimization problem. CP uses multiple algorithms to find feasible solutions to constraint satisfaction and optimization problems. In CP problems, search strategies are defined to model how the decision variables change with iterations to satisfy the constraints. Numerous algorithms are found in the literature; among the most relevant to the constrained VRP in emergency situations are domain reduction and constraint propagation.

Each decision variable in the CSP has its domain. The domain reduction algorithm works to remove these values from the domain of variables that are apart from any feasible solution (i.e. violate the set of constraints). This process is performed for all the variables in each constraint. At some point, the algorithm may discover that eliminating some values from the variable domain results in an unfeasible solution, at this stage the previous domain value is retrieved. Constraint propagation determines how the domain reduction technique is conducted among several constraints.

CP appears to be a good approach for tackling real-world VRPs because of its ability to effectively address problem-specific constraints. Alshalalfah and Shalaby (2008) and Alshalalfah and Shalaby (2009) introduced a new algorithm for solving the flexible route transit scheduling problem, in which CP found to be an efficient tool to model specific customers and transit constraints. CP produces initial solutions that satisfy the constraint set; however, an improvement mechanism must be in place to enhance the first solution (Shaw et al., 2002). Local search techniques fit well into the evacuation problem to improve the initial routing plan by exploring neighbourhoods. Integrating constraint programming and local search techniques would synergize their potential utility in more challenging real-world VRPs. Such integration is particularly appealing in the case of emergency evacuation because side constraints, such as bus capacities, evacuation time window, and pick-up and delivery play a paramount role in the success of an evacuation plan.

The emergency evacuation routing and scheduling problem is solved in two stages. First, a model for the evacuation area is built and then the routing and scheduling problem is
solved/optimized. ILOG Dispatcher™ and Solver™ are two examples of integrated constrained programming software that are used to build and solve the problem, respectively.

### 4.4.1 Building the Evacuation Model

Building the model comprises four basic objects: *Nodes, Visits, Vehicles* and *Costs*. These objects are used to construct the network and model the evacuation problem in a CP environment.

- **Nodes**: Any transportation network is composed of a set of intersections or nodes with specific coordinates and a set of roads or arcs connecting these nodes. These $N$ nodes are then used to compute distances and times (and subsequently costs) between pairs of nodes $(i, j)$. A set of Arcs $A$ (links) is then defined to connect these nodes and subsequently construct the graph (network).

- **Visits**: A visit represents an activity that the vehicle has to perform within the specified time window ($T_{\text{Min}}$, $T_{\text{Max}}$). A visit is located at a single node and is performed by only one vehicle $v$. Visit locations are the evacuee pick-up points, $E$. Multiple visits may be created at the same node in case the demand for pick-ups is greater than the vehicle capacity, $\text{CAP}_V$. For example, a specific pick-up point $e$ along the graph may be visited more than once if the demand located at this point is greater than the capacity of the vehicle assigned to perform this visit. Visits have quantities, which can be weight, volume or numbers of objects. These quantities are the number of evacuees boarding and alighting at the pick-up and delivery points, respectively.

- **Vehicles**: Vehicles represent the supply that serves the demand of visits. A vehicle has a start and end visit and can have variable start and end times associated with each visit. Visits are the pick-up points from hazard areas $E$, and the safe shelter (destination) points $S$. Vehicles have limited capacities $\text{CAP}_V$. These capacities represent the total number of people a vehicle can carry at any point along the route. Therefore, one vehicle might have multiple runs given an assigned (or optimized) schedule.

- **Cost**: Cost attributes are objects closely associated with visits and vehicles. The most common costs are time and distance. Cost is used to model side constraints such as capacity, time windows, etc. For example, costs could be attributed to the distance between
nodes and projected travel time along arcs. In addition, waiting time of evacuees at the stop can be added as necessary to the cost function. Cost is also used to calculate the objective function.

4.4.2 Solving the Evacuation Routing and Scheduling Problem

Solving the evacuation bus routing and scheduling problem consists of finding a value for each decision variable while simultaneously satisfying the constraints and optimizing the objective function. Given an evacuation area, the evacuation demand and available buses, the routing and scheduling of transit vehicles is solved to minimize the total cost (travel time and/or waiting time) while satisfying bus capacity, time window constraints, and pick-up and delivery constraints. Search strategies and constraint propagation are utilized to solve the problem. Two types of constraint propagation are used: initial constraint propagation and constraint propagation during search. The initial constraint propagation removes all values from domains that will not take part in any solution. After initial constraint propagation, the search space is greatly reduced and search strategies are utilized to explore the search tree (the remaining part of the search space) for feasible solutions that satisfy the constraints and further improve the objective function.

The search strategy comprises two steps. First an initial solution is found using route construction techniques. A known routing heuristic insertion algorithm (Jaw et al., 1986) is used to provide the initial solution. Once the initial solution is obtained, the second step is to apply a local search procedure to improve the obtained solution by introducing small changes (called neighbourhoods) to the current solution. The new solution is tested again for constraint feasibility using constraint propagation during search, and its cost is computed. If the new solution is feasible and has reduced the cost, it is accepted as the new solution; otherwise, the algorithm back-tracks along the search tree and tries a different set of values for the variables. Back-tracking offers the flexibility to retrace the search moves that may turn out to be wrong. Alternatives can be tried and if they do not succeed they can be reversed. In this way, only moves to feasible and improved solutions are accepted. Solutions that do not improve the cost or violate the constraints are rejected. This is known as a greedy improvement algorithm, as it only makes changes to the solutions that improve the cost. This process is repeated to the point at
which a certain stopping criterion is met. The stopping criterion is met when no further changes can be found. The solution procedure is illustrated in Figure 4.3.

4.4.2.1 Route Construction Heuristic to Find Initial Solution

Route Construction heuristics select visits sequentially until a feasible solution is found. Sequential methods construct one route at a time, while parallel methods build several routes simultaneously. The insertion algorithm (Jaw et al., 1986), one of the most commonly used route construction heuristics in vehicle routing problems, is used to route emergency evacuation buses.

Figure 4.3 Framework for Integrating Constraint Programming and Local Search Techniques for MDTCPD-VRP Emergency Evacuation
**Insertion Heuristic**

The insertion heuristic works sequentially by inserting evacuees \((E)\) at a time into the schedule of the available vehicles \((V)\). The insertion is conducted according to the order in which the visits were created at the best possible place, in terms of cost. The heuristic comprises the following steps (see also Figure 4.4):

1. Consider a fleet of evacuation buses to be assigned to routes
2. Construct a list \(L\), of unassigned pick-ups and drop-offs for evacuees \((E, S)\)
3. Insert evacuees \((E)\) in a route at a feasible position where the least increase in cost is achieved and the constraints are satisfied
4. Remove the drop-off visit of the same trip \((S)\) from \(L\)
5. Check if there are any un-routed evacuees in the list \(L\); if yes go back to step (3)

![Figure 4.4 Insertion Heuristic](image)

**4.4.2.2 Local Search Methods to Improve Initial Solution**

The route construction heuristic finds an initial feasible routing plan. The first solution obtained by the insertion heuristic can be further improved using Local Search techniques. Local search methods explore neighbourhoods iteratively to improve the initial solution. To design a local search algorithm, the improving mechanism and the stopping criterion have to be defined. The improving mechanism typically works by changing one attribute or a combination of attributes (for example, arcs connecting sets of customers) of a given solution, resulting in a new solution. The new solution is then compared with the current solution and, if the new solution performs better, it replaces the current solution and so the search continues.
The computational effort associated with various solution algorithms is an important consideration when exploring optimal routing and scheduling solutions for large-scale evacuation problems. Virtually all vehicle routing and scheduling problems belong to the class of NP-hard problems (Garey et al., 1979) in which solving small problems to optimality is difficult with reasonable computational effort. Consequently, in the case of solving real-life problems such as evacuation by mass transit, one should not insist on obtaining the optimal routing and scheduling plan. A good feasible solution within a reasonable amount of computation time is acceptable.

As shown in Figure 4.5, a local search algorithm starts from an initial routing plan $R_0$ and continues replacing $R$ with better solutions from its neighbourhoods $N(R)$ until no further improvement is obtained. The central idea of neighbourhoods is to explore a set of solution changes that represent better potential solutions that would, in our case, reduce the total evacuation time.

Figure 4.5 Local Search Concept
Typically, iterative local search methods that have been applied to the VRP are based on edge-exchange algorithms\(^9\) (Braysy and Gendreau, 2005). Logically, the difference between neighbourhood algorithms depends on how many routes are involved in the improvement process and how many arcs (edges) within each route are exchanged. The improvement mechanism assumes that the cost is proportional to the route cost (e.g. length, travel time, or combination of both) and that, therefore, a shorter routing plan is less costly (e.g. a tour that minimizes the total evacuation time), consequently the initial solution is improved. The local search methods proposed by Braysy and Gendreau (2005) are utilized in this research; these include: 2-Opt Neighbourhood, Or-Opt Neighbourhood, Relocate Neighbourhood, Cross Neighbourhood and Exchange Neighbourhood.

---

\(^9\) Edge-exchange neighbourhoods: a set of tours that can be obtained from an initial tour by replacing a set of \(k\) of its edges by another set of \(k\) edges (Braysy and Gendreau, 2005)
5 Multimodal Evacuation Framework

After the terrorist attacks of Sept 11, 2001, and the massive destructions of the hurricane season in 2004 and 2005, US attention has focused on harnessing the capacity of transportation systems to efficiently respond to emergencies and to evacuate the population in a timely and safe manner. Lately, it has been recognized that public transportation systems play a significant role in emergency evacuation. As discussed in section 1.1, the history of disasters in Canada is not comparable to that of the US; however, the latter has contributed to shaping the emergency management systems in Canada and worldwide. Despite the fact that each disaster or emergency event has its own circumstances, lessons learned from major disasters have shaped the emergency management systems in Canada and US (Litman, 2006; Lindsay, 2009). Numerous lessons were learned after studying the consequences of each disaster; therefore, a special committee was formed by the Transportation Research Board (TRB) to examine how the potentially critical role of public transit can be fulfilled in emergency evacuation and produced a special report entitled “The Role of Transit in Emergency Evacuation”. The committee reviewed the literature and analyzed the emergency response and evacuation plans of the 38 largest urbanized areas in the US (TRB, 2008).

This chapter summarizes the relevant lessons and findings of this report, to build on these lessons while integrating multiple modes in the presented framework. It then highlights the modes that can be potentially incorporated in emergency evacuation plans. It ends with a description of a multimodal evacuation framework that integrates the two optimization platforms discussed in Chapters 3 and 4.

5.1 Role of Transit in Emergency Evacuation: Lessons Learned

Numerous factors affect the transit system in emergency evacuation; these factors include internal factors (e.g. transit system, nature of emergency, evacuee behaviour) and external factors (e.g. urban area characteristics). As envisioned by the author, these factors are charted in Figure 5.1 so as to highlight the interaction between them. As shown in the figure, these factors are
interrelated; for example, in dense urban areas that have connected and large transit systems, people are typically more trained and knowledgeable about emergency evacuation plans resulting in less chaos and disruption to the system in cases of emergency, as opposed to the limited transit systems in small urban areas that can serve only a small portion of the population leaving more transit-dependent people at risk.

Figure 5.1 Factors Affecting the Role of Transit Systems in Emergency Evacuation

Considering the above structure of external and internal factors that shapes the role of public transit in emergency evacuation, the examination and review of the plans of the 38 largest urbanized areas in the US led to the following conclusions:

- Most evacuation plans were found inadequate to manage catastrophic events.
- Failure to incorporate all available modes of transportation, including transit, in evacuation plans is found to be a major concern.
- Failure to identify and accommodate special-needs populations and those who are transit-dependent is a major weakness of most evacuation plans.
- Capacity and congestion issues caused by the surge in demand of automobile traffic prevented the travel of transit vehicles on the highways of urban areas.

Therefore the transportation infrastructure (including transit infrastructure and automobile infrastructure) may come to a halt because of improper planning for emergency evacuation.
5.2 Modes in Evacuation

Automobiles have been the dominant mode in emergency evacuation in most cities and urban areas. In the US, only a small fraction of states have incorporated multiple modes of transportation into their emergency evacuation plans (TRB, 2008). Also, the lack of coordination between transit agencies and traffic operators may further hinder the potential for integrating multiple modes into emergency evacuation.

Optimized multimodal evacuation is still largely missing from most emergency evacuation plans. However, utilizing the readily available transit capacity can significantly improve the evacuation process and lessen evacuation casualties. During evacuation events, available buses (e.g. public transit, commuter service, school) in the area can be used and routed to shuttle evacuees to safety, a process that can be explicitly optimized as presented in Chapter 4. A single bus-only highway lane can carry up to six times as many passengers as a passenger car-only highway lane (Litman, 2006). In addition, standard buses, LRT, and Rapid Transit (subway or metro) can carry up to 5400, 28,800, and 72,000 spaces/hour respectively (Vuchic, 2005). Therefore, transit services afford huge capacity that can significantly reduce the clearance time in the case of evacuation. In addition, when a disaster hits, many evacuees do not have access to their vehicles and hence must be evacuated by transit. Transit has played a vital role in the no-notice evacuation of both New York City and Washington D.C. that followed the event of Sept 11, 2001.

5.3 The Overall Framework: Multimodal Evacuation Plan

The framework attempts to optimize the use of multiple modes during emergency evacuation. Figure 5.2 illustrates the steps involved in achieving this goal by combining OSTE (refer to Chapter 3) and MDTCPD-VRP (refer to Chapter 4) in one platform. The framework starts by estimating the evacuation demand using a regional demand survey (e.g. TTS) and a representation of traffic analysis zones. The output of this step is a representation of the spatial and temporal distribution of the population and their modes of travel. OSTE plans are then generated for the vehicular demand using genetic algorithms as a global optimization technique and a dynamic traffic assignment tool. OSTE generates an optimal evacuation schedule, optimal destination choices if requested and optimal routes to destinations. It also produces link travel
times that are used as input for the optimal routing and scheduling of transit vehicles. The routing and scheduling of transit vehicles is then solved using constraint programming. The automobile OSTE plan and the transit optimal routing and scheduling plan are finally combined for dissemination to evacuees. It is to be noted that currently the framework does not attempt to loop back from the transit assignment component to the traffic assignment component. However, while extracting the travel times from the DTA model to form the input to the MDTCPD-VRP, the most congested travel times are used as a worst case for buses while travelling through the network. Although it may overestimate the travel times for buses, it should compensate for the uncertainty of travel times for such heavily utilized transit vehicles; i.e. if the process errs it does so on the conservative side.
Figure 5.2 Framework for Optimization of Multimodal Evacuation
Prototype Implementation - Toronto Waterfront Application

This chapter demonstrates the applicability and feasibility of the approaches proposed in Chapters 3 and 4 by implementing a prototype. This prototype shows the main principles of the proposed evacuation planning model; it does not, however, represent a large-scale implementation of the proposed framework. It is worth noting that the reported results reflect the hypothetical evacuation event presented by the prototype. The application of OSTE and MDTCPD-VRP to a large-scale case study that necessitates the evacuation of the entire City of Toronto is presented in Chapter 8.

6.1 Application Context

A hypothetical evacuation scenario for a very busy part of Toronto, the financial district in the downtown area, is used as a case study in this investigation. This area has one of the highest concentrations of economic activity in Canada. It is the most densely built-up area in Toronto and has around 100,000 commuters entering and leaving the financial district every working day. Therefore, it is perhaps one of the most difficult to evacuate in the case of an emergency. Evacuation is modelled during the noon period which is a worst case scenario because most commuters will be at work. Two network representations are modelled, one with a fixed destination choice and another with a flexible destination choice (super-zone). This case study is a moderately challenging problem that demonstrates the applicability, capabilities and usefulness of OSTE and MDTCPD-VRP as emergency evacuation planning and optimization tools.

Capturing the evacuation dynamics in both time and space is central to OSTE. Examples of dynamic simulation models include DYNEV (Pidd et al., 1996) and MASSVAC (Southworth and Chin, 1987), at the macroscopic level, and CORSIM (Theodoulou and Wolshon, 2004), VISSIM (Williams et al., 2007), and PARAMICS (Chen and Zhan, 2006) at the microscopic level. It should be noted that microscopic simulators are computationally demanding, limiting their use to relatively small networks. Mesoscopic simulators offer a good compromise between macro and micro simulators, with detailed representation of vehicles and fast computational speed. At the mesoscopic level, DynusT (Chiu and Mirchandani, 2008), and DynaSmart-P
(Sbayti and Mahmassani, 2006), can potentially be utilized in the area of emergency evacuation with further modification to account for evacuation modelling.

The transportation analysis zones (TAZs) are defined for the study area based on the City of Toronto zoning system. Evacuation demand is estimated from arrivals to the study area during the morning peak (6–9 am) of a typical week day in Downtown Toronto. This estimation resulted in a total evacuation demand of 20,000 vehicle trips and 47,000 transit trips including automobile passengers and all transit trips (cyclists, taxi passengers and walking). The vehicular traffic is assumed to be the reverse of the morning peak while the transit demand is estimated using TTS survey data (DMG, 2003). In this application, the focus is on the evacuation process irrespective of the source of the evacuation demand. However, a more accurate and comprehensive demand estimation model is presented in Chapter 7. Vehicular traffic originating from the evacuation zones (origins) use automobiles as the primary mode to reach safe shelters (destinations); whereas, transit-dependent evacuees are geographically distributed to specific pick-up points. The maximum acceptable walking distance for transit evacuees with respect to each zone centroid is defined. It is found that around 60% of transit users in Toronto live within an average of 300 m from a transit stop (Alshalalfah and Shalaby, 2007). Therefore, a procedure is applied using ArcGIS to locate transit pick-up locations such that maximum walking distances are less than 400 m from the centroid of each zone. Larger zones are accordingly divided into smaller subzones.

