1. Report No. SWUTC/07/473700-00077-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Commuter Rail Circulator Network Design Problem: Formulation, Solution Methods, and Applications		5. Report Date August 2007
		6. Performing Organization Code
7. Author(s) Nicholas E. Lownes and Randy B. Machemehl		8. Performing Organization Report No. 473700-00077-1
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTRS95-G-0006
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135		13. Type of Report and Period Covered
		14. Sponsoring Agency Code
15. Supplementary Notes Supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program		

16. Abstract

Commuter rail is increasingly popular as a means to introduce rail transportation to metropolitan transportation systems. The long-term benefits of commuter rail include the addition of capacity to the transportation system, providing a quality commute alternative, and shifting land use toward transit-oriented development patterns. The success of a commuter rail system depends upon cultivating a ridership base upon which to expand and improve the system. Cultivating this ridership is dependent upon offering a quality transportation option to commuters. Characteristics of commuter rail systems in the United States present challenges to offering quality service that must be overcome. Commuter rail has been implemented only on existing rail right-of-way (ROW) and infrastructure (depending upon condition) in the United States. Existing rail ROW does not often coincide with current commercial and residential demand centers and necessitates the use of a circulator system to expand the service boundary of commuter rail to reach these demand centers. The commuter rail circulator network design problem (CRCNDP) addresses a particular aspect of the commuter rail trip, seeking to improve the performance of the entire system through accurately modeling the portion of the trip from rail station to the final destination. This final leg includes both the trip on the circulator vehicle and the walking trip from the circulator stop to the final destination. This report seeks to provide an innovative mathematical programming formulation and solution methodology for the CRCNDP and apply this method to a case study.

17. Key Words Public Transportation, Commuter Rail, Route Design, Transit, Accessibility	18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161		
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 145	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

The Commuter Rail Circulator Network Design Problem: Formulation, Solution Methods, and Applications

by

Nicholas E. Lownes
Randy B. Machemehl

Research Report SWUTC/07/473700-00077-1

Southwest Region University Transportation Center
Center for Transportation Research
University of Texas at Austin
Austin, Texas

August 2007

ACKNOWLEDGEMENTS

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

Commuter rail is increasingly popular as a means to introduce rail transportation to metropolitan transportation systems. The long-term benefits of commuter rail include the addition of capacity to the transportation system, providing a quality commute alternative, and shifting land use toward transit-oriented development patterns. The success of a commuter rail system depends upon cultivating a ridership base upon which to expand and improve the system. Cultivating this ridership is dependent upon offering a quality transportation option to commuters. Characteristics of commuter rail systems in the United States present challenges to offering quality service that must be overcome. Commuter rail has been implemented only on existing rail right-of-way (ROW) and infrastructure (depending upon condition) in the United States. Existing rail ROW does not often coincide with current commercial and residential demand centers and necessitates the use of a circulator system to expand the service boundary of commuter rail to reach these demand centers. The commuter rail circulator network design problem (CRCNDP) addresses a particular aspect of the commuter rail trip, seeking to improve the performance of the entire system through accurately modeling the portion of the trip from rail station to the final destination. This final leg includes both the trip on the circulator vehicle and the walking trip from the circulator stop to the final destination. This report seeks to provide an innovative mathematical programming formulation and solution methodology for the CRCNDP and apply this method to a case study.

EXECUTIVE SUMMARY

The United States has experienced tremendous population and economic growth over the past half-century. Since the 1950's the population has more than doubled and our economic productivity has allowed the ever-increasing number of Americans to enjoy a standard of living unparalleled in human history. This growth trend, based on U.S. Census projections, is expected to continue through the middle part of the 21st century. Millions of Americans will increasingly choose to live in metropolitan areas as these areas are increasingly providing more economic incentive.

Our nation's urban transportation systems are already crowded. This is not expected to improve with the current demand for urban transportation far exceeding what can be provided strictly in terms of increased highway capacity. The current state of congestion in metropolitan areas necessitates that metropolitan planners and engineers consider both roadway capacity additions as well as providing alternative means of travel within the metropolitan region. A significant portion of the travel demand and congestion in any metropolitan region occurs during the peak commuting hours. While many large metropolitan areas experience peak periods that spread throughout the day, smaller areas experiencing significant population growth have large spikes in demand and congestion during the morning and afternoon commute periods. The need to provide commuters an alternative means of travel during these heavily congested periods is an opportunity to introduce commuter rail into the transportation system.

It is important to note that congestion relief should not be the only goal of a commuter rail system. In the short term, commuter rail provides a safe, reliable, high-quality option for commuters to travel to work. In the long term, commuter rail provides a metropolitan area the means to more effectively manage its growth and character. By establishing commuter rail, the possibility of future rail and transit modes being implemented improves; future modes that collectively provide relief for a potentially ever-increasingly congested system.

Though the potential benefits of commuter rail are great, the best method to implement commuter rail is far from decided. Rail infrastructure is ubiquitous in most urban areas in the United States as access to regional and national transportation has always been a vital component to the growth and success of an area and has historically been fundamental to the growth of metropolitan areas. Deregulation of the rail industry led to many sections of track being abandoned or sold in an effort to optimize the freight rail industry by shedding unproductive infrastructure. The fate of this right-of-way (ROW) varies. The benefits of such ROW and infrastructure is obvious; the upfront capital costs of implementing a rail system in the area are significantly reduced by eliminating much of the ROW acquisition costs, reducing much of the environmental red tape (as the new construction footprint is minimal), and possibly reducing the cost of construction through the use of existing infrastructure, depending upon its condition.

If the majority of commuter rail users do not live or work within walking distance of existing rail or proposed rail stations, some means of accessing commuter rail and final destinations are needed if commuter rail is to be considered a viable option for commuters. The means of access are circulator systems, be they fixed guideway or rubber tire. Circulator systems are a means to provide access to commuter rail, and provided that commuter rail compares favorably with other modes of travel, represent a

way in the short term to help cultivate and retain ridership until development near the rail stations catches up to enhance the stability of the system and its ridership.

There is no silver bullet for making a public transportation system work well. A public transportation system can flourish or perish with or without a robust circulator system. However, a public transportation system will certainly fail if it is unable to attract passengers. The commuter rail circulator network design problem (CRCNDP) is one tool to enable transit authorities to be able to better cultivate and retain public transportation passengers.

The CRCNDP is seeking to better model the accessibility of public transportation, and this improved accessibility is intended to have a positive impact on commuter rail ridership. There is difficulty in accounting for all of the variables that impact any traveler's decision to ride or not to ride public transportation. There is even greater difficulty in gathering data to calibrate and validate any model intending to estimate the relationship. An ideal CRCNDP formulation would include a model that accounts for the relationship between ridership and accessibility and is certainly a worthy extension of this work.

Another obvious extension of this work is the accommodation of multiple routes directly in the formulation and solution methods. The current methods allow for multiple routes only by choosing, for example, the three "best" unique routes that the solution procedure produces. This method has several flaws, not the least of which is the lack of accounting for the impacts on unserved demand that a second or third unique route would present. Configuring the algorithms to account for multiple routes require significant effort, as the inclusion of multiple routes increase the complexity, performance and computational resource requirements of any solution method.

The CRCNDP is a relevant problem for today's public transportation authorities and analysts. American cities have not developed in the dense, transit-friendly manner of many other cities throughout the world where transit flourishes. Because of American prosperity and reliance upon the automobile, a development pattern has been propagated over the past 50 years that is not ideally suited for public transportation.

As rail is viewed as the most desirable of public transportation options, it is rail that is looked to as an expedient means of increasing the transit capacity of a region. Commuter rail service in these regions is often chosen because service during peak hours will provide the largest number of potential passengers in a cost-effective manner. Commuter rail is implemented on existing rail right-of-way and infrastructure; improving the cost effectiveness yet hurting the overall accessibility of the system. It is in improving access to commuter rail that the needs for circulators and the CRCNDP are most evident. Conditions are ripe for improving public transportation in America, and the investment in public transportation made today needs to be an intelligent one; developing an accessible, sustainable system that will perform well and achieve long-term goals.

A future direction for CRCNDP work would be the development of an interface that would allow the CRCNDP to be solved without the aid of an expert researcher. Such a tool would have great utility as a GIS tool for which an analyst could select a particular network surrounding a station, identify candidate stop locations and TAZ centroids, and the CRCNDP tool would provide an optimal (or very good) route design.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The United States has experienced tremendous population and economic growth over the past half-century. Since the 1950's the population has more than doubled and our economic productivity has allowed the ever-increasing number of Americans to enjoy a standard of living unparalleled in human history. This growth trend, based on U.S. Census projections is expected to continue through the middle part of the 21st century resulting in a population of approximately 420 million American citizens (U.S. Census Bureau 2004). These 420 million Americans will increasingly choose to live in metropolitan areas as these areas are increasingly providing more economic incentive.

Our nation's urban transportation systems are already crowded. The oft-cited TTI Mobility Report 2005 (Schrank and Lomax 2005) estimates that urban congestion cost Americans approximately $63 billion in 2003 alone. This is not expected to improve with the current demand for urban transportation far exceeding what can be provided strictly in terms of increased highway capacity. Restrictions in state and federal budgets, environmental and political factors, and the lack of available right-of-way are often defining the ceiling on the capacity that current highways can provide under current operating conditions.

Metropolitan areas are now looking at alternative means to provide the most reliable, safe, and efficient transportation systems. Many areas are evaluating and implementing ITS technologies to provide reliable, real-time information to aid drivers in better route choice and emergency service personnel in better incident management. ITS technologies are also being developed and tested that could allow vehicles to travel at higher speeds with smaller headways as part of the USDOT Vehicle Infrastructure Integration (VII) initiative. Reducing the space between vehicles traveling at highway speeds through improved vehicle-to-vehicle and vehicle-to-infrastructure communication would significantly increase the capacity of the existing lane-miles of the system. However, these technologies are in their relative infancy and many social, political, and liability issues will need to be addressed before widespread deployment of the most significant capacity-increasing technologies.

Metropolitan areas are increasingly looking to the past, to a time when rail dominated and shaped the transportation landscape, to provide an alternative means of traveling the congested transportation system. Commuter rail has seen a surge in popularity in the past decade and continues to be on the short list of potential transportation improvements for many metropolitan areas. There are many factors that influence the decision to consider commuter rail as a transportation mode for metropolitan areas. These factors will be discussed in significantly more detail. Additionally, a method by which the implementation of commuter rail can be optimized is developed and presented later in this report.

MOTIVATION FOR STUDY

The current state of congestion in metropolitan areas necessitates that metropolitan planners and engineers consider both roadway capacity additions as well as providing alternative means of travel within the metropolitan region. A significant portion of the travel demand and congestion in any metropolitan region occurs during the peak commuting hours. While many large metropolitan areas experience peak periods that spread throughout the day, smaller areas experiencing significant population growth have large spikes in demand and congestion during the morning and afternoon commute periods. The need to provide commuters an alternative means of travel during these heavily congested periods is an opportunity to introduce commuter

rail into the transportation system. Evidence that a new commuter rail system can immediately reduce highway congestion is not currently reliable. However, the 2005 Mobility Report (Schrank and Lomax 2005) details the amount of delay averted through public transportation, supporting the case for public transportation's long term benefits to a transportation system. The American Public Transportation Association (APTA) lends further evidence to the benefits of public transportation, stating that current public transportation usage reduces U.S. gasoline consumption by 1.4 billion gallons per year, and would provide individual households significant economic benefit (Bailey 2007).

Litman (2005) states that commuter rail will impact congestion both directly and indirectly. The direct impacts are realized when a commuter rail trip is substituted for an auto trip, though it is likely that the number of trips diverted from the highway system will be insignificant with respect to the volume of highway traffic in the system in the short term. The indirect impacts are derived from more accessible land use and public transportation gradually changing the travel patterns of a larger number of travelers. Litman further states that while the indirect impacts are difficult to observe, studies do support the hypothesis that these indirect impacts are significant.

It is important to note that congestion relief should not be the only goal of a commuter rail system. In the short term, commuter rail provides a safe, reliable, high-quality option for commuters to travel to work. In the long term, commuter rail provides a metropolitan area the means to more effectively manage its growth and character. By establishing commuter rail, the possibility of future rail and transit modes being implemented improves; future modes that collectively provide relief for a potentially ever-increasingly congested system. As available land for capacity expansion dwindles, having an established transit system will no doubt pay dividends.

Though the potential benefits of commuter rail are great, the best method to implement commuter rail is far from decided. Rail infrastructure is ubiquitous in most urban areas in the United States as access to regional and national transportation has always been a vital component to the growth and success of an area and has historically been fundamental to the growth of metropolitan areas. Deregulation of the rail industry led to many sections of track being abandoned or sold in an effort to optimize the freight rail industry by shedding unproductive infrastructure. The fate of this right-of-way (ROW) varies. Many sections are simply abandoned in place and some are converted to bicycle or walking trails. Texas has converted 174 miles of abandoned or unused rail corridors into public trails and 712 additional miles are planned as of September 2006 (Rails-to-Trails 2007). and some have been purchased by local transit authorities for the occasion that some means of rail service is warranted and desirable for their service area. The benefits of such ROW and infrastructure is obvious; the upfront capital costs of implementing a rail system in the area are significantly reduced by eliminating much of the ROW acquisition costs, reducing much of the environmental red tape (as the new construction footprint is minimal), and possibly reducing the cost of construction through the use of existing infrastructure, depending upon its condition.

Need for Circulators

A detriment of using existing track, however, is that the track layout existed prior to the metropolitan area maturing to its present configuration and it was not selected because of its convenience for 21st century suburban commuters. Grava (2003) discusses the issue of commuter rail placement:

A basic issue related to the use of existing rail alignments is their placement. They were usually established more than 100 years ago to serve a completely different city configuration and respond to the needs of that time. They are not necessarily central to the current corridors of residential and commercial activity.

Much of the rail infrastructure was focused by necessity on freight traffic serving industrial centers. Citizens are generally not inclined to reside near industrial demand centers, nor necessarily work near these centers as the scale of rail operations generally implies low-density warehousing, storage, or holding areas. Additionally, people do not tend to live near freight rail lines for fear the noise and vibration that often accompany rail transportation will make a residence unpleasant and devalue a property if not built to mitigate the nuisance impacts of rail traffic.

To summarize: people do not tend to live within walking distance (as will be properly defined later) of existing rail lines or the places that the existing lines are intended to serve. Additional ROW is not generally available to construct rail infrastructure in a more favorable location in urban areas. Even if it were, moving a rail line to better accommodate residential areas within walking distance is likely to meet stiff opposition from residents who do not care to have trains, passenger or freight, rolling through their backyard.

Where does this leave the transit authority? If the majority of commuter rail users do not live or work within walking distance of existing rail or proposed rail stations, some means of accessing commuter rail and final destinations are needed if commuter rail is to be considered a viable option for commuters. The means of access are circulator systems, be they fixed guideway or rubber tire. Circulator systems are a means to provide access to commuter rail, and provided that commuter rail compares favorably with other modes of travel, represent a way in the short term to help cultivate and retain ridership until development near the rail stations catches up to enhance the stability of the system and its ridership.

The two ends of the commute trip have different characteristics. The home end of the trip has been addressed in many commuter rail and suburban transit systems through the use of park and ride facilities, a method that appears to be the preferred method of aggregating passengers on new commuter rail systems in the United States. These facilities take advantage of available land in suburban regions to accommodate transit users' vehicles and help when transit usage is being first introduced to a region. In more mature metropolitan areas that implemented park and rides long ago, problems are arising when parking space supply is falling short of demand and the availability of land has dwindled as other development has occurred. In these mature areas innovative methods of park and ride management coupled with additional modes of accessing the transit facility are needed to support the increasing demand.

This report will focus on new commuter rail systems and will make the assumption that at the home end, park and ride facilities will provide access to the commuter rail line. The method presented focuses on the destination end of the commute trip to a downtown or business district. It is highly unlikely that a commuter will have a personal vehicle at the destination station, and as mentioned previously, it is also unlikely that the commuter will work within walking distance of the destination station, unless that station resides in the central business district. A method is therefore needed to optimally design the circulator network at the destination end of the commuter rail trip, considering all aspects of the trip: transfer, wait, ride, and walk times. This method should also take a big picture view of how an optimal circulator route will impact the overall mode share and how the circulator system will evolve with changing land use and how the system itself will impact the changing development and land use.

Existing and Planned U.S. Commuter Rail Systems

Table 1 provides a comprehensive list of existing commuter rail services in the United States, according to the American Public Transportation Association (APTA), current as of March 2007 (APTA 2007). These systems adhere to the definition of commuter rail put forth in this report to varying degrees. The primary source of this variation is the operation of commuter rail services in non-peak periods. The definition used in this report (discussed in subsequent section) limits commuter rail service to peak hours only, the APTA list allows for flexibility in this particular definitional constraint.

TABLE 1: APTA LIST OF EXISTING COMMUTER RAIL SERVICES IN THE UNITED STATES

Albuquerque, NM RailRunner (New Mexico Rail Runner) *√†
Alexandria, VA VRE (Virginia Railway Express)   √†
Anchorage, AK ARC (Alaska Railroad Corporation) --
Baltimore, MD MARC (Maryland Transit Administration, MTA) √†
Boston, MA MBTA (Massachusetts Bay Transportation Authority) †
Chesterton, IN NICTD (Northern Indiana Commuter Transportation District) †
Chicago, IL METRA (Northeast Illinois Regional Commuter Railroad Corporation) †
Dallas, TX TRE (Trinity Railway Express) *†
Harrisburg, PA PennDOT (Pennsylvania Department of Transportation)(unofficial) --
Los Angeles, CA Metrolink (Southern California Regional Rail Authority) †
Nashville, TN Music City Star (Regional Transportation Authority, RTA) *√
New Haven, CT SLE (Connecticut Department of Transportation Shore Line East) *√
New York, NY LIRR (MTA Long Island Rail Road) MNRR (MTA Metro-North Railroad) †
Newark, NJ NJT (New Jersey Transit Corporation) †
Oceanside, CA Coaster (North County Transit District, NCTD) *
Philadelphia, PA SEPTA (Southeastern Pennsylvania Transportation Authority) †

Pompano Beach, FL Tri-Rail (South Florida Regional Transportation Authority) *†
San Carlos, CA CALTRAIN (Peninsula Corridor Joint Powers Board) *
Seattle, WA Sound Transit (Central Puget Sound Regional Transportation Authority) √†
Stockton, CA ACE (Altamont Commuter Express) *†
Syracuse, NY On Track --
* Uses circulator/shuttle √ Meets definition of Commuter Rail (peak hours, peak direction only) † Connects with existing transit -- insufficient information for consideration

At first glance, the majority of commuter rail systems do not seem to utilize circulator, or shuttle, systems to provide passengers access to their final destinations. Most appear to use existing transit, and looking at Table 1, about 67% use predominantly existing transit. However, if one is to apply the definition of commuter rail used in this report, that is, rail service only during peak hours and in the peak direction, the list is reduced considerably (shown below in Table 2), and the percentage of systems relying exclusively on existing transit drops to 50%.

It is true that all systems attempt to provide connectivity to existing transit systems, this is good practice and should help improve the overall mobility in the region. However, three systems rely predominantly on circulator systems: Albuquerque, Nashville, and New Haven. The average population of these three cities is approximately 950,000. The three cities that meet the commuter rail definition yet rely on existing transit connections extensively, Alexandria, Baltimore, and Seattle, have an average population of 3,460,000. This finding should not be surprising and highlights the applicability of this report work. Smaller cities will generally have less developed transit systems and have lower density development and therefore greater need for additional resources dedicated to improving the connectivity of the commuter rail system.