The network, shown in Figure 6.1, is geographically coded in DynusT (Chiu et al., 2008) and ILOG Dispatcher, with detailed representation of surface streets, expressways, on-and-off ramps as well as traffic control systems. The evacuation zones, buffer zones and pick-up points are illustrated in the figure. For the purpose of analysis, an available fleet of 50 shuttle buses is assumed, each with a capacity of 50 passengers. Shuttle buses are assumed to be stored at two depots as shown in the figure.
6.2 Design of Experiments

Three phases of testing are designed to examine the effects of several factors on the proposed evacuation plan. Phase A explores the effect of various demand levels and time pressures on
emergency evacuation. Phase B investigates the effect of various objective functions in modelling the emergency evacuation problem. Phase C examines the effect of mode shift on the performance of the system overall. The following sections detail the development and testing of each phase as well as the interpretation of results.

6.3 Phase A: Evacuation Scheduling and Destination Choice under Various Levels of Demand and Time Pressures

6.3.1 Scenario Design

Three evacuation scenarios are investigated in Phase A. Each scenario differs from the other in one or more of the following:

- Destination choice/assignment (fixed outbound destinations or one super-zone destination)
- Evacuation demand scheduling (simultaneous evacuation or optimized evacuation demand scheduling)
- Demand loading level: DL = 20,000, 25,000 and 30,000 vehicles are released over the mobilization horizon (defined next), larger numbers of evacuees mimic higher evacuation demand due to higher population densities in the future and/or more severe evacuation events.
- Mobilization horizon level: ML = 60, 120 and 180 min. The mobilization horizon is the time span within which all evacuees must start their journey, i.e. leave their origin. How long it will take to reach safe destinations is an outcome of OSTE which is a function of the level of congestion in the system. The latter is a function of many factors including demand level, road capacity and mobilization horizon. In risky evacuation events such as spreading of chemical fumes, exposure is detrimental to health and this puts time pressures on the evacuation process. Note that very fast mobilization is not necessarily desirable because it may cause gridlock and lead to longer network evacuation times. On the other hand, a very slow mobilization process may unnecessarily prolong the evacuation process. Therefore, an optimum in between the two extremes is sought.

Simultaneous Evacuation (SE)
The simultaneous evacuation scenario is considered the baseline scenario. In simultaneous evacuation, evacuees are advised to evacuate the hazard zones immediately and they all start to leave at once.

*Optimal Temporal Evacuation (OTE)*

In this scenario, the scheduling vector is optimized and evacuees are advised when to start the evacuation process. This scenario assumes fixed destinations, i.e. evacuees are directed according to their predefined OD matrix (where they came from).

*Optimal Spatio-Temporal Evacuation (OSTE)*

This scenario is similar to the previous one except for relaxing the assumption of fixed destinations. All gateways to safety are amalgamated into one hypothetical super-zone which all evacuees try to reach quickly. Evacuees are advised of the time when to start heading towards the super-zone. Therefore, in the model, evacuees are dynamically assigned to their best routes, which can be optimized using user optimal, system optimal or mixed assignment criteria. The outcome from the model reveals the optimal exit zone. Evacuees, in reality, would be advised of the time when to leave and where (which zone) to go to.

The effectiveness of OSTE and OTE relative to the baseline scenario (SE) is examined for each of the above demand and risk levels resulting in 18 scenarios (2 methods (OSTE and OTE)*3 demand levels *3 mobilization levels). Simultaneous evacuation is also examined for the three demand levels raising the total number of scenarios to 21.

Two measures of effectiveness (MOEs) are examined for each scenario: total/average trip time and total/average trip distance.

6.3.2 Phase A Results and Analysis

The optimal staging policies for OSTE and OTE evacuation scenarios are presented in Figure 6.2. Each figure shows two evacuation scenarios: one with optimal spatio-temporal patterns (OSTE), the bold curves; while the others are for optimal temporal patterns (OTE). For example, (ML = 60, DL = 20000) implies that 50 minutes after the start of evacuation, 93% of the
Evacuees have been loaded onto the network using the OSTE policy. The following key observations/conclusions can be drawn from the figures:

1. As shown in Figure 6.2, the OSTE control strategy results in more temporally spread out evacuation schedules compared to OTE. When destination choice is flexible, as in OSTE, demand is better spread over space and interestingly the evacuation load can be better spread across time as well.

2. As a direct result of the better spread of demand over space and time, OSTE results in faster evacuation than OTE, i.e. a higher evacuation percentage in less time. OSTE curves are generally higher and to the left of the corresponding OTE curves.

3. OSTE policies result in less time spent in the system and a smaller number of evacuees stuck in the network, as illustrated by the smaller area encompassed between the loading and evacuation curves. Therefore, the OSTE policy always outperforms the OTE policy in terms of evacuating people more rapidly and minimizing en-route evacuees.

4. As shown in Figure 6.2, in the case of low temporal risk levels (ML = 180 min), the percentage improvement in OSTE departure pattern relative to OTE is not as great as at high temporal risk levels (ML = 60 min). However, the OSTE evacuation curves still outperform the OTE evacuation curves. Therefore, it is always better to optimize the evacuation schedule and destination choice concurrently, especially under higher demand and time pressures.
Figure 6.2 Loading and Evacuation Curves for Different Scenarios at Different Risk Levels
6.3.2.1 Average Evacuation Time and Trip Distance

Two system-wide MOEs are considered to examine the effectiveness of the proposed system: the total travel time (TTT) and the total trip distance (TTD). Note that the mean values are simply the totals divided by the number of evacuees. The effectiveness of OSTE relative to OTE and SE, in terms of TTT, is illustrated in Figure 6.3, Figure 6.4, and Figure 6.5 for the three mobilization levels, ML = 60, 120 and 180 min, respectively. Figure 6.3-b, Figure 6.4-b, and Figure 6.5-b zoom into the difference between OSTE and OTE. The relative effectiveness of OSTE to OTE and SE in terms of TTD is illustrated in Figure 6.6 at a mobilization level ML = 60 min.

Figure 6.3 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 60 min)

Figure 6.4 Total Travel Time Relative Effectiveness of OSTE, OTE and SE (ML = 120 min).
The following key findings are drawn based on an analysis of the model results:

1. As shown in Figure 6.3, Figure 6.4, and Figure 6.5, OSTE steadily outperforms OTE and SE in terms of TTT. For instance at ML = 60, the percentage improvements of OSTE relative to OTE and SE are 17.4% and 53.8% respectively. OSTE slightly outperforms OTE in terms of TTT.

2. The total travel time increases as the risk level increases (i.e., mobilization decrease and/or demand increase). This is evident from comparing the travel time patterns for various demand and mobilization levels. Figure 6.3-b, Figure 6.4-b, and Figure 6.5-b illustrate that the TTT increases with the level of demand (seemingly exponentially) and inversely with the mobilization horizon, i.e. higher demand levels and tighter loading horizons lead to longer travel times and vice versa.

3. As the evacuation risk increases (with an increase in demand level) the effectiveness of OSTE relative to OTE is stronger. This is manifested in Figure 6.3-b, Figure 6.4-b, and Figure 6.5-b. This conclusion emphasizes the importance of relaxing the fixed OD demand matrix in cases of evacuation and considering OSTE as an alternative, while not denigrating the already efficient OTE.

4. OSTE significantly surpasses OTE and SE in terms of total travel distance. For instance at ML = 60, the percentage improvement of OSTE relative to OTE is 74.3%. This improvement is attributed to the destination choice optimization, where evacuees are guided to the nearest safe exit. It is note-worthy that OSTE outperforms OTE both in
terms of TTT and TTD. However, the impact of OSTE is substantial in terms of TTD which is intuitive.

5. The TTD increases linearly as the demand level increases. As shown in Figure 6.6, the slope of the TTD line in the case of OTE and SE is steeper than that of the OSTE case (For instance, at ML = 60, Slope_{OTE} / Slope_{OSTE} = 3.5). In effect, this means that for two evacuation scenarios (OSTE and OTE) with the same evacuation time pressure, OSTE could create more room for evacuees to travel relative to OTE and SE. This shows that the flexible destination procedure can artificially enhance network resilience in the case of emergency situations. Moreover, OTE results in almost the same trend as SE since both network representations have fixed destination assignment.

![Figure 6.6 Relative Effectiveness of Total Trip Distance for OSTE, OTE and SE for ML = 60 min](image)

6.4 Phase B: Multiple Objective Optimization for Multiple Modes in Emergency Evacuation

6.4.1 Scenario Design

In Phase B, the focus is mainly on testing multiple objective functions for multimodal evacuation. Two main objective functions are formulated so as to compare their merits for emergency evacuation. In each scenario, OSTE is utilized to solve the same objective function as MDTCPD-VRP. OSTE optimizes the traffic conditions which are then fed into the MDTCPD-VRP as an input that constitutes the network travel times for mass transit routing. The first objective is to minimize the in-vehicle travel time (en-route time) regardless of the waiting time at the origin before evacuees start their trips. The second objective is to minimize the waiting time of evacuees and their travel time. In addition to the two main objectives, a third objective is
modelled in the MDTCPD-VRP to account for the fleet size, this is the cost associated with using additional transit vehicles. To account for the aforementioned objectives, five multi-objective evacuation scenarios are investigated. Each scenario differs from the other in one or more of the following objectives:

- Travel Time
- Waiting Time
- Fleet Cost

**Min Travel Time (TT)**

The most common objective function in the emergency evacuation literature is the minimum evacuation travel time $TT$. In this case, OSTE1 minimizes the travel time given a predefined loading horizon. As in phase A, three loading levels (60 min, 120 min, and 180 min) are investigated. The output of minimizing the $TT$ in OSTE1 is fed to the MDTCPD-VRP to constitute the baseline scenario which replicates the typical routing cost in the literature. In the $TT$ scenario, the routing problem is primarily optimized to achieve the minimum routing cost plan, which is the minimum total travel time by all vehicles in serving all evacuees within the available time window.

**Min Waiting Time (WT)**

This scenario is designed to minimize the waiting time for vehicular evacuees in OSTE only, in which all evacuees are advised to leave the hazard zone immediately without any pre-trip information or announced evacuation plan. It is worth noting that the absolute minimization of waiting time is equivalent to *simultaneous* evacuation in which all the population rushes into the transportation network at time zero, which might result in early gridlock and longer travel time.

**Min Travel Time and Waiting Time (TT.WT)**

The $TT.WT$ scenario better mimics the evacuation situation since it achieves a balance between the in-vehicle running time and the waiting time of evacuees. OSTE2 is utilized in this scenario to provide the minimum possible waiting and travel times for evacuees. In recent studies (Miller *et al.*, 2005; Roorda *et al.*, 2006) in Toronto, it is reported that transit travellers value the waiting
time at the stop three times more than the in-vehicle travel time. The same weighting scheme is used in the MDTCPD-VRP\textsuperscript{10}.

**Min Travel Time and Vehicle Cost (TT.VC)**

The $TT.VC$ scenario is similar to the $TT$ scenario with the addition that it is designed to account for the vulnerability associated with fleet capacity in the case of emergencies. The use of one extra vehicle is included in the objective function as a penalty that increases the cost significantly. Minimizing fleet size implies minimizing the operating and running cost of the transit vehicle, an ultimate goal for transit operators especially in the case of emergencies in which lack of drivers and shortage in maintenance crew is inevitable. In this scenario, a trade-off between minimizing the routing cost and vehicle cost is attained in the solution algorithm.

**Min Travel Time, Waiting Time and Vehicle Cost (TT.WT.VC)**

The $TT.WT.VC$ scenario is similar to the $TT.VC$ scenario with the addition of evacuee waiting time. This scenario adds more complexity to the problem since minimizing the fleet size implies increasing the routing cost and consequently increasing the waiting time, such a compromise is achieved with some interesting results shown in the subsequent sections.

6.4.2 Phase B Results and Analysis

6.4.2.1 Optimal Scheduling, Destination, and Routing

The output of OSTE is the concurrent optimal departure time, destination and path for each driving evacuee. The resulting link travel times are then fed as input to the MDTCPD-VRP. The evacuation curves for the $TT$ ($\text{OSTE1}$), $WT$ ($\text{SE}$), and $TT.WT$ ($\text{OSTE2}$) scenarios are illustrated in Figure 6.7. As shown in the figures, in the $TT.WT$ scenario the evacuation curve is shifted to the left compared to the $TT$ case and also the network clearance time is reduced. Table 6.1 shows the criteria for evaluating the efficiency of each scenario; these include: average travel time, waiting time, travel distance and network clearance time. The table incorporates the results obtained from the two optimization platforms; OSTE and MDTCPD-VRP. Examination of the results leads to the following conclusions:

\textsuperscript{10} It is worth noting that the relative weight of waiting time to travel time in case of emergency evacuation needs a caveat. This weighting scheme is assumed for the lack of better information.
The absolute minimization of waiting time (SE), as intuitively expected, results in severe congestion and long travel times. Compared to the \( TT \) case, an order of magnitude increase in average travel time and travel distance is reported. However, as expected, the waiting time for the \( TT \) and \( TT.WT \) scenarios is deliberately higher which is the essence of demand scheduling in evacuation.

Minimizing the travel time given a predefined loading horizon (\( OSTE1 \)) would improve \( A_1 \) explicitly and improve \( A_2 \) implicitly and may or may not improve \( T \) (refer to Figure 3.2 for definitions of \( A_1 \), \( A_2 \) and \( T \)). In the case of long mobilization times (\( T_{180}, T_{500} \)), the loading curve has been shaped to minimize only the area between \( L(t) \) and \( E(t) \); this is clearly shown by the resulting linear evacuation curve in Figure 6.7. As illustrated in Table 5, the percentage improvement in travel time (compared to SE) for the 60 min, 120 min, 180 min, and 500 min loading levels are 43%, 61%, 77% and 95% respectively. It should be noted that over stretching the evacuation process (\( T_{500} \)) will lead to very short travel times; however, the waiting time is significant. Also, over rushing the evacuees to the network (SE) will do exactly the opposite, resulting in very long travel times and almost no waiting time. A compromise between the two is achieved through \( OSTE2 \), in which the waiting and travel times are included in the objective function.

Minimizing the travel time and waiting time without a predefined loading horizon (\( OSTE2 \)) would explicitly improve \( A_1 \) and \( A_2 \) as well as \( T \). A compromise between waiting time and travel time is achieved in this scenario. As expected, the \( TT.WT \) scenario resulted in a slight increase in travel time compared to the \( TT \) scenario and longer waiting time compared to the \( WT \) scenario. As shown in Figure 6.7, the \( TT.WT \) scenario best mimics the evacuation case in which evacuees will be advised when to leave their origins to reach a safe destination in the quickest possible way. It is worth noting that the gain from decreasing the waiting time outpaces the loss associated with increasing the average travel time. Also, the \( TT.WT \) scenario results in the minimum possible network clearance time (117 min).
6.4.2.2 Optimal Routing and Scheduling of Transit Vehicles

The following results demonstrate optimal scheduling and routing of shuttle buses in each scenario. The MDTCPD–VRP is modelled and solved using constraint propagation and optimization techniques. ILOG Dispatcher and Solver are used to model and solve the problem simultaneously in a very efficient computation (ILOG, 2008). On a Dual Core CPU with 2 GB of RAM, it took 5.2 minutes to solve the optimization problem for each of the above-mentioned scenarios. The model generates the optimal routing and timetable for each evacuation bus as it shuttles between pick-up points and nearest shelters. It is to be noted that a bus does not necessarily return to the same shelter all the time. Rather, it heads to a series of pick up points until it is full then heads to the nearest shelter to drop off evacuees, then goes back into the
evacuation area, picks up evacuees from a series of nodes until full, heads towards the nearest shelter from the point where it became full, and so on. The process is similar for all buses and continues until all evacuees reach safety in the given scenario. The model is applied to all scenarios in the same manner. For the sake of conciseness, the detailed routing and scheduling of one bus is shown in Figure 6.8. As shown in Figure 6.8, the arrival and departure times of the bus at each pick-up point (visit) are extracted and the associated travel time and distance travelled are illustrated. The optimal route is a sequence of nodes starting from pick-up point to shelter to next pick-up point.

For example, one bus (Vehicle 50) is scheduled to start from the depot (Depot 2) at time 0, picks up 50 passengers (visit 919) and travels along the optimal route as shown in Figure 6.8. The vehicle then drops off the evacuees at shelter1 and continues to pick up evacuees located at visit 916 and drops them off at shelter2. The routing and scheduling plan continues until all vehicles accomplish the assigned tasks. At the end of the evacuation, all vehicles return back to the shelters (terminal) as in the case of bus #50 shown in the last row of Figure 6.8.
Figure 6.8 Example of Routing and Scheduling of Transit Vehicle
The output of MDTCPD-VRP is the optimum routing and scheduling of buses\textsuperscript{11}. Each scenario is evaluated based on the criteria defined in Table 6.1. Figure 6.9 demonstrates the effectiveness of each scenario in light of the MOE.