TABLE 2: COMMUTER RAIL SERVICES CONFORMING TO REPORT DEFINITION

Albuquerque, NM RailRunner (New Mexico Rail Runner) *√†
Alexandria, VA VRE (Virginia Railway Express)   √†
Baltimore, MD MARC (Maryland Transit Administration, MTA)   √†
Nashville, TN Music City Star (Regional Transportation Authority, RTA) *√†
New Haven, CT SLE (Connecticut Department of Transportation Shore Line East) *√†
Seattle, WA Sound Transit (Central Puget Sound Regional Transportation Authority) √†

Additionally, the three circulator-dependent lines are relatively new, with Albuquerque coming on line in 2003, Nashville in 2006, and New Haven in 1990. In contrast, Alexandria began service in 1992, Baltimore in 1983, and Seattle more recently in 2003. Obviously, New Haven and Seattle break the mold, but a trend is obvious in which newer commuter rail systems tend to rely more heavily upon circulators.

These two observations support the application of the method developed in this report, and highlight the need for such a method, in relatively small, new commuter rail systems. The APTA list also contains 21 planned commuter rail systems, cities which tend to be smaller than those with existing commuter rail systems. Table 3 provides a comparison between metropolitan areas with existing commuter rail systems and those which are planning new systems. The not surprising observation is that cities with existing systems tend to be larger than those currently planning systems.

TABLE 3: POPULATION COMPARISON OF COMMUTER RAIL METROPOLITAN AREAS

	Commuter Rail Metropolitan Area			
	Average Population	Max Population	Min Population	Standard Deviation
Existing System	4,244,254	18,323,002	319,605	4,769,849
Planned System	1,563,935	4,247,981	124,279	926,268
Total	2,801,005			

The histogram in Figure 1 helps further illuminate the population distribution of these 21 planned commuter rail lines.
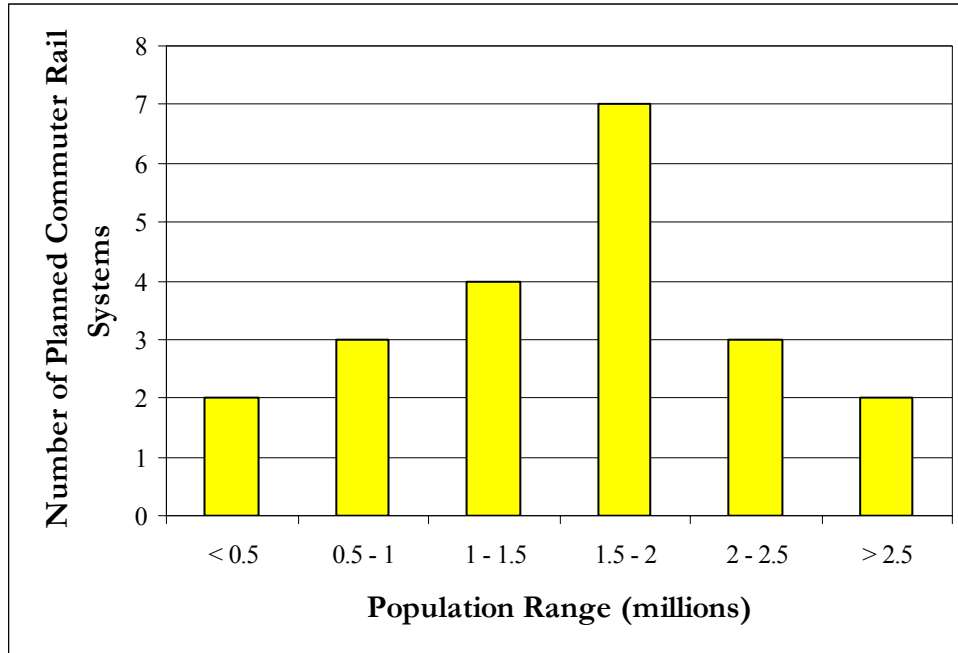
FIGURE 1: METROPOLITAN POPULATION OF PLANNED COMMUTER RAIL SYSTEMS

Inspecting the left side of the histogram in Figure 1 (which is approximately normally distributed), one sees that there are nine lines currently planned that are similar in size to the three circulator-dependent commuter rail systems described previously (Albuquerque, Nashville, and New Haven). These nine metropolitan areas are almost half of all currently planned systems, suggesting that there is currently a market for circulator systems serving commuter rail.

TABLE 4: METROPOLITAN AREAS WITH >600,000 POPULATION AND NO RAIL SERVICE

MSA	2003 Population
Indianapolis-Carmel, IN	1,525,104
New Orleans-Metairie-Kenner, LA	1,316,510
Memphis, TN-MS-AR	1,205,204
Oklahoma City, OK	1,095,421
Bridgeport-Stamford-Norwalk, CT	882,567
Honolulu, HI	876,156
Tulsa, OK	859,532
Dayton, OH	848,153
Albany-Schenectady-Troy, NY	825,875
New Haven-Milford, CT	824,008
Fresno, CA	799,407
Omaha-Council Bluffs, NE-IA	767,041
Oxnard-Thousand Oaks-Ventura, CA	753,197

Worcester, MA	750,963
Grand Rapids-Wyoming, MI	740,482
Allentown-Bethlehem-Easton, PA-NJ	740,395
Baton Rouge, LA	705,973
Akron, OH	694,960
Springfield, MA	680,014
Bakersfield, CA	661,645
Toledo, OH	659,188
Syracuse, NY	650,154
Columbia, SC	647,158
Greensboro-High Point, NC	643,430
Poughkeepsie-Newburgh-Middletown, NY	621,517
Knoxville, TN	616,079
Little Rock-North Little Rock, AR	610,518
Youngstown-Warren-Boardman, OH-PA	602,964

In addition, there are nearly 30 additional U.S. metropolitan areas over 600,000 in population and less than roughly 1.5 million, according to the U.S. Census Bureau (2003) with no current rail service that may make excellent candidates for commuter rail in the future, shown in Table 4.

Table 4 is provided to serve as a useful demonstration tool of the applicability of this research surrounding commuter rail systems. However, this should not be viewed as a statement that commuter rail is incompatible with other forms of rail or rubber-tire transit. This is merely intended to identify metropolitan areas that may most likely be considering a commuter rail addition to their transportation system that would benefit from the results of this report. These areas will tend to be those with a smaller, less dense population and less existing infrastructure to support a rail system (those areas with no commuter passenger rail service).

Commuter Rail Circulator Route Network Design Problem

The Commuter Rail Circulator Route Network Design Problem (CRCNDP) has not been studied extensively. This is likely due to the fact that increasing interest in the CRCNDP is a function of increasing interest in commuter rail and the Bus Transit Route Network Design Problem (BTRNDP), which considers an entire transit network and has received much focus and effort over the years. Studies that have addressed the CRCNDP specifically and those that impact it indirectly will be discussed in detail in Chapter 2. For the purposes of this Chapter, the shortcomings of the work to date in this field will be summarized in an effort to better describe the motivation for this study.

Relationship with Mode Choice

Nearly all current work addressing or related to the CRCNDP assumes that peak hour commuter origin-destination (O-D) data is known and fixed. In fact, this assumption is not limited to the CRCNDP as Dial (1996) discusses in his work addressing traffic assignment and mode choice. This assumption is made primarily to make an already difficult problem somewhat tractable and solvable with available technology and data. It is unreasonable to expect that large

metropolitan areas have hour-by-hour O-D tables that are both available and have a reasonable degree of accuracy. Some existing work assumes a fixed overall commuter O-D table with a variable transit mode share, making the problem more tractable while still allowing for a degree of demand variability.

Even though disaggregate O-D data is not often readily available; the relationship of optimal routing and mode choice should not be ignored. A metropolitan area, while not having O-D data for each hour of the day, is likely to have some data available for the peak hours which are primarily dominated by commute travel. This fixed O-D data can serve as a baseline for commute travel and a mode share of this fixed baseline can be predicted. As travel time and travel time reliability are two key factors in any mode choice model, the routing design of circulator routes servicing commuter rail could potentially play a significant role in commuter rail mode share by affecting the population fraction for which commuter rail is deemed accessible.

Accounting for Walking Portion of Trip

All trips via commuter rail are going to require at least some portion of the journey be completed on foot. Whether the final destination is within walking distance of the train station or a circulator stop, a nonzero distance will need to be traversed on foot. This distance will be of much importance in influencing the use of commuter rail as evidence has shown over the years that most people are unwilling to walk more than about ¼-mile for bus transit, and ½-mile for rail (Grava 2003). Also, walking distance should be considered in the context of the entire commute trip. A worker choosing commuter rail is likely to be facing four travel modes during the commute trip: their personal vehicle to travel to the park and ride, the train, the circulator, and finally as a pedestrian to their final destination. In light of the complexity of the trip the walking distance may prove a crucial factor in attracting and maintaining ridership on commuter rail.

Accounting for Bus Stops

Most current work considers demand being produced at zone centroids or uniformly distributed throughout a "coverage" corridor. This assumption ignores two aspects of the commuter rail trip that will be of importance, bus stop location and the previously discussed walking portion of the trip. Because walking trips may have such an important impact on the choice of commuter rail as a commute mode, it is necessary to remove obstacles to accurately portraying this portion of the trip.

The current practice of considering demand originating and being destined for zonal or block centroids reduces the impact of the walking portion of the trip. To accommodate the walking portion of the trip, bus stops should be modeled to exist at mid-block locations or at intersections. In defining bus stops in this realistic manner, one can better account for the walking distance that must be traversed to the final destination. In short, defining origins and destinations in the more accurate manner at the edges of blocks is necessary if the walking portion of the trip is to be more accurately modeled.

9

ROLE AND FUNCTION OF COMMUTER RAIL

Several sources addressing the broad topic of commuter rail (Grava 2003, Gray et al. 1992, Vuchic 1981, Vuchic 2005, and Edwards 1992) were used to develop the following list of distinctions regarding commuter rail:

1. Commuter rail operates on existing ROW and track (in practice if not by definition)
2. Commuter rail, by definition, operates only during peak commuting hours.
3. Commuter rail has a transitional nature.

These distinctions make commuter rail a unique mode separate from close relatives heavy rail, regional rail, subway, and light rail. These distinctions are further examined and discussed in the following sections.

Commuter Rail Distinctions

Following is a discussion of the distinct characteristics of commuter rail relative to other urban rail modes. There are likely further distinctions that could be made from a rolling stock or platform design perspective, but for this research a broader operational perspective is adopted.

Use of Existing Track and ROW

Maintaining the definition of commuter rail held by Vuchic (1981), where commuter rail operates only in peak periods, in the United States commuter rail has been implemented only on existing ROW and depending upon condition, existing track infrastructure (Grava 2003). No new routes have been constructed specifically for commuter rail. Grava also outlines the strengths and weaknesses of the current commuter rail practice in the U.S.:

Strengths

- Efficient (low energy) operations,
- Fast and comfortable service (function of station spacing and engine power),
- More reliable and safe than other modes,
- Uses existing resources,
- Service can be implemented quickly, and
- Commuter rail has an overall good public image.

Weaknesses

- Little location flexibility,
- Coordinating with freight rail may be necessary,
- Infrastructure may have to be rehabilitated,
- Rolling stock will need to be acquired,
- Needs good ridership to justify capital investment,
- Environmental concerns (air, noise, vibration),
- Safety issues for pedestrians at grade crossings, and

- Residents rarely want a rail line adjacent to their property.

Certainly, some of these strengths and weaknesses of commuter rail will exist regardless of whether rail is implemented on existing ROW. However, using existing ROW reduces initial capital costs, enabling the transit authority to begin service earlier Additionally, existing ROW reduces the environmental impact footprint of the project and thus should reduce the burden of the National Environmental Policy Act (NEPA) process. Since the existing trackage has been in place longer than neighboring development, it is less likely that existing ROW will traverse as many areas sensitive to the noise and vibration associated with commuter rail, especially if freight service currently operates on the line. Existing infrastructure may also provide the benefit of at-grade crossing and pedestrian crossing infrastructure.

The use of existing ROW and infrastructure that is currently under transit authority ownership, or is possible to acquire is a practice that will likely continue in the future. Any commuter rail system planning efforts need to acknowledge that commuter rail implementation is likely going to continue along existing ROW.

Peak Hours Only

Commuter rail is distinct from heavy and regional rail modes in that it is in service only during peak hours, the distinction put forth by Vuchic (1981). All communities may not experience the same peak periods or hourly distribution pattern. In fact, the hourly distribution pattern can vary between roads within the same community, as demonstrated in the Highway Capacity Manual 2000 (TRB 2000). Regardless of the location of the particular community, there are usually still distinct commuting peak hours in the morning and evening, and these are the hours that commuter rail is designed to service. If an all-day, non-peak hour service is being implemented, then it is not considered commuter rail for the purposes of this report.

Peak Direction Only

Commuter trips are historically directionally distinct in the morning and evening commute periods. Suburban development patterns created bedroom communities on the periphery of the city that were primarily single family residences that commuted to the downtown area for the work day. This directional distribution has lessened in recent years with more commercial development in the suburbs and the resulting inter-suburban commuting. However, a dominant portion of the commute trips still maintain the traditional suburb-to-city center in the morning and the reverse in the evening. It is these trips that commuter rail is designed to serve, with service in only the peak direction in the morning and evening.

Servicing only the peak direction will create some issues with vehicle storage at the downtown end of the trip during the day, and these issues must be addressed so that the trains are not interfering with other modes of downtown traffic or other users of the rail line. This highlights another benefit of using existing infrastructure, as old rail lines are likely connected to old rail yards providing an area for equipment storage during non-operational hours and for the routine repair, cleaning, painting, refurbishing, and overhaul of the equipment (Grava 2003).

Transitional Nature

Vuchic (2005) notes that "the trend has been to upgrade commuter rail into regional rail lines or networks", a sign that commuter rail is not usually considered to be a permanent mode.

Commuter rail often appears to be a means of introducing rail as a viable, useful component in a transportation system. As commuter rail becomes more ingrained in a community; ridership improves, development patterns adjust to the rail service, and demand for non-peak service increases and commuter rail transitions into a regional or heavy rail role. This transition includes non-peak service, bidirectional service, expansion of the current service, or even an upgrade to subway service. There is no single transitional pattern to commuter rail; the pattern emerges as needs arise in a community. However, it is clear that commuter rail is not destined to stay commuter rail. Commuter rail serves a short term goal of generating ridership, acceptance, and excitement for rail service and adjusts accordingly as these goals are accomplished and subsequently expanded.

Cultivating and Retaining Ridership

Several times the goal of commuter rail has been defined as "cultivating and retaining ridership". The introduction of commuter rail in a community should be undertaken with long-term goals in mind. Political maneuvering and implementation costs are short term and need to be addressed as such. In the long term, commuter rail is essential to the long-term viability of a transportation system. Commuter rail is being used as a means to guide development and land use along rail lines promoting denser, walkable neighborhoods that are less reliant on the personal automobile. In the long term rail may carry a significant portion of trips in a community, such as New York City, but this cannot happen until the commuter rail line helps make rail an integral part of a younger transportation system. In the short term, what commuter rail can do is create a ridership foundation and work towards the positive perception of commuter rail and its necessary role in a community.

These distinctions are important since defining the goal of a system significantly impacts the objective one may use to optimally configure the system. If a long-term goal is given the immediate focus, such as congestion reduction, maximizing profit or influencing land use, the optimal commuter rail system and therefore the essential circulator system will look different and perhaps be inappropriate in the short term and may actually hurt the chances of accomplishing the long-term goals. A short-term focused design is essential for commuter rail circulators as the impacts commuter rail will have on a system are very hard to predict. Focusing on short-term reachable goals, such as cultivating steady ridership by providing quality service to as many people as practical and maintaining circulator system flexibility to changes in the overall transportation system will help ensure successful implementation of rail into a community.

Uniqueness of Commuter Rail Circulators

An interesting and potentially useful aspect of circulator systems designed for servicing commuter rail stations is that for a new commuter rail line no routes exist to serve this specific purpose. As no station exists to be serviced currently, any bus routes that currently operate in the area are designed to serve some other set of origins and destinations. Provided that resources are not diverted from the existing system, the optimal routing of the circulator system can be pursued without considering a cost of eliminating existing service. The routes can be designed and implemented in an optimal manner from a ridership and operator perspective with a lesser risk of political and social backlash.

Challenges for Commuter Rail

The success of commuter rail in a given region is certainly not guaranteed and is subject to many operational, economic, political, social, and environmental aspects of the region. Some of the operational challenges are discussed in the following sections. While the weight of these operational challenges relative to the other categories of challenges depends on the region, addressing the other challenges is beyond the scope of this report and will not be dealt with explicitly, though every attempt will be made to properly consider or at least acknowledge non-operational issues and challenges where appropriate.

Travel Time

A big benefit of commuter rail, provided that the number of at-grade crossings is limited, is the reliability of the travel time from point A to point B. Commuter rail, while perhaps not providing equivalent total travel time with the personal auto, does provide a more reliable travel time. For example, a home to work trip may average 40 minutes via auto and 50 minutes via commuter rail. The personal auto trip may also have a high variance associated with it, resulting in several trips per month (or week) exceeding an hour. In contrast the commuter rail trip will likely always be near its average trip time, rarely approaching an hour of travel time.

An issue that often surfaces with all transit systems, commuter rail included, is the issue of transfer, wait, and walk time. Commuter rail has an image in the United States as a "premium service" (Grava, 2003) and therefore the in-train travel time may be considered more favorable than personal auto in-vehicle travel time as commuter rail affords the passenger the opportunity to read, sleep, or be otherwise distracted from the roadway.

A major issue in transit systems is the number of transfers, the wait time at transfers or stops, and the walking portion of the trip. Wait time is generally considered more onerous than in-vehicle travel time and transfers often imply a cost greater than the wait time alone (Newell 1979). Even if commuter rail can provide a more reliable travel time at a premium level of service, much attention should be paid to the transfer, wait, and walk time.

The Typical Commuter Rail Trip

The typical commuter rail trip on new systems in the United States is likely much more complex than a commuter rail or other rail trip in a system with more developed rail transportation. In a well-developed rail system, such as New York's subway system, the rail portion of the trip is book ended by walking trips. In a new commuter rail system a home to work trip will often begin with an auto trip from the home to a park-and-ride or a trip on a separate collector system from home to the rail station. The rail portion of the trip will commence from home station to destination station and the final leg of the commuter rail trip will likely include some sort of circulator system, followed by a walking trip.

The current land use of many station areas can influence this aspect of commuter rail. Existing, abandoned freight rail lines are not typically attractive to business and residential development and many planned stations are in areas that currently have little work trip generating development within walking distance of the station (with the exception of CBD stations). At the destination end of a commuter rail trip it is likely that to cultivate the desired ridership, commuter rail planners will have to service commuter trips that are destined for places outside of the walking distance of the station.

This trip encompasses four modes, two transfers, and a walking trip. This trip has several potential aspects that can discourage ridership, the most obvious being the transfers and the walking trip. Focusing on avoiding the pitfalls of long transfers and long walks will go a long way in helping to cultivate the ridership that a new commuter rail line needs.