### Table 6.1 Comparative Analysis of Evacuation Scenarios

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Average Waiting Time (min)</th>
<th>Average In-Vehicle Travel Time (min)</th>
<th>Average Distance Travelled (m)</th>
<th>Network Clearance Time (min)</th>
<th>* Average Number of Runs/Vehicle</th>
<th>** Number of Vehicles Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE</td>
<td>WT</td>
<td>0</td>
<td>44</td>
<td>4158</td>
<td>291</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>TT\textsubscript{60}</td>
<td>29</td>
<td>25</td>
<td>1086</td>
<td>396</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>TT\textsubscript{120}</td>
<td>48</td>
<td>17</td>
<td>1050</td>
<td>306</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>TT\textsubscript{180}</td>
<td>81</td>
<td>10</td>
<td>1038</td>
<td>222</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>TT\textsubscript{500}</td>
<td>242</td>
<td>2</td>
<td>1047</td>
<td>500</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>TT.WT</td>
<td>13</td>
<td>14</td>
<td>1252</td>
<td>117</td>
<td>--</td>
</tr>
<tr>
<td>Transit Evacuation (MDTCPD-VRP)</td>
<td>TT</td>
<td>60</td>
<td>3.60</td>
<td>2310</td>
<td>123.6</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>TT.WT</td>
<td>45</td>
<td>3.74</td>
<td>2451</td>
<td>76.7</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>TT.VC</td>
<td>61</td>
<td>3.87</td>
<td>2607</td>
<td>123.8</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>TT. WT.VC</td>
<td>51</td>
<td>3.75</td>
<td>2457</td>
<td>86.7</td>
<td>22</td>
</tr>
</tbody>
</table>

\* Average Number of Runs/Vehicle: The average number of runs each bus has to travel to serve all evacuees.

\** Number of Vehicles used to accomplish the plan and its depot location.

\textsuperscript{11} It is noteworthy that the notion of *optimum* herein refers to the best attainable values for the decision variables while solving the optimization problem; however, a global optimal solution is not guaranteed given the heuristic nature of the solution algorithm.
Figure 6.9 Effect of Evacuee Waiting Time and Vehicle Cost on In-Vehicle Travel Time

Analysis of the results leads to the following conclusions:

- Including the waiting time in the objective function has evened out the average in-vehicle travel time. Although the total in-vehicle travel time appears to be the same in both scenarios (TT and TT.WT), the pattern is different. This observation confirms that workloads assigned to each vehicle/driver are balanced as shown clearly by the consistent pattern in travel times in Figure 6.9-a. The results from these runs are consistent with our expectations. Including the waiting time in the objective function has increased the in-vehicle travel time and the total travel distance. However, it is worth noting that the average in-vehicle travel time...
increased by only 4% and the average travel distance by only 6%, yet the network clearance
time is reduced by a full 40% and the average waiting time of evacuees dropped by 25%.

- Minimizing the fleet size is crucial in emergency evacuation due to the lack of resources to
  run the fleet at full capacity. Figure 6.9-b demonstrates the effect of including the vehicle
cost in the objective function. As clearly shown in Figure 6.9 and Table 6.1, the number of
vehicles required to carry out the plan in the TT.VC scenario is reduced and the in-vehicle
travel time and travel distance are increased compared to the TT scenario. However, it is
interesting to note that the average in-vehicle travel time increased by 7.5% and the average
travel distance by 13%, yet the number of vehicles utilized is reduced by a full 34%.

- The TT.WT.VC scenario is found to be the most challenging because the vehicle cost
  objective conflicts with both the routing cost and waiting time objectives. It is shown in
Figure 6.9-b that including the waiting time in the TT.VC scenario (TT.WT.VC) not only
evens out the number of runs to be carried by each vehicle, but also redistributes the number
of vehicles among the two depots. As shown in Table 6.1, the total number of vehicles in
TT.WT.VC compared to TT.VC increased by 42%; however, the network clearance time is
decreased by 32%; the average in-vehicle travel time decreased by 3%; the average travel
distance decreased by 6%; and the average waiting time decreased by 17%. These
improvements are due to the redistribution of vehicles between the two depots.

- On the other hand the relative effectiveness of the TT scenario compared to the TT.WT.VC
scenario is manifested by several reductions: number of vehicles down by 12%, network
clearance time down by 32%, and average waiting time down by 15%, outpacing the loss
from an increase in the average in-vehicle travel time of only 4% and average travel distance
of only 6%. This is due to the spatial redistribution of the fleet between the two depots and
the temporally spread out scheduling pattern of vehicles which minimizes the waiting time
for evacuees.

- From Table 6.1, the network clearance time is minimized (117 minutes) for automobile
evacuees when optimizing both travel time in the network and mobilization time at origin.
All transit-evacuees clear the network after only 76 minutes, which suggests that mass
transit can still provide extra capacity to expedite the evacuation process and alleviate traffic
congestion. Abdelgawad et al. (2010) proposed the concept of “equilibrium mode-split” in
evacuation situations, where travel times per mode are equal and minimized.
6.5 Phase C: Impact of Traffic and Transit Mode Split on Evacuation Time

In congested urban areas, transit can play a paramount role in emergency evacuation. During emergencies, evacuees have several options such as driving, transit, walking and biking if feasible. Driving and transit are obviously the dominant choices. If all evacuees drive, severe congestion will hinder the evacuation process and some evacuees may not have cars at all. On the other hand, it may not be practical to expect drivers to leave their cars, which are handy to them, and seek transit, and therefore evacuation using transit alone may neither be practical nor always possible. It seems plausible to examine combining the use of private vehicles and transit to use all available resources (cars, buses) in the evacuation process. However, the impact of mode split between transit vehicles and vehicular traffic has not been adequately explored or quantified in the literature, i.e. how the mean and total evacuation times change with mode split values in between the extremes of 100% vehicles and 100% transit? The impact of mode split is investigated in this phase. Five vehicular percentages that constitute five scenarios are explored in an attempt to guide decision-makers and emergency planners with regard to the impact of splitting the evacuation demand amongst available modes during evacuation, or at least shed some light on the level at which vehicles should be replaced by transit vehicles in an emergency evacuation to avoid risking excessive breakdown.

Figure 6.10 shows the relative performance of transit vehicles to vehicular traffic. In terms of total travel time as shown in Figure 6.10a, the transit system can evacuate the whole population in the lowest total travel time (point T in Figure 6.10a). It is also clear that the higher the transit usage the better the system evacuation time which is best at point T. However, if some balance between total car evacuation time and total transit evacuation time is desired, the point at which such balance between mode total travel times occurs is the point that minimizes the total area under the transit and vehicular traffic curves. For instance, the point at which the total travel time is equal for transit users and vehicular drivers is found to be the 25% mode split as shown by point A and the hatched area in Figure 6.10. Any other mode split percentages will cause the total area under both curves to increase (conceptually similar to the user equilibrium in traffic assignment where the total area under the link performance functions is minimized) and hence one of the two modes will have total travel time higher than the other. Line C in Figure 6.10a represents the baseline case of 20,000 vehicles (35% of evacuees) and 47,000 transit trips (65%
of evacuees). Therefore, in order to have all evacuation modes experience the same total travel time, 75% of the evacuation demand needs to be evacuated by transit vehicles and 25% by vehicular traffic, i.e. a 10% mode shift to transit is required. The impact of mode split on the average evacuation time per evacuee is shown in Figure 6.10b which depicts trends similar to Figure 6.10a. The average evacuation time is an important bottom line question to individual evacuees, i.e., how long it will take someone to reach safety using car vs taking transit shuttles. In Figure 6.10b the mode split that results in equal average travel times is 15%. It is noteworthy that the above average numbers are pertinent to in-vehicle travel only and do not include wait time before the trip begins; however the mean wait time for car users is 48 min while the mean wait time for transit users is 45 min, which is almost identical. Therefore, in-vehicle travel time is indicative of the performance of both modes.

Figure 6.10 is very interesting from several different perspectives. First, the system may be evacuated faster if some car-based evacuees shift to transit shuttles if available. This is to be taken into consideration by evacuation planners who should make every effort to arrange for assembling shuttle bus fleets quickly in cases of emergency. This is to be taken into consideration as well by evacuees, who may find transit to be a sensible option, if they are properly educated and informed. From an ITS perspective, traveller information systems can be deployed to affect the desired mode split or at least influence evacuation mode choice. Lastly, planners, evacuees, and traveller information providers should all note that transit may be faster up to a limit beyond which some will find it faster to drive (left to point A’ in Figure 6.10b). In conclusion, evacuating 100% of the population by transit will achieve the least total evacuation time for the whole network (point T). However, under any mode split to the left of point A’ in Figure 6.10b, it may be harder to persuade drivers to abandon their cars and take a transit because they can reach safety faster by car. Therefore, from a system perspective, it seems plausible to achieve a mode split that equilibrates total mode travel time (point A in Figure 6.10a), while from a user perspective a mode split around point A’ in Figure 6.10b is more sensible.
Figure 6.10a Total Travel Time of Transit and Vehicular Evacuation vs. Percentage of People Evacuated by Car

Figure 6.11b Average Travel Time of Transit and Vehicular Evacuation vs. Percentage of People Evacuated by Car
Emergency Evacuation Demand Estimation from Regional Travel Survey

An accurate description of the spatial distribution of the population, by time of day and mode of travel, is essential to realistically model major population evacuation. Unlike day-to-day travel patterns, planning for emergency evacuation has a unique demand distribution that should be carefully examined in order for the model to produce accurate evacuation performance measures. Typically, travel demand modelling in evacuation is based on post-survey data after disasters or based on trip generation and participation rates of geographic areas. However, many cities lack such information due to the rare occurrence of major disasters. Therefore, a demand estimation model for emergency evacuation is paramount to produce realistic and meaningful results in case of large-scale evacuation as described in Chapter 8. This Chapter presents an evacuation demand estimation model, in which not only the value of the evacuation demand per Traffic Analysis Zone (TAZ) is determined but also the spatio-temporal distribution of demand is estimated.

7.1 Data Source: The Transportation Tomorrow Survey (TTS)

The Transportation Tomorrow Survey (TTS)\(^\text{12}\) is the largest and most comprehensive travel survey in Canada and is conducted once every five years. The TTS covers 5% of all households in the Greater Toronto Area (GTA) and surrounding areas, selected at random. The TTS was first conducted in 1986 and since then it has been carried out once every five years (1991, 1996, 2001, and 2006). The data used in this application are the TTS records collected for the year 2001. The 2006 survey data were still undergoing final refinements at the time of conducting this research and hence were not used. The demand estimation model includes the entire GTA which is divided into six regions; namely, Toronto, Durham, York, Peel, Halton and Hamilton, see Figure 7.1.

Data reported by the TTS include two sets of data: demographic characteristics such as age, gender, household size, dwelling type, to name a few; and travel patterns such as trip purpose

\(^{12}\) It is worth noting that the TTS does not capture tourism and business visitors trips.
and mode of travel data. While the survey was designed to cover only 5% of the households, expansion factors are used to expand the collected data to represent the total population of the survey area in the year of the survey. The expansion factors are determined based on geographical areas and verified based on Canada Census data that are used as the control total for calculating the expansion factors (DMG, 2003). For example, in the year 2001, across the GTA, 29,345 individuals reported that they started their trips at 7:30 AM using a driving mode; this is expanded to 568,047 driving trips across the GTA for a weekday giving an average expansion factor of 19.35.

![Greater Toronto Area](http://www.jpint.utoronto.ca/gta01/GTA.html)

Figure 7.1 GTA Regions for the year 2001 (source: [http://www.jpint.utoronto.ca/gta01/GTA.html](http://www.jpint.utoronto.ca/gta01/GTA.html))

Typically, the zone system used in the TTS surveys is called the GTA zone system; however, it extends to the areas external to the GTA. For example, in the 2001 TTS, the 1996 GTA zone numbering system has been extended and modified to include non-GTA areas included in the 2001 survey and to break down large zones into smaller zones in some regions. Each region is identified with a range of zone codes that represent the number of zones within each region. The zoning system used in this application is the 1996 zone numbering system which covers the same...
geographical area. For example, zone-codes in the City of Toronto range from 1 to 463 (see Table 7.1).

Table 7.1 Survey Area Zone Numbering for GTA Regions

<table>
<thead>
<tr>
<th>Agency</th>
<th>Zone-Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>1 – 463</td>
</tr>
<tr>
<td>Durham</td>
<td>501 – 765</td>
</tr>
<tr>
<td>York</td>
<td>1001 – 1353</td>
</tr>
<tr>
<td>Peel</td>
<td>1501 – 1749 (no 1716)</td>
</tr>
<tr>
<td>Halton</td>
<td>2001 – 2179</td>
</tr>
<tr>
<td>Hamilton</td>
<td>2501 – 2670 (no 2657)</td>
</tr>
</tbody>
</table>

7.2 Data Management Group (DMG) at the University of Toronto

The DMG of the Department of Civil Engineering at the University of Toronto is the custodian of the data sets reported and derived from the Transportation Tomorrow Survey. Although a useful and rich data retrieval system, the iDRS does not provide detailed data records for each person in the household during the day. Therefore, detailed information was obtained from the TTS survey thorough the data management group program (http://www.utrac.utoronto.ca/).

7.3 Demand Estimation Method

The detailed records of each person in each household are tracked during the course of a 24 hour period. The following attributes are used to construct a query to extract the demand data in each half hour during the entire day.

- Household sample number
- Person number within household
- Start time of trip (24 hour clock) 400-2800 (4 a.m. on the trip day to 4 a.m. the next day)
- Primary mode of trip\(^\text{13}\)
- GTA zone of trip destination
- GTA zone of household

The estimation process includes the following steps for each time interval:

\(^{13}\) In this application, modes are categorized as Drive (auto driver) and NonDrive modes (auto passenger, local transit, GO train, walk & cycle, other).
- Group people according to the start time of their trip (discretized into 30 minute intervals).
- Identify people who drive and identify their home location (zone). This results in an OD Matrix in which origins are the current location of people who drive and their default destinations, in the case of evacuation, are their homes (unless another safe destination is suggested by the destination choice module).
- Identify people who do not drive and identify their home location (zone). This results in an OD matrix in which origins are the current location of non-automobile people and their default destinations, in the case of evacuation, are their homes.
- Identify people who returned home.
- Identify people who have not yet made a trip.
- Identify people who are at their homes by combining the previous two steps.

The output of the demand estimation process are the spatial and temporal characteristics of the trips that are made by each of the three classes of people (travellers using automobiles, travellers using other modes, and travellers who are still at home or returned home). For those who are travelling when the crisis hits, their home locations are known and assumed to be their default destinations in the absence of a better destination choice. Ultimately, this method identifies where people are located by time of day and by mode of travel for the City of Toronto and the GTA.

### 7.4 GTA Demand Estimation Results

Figure 7.2 shows the aggregation of people across the GTA zones based on a 24 hour clock (400-2800) that starts at 4 AM on the trip day to 4 AM the next day. As shown in the figure, three groups of people are considered in the case of emergency evacuation: people who commute with the Drive mode, people who commute with the NonDrive mode, and people Resident at home. Figure 7.2 shows a wide peak activity period that starts from 8:00 AM in the morning and ends at 4:30 PM in the afternoon. The mode split is also clearly shown in Figure 7.2 where the vehicular trips across the GTA outpace all the other modes. An example for the spatial-temporal distribution of evacuation demand in the Greater Toronto Area is shown in Figure 7.3, Figure 7.4, Figure 7.5, and Figure 7.6 for the Total Population, Drive, NonDrive and Resident, respectively. Only two time intervals were selected for the sake of illustration; these are 6:00 AM
and 12:00. A complete set of maps for the City of Toronto and the GTA for each time interval (30 min) and for each demand category (Drive, NonDrive and Resident) is produced using ArcGIS, which brings the total number of maps to 288; only few of them is presented in the thesis for conciseness.

Figure 7.2 Drive, NonDrive and Resident Distributions in the Greater Toronto Area
Figure 7.3 Total Population Spatial and Temporal Distribution in the Greater Toronto Area
Figure 7.4 Drive Population Spatial and Temporal Distribution in the Greater Toronto Area
Figure 7.5 *NonDrive* Population Spatial and Temporal Distribution in the Greater Toronto Area
Figure 7.6 Resident Population Spatial and Temporal Distribution in the Greater Toronto Area
7.5 City of Toronto Demand Estimation Results

The same procedure is followed to extract the evacuation demand for the city of Toronto. The City of Toronto comprises 463 zones (16 planning districts) with a total population of 2.37 M. Figure 7.7 illustrates the temporal distribution of total people located in the City of Toronto. A similar trend to the GTA pattern pertained in the city of Toronto with the exception that the mode split is reversed where commuters without automobiles are higher than automobile commuters. This is attributed to the city characteristics and the connected public transit system, especially in the downtown area. Also, congestion pricing schemes have contributed to this shift.