Transfer Times and Walking Distances

The quality of service of the train portion of the trip is obviously important to the success of a commuter rail line, as is the availability of parking at park-and-rides. However, a mode that may require four modes per trip, including two transfers and a walking trip must focus on more than optimizing the in-vehicle travel time. Transfer times must be minimal. Walking trips must be as short and enjoyable as possible.

Seamless transfer is a term that refers to the transfer from the commuter rail to the circulator. The headways of a commuter rail line are known and are very reliable, as commuter rail operates on its own dedicated guideway and has few, if any, at-grade crossings. A circulator system should be employed that guarantees the commuter rail rider that the circulator vehicle is waiting at the destination station upon arrival of the train. This seamless transfer will provide minimal transfer times, equal only to the sum of the train alighting, walking, and circulator bus boarding times at the station. This sort of circulator system also provides practical boundaries to the magnitude of the circulator system to be provided. As a vehicle must be waiting at the station for the arrival of each commuter train at each station, the headway of the train and the number of vehicles available (or the budget to purchase circulator vehicles) will provide a solid boundary of service for the station that is adjustable according to the objectives of the system.

The walking portion of the trip can potentially influence the choice of commuter rail in ways other than simply distance. The number of crossings, the amount of shade, safety, and the general interest of the area can all influence the walk trip as suggested by Kuby et al. (2004) and their findings regarding the significance of weather on light rail transit (LRT) trips.

Land Use Impacts

There is no doubt that rail transit can impact the land usage around its stations and beyond. Increases in density, mixed-used development, property values, and walkability have been observed across the country. Dunphy et al. (2004) cite many examples of these observed impacts of commuter rail on property value and urban infill. However, the magnitude and nature of the impacts is very location-specific and is difficult to predict. The tripmaking impacts of station area development are equally difficult to foresee during preliminary planning stages.

These land use issues suggest that a flexible circulator system be initially employed in the early stages of commuter rail. This sort of flexibility will allow for the station area to develop and with periodic revisiting of the routing re-optimize the service to better serve the commuter rail travelers and the development that is likely to be spurred near the station.

Unserved Demand Cost

Another consideration in the design of a commuter rail circulator system is the cost to the commuter rail system and to its travelers of not serving all potential riders. This cost would include the cost of not providing circulator service within walking distance of a particular demand location, of providing poor service to a location, of requiring long walks to a location, or of a location simply being too far from a rail station. Bailey (2007) suggests the savings

experienced by two-adult households with access to public transportation (within ¾ mile) are on the order of $6,000 per year per household.

REPORT OBJECTIVES

There are five objectives of this study, each contributing individually to the understanding of this problem as a whole and each contributing cooperatively to the overall purpose of this report.

Investigate Current State of Knowledge

Current methods of commuter rail circulator route design are not well documented. Much can be learned from literature addressing the general bus transit network design problem (BTRNDP) and other transit routing literature. This objective seeks to develop a foundation of knowledge upon which the method proposed in this study can build.

Develop Procedure to Solve Commuter Rail Circulator Network Design Problem

A procedure to efficiently and systematically solve the commuter rail circulator network design problem (CRCNDP) will be developed. Heuristic and exact methods will be investigated as appropriate to the characteristics of the problem as they are defined by the problem formulation. This procedure will attempt to incorporate a process by which the impact of optimal circulator routing upon commuter rail mode share can be better understood and incorporated through the use of unserved demand costs.

Account for Walking Portion of Trip

As discussed earlier, the potential for the walking portion of the commuter rail trip to influence the commuter trip making choices is considerable. The development of a cost function to incorporate detail regarding the walking portion of the trip will be undertaken in conjunction with the formulation of the CRCNDP.

Investigate Formulation Performance

The formulation developed for this study will not represent the only way possible to formulate the CRCNDP. A sensitivity analysis and qualitative review of the formulation from both the user and operator perspective will be undertaken. It is the intention of this objective to identify areas for future formulation improvement, sources of potential bias, and ways to streamline the formulation or adapt to new solution techniques.

Case Study

A major source of inspiration for this research is the ongoing efforts of Austin, Texas and Capital Metropolitan Transportation Authority (Capital Metro) to introduce the first commuter rail line, MetroRail, to the Austin metropolitan area transportation system. This objective will seek to obtain relevant data about the Austin system and apply the solution methodology developed in this study to the proposed commuter rail line in Austin. This will be undertaken in

hopes of gaining a better understanding of the practicality of the method, providing useful input to Capital Metro, and discovering areas for future improvement of the method.

CONTRIBUTIONS

The objectives of this report lead to the following contributions to the transportation engineering profession and state of the art:

- An efficient and systematic solution procedure for the CRCNDP using the knowledge, skills, and technology available today.
- A formulation of the CRCNDP that accounts for unserved demand, walking trip variables, and a seamless transfer concept.
- A better understanding of the relationship between optimal commuter rail circulator design and commuter rail accessibility.
- A demonstration of the utility of the CRCNDP solution method through application to the Austin, Texas proposed commuter rail system.
- A review of the method from a user, operator, and practical perspective through sensitivity analysis, qualitative review, and review of the case study.

SUMMARY

This report is intended to thoroughly investigate the design and layout of commuter rail circulator systems which are vital to the success of a commuter rail system. This investigation will involve an in-depth review of existing literature, the development of an innovative formulation for the CRCNDP, the application of exact and/or heuristic solution procedures to the formulation for generated data sets, the application of the method to a case study in Austin, Texas, and a qualitative review of the method from several perspectives. It is intended that the results of this report will provide a better theoretical understanding of the CRCNDP and a practical tool that can be applied throughout the United States.

REPORT ORGANIZATION

Chapter 2 will provide an in-depth look at the literature that is available and relevant to the CRCNDP. The general role of commuter rail in today's transportation system will again be discussed. A significant amount of literature exists addressing the general transit route network design problem and bus transit network design problem, and this will be presented in some detail, focusing on previous efforts that are particularly relevant to the goals of this report. The evolution of formulations and solution methods will be discussed briefly, as well as some of the fundamental concepts that are to be used in the remainder of the study.

Chapter 3 will provide a step-by-step look into the development of the formulation for the CRCNDP and its objective function. In this section the general format of the formulation will be discussed, the constraint set constructed, and the objective function developed. Development of individual cost components of the objective function will be described in detail as it seeks to incorporate several concepts that have been discussed previously, such as walking cost functions, unserved demand, and implications of seamless transfer.

Chapter 4 will present the development and application of an enumerative solution method. Successes and shortcomings of this method will be discussed along with implications for subsequent methodological development. Chapter 5 presents the natural follow-up to Chapter 4, targeting the shortcomings of enumeration with improvements designed to efficiently

utilize computational resources. Preprocessing steps and stopping criterion will be presented and their performance displayed.

Chapter 6 describes the last solution method that is developed as part of this report. This method is based on the tabu search metaheuristic that has been employed in a wide variety of fields over the past 25 years. The method's performance and applicability are discussed in detail and comparisons with the previous two methods presented.

Chapter 7 describes a sensitivity analysis undertaken of the CRCNDP formulation parameters, focusing on unserved demand and walking threshold. The results presented here represent a significant contribution to the improved understanding of what "transit accessibility" means in a quantitative sense. Chapter 8 applies the knowledge gained throughout the process of studying the CRCNDP and applies it to a case study in Austin, Texas. The results are given in order to demonstrate the benefits of such an analysis and how the improved design of the CRCNDP can impact decisions.

This report concludes in Chapter 9 with a summary of the objectives and contributions of this work. A summary of important results will be given and explanation provided. The implications of this work in a broad sense, as well as directions for future study and application will conclude this work.

CHAPTER 2: LITERATURE REVIEW

A great deal of literature exists for addressing particular aspects of the commuter rail circulator network design problem (CRCNDP). Little deals specifically with the CRCNDP as a mathematical optimization problem, the focus of this report. Previous efforts have been hampered by limits in computing power and have not focused on commuter rail circulators specifically because the trend to use commuter rail as a means to introducing rail transportation into a transportation system has recently gained momentum. Much of the information presented will focus on areas other than the CRCNDP, four common areas of which are: circulator and feeder systems, the transit network design problem (TNDP), the transit equilibrium assignment problem (TEAP), and commuter rail case studies.

This chapter will first look at several case studies of commuter rail in the United States to serve as background for the CRCNDP. The role and function of commuter rail will then be briefly revisited in light of the case studies. Quantitative analysis papers will then be investigated in some detail, divided into three categories: analytical methods, mathematical programming, and transit equilibrium assignment. The chapter will conclude with a look at the gaps in the current CRCNDP knowledge.

CASE STUDIES

The American Public Transportation Association (APTA) lists 19 existing commuter rail systems (with sufficient information available) in the United States in addition to 26 proposed systems (APTA 2007). Beyond the existing and planned systems are the metropolitan regions poised to begin considering rail service as part of their transportation system. Grava (2003) describes the rule of thumb that a metropolitan area is ready to begin considering rail when the population reaches approximately 1,000,000 (Grava further discusses how this rule of thumb has no real theoretical basis, but stems from a desire to have order and rationality in the structuring of service systems).

Not all existing or potential services will be investigated here; however four will be examined in greater detail. Austin, Texas will be discussed as an example of a system that is in the planning and development stages and will display some of the goals and challenges associated with a new system. Nashville, Tennessee serves as an example of the type of metropolitan area that will most benefit from the methodological developments of this report. Dallas, Texas and the Trinity Railway Express (TRE) will be visited in an effort to display how a commuter rail system might look in the early years of development, its role in the transportation system and future directions for commuter rail. To display the transitional nature of commuter rail in many mature systems, the Fairmount Line in Boston, Massachusetts will be presented and will conclude the case studies.

Capital MetroRail – Austin, TX

The proposed Capital MetroRail system in Austin, Texas is slated to begin operations in 2008. Beginning with a 2004 referendum on commuter rail, the initial rail component of the Austin transportation system has made steady progress toward the 2008 operating goal. The 32-mile layout runs exclusively on existing tracks and will begin service as peak hour-only, thus qualifying it for the definition of commuter rail. Figure 2 displays the layout for the proposed commuter rail line in Austin, dubbed "The Red Line".

The Austin Commuter Survey report by Bhat et al. (2005) describes the positive image that commuter rail transit (CRT) possesses in Austin compared to the bus mode. In fact, all else being equal, commuters are willing to travel via CRT even if CRT takes up to 20 minutes longer than the travel time by bus. Even with this positive image, CRT cannot be expected to initially induce a mode shift from the personal automobile. In fact, the same study by Bhat et al. suggests that if 10% of the commuting population has CRT as a commuting option then commuter rail will capture approximately 15% of the mode share in the corridor with commuter rail availability, or 1.5% of the overall mode share.
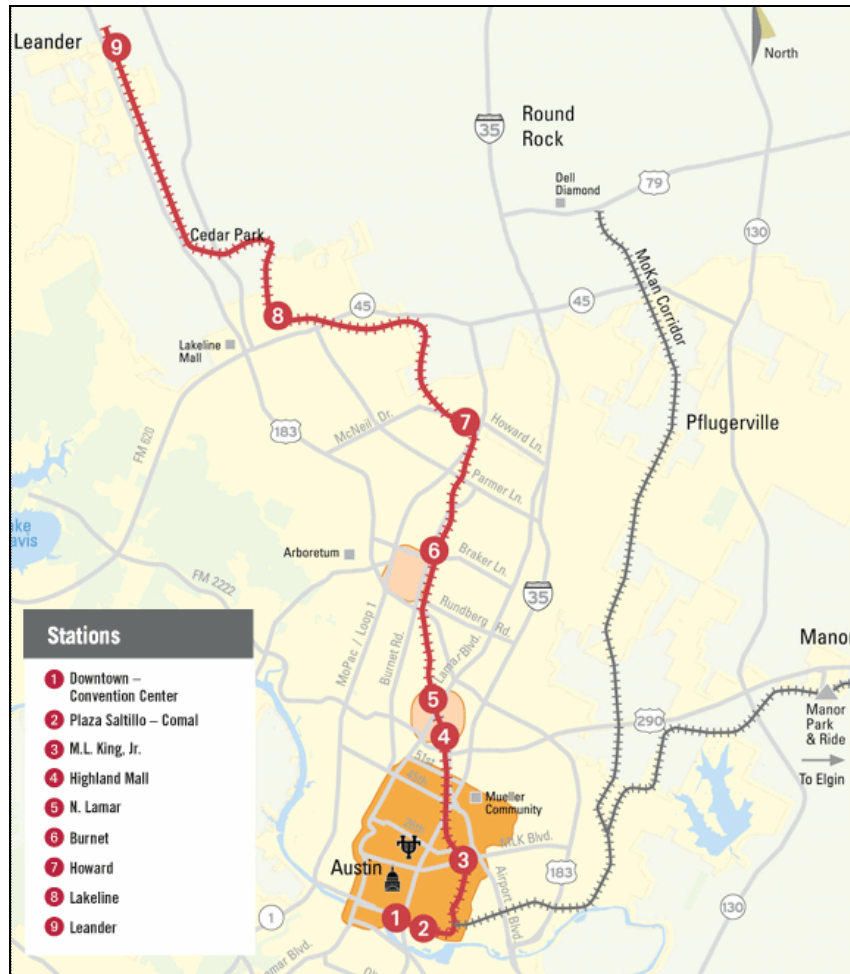
FIGURE 2: CAPITAL METRORAIL PROPOSED RED LINE MAP

Source: http://allsystemsgo.capmetro.org/capital-metrorail.html

This number must be considered in light of the definition of CRT availability in Bhat's study. CRT is considered available if the stations are within one mile of the commuter's residence and work place. For a starter system such as Austin's, on existing right of way and with the associated difficulties of existing rail right of way, the percentage of commuters having CRT available, even by this definition, may well be less than the assumed 10%.

Bhat et al. (2005) also suggest that benefits from CRT will not likely accrue until the system is well established and a higher percentage of the commuting population has CRT as an

20

option. The initial likelihood of commuters being within one mile of both origination and destination stations is low, evidence of the importance of circulator systems in a new commuter rail system. Increasing the availability of CRT is a key component to the success of MetroRail, as the transit authority, Capital Metro, acknowledges on their "All Systems Go!" long-range transit plan website (Capital Metro 2006).

Music City Star – Nashville, TN

The Nashville – Davidson – Murfreesboro, Tennessee metropolitan statistical area (MSA) boasts a population of approximately 1.3 million (U.S. Census Bureau 2003) and consists of 13 counties in the central Tennessee region. The largest city, Nashville constitutes approximately 600,000 of the MSA's residents. The population density of Nashville, the primary commuter trip attractor in the MSA, is relatively low at approximately 1200 persons per square mile (compare this to New York City, which is over 23,000 per square mile or Chicago at over 12,000 per square mile).

The Middle Tennessee Regional Transportation Authority (RTA) states that the Music City Star seeks to "present an alternative in the way we move about our community, increasing the choices, mobility, and independence of travel for each of our citizens" (RTA 2004). This goal is currently being pursued through the development of the initial phase of the Music City Star, the East corridor which will serve Davidson and Wilson counties (shown below in Figure 3).
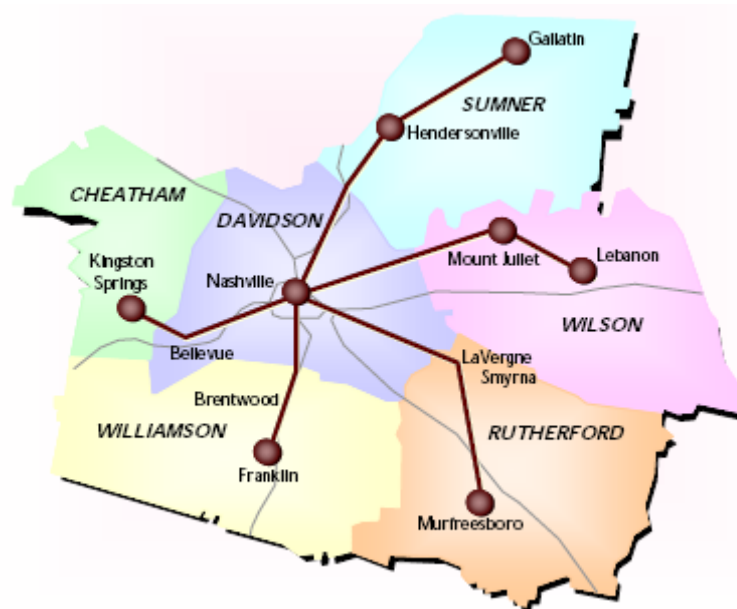
FIGURE 3: MIDDLE TENNESSEE COMMUTER RAIL NETWORK (MUSIC CITY STAR)

Source: RTA 2004

The East Corridor is currently operational and is providing service during the peak hours only in both directions (in and out of Nashville) along the line. This line relies heavily upon circulator systems to access all stations along the corridor as the population density is relatively low and there is not sufficient existing transit infrastructure to serve the commuter rail stations

reliably. It is in this aspect that the Nashville system well represents the type of commuter rail line that would benefit most from this report's methodological developments. Nashville is a relatively small city with no significant existing passenger rail infrastructure in the region, thus requiring significant effort in improving the accessibility of the system through the use of feeder and circulator systems.

Trinity Railway Express (TRE) – Dallas, TX

The Trinity Railway Express is a good example of commuter rail in Texas as a means to provide a beneficial service to a metropolitan area as well as establish commuter rail as a vital portion of the transportation system. The TRE is a 35-mile commuter rail line with 10 stations first established in 1996 as a 10-mile section servicing only the Dallas metropolitan region. The TRE now connects downtown Fort Worth, the Dallas/Fort Worth international airport, and the Downtown Dallas district. The suburban stations provide free commuter parking and function as park and rides while the urban stations serve as connection points for other urban transit services or convenient access to major attractions. There are three particular travel demands that Dallas rail system currently serves: the airport, special events at the American Airlines Center, and commuter access. Texas Instruments employees improve ridership, as TRE targets this large demand producer with a specific circulator system and connections with light-rail transit that serves a station near the Texas Instruments campus (Acken 2005). Figure 4 is an overview of the TRE current configuration and the layout of the stations served and relevant generators and destinations.