![Figure 7.7 Temporal Distribution of Evacuees in the City of Toronto](image)

The plots show the cumulative number of people by mode within the geographical bounds of the City of Toronto at any instant in time regardless of whether they are travelling within the city, heading out of the city, heading into the city, or simply present at home. In total, the number of people in Toronto peaks at 108% of the City’s population (residents at 4:00 AM). This increase is attributed to the high concentration of economic activity in the City of Toronto and particularly the business and financial district. It is also interesting to see a wide peak activity period that starts from 7:00 AM in the morning and ends at 6:00 PM in the afternoon. The peak
demand is found to be 2.56 M people and occurs at the 11:30 AM-Noon interval, which constitutes the worst case scenario for evacuating the City of Toronto. It is important to note that total trips in the GTA that are processed by the demand estimation process sums up to the total population in the GTA of 5.368 M (DMG, 2003), 2.56 M of which are present in Toronto around noon time and the rest are outside of the bounds of the city. An example for the spatial-temporal distribution of evacuation demand in the City of Toronto is shown in Figure 7.8, Figure 7.9, and Figure 7.10 for the Drive, NonDrive, and Resident populations, respectively. Only two time intervals were selected for the sake of illustration; these are 6:00 AM and 12:00. As discussed previously, the maximum total demand across the City of Toronto is found to be around noon, but it is worth identifying what is the most critical zone with respect to the total number of people. As shown in Figure 7.11, it is found that zone number 223 holds the maximum number of people (33,230) at 12:30 PM. Zone number 223 is bounded by Front street from the south, King street from the north, Bay street from the East, University Avenue from the West and Union Subway station is located in the middle (see Figure 7.12). This is not surprising given that Union Station is the central hub for all inter-city Transit in Toronto, serving approximately one quarter of a million passengers each day.
Figure 7.8 Drive Population Spatial and Temporal Distribution in the City of Toronto
Figure 7.9 *NonDrive* Population Spatial and Temporal Distribution in the City of Toronto
Figure 7.10 Resident Population Spatial and Temporal Distribution in the City of Toronto
Figure 7.11 Population Distribution within the City of Toronto Zoning System by Time of Day
Figure 7.12 Location of the Maximum Number of People in the City of Toronto
8 Large-Scale Application - Evacuation of the City of Toronto

This chapter documents the efforts of applying the OSTE model presented in Chapter 3 and the transit routing and scheduling model presented in Chapter 4. The methodologies outlined in the previous chapters are applied to develop a large-scale evacuation model of the City of Toronto (see Figure 7.1). The large-scale application demonstrates the essence of the proposed approach and builds on the lessons learned from the prototype implementation described in Chapter 6.

8.1 Application Context and Analysis Scope

The City of Toronto is located in the centre of the Greater Toronto Area (GTA). The City of Toronto is bounded to the South by Lake Ontario, to the West by the City of Mississauga and City of Brampton, to the East by Durham Region and to the North by York Region. Toronto is a unique city; it is the oldest, densest, most diverse area in the region. It contains one of the highest concentrations of economic activity in the country. The key economic activity in Downtown Toronto is the high value added office sector, particularly in financial and business services.

It is clear that the frequency and impact of natural/man-made disasters are increasing worldwide and Canada (and specifically the City of Toronto) is not immune to this trend. Among the most frequent disasters are earthquakes, hurricanes, tsunamis, forest fires, tornados, ice storms and severe rain storms. In 1998, the largest Canadian disaster, the ice storm, struck Quebec and Ontario during which more than 5 million people were affected by at least one power outage. The extensive damage suffered by municipalities, companies, homeowners and forest owners makes this disaster one of the harshest disturbances recorded in North America. The Insurance Bureau of Canada estimated that claims resulting from the storm exceeded $1.1 billion. Also, the immediate economic cost was estimated at $1.6 billion (McCready, 2004), which led to the largest insurance payout in Canadian history. In 2005, Toronto and the surrounding area were hit by a severe rainstorm and tornadoes that led to the second-largest insurance payout in Canada’s history (Insurance Bureau of Canada, 2007\textsuperscript{14}). Homes were damaged in the areas of Kitchener,

\textsuperscript{14} Source: http://www.ibc.ca/en/Natural_Disasters/
Guelph, and possibly Toronto, Ontario. The damage covered an area from Stratford, Ontario (20 km west of Kitchener), to Peterborough, Ontario, and along Georgian Bay near Collingwood.

8.2 Analysis Scope and Definitions

This application demonstrates the integrated OSTE, Transit Routing and Scheduling, and Demand Estimation models to optimally evacuate the entire city of Toronto in cases of emergency. The City of Toronto is a typical example of a large North American city with a population of 2.37 M. The City of Toronto is located in the centre of the Greater Toronto Area (GTA). It is the oldest, densest, and most diverse area in the region. Toronto’s financial and business district is the highest concentration of economic activity in Canada. It is also home to the University of Toronto. The scope and definitions of the analysis are outlined as follows:

- All the data used in the application reflect 2001 year conditions.
- All the zones used in the application reflect the 1996 GTA zoning system.
- A simulation model for the City of Toronto is developed to simulate execution the evacuation plan.
- A simulation model for the entire GTA is developed to represent noncompliant traffic.

8.3 Supply Modelling

8.3.1 The GTA and City of Toronto Roadway Networks

In this effort, the GTA and City of Toronto road networks are developed in DynusT© (Dynamic Urban System in Transportation), a mesoscopic DTA model that is well-suited for dynamic traffic simulation and assignment on a regional scale. Mesoscopic models simulate the movement of individual vehicles in the transportation network but move them in groups according to the diagrams of fundamental traffic theory. Mesoscopic models achieve a compromise between microscopic and macroscopic models; they do not suffer from the curse of dimensionality when problem/population size increases like microscopic models and they provide more insightful analysis and detailed results than macroscopic models.
In this implementation, data were obtained from the EMME/2 planning model\(^\text{15}\). The EMME/2 planning model is maintained by the Data Management Group (DMG) at the University of Toronto and University of Toronto researchers. Although a rich and comprehensive planning model, some essential data to develop the GTA and City of Toronto models at the mesoscopic level were not available. Therefore the following tasks were carried out to develop the simulation model:

1. Traffic Analysis Zones (TAZs) are created using the information available from ArcMap\(^\text{©}\). The vertices of each zone in the GTA are exported from the TAZ layer (see Figure 8.1) and then a series of programs are created to automate the importing process. The process resulted in 160,324 vertices imported into the simulation model. Special attention is paid to keeping the mapping consistent between the TAZ zoning system and the zone numbers created in the simulation model.

2. Unlike the centroid-connector method of generating traffic in planning models, existing roads (links) have to be defined to generate traffic in the mesoscopic simulation model. Based on the author’s knowledge and familiarity with the network and using available data, generation links are created to release traffic from parking lots, parking garages, residential areas, etc. Special attention is given to model generation links at intersections to avoid unrealistic congestion at the beginning or end of generation links. In addition, the traffic generated from these links is distributed across the road segments proportionally to the capacity and the length of these road segments.

3. The evacuation demand, resulted from the demand estimation model, is assigned to the TAZs according to the defined generational links in the form of a demand matrix. The demand matrix could take either the form of regular OD matrix in the case of pre-defined evacuation destination zones (i.e., SE and OTE scenarios) or the form of multiple origins-one-destination matrix in the case of superzone network representation (i.e., OSTE and SE-DC). It is worth noting that in case of emergency evacuation carpooling is anticipated; therefore, a high passenger occupancy factor is assumed (i.e., PEF =5).

\(^{15}\) Contact person is Susanna Choy, The Joint Program in Transportation, University of Toronto
4. The freeway system across the GTA is carefully modelled so as not to include any generation links or destination nodes. On and off ramps link categories are created to realistically model movements at the entrances and exits of freeways.

5. Traffic flow models have to be identified in the mesoscopic model as described previously. A typical Greenshield traffic flow model is used while constructing the speed-density function for different road segments according to their category (e.g. highway, freeway, collector, on and off ramps, etc). The uncongested portion of the fundamental traffic diagrams was set to approximately match the typical link-volume delay function for different road segments provided by EMME/2.

6. Traffic signals in the model are designed as fully actuated to automatically handle the anticipated traffic fluctuations in case of emergency evacuation. In reality, approximately 25% of traffic lights in Toronto use adaptive control (SCOOT). The rest are either actuated or pre-timed. The pre-timed signals would require evacuation-specific timing plans in reality, which was approximated by fully actuated operation in the model.
Figure 8.1 The GTA Traffic Analysis Zones

As discussed in Section 8.1, the evacuation area includes the City of Toronto. Therefore, a replica of the GTA model has been developed to include only City of Toronto. The GTA model is developed to estimate non-complaint traffic in emergency evacuation and to account for inter-city trips made to/from the City of Toronto. The GTA and City of Toronto simulation models are illustrated in Figure 8.2 and Figure 8.3 respectively. Table 8.1 illustrates the GTA and City of Toronto model characteristics.
Figure 8.2 The Greater Toronto Area Simulation Model

Figure 8.3 The City of Toronto Simulation Model
Table 8.1  The GTA and the City of Toronto Model Characteristics

<table>
<thead>
<tr>
<th>Model</th>
<th>GTA Model</th>
<th>City of Toronto Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Covered (km²)</td>
<td>8,573</td>
<td>628</td>
</tr>
<tr>
<td>Perimeter (km)</td>
<td>624</td>
<td>120</td>
</tr>
<tr>
<td>No. of Nodes</td>
<td>12,082</td>
<td>3,393</td>
</tr>
<tr>
<td>No. of Links</td>
<td>29,194</td>
<td>7,480</td>
</tr>
<tr>
<td>No. of Zones</td>
<td>1,677</td>
<td>463</td>
</tr>
<tr>
<td>Length of Roads (km)</td>
<td>23,303</td>
<td>3,610</td>
</tr>
<tr>
<td>Population (M)</td>
<td>5.38</td>
<td>2.37</td>
</tr>
</tbody>
</table>

8.3.2 Transit Infrastructure: The Toronto Transit Commission (TTC) System Fleet

The TTC fleet consists of bus, light rail transit (streetcar), and rapid rail transit services. In 2001, during the peak AM period, it is reported that the TTC operated up to 294 transit routes, these routes covered around 10,000 transit stops with a service frequency that ranged from high frequency transit service (headway = 2.5 minutes, e.g. Route No. 36, Finch East) to low frequency transit service (headway = 60 minutes, e.g. Route No. 162, Lawrence-Donway) (TTC, 2001). The fleet includes about 1500 vehicles (around 1320 buses plus 180 streetcars) during the peak AM period. The Rapid Transit service includes 4 subway (metro) lines (Bloor-Danforth, Yonge-University-Spadina, Sheppard and Scarborough RT) which cover a large area of the City of Toronto. The location of subway stops are extracted from the transportation network and placed on the subway lines as illustrated in Figure 8.4. The Rapid Transit Fleet characteristics are attached to Figure 8.4 (TTC, 2001). Subway lines are coded as double track lines in which each station is represented by two platforms; one for each direction. Also, additional constraints are added to model FIFO operation of subway trains while departing and arriving at subway stations. It is assumed that, in cases of emergency, the entire bus fleet will be available for our system to reschedule and reroute based on the needs of the evacuation process. Regular transit services will no longer be in effect. The whole bus fleet would operate as shuttles to the nearest safe zone. Streetcars are not used in this application.

\[16 \text{ It is worth noting that the 2001 TTS data does not include the Sheppard subway line. In this implementation, the Sheppard subway line is included to the existing transit system capacity for the sake of completeness.}\]
8.4 Demand Modelling

Unlike the typical day-to-day demand patterns, in this application the evacuation demand has the following unique characteristics:
- As discussed in Chapter 7, the worst time for a crisis to hit is around noon time. Trip start times were used to extract the demand data as discussed in Section 7.3. Based on the TTS, it is reported that the average automobile in-vehicle travel time in the GTA is 12 minutes and the average transit in-vehicle travel time is 25 minutes (Roorda et al., 2006). Therefore, in this analysis, the trip start times were discretized into 30 minute intervals.

- The demand estimation process resulted in three main categories of trips that are modelled in evacuating the City of Toronto. Table 8.2 shows the breakdown of trips according to the location of evacuees and their mode of transport at the onset of the evacuation event, i.e. at noon.
Table 8.2 Evacuation Demand in Each Zone in the City of Toronto by Mode at Noon

<table>
<thead>
<tr>
<th>Trip</th>
<th>No of Trips</th>
<th>Description</th>
<th>Spatial Distribution of Evacuees Trip Origins</th>
</tr>
</thead>
<tbody>
<tr>
<td>At Home Evacuees</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,107,353</td>
<td>Cumulative number of people located at their homes in the subject zone at the onset of the evacuation at noon. These trips include trips that are not yet started and trips that started and ended at home at the time of the evacuation event.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transit</td>
<td>791,264</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Auto</td>
<td>663,209</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cumulative number of people starting from any zone other than their homes at the onset of the evacuation and ending at the subject zone in the 11:30-noon interval using “NonDrive” mode. These trips include trips made by the following modes: passenger, transit, walking, cycling.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cumulative number of people starting from any zone other than their homes at the onset of the evacuation and ending at the subject zone in the 11:30-noon interval using “Drive” mode. These trips include trips made by automobile drivers.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It is interesting to show that the total number of evacuation trips at noon (2.56 M) is greater than the population of the City of Toronto (2.36 M). This increase (8.2%) is attributed to the high concentration of economic activities in the City of Toronto and particularly in the Waterfront area where key financial, office sector and business services are located. It is also important to note that this increase is the net difference between internal-external and external-internal trips made from/to the City of Toronto.

In the case of evacuation, it may be harder to persuade drivers to abandon their cars and take any other mode. Also, transit users and non-drivers are captive to transit modes because their choices are limited. Therefore, it seems practical to assume that evacuees will use the same mode of transport as when commuting to the City of Toronto, i.e. the Not-At-Home drivers will evacuate using their cars and the Not-At-Home non-Drive users will evacuate using transit modes. Available transit modes in this implementation include the Rapid Transit system (Subway Lines) and surface street buses used as shuttles. For At-Home evacuees, trips are assigned to modes based on the mode split reported by the TTS for trips made by residents of the City of Toronto (DMG, 2003). The overall results are 1,216,886 evacuation trips by automobile and 1,344,942 evacuation trips by transit.

8.4.1 Representation of Noncompliant Traffic

In emergency evacuation in general and hurricane evacuations in particular, a considerable percentage of evacuees seek out their homes and relatives first. In a comparison of trip generation models in hurricane evacuations, Wilmot and Mei (2004) reported that 50–70% of destinations were to relatives and friends. In our implementation, evacuees are advised to go to the optimal shelter; however, deviation from the evacuation “plan” is unavoidable. Therefore, a certain percentage of evacuees are assumed not to comply (Noncompliant Evacuees) with the optimal plan and seek their homes. The percentage of noncompliant evacuees is treated as a sensitivity parameter, i.e. exogenously supplied by the modeller. This study makes no attempt to estimate this percentage on the basis of evacuee behaviour. During the demand estimation process, some of the evacuees who are not at home are directed towards their homes based on the specified percentage. Figure 8.5 shows the distribution of trips originated from the City of Toronto at the onset of the evacuation and destined to their home locations. As shown clearly in the figure, a considerable number of trips originate from outside Toronto.
Two components form the total demand in the evacuation planning model as shown in equation (8.1): 1) compliant evacuation demand (super-zone demand), and 2) noncompliant demand:

\[
\text{Total Demand (D)} = \text{Evacuation Demand (super-zone demand)} + \text{Noncompliant Demand}. 
\]  

To the best of the author’s knowledge and based on the literature, evacuee compliance to guidance is rarely modelled or reported. Although it is challenging to model evacuee behaviour, stated-preference surveys and post-analysis surveys of certain evacuation scenarios may be plausible avenues for modelling such behaviour. In the absence of past-evacuation surveys in Ontario and Toronto, it is assumed that 25% of evacuees will not comply with the provided guidance and will seek their homes first. This assumed percentage has no scientific basis. It is only for the purpose of analysis and illustration until a better behavioural approach is available.

To represent noncompliant demand, a dynamic user equilibrium traffic assignment is executed at the GTA level, in which the origins are the locations of people at the onset of evacuation and the destinations are their homes. Then, an analysis is made to capture these trips at the gateways of City of Toronto. The output of this analysis is an origin-destination matrix that identifies the demand pattern of people who may return to their homes instead of going to the advised
destination according to the optimized evacuation plan. Finally, a certain percentage of this demand is assumed not to comply with the plan while the destination of the rest of the population is optimized.

8.4.2 Estimation of Transit Stops Demand

As discussed in Section 8.4, at noon the total number of people to be evacuated using Transit modes is found to be 1.34 M. The spatial distribution of evacuees within each TAZ is important when estimating the demand at each transit stop. Therefore, an algorithm is developed to randomly distribute evacuees in TAZs according to the demand in each zone. Figure 8.6 illustrates an example for the Geo-location of evacuees within the TAZ. The number of geo-points in the figure represents the number of evacuees destined for this zone using any NonDrive mode. In this implementation, Rapid Transit (Subway) and Buses are the two transit modes assumed to evacuate the transit-dependent population.

Figure 8.6 Geo-Locations of Evacuees within TAZ (partial view of the City of Toronto)

8.4.2.1 Demand on Subway Stations

An access distance buffer zone (1000 m) for transit evacuees is defined and evacuees within the accessible buffer zone are assumed to be carried by subway to safe destinations. The spatial distribution of evacuees within each TAZ is important when estimating the demand within the buffer zone. In this study, a uniformly distribution of evacuees within each TAZ is assumed for
the lack of better information. Safe destinations for subway travellers are subway terminals (e.g. Kipling, Finch Subway stations). Figure 8.7 illustrates the demand bounded in the buffer zone of the subway system, which is estimated to be 615,434 evacuees. Evacuees are then spatially assigned to the nearest subway station. Figure 8.8 shows the distribution of evacuees at each subway station. The average walking distance to the closest station is found to be about 460 m with the maximum demand located at King and Union Subway stations.

Figure 8.7 Buffer Zone Representation and Geo-locations for Evacuation Demand Assigned to the Subway
8.4.2.2 Demand on Bus Network Stops

A large portion of the transit-dependent population (54%) is assigned to the Rapid Transit system. This not surprising given the “self-containment” nature of the City of Toronto and the high density and transit-oriented development throughout the city, especially along the Yonge-Spadina Subway line (Miller and Shalaby, 2003).

Figure 8.9 shows the Geo-Locations of evacuees assigned to the shuttle bus system. Main bus stops are extracted from the EMME/2 planning model and the TTC bus routes are provided by the University of Toronto Map Library\textsuperscript{17} (see Figure 8.10). Then evacuees are spatially assigned

\textsuperscript{17} The University of Toronto Map Library (online source, \url{www.main.library.utoronto.ca})
to the nearest bus stop, resulting in an average walking distance to the nearest bus stop of 332 m. The demand distribution at each bus stop is shown in Figure 8.11.

Figure 8.9 Geo-Locations of Evacuees Assigned to Shuttle Buses
Figure 8.10 Major Bus Stops and TTC Routes for Bus Shuttle Evacuation
8.4.3 Estimation of TTC Initial Conditions

In this implementation, special attention is paid to estimating the initial conditions for operating the TTC system in the case of an emergency evacuation. As discussed in section 8.4.2, the primary modes for evacuating a transit-dependent population are the Rapid Transit system (Subway) and shuttle buses. The initial conditions for both systems are estimated based on available data and the operational characteristics of each system.

According to the reported average speed and average headway for each subway line, the locations of subway trains are identified and used as initial condition while applying the scheduling algorithm for subway trains. In this implementation, each subway line is modelled separately and transfers between subway lines are barred. This decision is made to minimize the chaos associated with subway operations in the case of emergency evacuation. It is also worth
noting that in the case of subway operation, the routing is predetermined due to the nature of the subway tracks and the determinant backtrack locations. The problem is rather a scheduling problem that is concerned with the sequencing of subway runs.

The spatial distribution of transit shuttle buses in the TTC at the onset of evacuation is crucial to realistically model start-up conditions for buses. To the best of the author’s knowledge and from the available data, the location of buses within the transportation network is not readily available. However, this data can be extracted using the output of a recent research effort at the University of Toronto where a dynamic transit assignment model (namely MILATRAS) was developed. MILATRAS is capable of simulating passengers and bus schedules at transit stations (Wahba, 2009). A post-analysis method is developed to obtain the location of buses at any point of time through the extracted transit route records from MILATRAS. Although MILATRAS results were reported for the AM peak only (6:00 – 9:00 AM), the model was run until noon, which coincides with the onset of the worst case evacuation scenario of interest. Therefore, a series of programs are developed to extract the bus locations (e.g. at transit stops) around noon for all TTC bus routes (inbound and outbound bus routes). It is found that the maximum number of transit vehicles per route is along the 504 King Streetcar route. This is not surprising since the 504 King Streetcar is the most congested transit route in the TTC. Figure 8.12 shows the spatial distribution of the TTC fleet in the City of Toronto at noon at the onset of the evacuation scenario.

---

18 Permission was obtained from Prof. Moahmed Wahba and Prof. Amer Shalaby to access the MILATRAS data.
8.5 Results and Discussion

8.5.1 OSTE for the City of Toronto

8.5.1.1 Genetic Optimization Results

The output of the genetic-based optimization process is a scheduling vector ($\mu$) that minimizes the pre-specified objective function (see Section 3.4) with a modified network representation to model the destination choice problem (super-zone). The OSTE framework, integrated with the GENOTRANS engine, outputs the optimized scheduling vector and the detailed routing plan for the evacuation scenario.

Designing genetic-based optimization is problem-specific and requires extensive testing to make selections among the many variables as discussed in section 3.5.2. Table 8.3 shows the GA methods and parameter values chosen based on these tests.
### Table 8.3 Genetic Algorithm Methods and Parameter Values

<table>
<thead>
<tr>
<th>GA Design Element</th>
<th>Method or Chosen Value</th>
<th>Description and Chosen Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Size</td>
<td>Population Size $\geq$ Chromosome Size</td>
<td>80</td>
</tr>
<tr>
<td>Initial Population</td>
<td>Random</td>
<td>Each gene in the decision variable vector is encoded as a real number with a range between 0 and 1. The chromosome size (number of genes) represents the number of departure time intervals.</td>
</tr>
<tr>
<td>Selection</td>
<td>Linear Ranking Selection</td>
<td>The individuals in a population of n chromosomes are ranked in descending order of fitness, with a rank of n points given to the best individual and a rank of 1 given to the worst individual.</td>
</tr>
<tr>
<td>Recombination</td>
<td>Real Blend Crossover</td>
<td>Exchanges the genetic information between the population individuals; it acts on two parents in the intermediate population by combining their traits to form two new offspring. This is not applied to all chromosomes but depends on the probability $P_c$ defined by the crossover rate. $P_c = 0.9$ (90 percent of the time the crossover operator is applied).</td>
</tr>
<tr>
<td>Mutation</td>
<td>Real Gaussian Mutation</td>
<td>Randomly chosen genes are mutated with a range of $\pm 5%$</td>
</tr>
<tr>
<td>Stopping Criterion</td>
<td>Number of generations with no decrease in fitness function value</td>
<td>25</td>
</tr>
<tr>
<td>Parallel GA</td>
<td>Fully Connected Topology</td>
<td>Number of demes: 4 Migration policy: good migrants replace bad individuals Migration rate (the number of individuals to migrate): 15% of population</td>
</tr>
</tbody>
</table>

The objective function (fitness function) values for various scheduling vectors and destination choices are reported. Conceptually, the fitness value in this application is a function of the scheduling vector and the destination choice; however, there is no closed-form equation that can represent this relationship. The optimization-by-simulation approach applied in this implementation enables such a relationship to be studied and optimized. Since it is a multi-dimensional genetic optimization problem, it is possible to obtain (in the final set of scheduling vectors) several chromosomes (solutions) that possess similar fitness function values. Then the optimal set of parameters can be identified with the smallest fitness function value (in the case of minimization problems).

As discussed previously, genetic-optimization approaches require careful identification of the set of parameters that best replicates the specific problem at hand. A good design for the GA parameters is one that captures the fundamental properties of genetic optimization which are...
survival of the fittest and evolution of the population from one generation (iteration) to another. This is clearly shown in Figure 8.13, where the fitness function values are plotted in descending order per generation. The first generation represents 20 randomly selected feasible solutions; then the second generation is created using the genetic operators (crossover and mutation) to produce better generations with respect to the previous generation based on the average fitness value. As shown in the figure, the procedure is repeated for the next generations and a steady decrease in average fitness function value is maintained. Ultimately the average fitness of a population of chromosomes will approach the optimal fitness function value as the number of generations goes to infinity. While it is practically infeasible to continue the generic-optimization procedure indefinitely, typically a stopping criterion is specified to bring the optimization process to an end. In this implementation, the optimization process is terminated if the value of the fitness function does not change by more than 1% in 25 generations. Figure 8.14 also illustrates the evolution of the minimum fitness function value with the number of generations across multiple demes.

In general, the unconstrained genetic-based optimization procedure is generic since it does not restrict the scheduling vector values except for the feasibility requirements. This implies that the relationship between the optimization parameters (genes of each chromosome) is not determined a priori, but rather is an output of the optimization process. In this application, in order to achieve the feasibility requirements, two constraints must be met: 1) the non-negativity constraint for all parameters (gene values); 2) the sum of all parameters must be equal to one; this is to ensure that 100% of evacuees are released from the transportation network.
Figure 8.13 Evolution of the Optimization Process using the Genetic Algorithm
An advantage of using a simulation-optimization approach is the explicit representation of network congestion and capacity constraints, an essential matter in emergency evacuation. Therefore, relatively poor scheduling and destination optimization parameters will lead to more congestion and result in low fitness function values. This is clearly captured in Figure 8.15 where the best and worst chromosomes of the genetic algorithm evolution are plotted with the associated fitness function value. A significant difference in the scheduling vector parameter values is shown with corresponding significant values in the fitness function (123 vs. 938 min). It is to be noted that the chromosome size is determined by the evacuation horizon and the discrete scheduling intervals, for example the evacuation horizon is set to 2000 minutes and the scheduling interval is set to 50 minutes; this results in 40 parameters in each chromosome. When the scheduling vector is spread across the entire evacuation horizon, this results in the worst fitness value due to the fact that the waiting time of evacuees increases significantly. On the other hand, the set of parameters with the best fitness (lowest total system evacuation time) demonstrates a good compromise between in-vehicle travel time and waiting times at the origins.

Figure 8.14 Fitness Function Value as a function of the Number of Generations
8.5.1.2 Effect of Parallelization and Distribution of GAs

While GAs are capable of finding good solutions to practical engineering and science applications (David, 1994) in a reasonable amount of time, in some cases such as large-scale evacuation, GAs may require hundreds of expensive fitness evaluations and, depending on the cost (time) of individuals’ fitness evaluation, GAs may take days or even months to find an acceptable solution (Cantu-Paz, 2000). Therefore, there have been numerous efforts to make GAs faster and one of the promising techniques, as discussed in section 3.5.2, is to use distributed implementations of GAs. Alongside distributed implementations of GAs, some argue that the parallel GA (e.g. multi-deme) better mimics the nature of the population compared to the simple GA with single population used by a serial GA (Sumida B.H. et al., 1990; Cohoon et al., 1991). Therefore, parallel GAs are expected to converge faster if each deme is assigned to separate processors, however the quality of the solution depends on the choice of the population size in each deme, deme topology and the migration mechanism (Hart et al., 1996; Kruchten et al., 2004).

The effect of distributing the GA population to multiple slaves on the performance of the GA is typically examined in terms of: 1) elapsed time, 2) distribution speedup, and 3) efficiency of the DGA (William et al., 1996; Cantu-Paz, 2000). The analysis of the elapsed time focuses on the master processor and the number of available processors. In a typical GA generation, the master
sends a fraction (or all) of the population to each of the slave processors, using communication
time $T_c$. This communication time is exhausted in creating the files/directories necessary for each
processor in the available slave list. Although the master consumes some time in the selection,
crossover and mutation processes, this time is typically ignored compared to the communication
and execution times. Due to the unprecedented sheer size of the input/output files in this
application, the master remains idle and waits for the results from the available processors. Next,
each slave (processor) evaluates a fraction of the population in time $\frac{nT_x}{P}$, where $T_x$ is the
execution time of one individual, $n$ is the population size, and $P$ is the number of available
processors. Therefore the elapsed time for one generation is given by equation (8.2).

$$T_e = PT_c + \frac{nT_x}{P} \quad (8.2)$$

An important concern when implementing large-scale problems is that the frequent
communications between master and slaves may offset the gain in computation time. Therefore,
the speedup ($S$) of the master-slave parallel GA is another measure of the effectiveness of a
PGA. The speedup is the ratio of the execution time of the SGA to the elapsed time of the PGA
as shown in equation (8.3). The greater the ratio of $T_x$ to $T_c$, the more linear will be the
speedup.

$$S = \frac{nT}{PT_c + \frac{nT_x}{P}} \quad (8.3)$$

Although using more slaves reduces the computation time significantly, the communication time
increases. Therefore, a third measure of the effectiveness of the DGA compared to a SGA is the
efficiency ($E$). The efficiency is defined as the speedup $S$ divided by the number of processors $P$
(see equation (8.4). The efficiency represents the utilization of processors. Ideally, $E$ would be
constantly one and $S$ would be equal to the number of processors used (i.e. linear speedup).
However, in reality, the cost of communications prevents this ideal case from happening;
therefore, the efficiency is chosen as a measure of the deviation from the ideal case.

$$E = \frac{S}{P} \quad (8.4)$$
By examining the evolution of the fitness function with each GA generations, it is clearly shown in Figure 8.16 that the PGA outperforms the SGA in terms of fitness function value and convergence. First, the PGA results in a better fitness function value at the termination point of the GA (generation 25) (123 vs. 152, i.e. 18% reduction). This means that for the same number of generations, it is found that the PGA provides higher quality solutions. Secondly, the PGA results in faster convergence when compared to the SGA. This is clearly shown in Figure 8.16 by looking up how many generations of the PGA will result in the same fitness value as the corresponding SGA: it is found that the PGA can produce the same quality as the SGA (fitness of 152) in 1/3 of the number of generations (see dotted line), i.e. the PGA is three times faster than the SGA. This means that for the same fitness function value, the PGA converges faster. Therefore, the use of a PGA has not only resulted in faster convergence but also in a higher quality solution (Hart et al., 1996).

Figure 8.16 Effect of Parallelization on GA Convergence and Quality

To judge the effectiveness of implementing the PDGA, elapsed time, speedup and efficiency are examined. Figure 8.17, Figure 8.18 and Figure 8.19 illustrate the change in elapsed time, speedup and efficiency versus the number of processors, respectively. As clearly shown in the figures, the elapsed time decreases, the speedup increases and the efficiency decreases as the number of processors increases. The rate of decrease in elapsed time reduces as the number of processors is increased, expect for few sudden drops/peaks in the elapsed time/speedup and
efficiency; this is primarily because of the increase in communication cost with number of processors. This trend is similar to the speedup which increases almost linearly but with a decreasing slope as the number of processors increases. These drops/peaks form when a divisor results from dividing the GA population size by the number of available processors. This phenomenon is confirmed by examining the slightly decreasing efficiency curve. With a low number of processors, the efficiency is 100% which means that the addition of one processor greatly affects speedup; however, with the addition of more processors the communication cost causes the efficiency to deviate from its optimal value. Similar to the speedup, few peaks are observed in the efficiency curve with the highest efficiency (90%) occurred when the number of available processors equals the population size.

![Figure 8.17 Elapsed Time vs. Number of Slave Processors](image1)

![Figure 8.18 Speedup vs. Number of Slave Processors](image2)
Figure 8.19 Efficiency vs. Number of Slave Processors

8.5.1.3 Traffic Assignment Outputs

The output of the genetic optimization process (the optimal scheduling vector corresponding to the best chosen chromosome) is evaluated using the traffic assignment model to produce the detailed routing plan and measures of effectiveness. The mesoscopic representation of the traffic simulation model provides sufficient details for the analysis of departure time, destination choice (shelters), and routing plan for each vehicle. An important factor in planning for emergency evacuations is the status of evacuees in/out of the network with the evolution of the evacuation plan. In this implementation, special attention is paid to studying the dynamic interaction between the loading and evacuation curves and the area between them. Three scenarios are evaluated: OSTE, OTE, and SE. Each scenario examines a certain level of integration for different evacuation strategies. While OSTE integrates evacuation scheduling and destination choice optimization, OTE examines the effectiveness of evacuation scheduling only; SE, on the other hand presents the do-nothing scenario where evacuees are rushed to the transportation network without any pre-announced schedule. In fact, the SE evacuation scenario is evaluated twice; once with destination choice optimization (SE-DC) and a second time without destination choice optimization (SE). The latter replicates the worst case scenario where evacuees immediately seek their preferred destination which is not necessarily the optimal one.

The optimal loading and evacuation curves are shown for the four scenarios (OSTE, OTE, SE, and SE-DC), respectively, in Figure 8.20, Figure 8.21, Figure 8.22, and Figure 8.23. The following measures of effectiveness are extracted to judge the efficiency of evacuation strategies
when compared to the baseline scenario: average waiting time, average travel time, average total system evacuation time, average trip distance, network clearance time, and average stop time.

Figure 8.20 OSTE Optimal Loading and Evacuation Curves
Figure 8.21 OTE Optimal Loading and Evacuation Curves

Figure 8.22 SE Loading and Evacuation Curves
It is clearly shown in Table 8.4 that the SE strategy performs the worst since it results in the longest network clearance time (NCT, end of evacuation curve) and the most congested travel times (area between the loading and evacuation curves). This is not surprising given the surge in demand in emergency evacuation. Simultaneous evacuation with destination optimization can reduce in-vehicle travel time but not to the level that any scheduling strategy can achieve, i.e. OTE and OSTE always result in less in-vehicle travel times compared to SE and SE-DC. The average stop time (time when vehicles are caught in congestion) is longest in the case of SE and SE-DC. In terms of NCT, OSTE performs the best; however, SE-DC performs slightly better than OTE; as OTE explores the optimal scheduling curve so as to minimize the total system evacuation time, this typically results in evacuees being held back from being rushed to the transportation network (in this case by up to 82 min). On the other hand, it should be noted that network stability is another performance factor to consider. SE and SE-DC may result in a network (infrastructure) that has no further room for manoeuvre and could come into gridlock in the case of further panic and/or secondary events (i.e. possibly less stable). Stability however is not explicitly tested in this research. This conclusion is hypothesized based on the observation.
that SE and SE-DC have the longest stop time, i.e. traffic is more often in a stop-and-go condition.