Acken (2005) highlights how TRE currently serves specific destinations and markets and suggests that this can be embraced to solidify the relationship with commuter, special event, and air traveler markets. This may serve as a base upon which to expand commuter service and connect the commuter service to other transit services in the region and improve the origin and destination station access. Reliability has been critical to cultivating steady ridership and gaining a positive image in the community. This reliability has been maintained and the service area of the TRE extended through coordination with the Dallas Area Rapid Transit authority which provides local bus, light rail, and circulator service to the TRE terminals.
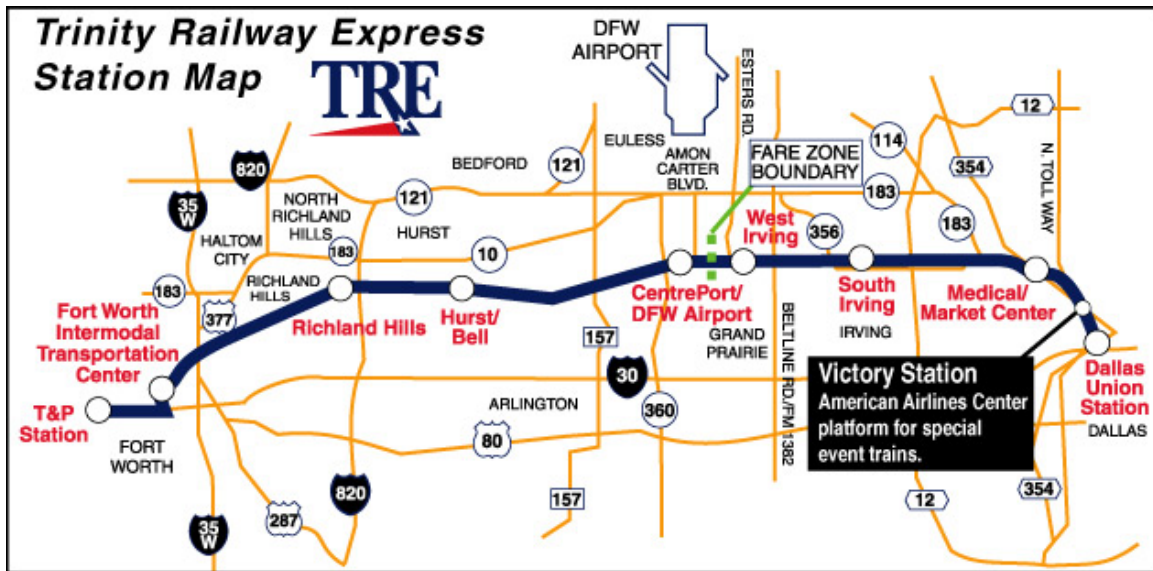
FIGURE 4: TRE SYSTEM MAP

Source: http://www.trinityrailwayexpress.org/map.html

MBTA Fairmount Line – Boston, Massachusetts

The Fairmount commuter rail line in Boston has a history extending back 150 years, though its current configuration dates to 1979 when it was restored to service as an alternative means of travel during construction on other lines in the Southwest Corridor of Boston. The Fairmount line is underperforming to a significant degree from a commuter rail perspective, though it is crucial to the operations of the Massachusetts Bay Transit Authority (MBTA) providing access to storage and overflow facilities as well as alternative routing for other commuter lines. The study undertaken by Nelson, Duse-Anthony, and Friemann (2005) was intended to make suggestions to improve Fairmount Line ridership, though for the purposes of this report it will serve as an example of commuter rail in a much more mature system, and the many evolving commuter rail roles.

The Fairmount Line is short, only 9.2 mi, and serves Boston only, no suburban communities. This alone distinguishes the Fairmount Line from many of the new CRT lines being built today which focus primarily on drawing the commuter from the suburban communities. The line has only weekday service, and as noted before it has consistently poor ridership. The poor performance is likely due to many demographic characteristics of the area served by the line, as the line serves an "archetypical inner-city, low-income, minority neighborhood" (Nelson, Duse-Anthony, and Friemann 2005). While low-income neighborhoods may usually tend to use transit more often, the nature of commuter rail service often did not meet the neighborhoods' travel patterns. Additionally, the original layout of the Fairmount Line intended to serve commuters outside the low-income neighborhoods, but these commuters have been unable to produce the ridership necessary to sustain the line. The existing stations in these neighborhoods were also in poor repair.

MBTA is looking to attract additional ridership from the low-income neighborhoods and the Nelson, Duse-Anthony, and Friemann (2005) study makes several strategic

23

recommendations. Several of their suggestions will make significant changes to the character of the line, so much so that the Fairmount Line may be better considered standard heavy rail transit or rapid rail transit. It is suggested that the line operate on weekends, operate for extended hours, and serve more stations than it currently does (and more than currently exist).

The Fairmount Line serves as an example of the transitional nature of commuter rail service. In older, mature cities and systems, the original intent and extent of commuter rail may not be sufficient to continue as a commuter-specific service. The transition to regular, full-service heavy rail may be a natural, steady progression or it may be the result of acute underperformance requiring immediate restructuring of the service goals. In either case, clearly commuter rail does not remain stagnant within a transportation system. Commuter rail certainly varies between localities, but over time a pattern of transition to heavy or regional rail is common to most systems.

Case Study Summary

The experiences in Austin, Nashville, Dallas, and Boston give a glimpse at the nature of commuter rail in metropolitan transportation systems and the important supporting role played by circulator systems. These experiences not only show the importance of carefully designing the initial commuter rail system but also display its transitional nature. The following section provides a summary of the role and function of commuter rail in general which reinforces the lessons learned from the case studies addressed previously.

ROLE AND FUNCTION OF COMMUTER RAIL RECAP

Several sources (Grava 2003, Gray et al. 1992, Vuchic 1981, Vuchic 2005, and Edwards 1992) stress three key components of commuter rail and its role and function in a transportation system and a community in general. These three components are echoed by the character of the four case studies investigated.

1. Commuter rail operates on existing ROW and track
2. Commuter rail, by definition, operates only during peak commuting hours.
3. Commuter rail has a transitional nature.

Other aspects of commuter rail to be considered include the fact that commuter rail is often viewed as a "premium" service Grava (2003). This perception implies certain creature comforts be provided to serve those who travel via commuter rail. This image may be tied into the transitional nature of commuter rail, as Vuchic (2005) notes that the "trend has been to upgrade commuter rail into regional rail lines or networks." In this stage of rail system development it is necessary to attract choice users to justify expenditures, further improve the image of rail as a means of travel, and increase the political and economic capital available for further expansions of the service and network. The additions and expansions to service will move the rail system out of the realm of commuter rail into a heavy rail transit or more general regional rail system.

TECHNICAL LITERATURE

There are three categories of formulation and solution that are common to the technical literature relating to the CRCNDP: analytical methods, mathematical programming, and the transit equilibrium assignment problem (TEAP). The analytical methods and mathematical

programming methods address network design specifically where the TEAP presents formulation and solution techniques that are concerned with how people travel on an existing capacitated transit route network. In the analytical method assumptions are made to ensure that the objective function is strictly convex. The analytical methods developed in the 1970's, 1980's, and early 1990's using first-order equations to solve a strictly convex objective function is an important step in the evolution of solution methods for the CRCNDP from mostly subjective expert opinion to mathematical methods providing a provable optimal solution.

Analytic Methods

Analytical methods of transit route optimization focus on developing a continuous convex objective function using assumptions that attempt to reasonably represent the transit network and finding first-order equations to solve for optimal stop spacing, headway, frequency, or other route characteristics. These methods tend to use simplified networks that conform to a geometry that can be managed with analytical techniques and make broad assumptions about the transit service and transit demand that are very restrictive. Ceder and Wilson (1986) suggest that analytic methods are useful for policy decisions but have little practical application in transit network design.

George Newell (1979), in an early work on the optimal design of bus routes sets the tone for much of the work in transit network design using analytical methods. A particular aspect of the problem that Newell discusses is the nonconvexity of the objective function for most transit optimization, that is, the higher the demand for trips on a route, the better the service one can provide. He suggests that the nonconvexity can be attributed to the waiting and transfer costs, which are not associated with the links of the bus route system. Newell assumes large, uniformly distributed origin and destination zones. He states that the purpose of his paper is not to determine optimal routes, but to illustrate the sensitivity of optimal geometry to the nature of trip distribution.

Newell also describes the general data and costs that must be accounted for (and known) to use his analytical method:
- Rates at which people want to make trips between various origins and destinations.
- Cost per mile operating the bus
- cost per person-mile of access(walking)
- cost per person-mile of riding the bus
- cost per unit time of waiting for the bus
- transfer cost (exclusive of waiting cost)

Newell makes an interesting point regarding the walking portion of the trip. He comments that passengers tend to prefer moving continuously toward their destination. This suggests that perhaps costs associated with walking from a stop to a destination should accommodate the preference to have a route that is continuously moving toward the destination. That is, the cost of the route should capture the preference of a rider to get off at an earlier stop to walk to the destination (say, for 6 blocks) rather than riding on the bus past a destination, then backtracking for 4 blocks, even if the backtracking route results in an overall lower travel time. This particular preference certainly has its limits and may be difficult to accurately capture.

Kocur and Hendrickson (1982) seek to address the problem of optimal transit route design without assuming that demand is fixed and is instead sensitive to the level of service provided by the transit system. They seek to better utilize the limited resources of the transit

service provider, that is, the objective of the work is to provide optimal service subject to the fleet or budget constraints of the transit operator.

There are restrictive assumptions associated with their approach, the most notable being the assumption of an infinitely fine rectangular grid network. This assumption is necessary to apply this method which uses calculus to find unconstrained optima and then applies Lagrange multipliers when certain constraints are known to hold.

The authors use a linear demand model as an approximation of a logit model for computational tractability. This works well to maintain the convexity of the objective functions that are created. Additionally, the authors present three types of objective functions: maximizing operator profit (minimizing operator deficit), the maximization of net user benefit, and a combination of the two.

Kuah and Perl (1988) offer another paper that applies a differentiation-based method of bus route and stop spacing optimization. Many of the same restrictive assumptions about the network layout are made to make the problem solvable with analytical methods. Interestingly, this paper addresses specifically the problem of designing an optimal feeder bus network to access an existing rail line. Three design variables are solved for: route spacing, operating headway, and stop spacing. Route spacing is relevant only in the context of the assumed geometry that is common to this method of feeder bus route optimization. This method seeks to define the optimal spacing of feeder routes, given that the routes are perpendicular to the commuter rail corridor. Also, the routes in these analytical methods are usually limited to traveling the same links outbound and inbound, circular routes or deviations from the outbound route are not permitted.

The authors make a good point that the assumption of fixed demand for the feeder bus system is applicable when the demand is affected primarily by the level of service and the ridership of the rail service. This is true to a certain extent, but the definition needs to be expanded to incorporate how the feeder bus itself will play a part in determining the level of service of the commuter rail. Access to and from the train portion of the commuter rail trip must be considered.

Kuah and Perl summarize the contribution of this paper: "With little computational effort, the proposed model can provide approximate values for the design variables as a function of system parameters and demand density." This is true, but one must be willing to accept the assumptions of the work. The values obtained through this method will be rough approximates at best; as with nearly any solution method, one trades simplicity for reliability.