Table 8.4 Comparative Analysis of Evacuation Strategies

<table>
<thead>
<tr>
<th>MOE</th>
<th>Scenario</th>
<th>Average Waiting Time (min)</th>
<th>Average In-Vehicle Travel Time (min)</th>
<th>Average Total System Evacuation Time (min)</th>
<th>Average Trip Distance (km)</th>
<th>Network Clearance Time (min)</th>
<th>Average Stop Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE</td>
<td>0</td>
<td>412</td>
<td>412</td>
<td>11</td>
<td>1815</td>
<td>380</td>
</tr>
<tr>
<td></td>
<td>SE-DC</td>
<td>0</td>
<td>175</td>
<td>175</td>
<td>11</td>
<td>800</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>OTE</td>
<td>82</td>
<td>112</td>
<td>194</td>
<td>14</td>
<td>940</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>OSTE</td>
<td>65</td>
<td>50</td>
<td>115</td>
<td>12</td>
<td>445</td>
<td>33</td>
</tr>
</tbody>
</table>

It is also obvious that OSTE outperforms all other strategies since it synergizes the scheduling and destination choice in a one shot optimization. The improvements are remarkably clear, especially in terms of NCT and average stop time. This reflects the essence of a wise evacuation strategy that holds evacuees back up to a point where if they are released into the network and seek the dynamic optimal destination, they will clear the network promptly with the minimum stopping time, encounter less congestion, and contribute to a shorter overall network clearance time. An order of magnitude savings in NCT (75%), total system evacuation time (72%), average stop time (92%) and in-vehicle travel time (89%) are reported when compared to the do-nothing case. This means that using OSTE, the network can be cleared four times faster than the do-nothing strategy; evacuees travel eight times less than in the do-nothing strategy, and finally evacuees stop eleven times less than in the do-nothing strategy, and finally the total system evacuation time can be reduced to one quarter.

The resulting optimal departure time, when vehicles are allowed to enter the network to minimize congestions effects and clear the network promptly, is interesting from several perspectives. First, it replicates the network breathing concept, which is an analogy for the breathing process where vehicles are inhaled by the network and dissipated by the network (Dixit and Radwan, 2009). This is clearly shown in Figure 8.24 in the network departure time pattern where vehicles are released in optimal steps to allow for network breathing.
Second, it captures the concept of *reserve capacity*, which defines the additional demand that a network can accommodate without changing its physical characteristics (Yang and Bell, 1998). This phenomenon is captured when comparing the actual number of vehicle that a network can accommodate under different demand patterns. Over rushing the network as in the SE strategy, results in a critical transportation network with no room for any additional demand, in other words it results in early gridlock in the network that lasts for days before being released. This is clearly evident in Figure 8.25 when comparing the number of vehicles in the network with the time for the OSTE and SE strategies. Although in SE 100% of the demand is released at time zero, it takes almost 12 hours to be completely released into the network, meaning that traffic was backed up because the network capacity could not accommodate such a surge in demand.
On the other hand, in OSTE, the same number of vehicles entered the network with the last vehicles being released after almost four hours which resulted in more room for routing options and a less congested transportation network. The same conclusion can be drawn from Figure 8.26 by comparing the density levels in the transportation network for SE and OSTE at the onset of the evacuation. One can easily identify which scenario has more room for additional demand if required. Thirdly, it captures *network breathing* at the destination level where destination choice changes dynamically with the evolution of the evacuation process depending on the capacity of the routes leading to safe destinations. While in OSTE optimal destinations depend on the departure time of vehicles and the congestion level (capacity) of routes, in SE destinations are typically the nearest safe shelter or homes regardless of the route conditions leading to these destinations. This can be illustrated by displaying the paths from an origin zone (e.g. Zone 229 in the middle of downtown Toronto) to the resulting optimal destination with the corresponding departure time for OSTE and SE (see Figure 8.27). In SE, there is only one *fixed* path that is the shortest path assigned at the beginning of the simulation and vehicles typically stick to that path until reaching a safe destination. It is clearly seen how different paths are taken to different destinations from the same origin in a way that depends greatly on the network conditions.
Figure 8.26 Density Levels in the Transportation Network for SE and OSTE Strategies
Figure 8.27 Examples of Optimal Destinations with the Corresponding Departure Time
8.5.2 Optimal Routing and Scheduling of Transit Vehicles

In general, transit services are designed based on the seating and standing capacity of transit vehicles; however, in emergency evacuation, evacuees are expected to tolerate more crowding and make the best use of any available space in the transit vehicle. Therefore, it is plausible to consider the crush load/capacity of a transit vehicle when planning for emergency situations. For a transit bus it is assumed that one bus can carry at most 90 passengers, a subway vehicle is assumed to carry at most 330 passengers; that is a subway train with six vehicles can carry up to 2000 passengers. Dwell times that reflect the physical and operating characteristics of transit units are calculated at transit stops (Vuchic, 2005). It should be noted that dwell times are mode-dependent; therefore, dwell times for buses are different from dwell times for subway trains.

The output of the MDTCPD-VRP is the optimum routing and scheduling of each transit vehicle. In addition to the optimized routing plans, the model provides extensive and detailed vehicle-by-vehicle output for each scenario; this is beneficial in many ways. The analyst can examine the scheduling of transit vehicles in the reported data at the transit stop level, individual bus level, or route level. For instance, at the stop level, the analyst can identify stop ID, bus arrival times, departure times, onboard passengers upon arrival, onboard passengers upon departure, alighting passengers, and boarding passengers. At the vehicle level, the analyst can observe and assess the fleet capacity requirements by examining vehicle ID, total number of passengers transported by that vehicle, travel distance, and travel time. At the route level, the analyst can construct the route that a given transit vehicle takes by examining vehicle ID, first stop ID, sequence of nodes that construct a route, pick-up points along that route, and destination stop ID.

---

19 The crush load (maximum people per vehicle) is typically more than 150% of a bus’s seating capacity. Such loads are unacceptable to regular day-to-day passengers. In typical day-to-day operation of buses in rush hours, crush loads might prevent the circulation of passengers at intermediate stops and therefore result in delays and reduce the overall vehicle and system capacity (TCRP, 2003). However, in the case of bus shuttling in emergency evacuation situations, buses are not allowed to stop at intermediate stops; typically, the buses are loaded to capacity (due to the huge surge in demand that exceeds the supply of buses) and stop only at safe shelters.

20 The Characteristics of Common Bus Transit Vehicles – United States and Canada are reported in (TCRP, 2003). A typical low floor 40’ transit bus has a passenger capacity that ranges from 55–70; therefore, it is plausible to assume a slightly higher crush capacity in the case of emergency situations, or a maximum of 90 passengers per transit bus.
8.5.2.1 Subway

As discussed in section 8.4.2.1, subway trains are scheduled to transport evacuees who are within the buffer zone to safe shelters (subway terminals). The model generates the optimal scheduling and timetable for each train. It should be noted that a train does not necessarily stop at each station; rather, it picks up evacuees from selected stations to achieve a certain objective function until its capacity is reached. Typically, the demand at stations is larger than the capacity of a train; therefore, trains travel in cycles until all the demand is exhausted.

Each subway line operates independently according to the operational characteristics defined in section 8.3.2; that is, no transfers are allowed between lines and evacuees are assumed to ride the subway from one station and stay until they reach safe destinations. This assumption is made to minimize the service disruptions and chaos associated with emergency evacuation.

The demand assignment process described in section 8.4.2.1 resulted in evacuees being carried by the closest subway line. The output is illustrated in Figure 8.8 where each subway station has a certain demand of evacuees that need to be transported to safe destinations. The Yonge and Bloor subway lines are the major lines that cover large geographical areas within the City of Toronto and pass through the dense core of the city; therefore, the majority of the demand is attracted to these lines. It is not surprising to find that the Yonge line has the highest number of evacuees, it carries around 373,360 evacuees which is almost double the demand assigned to the Bloor line (176,189). This is because of the denser land use around the Yonge line and the transit-oriented development within this area. The Scarborough and Sheppard lines carry 20,006 and 9,692 evacuees, respectively.

Two scenarios that represent two objective functions are evaluated in this implementation; the first objective minimizes the travel time (routing time) for transit vehicles (TT), while the second minimizes the travel time of transit vehicles and the waiting time of evacuees at stops (TT.WT); that is minimizing the total system evacuation time. Each subway line is evaluated based on the following measures: average number of runs per transit vehicle, average in-vehicle travel time, average total in-vehicle travel time, average travel distance, and average waiting time.
Table 8.5 demonstrates the MOEs for each subway line in each scenario. Figure 8.28, Figure 8.29, Figure 8.30, and Figure 8.31 demonstrate the MOE in both scenarios for the Yonge, Bloor, Scarborough, and Sheppard lines, respectively.

Table 8.5 Measures of Effectiveness of Rapid Transit Lines *

<table>
<thead>
<tr>
<th>Line/Mode</th>
<th>Bloor Line</th>
<th>Yonge Line</th>
<th>Scarborough Line</th>
<th>Sheppard Line</th>
<th>Shuttle Buses</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOE</td>
<td>TT</td>
<td>TT.WT</td>
<td>TT</td>
<td>TT.WT</td>
<td>TT</td>
</tr>
<tr>
<td>Average No of Runs**</td>
<td>2.1 (1.2)</td>
<td>2.1 (0.8)</td>
<td>4 (3)</td>
<td>4 (0.6)</td>
<td>1.6 (0.9)</td>
</tr>
<tr>
<td>Average Total In-Vehicle Travel Time (min)</td>
<td>96 (70)</td>
<td>90 (35)</td>
<td>228 (117)</td>
<td>216 (35)</td>
<td>47 (29)</td>
</tr>
<tr>
<td>Average In-Vehicle Travel Time (min)***</td>
<td>42</td>
<td>43</td>
<td>55</td>
<td>56</td>
<td>27</td>
</tr>
<tr>
<td>Average Travel Distance (km)</td>
<td>20.5</td>
<td>21</td>
<td>27.4</td>
<td>27.8</td>
<td>15.17</td>
</tr>
<tr>
<td>Average Waiting Time (min)</td>
<td>72 (63)</td>
<td>40.6 (36)</td>
<td>178 (157)</td>
<td>98 (70)</td>
<td>27 (27)</td>
</tr>
</tbody>
</table>

* Numbers between parentheses represent the standard deviation of the MOE across the number of vehicles.
** Average Number of Runs: The average number of runs each transit vehicle travels to serve all the evacuees.
*** In-vehicle travel time: the total in-vehicle travel time divided by the number of runs per vehicle.
Figure 8.28 MOE of Yonge Line in Both Scenarios
Figure 8.29 MOE of Bloor Line in Both Scenarios
Figure 8.30 MOE of Scarborough Line in Both Scenarios
Figure 8.31 MOE of Sheppard Line in Both Scenarios
Analysis of the results leads to the following conclusions:

- The Yonge line is the busiest line with each transit unit making, on average, four runs to transport evacuees to safe destinations. Including the waiting time in the objective function (as shown in the TT.WT Scenario) has evened out the average in-vehicle travel time. This is clearly shown by the significant drop (70%) in the standard deviation of the total in-vehicle travel time averaged over transit units (see Figure 8.28). Although the average in-vehicle travel time appears to be the same in both scenarios (TT and TT.WT), the pattern across transit units is significantly different. This observation confirms that the workloads assigned to each transit unit/driver have been balanced as clearly shown by the consistent pattern in travel times for the TT.WT scenario in Figure 8.28. Including the waiting time in the objective function has increased the in-vehicle travel time and the total travel distance. The average in-vehicle travel time increased by only 2% and the average travel distance by only 2%, yet the average waiting time of evacuees dropped by 45%. Furthermore, the standard deviation of the average waiting time for evacuees dropped by 55% when including the waiting time in the objective function. This demonstrates that the TT.WT scenario provides a more reliable service for evacuees in emergency situations. However, this is contrary to what people and transit authorities may be inclined to expect, which is that all trains stop at all stations.

- The Bloor line is the second busiest line where each transit unit makes, on average, two runs to transport evacuees to safe destinations. Including the waiting time in the objective function (as shown in the TT.WT Scenario) has evened out the average in-vehicle travel time. This is clearly shown by the significant drop (50%) in the standard deviation of the total in-vehicle travel time averaged over transit units (see Figure 8.29). Although the total in-vehicle travel time appears to be the same in both scenarios (TT and TT.WT), the pattern across transit units is different. Including the waiting time in the objective function has increased the in-vehicle travel time and the total travel distance. However, it is worth noting that the average in-vehicle travel time increased by only 2.5% and the average travel distance by only 2.5%, yet the average waiting time of evacuees dropped by 43%.

- Unlike the Yonge and Bloor lines, the Sheppard and Scarborough lines are not heavily utilized. Nevertheless, the same conclusions are drawn when comparing scenario T.T to
WT.TT by examining Table 8.5, Figure 8.30 and Figure 8.31. For the Scarborough line, including the waiting time caused the average in-vehicle travel time to increase by only 7.4% and the average travel distance by only 9.5%, yet the average waiting time of evacuees dropped by 15%. For the Sheppard line, including the waiting time gave the same average in-vehicle travel time and average travel distance, yet the average waiting time of evacuees dropped by 25%.

### 8.5.2.2 Shuttle Buses

The following results demonstrate the optimal routing and scheduling of shuttle buses in each scenario. The MDTCPD–VRP is modelled and solved using constraint propagation and optimization techniques as discussed in Section 4.4. ILOG Dispatcher and Solver are used to model and solve the problem, respectively (ILOG, 2008). It should noted that the VRP and its extensions (e.g. MDTCPD–VRP) are NP-hard problems; that is, solving small problems to optimality is difficult with reasonable computational effort (Garey et al., 1979). Consequently, in the case of solving large-scale problems such as the evacuation by mass transit of the City of Toronto, it may take days for the algorithm to obtain a reasonable “near-optimal” solution.

The model generates the optimal routing and timetable for each bus as it shuttles between pick-up points and shelters. It should be noted that buses are initially assigned to pick-up points according to their location at the onset of the evacuation as described in section 8.4.3. After the initial pick-up, buses shuttle to the nearest shelter then go back into the system but not necessarily to the same pick-up points. This means that each bus seeks the best pick-up point in order to achieve a certain objective function. Shuttle buses loop between the optimal pick-up points and safe shelters until all the demand is exhausted and finally head back to safe shelters. Due to the unprecedented sheer size of the problem and the huge search space, a feasible initial solution is attainable within a few hours of CPU time; however, the improvement process may take up to a few days of CPU time depending on the objective function being optimized. For example, to minimize the in-vehicle travel time for buses, on a Quad Core Machine with 8 GB of RAM, it took 5 hours to obtain an initial feasible solution and 3 days to improve the solution and completely solve the optimization problem.
Similar to the Rapid Transit scenarios, two scenarios that represent two objective functions are examined, minimizing total travel time (TT) and minimizing total travel time and waiting time (TT.WT). Examination of the results leads to the following conclusions:

- On average, transit shuttle buses make around 6 runs to transport evacuees to safe destinations. Including the waiting time in the objective function (as shown in the TT.WT Scenario) has evened out the average in-vehicle travel time and the average number of runs/vehicle. This is clearly shown by the significant drop (44 vs. 425) in the standard deviation of the total in-vehicle travel time averaged over transit units (see Figure 8.28). The average in-vehicle travel time is found to be 23 min in the TT scenario and 24 min in the TT.WT scenario; although close to each other, the pattern across transit units is significantly different. Including the waiting time in the objective function has increased the in-vehicle travel time and total travel distance. However, it is worth noting that the average total in-vehicle travel time increased by only 7%, yet the average waiting time of evacuees dropped by an order of magnitude (53 vs. 914). Furthermore, the standard deviation of the average waiting time for evacuees dropped by an order of magnitude (43 vs. 1844) when including the waiting time in the objective function. This demonstrates that the TT.WT scenario provides a more reliable service for evacuees in emergency situations. It should be noted that a careful examination of the absolute average of in-vehicle travel times and travel distance for the TT and TT.WT scenarios is required. Because these averages are calculated based on the average number of runs which is dependent on the objective function (TT vs. TT.WT), they already take into account the variability in the load assigned to each vehicle.
No of Runs/ Vehicle

Average In-Vehicle Travel Time

Average Distance Travelled

Figure 8.32 MOE of Shuttle Buses in Both Scenarios
For the sake of conciseness, the detailed routing and scheduling of one bus is shown in Figure 8.33 for scenario TT.WT. As shown in Figure 8.33, the arrival and departure times of the bus at each pick-up point (visit) are extracted and the associated travel time and distance travelled are illustrated. The optimal route is a sequence of nodes starting from a pick-up point to shelter to the next pick-up point. For example, one bus (Vehicle 401) is scheduled to start from the depot (location of the bus at the onset of the evacuation) at time 0, picks up 90 passengers (visit 6075) and travels along the optimal route as shown in Figure 8.33. The vehicle then drops off the evacuees at shelter1 and continues to pick up evacuees located at visit 5826 and drops them off at shelter2. The routing and scheduling plan continues until all vehicles accomplish the assigned tasks. At the end of the evacuation, all vehicles return back to the shelter (terminal) as in the case of bus #401 shown in the last row of Figure 8.33.
<table>
<thead>
<tr>
<th>Visit</th>
<th>Arrival Time (min)</th>
<th>Departure Time (min)</th>
<th>Evacuees on Board</th>
<th>Travel Time (min)</th>
<th>Travel Distance (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit6075</td>
<td>0.4</td>
<td>1.15</td>
<td>90</td>
<td>0.4</td>
<td>328.6</td>
</tr>
<tr>
<td>visit5826</td>
<td>34.3</td>
<td>35.06</td>
<td>90</td>
<td>33.9</td>
<td>2000.3</td>
</tr>
<tr>
<td>visit5761</td>
<td>67.7</td>
<td>68.37</td>
<td>90</td>
<td>33.3</td>
<td>2877.2</td>
</tr>
<tr>
<td>visit4786</td>
<td>103.2</td>
<td>103.90</td>
<td>90</td>
<td>35.5</td>
<td>30412.9</td>
</tr>
<tr>
<td>visit2270</td>
<td>149.1</td>
<td>149.84</td>
<td>90</td>
<td>45.9</td>
<td>41710.0</td>
</tr>
<tr>
<td>visit1268</td>
<td>189.8</td>
<td>190.52</td>
<td>90</td>
<td>40.7</td>
<td>32600.0</td>
</tr>
<tr>
<td>visit1082</td>
<td>197.0</td>
<td>197.68</td>
<td>90</td>
<td>7.2</td>
<td>6006.0</td>
</tr>
</tbody>
</table>

| Route from pickup point to shelter to next pickup point (sequence of nodes) |
|------------------|------------------|------------------|------------------|
| Visit 6075 | Depot | Visit 6075 | 10447 | 13116 | Shelter1 | 10000 | 13021 | 10091 | 13023 | 10095 | Visit 5826 |
| Visit 5826 | Shelter2 | 10441 | 10433 | 11589 | Shelter2 | 10000 | 13021 | 10091 | 13023 | 10095 | Visit 5761 |
| Visit 5761 | Shelter3 | 10433 | 13021 | 10091 | Shelter3 | 10000 | 13021 | 10091 | 13023 | 10095 | Visit 4786 |
| Visit 4786 | Shelter4 | 10336 | 13233 | 10330 | Shelter4 | 10001 | 10018 | 13002 | 10019 | 10020 | Visit 2270 |
| Visit 2270 | Shelter 5 | 11013 | 11012 | 10494 | Shelter5 | 11258 | 11257 | 11265 | 11255 | 13058 | Visit 1268 |
| Visit 1268 | Shelter 6 | 11674 | 11689 | 11690 | Shelter6 | 11241 | 11477 | 11476 | 10242 | 11475 | Visit 1082 |
| Visit 1082 | Terminal | 10236 | 10235 | 11475 | 11476 | 11477 | 11241 | Terminal |

Figure 8.33 Example of Routing and Scheduling of Transit Vehicle
9 Putting the Pieces Together: Towards a Complete Evacuation Modelling and Management Process

This chapter provides a new vision for a comprehensive evacuation planning model. It begins by contrasting the typical four stage planning model with the current state-of-the-practice emergency evacuation planning models. It ends with a description for the presented planning model and how it could be extended into a closed-loop evacuation control system.

Day-to-day travel patterns are typically modelled with conventional urban transportation planning models. These models, in a variety of ways, assess trip generation, trip distribution, mode split and trip assignment either sequentially or concurrently (e.g. combining destination and mode choice for instance). Under emergency evacuation scenarios the behaviour of the transportation system is vastly different from day-to-day travel patterns. Such emergency situations are characterized by sudden sheer non-discretionary demand, vulnerable supply (infrastructure), and poor system performance in the form of longer travel times, chaos, severe congestion, uncertainty, and destination vulnerability, to name a few. Moreover, travellers themselves may act differently in emergency situations compared to their usual daily travel. For instance, evacuees may be more likely to follow directions from officials as to which route to use instead of their habitual routes (Fu and Wilmot, 2004). Also, in daily travel, trip makers decide their trip start times to maximize the utility of travel; however, in the case of an emergency, evacuees might be urged to follow an evacuation schedule and or get directed to safe shelters that are not necessarily their pre-planned destination choices.

Despite the above unique characteristics of travel under emergency situations, a complete set of integrated tools for assessing demand generation, distribution, mode choice, destination choice and route choice was lacking until this research effort is conducted. For instance, a tremendous body of recent literature on evacuation planning, although extremely useful, assumes that demand is known or given, or focuses on one mode of evacuation (predominantly cars) with little attention to multi-modal evacuation using both cars and mass transit (Sbayti and Mahmassani, 2006). Although widely used transportation planning approaches comprehensively cover all aspects of travel starting from the generation of demand to dynamic travel assignment and link
travel times, this is not yet the case in emergency evacuation planning and this motivates the
author to close this gap.

9.1 Evacuation Planning vs. Transportation Planning Models

This section compares the evacuation planning models to the well established transportation
planning models. In a variety of ways, traditional planning models span four main stages: trip
generation, trip distribution, mode split and trip assignment, as shown in Figure 9.1. These
planning models capture reasonably well the typical daily origin-destination patterns; however,
they are not applicable to emergency evacuation modelling due to the vastly different spatio-
temporal travel patterns.

Figure 9.1 State-of-the-Art Four Stage Model

Figure 9.2 illustrates a summary of the five stage evacuation modelling process in a manner
analogous to the well-known four stage transportation planning process. A fifth layer is added to
the process to account for the departure pattern during evacuation, i.e. evacuation schedule. In
emergency situations, the mobilization pattern of evacuees plays a paramount role in the
performance of the system and in the success or failure of the evacuation process. Despite this
importance, mobilization curves are typically assumed as previously discussion. Only in recent
years efforts have emerged that focus on optimizing the mobilization pattern of evacuees so as to
minimize or maximize an objective (Sbayti and Mahmassani, 2006). Trip distribution is typically
assumed on the basis of past emergency events. Recent research has started to address the potential of optimizing evacuee destinations (Chiu et al., 2006; Yuan et al., 2006). Most evacuation modelling studies focus on automobile-based evacuation. Therefore, mode split is rarely modelled or even realistically assumed and is certainly not optimized (TRB, 2008).

**Figure 9.2 State-of-the-Practice Evacuation Planning Models**

**9.2 Evacuation Planning Modelling Process**

Despite traditional transportation planning models and state-of-the-practice evacuation models that have contributed significantly to improving the evacuation process, an integrated evacuation model is only made available through the presented multimodal evacuation planning model. Most evacuation planning models deal with each layer (stage) separately, while the decision-making stages of evacuees are closely interrelated. For example, the departure time of evacuees may influence their destination choice and their destination choice maybe be affected by congestion on the routes to the chosen destination. This is assuming that the mode choice is known *a priori*, i.e. how many evacuees own and/or have access to cars and how many are transit-captives. It is indeed clear that more effort is needed to synergistically integrate some or
all of these decision elements to further improve the efficiency of the evacuation process. The presented approach combines evacuation scheduling (departure curve), destination choice (trip distribution) and route choice (trip assignment) into a single comprehensive solution.

Also, an accurate representation of the spatial distribution of population, by time of day and mode of travel is essential to realistically address major population evacuation. Unlike day-to-day travel patterns, emergency evacuation has a unique non-recurrent demand distribution that depends on the time that emergency strikes and how the population is distributed at that time. The presented approach attempts to carefully assess evacuation demand based on a knowledge of the likely location of people at different times of the day, which is important for the model to produce accurate evacuation management measures (refer to Chapter 7).

Furthermore, automobile evacuation has received the most attention; consequently multimodal evacuation is still largely missing from most emergency evacuation studies. A significant portion of the population in cities like Toronto use public transit particularly within, towards, and out of the downtown core. This portion of the population does not have access to their automobiles during the day, regardless of whether they own one. Utilizing the readily available transit capacity is therefore essential to not only shuttle the transit captives to safety but also to expedite the overall evacuation process and reduce network clearance time by moving people en mass. Therefore, the presented approach explicitly optimizes mass transit-based evacuation (refer to Chapter 5).

In summary, the presented approach (see Figure 9.3) considers the following elements to be essential in order to realistically plan for emergency evacuation:

- Accurate estimation of the spatial and temporal distribution of population (*trip generation*).
- Accurate identification of available modes and population captive to certain modes (*mode split*).
- Integrated framework that accounts for various evacuation strategies such as evacuee scheduling, destination choice and route choice simultaneously (*departure curve, trip distribution, and trip assignment*)..
- Multimodal evacuation strategies that synergize the effect of multiple modes.
- Robust and extendible optimization and solution algorithms that can tackle such multi-dimensional non deterministic problems.

9.3 Open vs. Closed Loop Evacuation Management: Future Directions

Although planning for emergency evacuation is paramount to ensure public safety from man-made and natural disasters, prediction of evacuation scenarios is challenging due to the highly dynamic and stochastic nature of emergency situations. No matter how well an evacuation plan is scrutinized or optimized, actual evacuation patterns will almost definitely deviate from the plan. Therefore, actual system behaviour needs to be monitored and managed in real-time and the evacuation plan will need to be re-optimized in accordance with the measured state of the system in a rolling-horizon closed-loop control fashion. Very few studies have addressed real-time traffic management or emergency evacuation in real-time (Liu et al., 2007). Numerous factors can contribute to considerable deviation of evacuation evolution from offline optimized evacuation plans. For instance, potential chaotic behaviour of evacuees, transportation network vulnerability to any disruption (e.g. incidents), and potential noncompliance of evacuees to announced plans can all cause such deviations. Therefore, there is a strong need to close the evacuation control loop by feeding back actual system conditions measured in real-time into online evacuation optimization engines. The closed loop evacuation control system can guide and drive the transportation network towards optimal performance despite unexpected disturbances or deviations from original plans.

The envisioned closed loop evacuation control system is illustrated schematically in Figure 9.4. The system starts by disseminating an offline optimized plan from a system such as OSTE. An online monitoring system reports real-time information about the status of evacuees and current road network conditions. The monitoring system also updates the status of the transit system in the form of the current location of transit vehicles and the number of passengers already evacuated. Numerous ITS technologies can be utilized in the process including, for instance, Global Positioning Systems (GPS), mobile devices, automatic passenger counters, etc. Once the real-time status of the system is updated, a new time-dependent origin-destination matrix can be estimated on the basis of measured flows in the network (Kattan and Abdulhai, 2006). The estimated demand matrix is then input to the two optimization engines (OSTE and MDTCPD-VRP) to generate new plans for the next horizon.
In large-scale applications such as evacuating a large city like Toronto, the real-time implementation of such a closed loop approach can be challenging and will require considerable computer processing power or even parallel processing as described in Chapters 3 and 8.
Figure 9.3 Proposed Emergency Evacuation Planning Model
Figure 9.4 Closed Loop Evacuation Control System
Conclusions

This thesis focuses on the development of emergency evacuation plans, in particular synergizing the effect of multiple evacuation strategies and modes in emergency evacuation situations. The study considers multiple dimensions of the emergency evacuation problem: evacuation scheduling, destination choice (optimization), route choice, multiple modes, and evacuation demand estimation. Such dimensions are either simplified (e.g. demand estimation) or ignored (e.g. multiple modes) in existing emergency evacuation plans. The developed framework is capable of modelling with-notice and no-notice evacuation events. Furthermore, it represents a coherent, integrated, multimodal framework which deals with the issue of transporting transit-captive and car-less populations in emergency situations. It addresses many limitations of existing emergency evacuation planning models by exploiting emerging technology already established in the area of automobile and public transit evacuation. Such emerging approaches include the dynamic traffic assignment of transportation networks, the routing and scheduling of mass transit vehicles, parallel distributed Genetic Algorithms, Constraint Programming and Local search methods, High Performance Computing (HPC) and the adoption of Geographical Information Systems (GIS) to capture the spatial and temporal distribution of evacuees. It provides emergency evacuation planners as well as emergency management operators with a platform for experimenting with ‘what if’ evacuation scenarios in medium and large-scale applications.

Emergency evacuation is the collective movement of people using multiple modes of transport from a hazard area to safe shelters via specific routes. This collective movement creates travel demand and travel patterns that are drastically different from everyday regular traffic and transit volumes. Therefore, the transportation network performance can severely deteriorate when such drastic changes in demand patterns occur as a result of man-made or natural disasters. The inevitable loss of capacity due to the disaster and associated incidents can further complicate the matter. Therefore, when a disaster or hazardous event occurs, the goal is to coordinate, control, and possibly optimize the utilization of the existing transportation network capacity. The proposed framework therefore is geared towards investigating the effect of synergizing and...
optimizing multiple evacuation strategies (e.g. evacuation scheduling, destination choice, and route choice) on the performance of the transportation network. The framework also endeavours to optimize the use of available transit capacity to expedite the evacuation process and to provide a viable option for a car-less population and transit-captives. An operational prototype of the proposed modelling framework has been developed and tested. The purpose of the prototype is to demonstrate the feasibility and applicability of the proposed framework – refer to Chapter 6. A large-scale implementation for the evacuation of the City of Toronto is conducted as a case study – refer to Chapter 8.

The platform is based on two optimization modules: an Optimal Spatio-Temporal Evacuation (OSTE) module for optimizing the automobile evacuation and a Multiple Depot Time Constrained Pick-up and Delivery Vehicle Routing module (MDTCPD-VRP) to optimize the transit evacuation. The system incorporates an estimation of evacuation demand using a regional demand survey (e.g. TTS) and a representation of traffic analysis zones. OSTE plans are then generated for vehicular demand using genetic algorithms in a global optimization technique and dynamic traffic assignment tool. The routing and scheduling of transit vehicles is then solved using constraint programming. The automobile OSTE plan and the transit optimal routing and scheduling plan are finally combined for dissemination to evacuees as a multimodal evacuation plan.

The proposed approach, based on a mesoscopic representation, is capable of providing various types of information regarding system performance to traffic and transit operators. On the traffic side, it can provide detailed information about the departure times of vehicles, optimal destinations, routes to destinations, and travel times to destinations. On the transit side, the model provides detailed output at the stop level (e.g. stop ID, bus arrival times, departure times, alighting/boarding passengers), at the vehicle level (e.g. vehicle ID, total number of passengers transported by that vehicle, travel distance and travel time), and the route level (e.g. first stop ID, sequence of nodes that construct a route, pick-up points along that route and destination stop ID). The system also captures the dynamic interaction between demand and supply by examining the loading and evacuation curves with the evolution of the evacuation plan.
The analysis of the prototype implementation (refer to Chapter 6) leads to the following conclusions:

- Considering only the travel time of evacuees in emergency evacuation results in excessive waiting time of evacuees especially in the case of no-notice evacuation. Also, minimizing the waiting time will ultimately lead to simultaneous evacuation which results in early gridlock and severe congestion. A good compromise is to account for both evacuee waiting time at origins and travel times through the transportation network, using the proposed multi-objective framework.

- Transit systems can substantially improve the evacuation process due to the readily available capacity of transit services. Moreover, due to the typical limitations of transit system operators in providing sufficient capacity to cope with the surge in demand during evacuation, it becomes crucial to optimize the usage of the transit fleet in evacuation situations.

The analysis of the evacuation demand estimation model (refer to Chapter 7) leads to the following conclusions:

- An accurate demand estimating model has the potential of identifying the evacuation demand by *mode* over *time* and *space*. These attributes are essential in order to realistically plan for regional evacuation scenarios.

- The proposed demand estimation model demonstrated that the noon period is the time when the maximum number of people is present in the City of Toronto, totalling 2.56 million people.

- The demand estimation model is designed to identify the number of people who are captive to transit modes during evacuation, which is found to be equivalent to 1.34 million evacuation trips, with the rest of the trips (1.22 million) evacuating in automobiles as their primary mode.

- The number of people in Toronto peaks at 108% of the City’s population (residents at 4:00 AM). This increase is attributed to the high concentration of economic activity in the City of Toronto and particularly the business and financial district.
- It is interesting to find a wide peak activity period in the City of Toronto. This peak in trips starts at 7:00 AM in the morning and ends at 6:00 PM in the afternoon.

The analysis of the large-scale application to the City of Toronto (refer to Chapter 8) leads to the following conclusions:

- Similar to the conclusion drawn from prototype testing, a good compromise when modelling emergency evacuation is to account for both the waiting time of evacuees at the origins and their travel times through the transportation network, using the proposed multi-objective framework. By solving the multi-objective optimization problem to evacuate the City of Toronto, it is found that OSTE clears the network four times faster than the do-nothing strategy (SE), evacuees travel eight times less than in the do-nothing strategy, and finally evacuees stop eleven times less than in the do-nothing strategy; in this case, the automobile-based evacuation time averaged over 1.22 million people is about 2 hours and the NCT is about 8 hours.

- Transit systems (Rapid Transit and Buses) can substantially improve the evacuation process due to the ready availability of capacity in transit services. The Toronto Transit Commission (TTC) fleet is capable of evacuating the transit-dependent population (1.34 million) within two hours on average. The 4 subway lines of the City of Toronto carry around 0.62 million people and can evacuate these people in 154 min on average; the busiest line is the Yonge Line. The available TTC shuttle buses (1320 vehicles) can evacuate the remainder of the transit-dependent population (0.72 million) in about 84 minutes on average.

10.1 Summary

Chapter 1 presents the motivation behind this research effort alongside the research objectives. It then discusses the challenges that faced early emergency management systems while highlighting the major limitations of existing approaches. After reviewing the literature in Chapter 2, the author summarizes the major challenges and gaps in the existing literature in section 2.8. Lack of efficient multimodal evacuation was identified as a major gap in the literature. Chapters 3 and 4 endeavour to bridge this gap.

Chapter 3 discusses the development of a new modelling framework for automobile evacuation, namely the Optimal Spatio-Temporal Evacuation – OSTE. The proposed modelling framework
is sensitive to changes in time-dependent traffic routing through the transportation network (supply). It also models evacuee departure times and destination choices (demand). Through an explicit representation of the loading and evacuation curves, it captures the interaction between traffic network performance (e.g. congestion) and the scheduling and destination choices of evacuees. Through the multiple objective function design of OSTE, it is capable of modelling with-notice and no-notice evacuation scenarios. OSTE considers the multiple dimensions of the evacuation problem: evacuation scheduling, destination choice, and route choice. OSTE is founded upon a simulation-optimization method that results in the optimal scheduling vector and destination choice in a DTA environment. The output of OSTE is guidance information for evacuees on when to evacuate (departure time) and where to go (optimal destination) as well as how to get there (optimal route).

Chapter 4 discusses the development of a new variant of the vehicle routing problem (VRP) to model the routing and scheduling of transit vehicles during emergency evacuation. The model extends the well-established VRP in the following dimensions: Multiple Depots, Time Constrained, Pick-up and Delivery, hence MDTCPD-VRP. The model is capable of modelling the routing and scheduling of transit vehicles such as the movement of shuttle buses between shelters and pick-up points as well as the scheduling of dedicated rights of way for services such as rapid transit vehicles (e.g. subway). The MDTCPD-VRP is founded upon a network structure module to model the transportation network and a constraint programming and local search module to solve the constrained-optimization problem. The output of MDTCPD-VRP is a detailed routing plan and timetable for each transit vehicle to evacuate.

Chapter 5 presents the overall framework that integrates the automobile evacuation module (OSTE) discussed in Chapter 3 and the transit evacuation module (MDTCPD-VRP) discussed in Chapter 4.

Chapter 6 demonstrates the feasibility and applicability of the proposed framework through a prototype implementation on a medium-size evacuation scenario in which a portion of Toronto’s waterfront is evacuated. Although the evacuation demand was estimated by aggregate regional travel demand survey data, the prototype implementation is intended to focus on model development and testing irrespective of the source of the evacuation demand.
Chapter 7 presents a demand estimation model from a regional travel demand survey (Transportation Tomorrow Survey, TTS). The model is capable of providing not only the evacuation demand per traffic analysis zone (TAZ) but also the spatial and temporal distribution of demand by mode of travel. This model provides an accurate description of the spatial distribution of population by time of day and mode of travel that is essential to realistically plan for large-scale population evacuation.

Chapter 8 documents efforts to make the OSTE conceptual framework presented in Chapter 3, the MDTCPD-VRP presented in Chapter 4 and the demand estimation model described in Chapter 7 operational. It presents a large-scale application to an evacuation of the entire City of Toronto in the case of an emergency situation. The total number of people present in the City of Toronto in the worst case scenario was more than 2.56 million people, resulting in an evacuation scenario that is more challenging than the evacuations caused by Hurricanes Katrina and Rita. Synergizing numerous emerging approaches (e.g. parallel distributed genetic algorithm, dynamic traffic assignment, mesoscopic simulation, constraint programming and local search, ArcGIS© Visual Basic Application) has led to the development of an efficient emergency framework that can clear the City in 8 hours with an average total system evacuation time of 2 hours.

Chapter 9 envisions a rolling-horizon closed-loop emergency evacuation control system. The system requires an online monitoring system to report real-time information about the status of evacuees and current network conditions. Then, a new time-dependent origin-destination matrix can be estimated on the basis of measured flows in the network that will form the input to the two evacuation modules (OSTE and MDTCPD-VRP) to generate new plans for the next horizon.

### 10.2 Contributions

The research presented in this thesis represents a significant contribution in five main areas of research. First, a major in-depth demand data analysis effort has been undertaken which has provided groundbreaking information on essential elements in emergency evacuation. This was the first attempt to estimate the evacuation demand from a regional travel survey in the absence of historical data or post-surveys from previous disasters. The purpose of the demand estimation method was to examine the spatial and temporal distribution of the population and their mode of travel. Capturing the time, space, and mode dimensions of the evacuation demand is, to the
author’s knowledge, unprecedented. Some simple evacuation demand models and other behavioural methodologies have been typically used in the literature (Alsnih, 2004; Fu, 2004; Mei, 2002; Southworth, 1990). These demand estimation models were thwarted by the lack of identification of evacuee transportation modes. The demand estimation model presented provides rich information that is essential to multimodal emergency evacuation. The presentation of the demand estimation model was made available through the integration of GIS with the transportation network (road and transit networks) and the Traffic Analysis Zones in which the GIS spatial-capabilities alongside with the VBA ArcGIS tools have been utilized.

The second major contribution is the development of an Optimal Spatio-Temporal Evacuation model that synergizes evacuation scheduling, destination choice and dynamic route choice. The model generates optimal guidance information in the form of departure times, destinations and routes. The combination of evacuation scheduling and destination choice resulted in significant system performance improvements when compared to the effect of each strategy independently. When compared to the do-nothing case, an order of magnitude improvement in system performance is achieved. The model is based solely on the interaction between a dynamic traffic simulation model and a global optimization heuristic method (Genetic Algorithm). A Genetic Algorithm is guaranteed, if designed properly, to produce a near optimal solution compared to other simple deterministic line search techniques. As the problem size or application size grows, the implementation of a GA becomes infeasible due to the long execution times. Therefore, a parallel distributed genetic algorithm (PDGA) is designed to tackle large-scale emergency evacuation problems. The design and implementation of the PDGA is based on IBM WebSphere technology that is intended to be generic to assure the transferability of the optimization-simulation technique. The effectiveness of GA parallelization and distribution is examined. The parallel GA proved to provide higher quality solutions and faster convergence. The distributed GA gave a linear speedup when compared to the single GA.

The third major contribution is the development of a new variant of the well-established vehicle routing problem (VRP) to model transit vehicle routing and scheduling during emergency evacuation. The new variant includes: (i) Multiple Depots to account for dispersed transit vehicles in the transportation network; (ii) Time Constraints to account for the desired evacuation time window; and (iii) Pick-up and Delivery locations for evacuees to allow evacuees
to be picked up from dispersed stops to avoid excessive walk distances. The Multi-Depot Time-Constrained Pick-up and Delivery VRP (MDTCPD-VRP) is known to be an NP-hard problem; therefore, a good feasible solution within a reasonable amount of computation time is acceptable. Due to the numerous constraints incorporated in the problem, constraint programming and local search techniques are found to be viable tools to tackle this problem. The model requires an accurate representation of the transit network. The model is generically designed to account for additional constraints and multiple objective functions to model more complex routing and scheduling problems as in the case of emergency evacuation. The disaggregate output of the model enables the analyst to examine a myriad set of measures that are essential to the performance of emergency evacuation plans. Also, the output format is compatible with GIS network representations so as to geographically display the routing plan of transit buses while shuttling evacuees from pick-up points to safe destinations (shelters).

The fourth major contribution is the explicit modelling of different time structures in emergency evacuation with the evolution of the evacuation plan. This concerns the modelling of the dynamic interaction between the loading and evacuation curves. This leads to the incorporation of multiple objective functions while planning for emergency evacuation. Objectives include minimization of travel time, minimization of waiting time, minimization of waiting and travel times, minimization of fleet cost. Achieving a compromise between multiple objectives is challenging, in which case a set of Pareto solutions may result.

The fifth contribution is the amalgamation of the former discussed approaches and modelling techniques in one framework that is geared towards producing efficient and realistic multimodal emergency evacuation plans. The framework is tested for feasibility in a prototype implementation and the results were found very encouraging. The model is then tested on a large-scale application which covered the evacuation of the entire City of Toronto. To the author’s knowledge, this is the largest emergency modelling effort investigated in the evacuation literature. The smooth migration from the prototype implementation to the large-scale application demonstrates the transferability and scalability of the model.
10.3 Future Research

As with any research effort, while this thesis is meant to answer some research questions about emergency evacuation, it triggers even more. Abundant opportunities for improving and extending state-of-the-art emergency evacuation planning emerge from the work presented in this thesis. The prospect of extending this research is charted in the following sections:

1. **Spatially Phased Evacuation**: One can also envision the administration of evacuation scheduling on the basis of priority and exposure to risk which can be described as phased/zoned evacuation. Phased evacuation could be modelled by imposing variable risk factors in the objective function based on the proximity of the hazard to certain zones. Although phased evacuation could plausibly be implemented, user equity and compliance may remain as issues.

2. **Spatio-Temporally Staged Evacuation**: One potential improvement is to synergize the effect of evacuation scheduling and evacuation phasing in one platform. While evacuation scheduling is concerned with determining the percentages of evacuees that should evacuate their zones at each evacuation interval, evacuation phasing is concerned with determining the sequence of zones to evacuate. The ultimate goal is to determine the optimal zone sequencing (phased evacuation) and for each evacuation zone an optimal schedule is advised (scheduled evacuation). Although an appealing synergy, the simultaneous optimization of evacuation scheduling and phasing is a demanding problem that requires ample computational resources, especially for large-scale evacuation problems. However, for use as a planning model, computation time might be alleviated by utilizing high performance computing clusters as well as macroscopic simulation tools.

3. **Evacuee Behaviour**: As presented in the previous two lines of potential research, further investigation is required to assess the likely compliance of evacuees to the recommendations provided. Behaviour can be potentially assessed by analyzing evacuee behaviour during past real evacuation scenarios, via stated preference surveys or through mixed reality simulations by exposing actual subjects (people) to simulated evacuations in a virtual environment and watching their behaviour.
4. **Information Communication/Dissemination:** To ubiquitously disseminate evacuation recommendations (scheduling, phasing, destination, etc.) to the public, a communication mechanism must be in place that is capable of considering policy-oriented issues such as coordination across jurisdictions, public awareness, evacuation type, education/involvement of the transportation engineering community, inter-agency communication and data sharing (Wolshon *et al.*, 2005). Although this research effort makes no attempt to address all the policy-oriented issues, it provides a set of tools and modelling procedures that can help decision-makers and planners address most of these issues.

5. **Multimodal Evacuation:** Despite the fact that the thesis is the first to address multimodal evacuation by using public transit shuttle buses, rapid transit vehicles and automobiles in a single platform, there are still other modes of evacuation such as walking and cycling that could be integrated into the framework. Although, the concept of “equilibrium mode-split” in emergency evacuation has been briefly discussed in this research, more analysis and modelling is needed to study the optimal mode shift to fully utilize the available transportation network.

6. **Iterative Traffic and Transit Assignment:** The current platform does not loop back from the transit assignment component to the traffic assignment component, a potentially fruitful step to explore in future research. In this research, while extracting the travel times from the traffic model to form the input to the transit model, the most congested travel times are used as a worst case for buses while travelling through the network. Although it may overestimate the travel times for buses, it should compensate for the uncertainty of travel times for such heavily utilized transit vehicles; i.e. if the process errs it does so on the conservative side. This may be acceptable at this level of analysis; however, when examining the optimal mode shift as discussed previously, this interaction should be properly designed.

7. **User Equity and Compliance:** While the findings of this research are very encouraging in terms of system performance, one potentially critical dimension that needs further research is user equity, which may also have a direct bearing on the compliance of evacuees. Since evacuation scheduling, particularly for automobile-based evacuation, implies holding back some evacuees and releasing others, the question arises as to who to
let go first. If evacuees perceive that it is risky or unfair that they are being held back, they may simply ignore the guidance being provided which defies the purpose of scheduling. User equity and compliance has direct connection to the method of traffic assignment. Whether user optimal or system optimal traffic assignment methods should be used in evacuation planning is still an open research question. System optimal assignment is an appealing system-wide target; however, it implies that some evacuees will experience longer travel times than their best attainable travel time. It seems unrealistic to expect evacuees to offer such sacrifice during evacuation in order for the system to be optimized, not to mention the tort liability of the system operator. On the other hand, in user optimal traffic assignment with no guidance, travellers follow the routes that minimize their own (explicit or perceived) travel times based on their prior experience which implies past iterative learning of route travel times in emergency situations. However, such iterative learning and assignment processes require a long time to evolve which is not the case in emergency evacuation if explicit guidance is not provided. Therefore, in this research, it is more plausible to assume that evacuees will start their evacuation journey on experience-based user optimal routes. With the progress of the evacuation journey and as they encounter congestion, evacuees will continue to adjust their route selection and destination choice based on their knowledge of the network and their observation of congestion. Having stated this, OSTE might accommodate any traffic assignment method, a potential area for future investigation.

8. **Network Vulnerability and Reserve Capacity**: Incidents and the transportation-related consequence of link failures including congestion and overall system deterioration are unavoidable in emergencies. Therefore, the vulnerability of the transportation network should be assessed. Such assessment would provide valuable input to emergency management response systems, so that they can identify the most vulnerable locations in the system and then dedicate their personnel and emergency management vehicles in the proximity of these locations. Also the concept of reserve capacity, which define critical areas in the network that require capacity expansions to maximize the network spare capacity (Yang and Bell, 1998), can be potentially explored to maximize the use of the transportation networks especially in the case of emergencies.
9. **Control and Supply Management:** Although the main strategies adapted in this research focus on the demand side of the problem (i.e. scheduling, destination choice and route choice), the scope could be extended to include supply management and optimal control strategies such as contra-flow, traffic signal control, ramp metering, and variable message signs. Such emerging strategies have the potential to expedite the evacuation process and helping in implementing the presented plan.

10. **Closed-Loop Evacuation System:** In recognition of the essence of a comprehensive evacuation framework, the closed-loop evacuation proposition in Chapter 9 is presented qualitatively. It is also important to emphasize here that such a holistic framework requires a multidisciplinary research effort that masters the following areas of research: communications, control, behaviour, optimization, and decision logic.

11. **Rolling Horizon Optimization for Potential Real-Time Implementation:** To alleviate the computation time requirements for large-scale emergency evacuation, a feedback rolling horizon genetic optimization could be developed. Using a current optimized evacuation plan, one could easily perturb the solution to reflect current/projected network conditions (or better off predict/estimate current network conditions), then through a rolling horizon technique use the latest optimization results to re-optimize the evacuation again in real-time. This technique would require far fewer resources than the base optimization owing to its starting point.

12. **Stochastic Chance-Constrained and Percentile Optimization for Uncertainty in Evacuation Planning:** A major difficulty of the presented deterministic formulation for the optimization of multimodal evacuation plan is that (optimal) decisions have to be taken prior to the observation of the system state. Therefore, in case of emergency evacuation, one can hardly identify any decision/plan which would definitely exclude later constraint violation caused by unexpected behaviour and chaos of evacuees. Chance-constrained and percentile optimization is a method to balance such constraint violation by compensating decisions/plans taken in the form of a probability assigned to the optimization output as opposed to a specified fixed window. For example, if the objective is to minimize the total evacuation time (ET), then the problem can be formulated as \( \text{Minimize ET, subject to } \Pr(F(X) \leq ET) \geq \eta \), where \( X \) is the decision variables and \( F(X) \) is the function that is represented herein by the simulation model.
output, and $\eta$ is the probability (or confidence level) of achieving such minimum evacuation time. This potential profitable extension of the current research would also help cities and officials specifying reasonable target for evacuating large cities, for example, in the presented large-scale evacuation of the City of Toronto, it can be stated that the goal is to minimize the total evacuation time and it is found that the probability of evacuating the city within 8 hours is greater than 95%. Though, numerous factors can affect the specified probability, such as level of awareness of evacuees, planning methods, consistency of data, etc.

13. **Public Awareness and Evacuation Drills:** Building on the lessons learned from major disasters, holding public training sessions and information dissemination in neighbourhoods and districts have strong potential to increase the familiarity of residents with the process of emergency evacuation and to raise their level of awareness in this area. Methods for conducting evacuation drills and increasing public awareness and preparedness in case of emergencies are important research subjects. Exploring these methods has the potential to not only increase the level of trust in public officials and organizations but can also increase the level of confidence among the population and improve their capabilities in mitigation, preparedness, response and recovery plans.

14. **Accelerated Subway Operations:** In case of emergency evacuation, trains can skip some stops/stations; therefore, the trade-off between the area coverage and travel speed on subway lines does not exist. In such case the only way to increase the line speed is to introduce the so called *accelerated operations*. Accelerated operations can include skip-stop operations, zonal operations, and express/local operations. Although the scheduling output of the subway lines can be approximated to the skip-stop operations, the scheduling algorithm doe not explicitly optimize the subway method of operation in case of emergency; a potentially plausible step to explore in future research

15. **Bus Shuttling on Regular Transit Routes:** One potential area to extend the bus shuttling framework in emergency evacuation is to route the fleet of buses on regular bus routes. This method of operation will minimize the chaos to both drivers and travellers. In which case drivers continue on their predefined routes without deviating for any pickup point, but their destination should be shifted to the nearest exit at the boundary of the hazard area. For evacuees, they will be directed to the nearest bus stop according to their
familiarity with the existing bus routes; in such case, the walking distance of evacuees is expected to increase as well as their waiting time. Therefore, a compromise is needed to evaluate the inconvenience incurred to evacuees and the ease of operation to the transit service operator.
REFERENCES