Chang and Schonfeld (1991a) attempt to optimize bus service with time dependence and elasticity in their demand characteristics. The term "time-dependent" refers to different bus departure times. They vary demand by time of day, optimizing routes for peak and off-peak periods. They also consider four different types of demand conditions: steady fixed, cyclical fixed, steady equilibrium, and cyclical equilibrium.

The geometry assumed in this paper is not ideal for modern vehicle routing optimization, as it does not correlate well with the street network. A problem with the method is that it requires the engineer to partition the network into defined service areas that will be serviced by only one route prior to analysis. This practitioner judgment can certainly be useful in developing solutions that incorporate aspects of the problem that are difficult to include in a model, however, this subjectivity may introduce biases that yield a suboptimal solution.

The authors assume a linear demand function with parameters that are intended to serve as a measure of the elasticities; however the authors state that they are not direct measures of the elasticities of such a linear model. The authors state that "profit maximization" and "social welfare maximization" are the most appropriate objective functions for private operators and

public agencies, respectively. Social benefit is defined as the users' willingness to pay minus the total cost the users actually pay. Headway and route spacing are optimized in the fixed demand situations using differentiation.

Spasovic and Schonfeld (1993) continue work in the analytical arena under different assumptions to maintain the convexity of the objective function. This work focuses on determining optimal length of transit routes extending radially into the suburbs. Optimal route spacing, headway, and stop locations are also determined subject to the definitions and restrictions related to the hypothetical transit network.

The assumptions in this work are shown below and are illustrative of the type of assumptions that tend to accompany analytical methods. Many of the assumptions are reasonable, especially with the computational improvements they provide. However, the assumptions about the location, behavior, and distribution of transit demand make any results subject to much qualification. An assumed uniformly distributed demand is much more likely to result in uniformly distributed bus stops, something that may not particularly coincide with the demand distribution in reality. Common analytic method assumptions are:

- Corridor is served by transit system of n parallel routes of uniform length L, separated by lateral spacing M,
- Routes extend from CBD outward,
- Transit demand is uniformly distributed throughout corridor, time, and is insensitive to service,
- Commuter travel pattern is many to one or one to many, with the one being CBD,
- Alight and exit only at stops,
- Dense rectangular grid network,
- All vehicles serve all stations,
- Walking is only access mode, access speed is constant,
- Wait time = ½ headway. Headway is uniform along parallel routes,
- No infrastructure costs,
- Demand is less than capacity, and
- No limit on fleet size.

This paper develops a nonlinear objective function for the assumed network and problem properties listed above. A capacity constraint (only) is included in the constrained optimization problem. Authors propose an algorithm to solve the problem in a method superior to a "penalty method". Their algorithm fixes three decision variables and optimizes the problem for the one remaining free variable. The optimal values of the free variable are then fixed and applied to optimizing a different variable. This is repeated "until it converges on an optimal solution".

Chien and Schonfeld (1998) attempt jointly optimizing both a rail transit line and a feeder bus system designed to serve this rail transit line. The necessary assumptions are made regarding demand distribution and geometry to maintain the convexity of the elaborate objective function that is developed. The optimization method developed for this particular problem solves for rail line length, rail station spacing, bus headway, bus stop spacing, and bus route spacing. As with much analytic optimization work, these measures suffer from ambiguity.

Chien et al. (2001) later applied analytic methods to a comparison of conventional bus systems (fixed route) and subscription bus service (dial-a-ride). Their work optimized a strictly convex objective function for service zone size and vehicle size. Interesting cost functions are employed, including a linear function of bus seats for bus operator cost and a nonlinear function for nonadditive travel time. The concept of nonadditive travel time (a topic detailed in Gabriel

and Bernstein (2000)) accounts for the tendency for travelers to view a single 10-minute period of time differently than 10 1-minute waiting periods, that is, nonlinear cost functions. These cost function may play a role in the CRCNDP which can be formulated to accommodate varying levels of modeling complexity. This nonadditive cost and representing demand as an average and associated standard deviation are what distinguish this later work from Chang and Schonfeld (1991b), in which fixed demand and additive costs are used.

Mathematical Programming Methods

Mathematical programming methods of addressing transit network design problems (TNDP) have evolved over the past 30 years due to improvements in programming techniques, heuristic techniques, and computational power. The TNDP that are addressed in the following papers are nearly all NP-hard problems and the papers discussed present different heuristic methods of solving the problem. The few papers that have addressed the CRCNDP specifically have been able to solve small networks to optimality, but these solutions are often limited by the formulation used to represent the CRCNDP.

The most appropriate place to start discussing mathematical programming methods is at the genesis of a key concept in network design problems. Wardrop (1952) proposes the concept now known as "Wardrop Equilibrium". This equilibrium addresses the situation in which alternate routes exist for a traveler and the means by which one can theoretically distribute traffic flow between the alternatives.

Wardrop proposes two criteria based on journey times that can be used to determine the distribution on the routes, which provide succinct definitions of the objective function options available to transportation network analysts:

- The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route – User Equilibrium (UE)
- The average journey time is minimum – System Optimal (SO)

Wardrop does not specifically address TNDP, circulators, or commuter rail in any manner. However, this seminal work is critical in the development and understanding of traffic network design and traffic assignment algorithms that are commonly used today.

Moellering, Gauthier and Osleeb (1977) provide an early bus transit route planning effort using early-generation geographic information tools and mathematical programming. Their interactive computer graphics-based system acknowledged that attracting choice users is a key difficulty of a transit planning effort and that providing information to the potential riders is crucial to the success of the system. Additionally, they incorporated a walking distance parameter in their formulation, recognizing that the value of time placed on the walking portion of the trip is different than in vehicle time (it also varies between users).

Stein (1978) discusses how a bus routing problem is a variant of the TSP, just as the CRCNDP. However, Stein is addressing dial-a-ride service with randomly generated origin and destination demand. His problem is many-to-many over a predefined service area. Stein presents an heuristic method of addressing this problem by breaking the coverage area into many smaller subregions, finding optimal tours for these regions and then connecting these tours to other regions. His method of using smaller subregions has applicability in developing good solution methods for the CRCNDP.

Dubois, Bel, and Llibre (1979) addressed the TNDP focusing on bus routes in a medium sized city. Their approach is typical of many efforts to date and displays the obstacles and challenges that must be overcome in any mathematical programming approach. Their approach

uses three sub-problems: choosing the set of streets, selecting the set of bus lines to serve the street network, and then defining optimal service frequencies for these bus lines. The candidate street set and bus line set are developed with greedy heuristic procedures.

Limitations of this work arise as demand is assumed generated from centroids and travel time is evaluated based purely on the physical length of the street sections, though these are not overly restrictive assumptions. The authors also consider a particular budget for modifying the existing routes suggesting that this method is not necessarily tailored to designing a new route layout. In 1979, the technology to solve even a small network problem was difficult to procure, and the authors discuss the futility of trying to optimally solve a network with even 50 nodes.

The authors define a set of shortest paths that represent a set of bus lines that provide a direct route for every O-D pair with no transfers. This set is then reduced by creating subsets that combine very short paths into a single path, and then removing duplicate paths. They then propose to evaluate the travel times of passengers using lines in the network in comparison with the optimal path. This heuristic procedure selects bus lines in a logical, but somewhat qualitative, manner (accomplished by an engineer or planner who considers resource limitations) to add to the proposed network and then computes the difference in travel time between the proposed route network and the travel times if each O-D pair were able to travel via its shortest path. The number of transfers needed for the particular proposed route layout is also computed. The engineer or planner continues to modify the route network until some threshold values of travel time sub-optimality and number of connections is reached.

The assignment portion of the work (to optimize frequencies) assumes that a traveler is choosing only between bus and walking modes. The work also assumes constant travel times for each. This work does attempt to capture the effects of mode choice; however the restrictive assumptions and the particular method used create limitations on the applicability of the results. This early work contains several elements that are still sources of challenges for today's analyst:

- Choosing candidate routes out of a near-infinite population,
- Manageably representing the street network and mapping routes to the network, and
- Representing the variability of modal travel time and the impact of this variability on demand/ridership and the optimal routing.

These challenges continue to surface in later work and they also highlight some of the challenges that are being addressed in this report.

Ceder and Wilson (1986) focus on the bus network design problem (BNDP) and look at previous work while proposing an "easier to implement" method in terms of data acquisition and analytical complexity. The authors consider both operator and passenger interests in redesigning an existing bus transit system when developing the objective functions and constraint sets.

The authors cite the lack of mathematical programming work directed at the BNDP and its relationship with transit demand in 1986. The paper mentions that the majority of the work prior to 1986 is focused on driver and vehicle scheduling, suggesting that researchers were attempting to help the transit authorities maximize their revenue, a worthy, if elusive, objective. However, it was apparent to the authors (and is still today), that there is little chance that a transit agency is going to be a profit-generating entity. Therefore, the transit agency's motive is not then to generate revenue, it is to provide the very best service within a certain budget; a more passenger/customer focused approach.

They discuss that using a fixed demand table, and updating the demand after each network design iteration, is the most reasonable means of network design in a practical, implementation sense. Authors state that variable demand in the formulation may cause the optimal solution to the variable problem to be one with little, if any, bus service. The paper also

addresses several aspects of the BNDP that are specific to large scale bus system management. One example is the two-stage partition of the optimization work, looking at frequency determination and route design separately. This is out of practical necessity for large networks with much leeway in headway times. For the CRCNDP with seamless transfer to circulator routes, the question of frequency is already fixed by the frequency of the commuter rail, and the problem is therefore reduced to the one stage: route design.

Kuah and Perl (1987) provide a mathematical programming approach to the Feeder-bus Network Design Problem (FBNDP), characteristics of which are:
- It seeks to minimize operator and user costs,
- Many-to-one demand pattern generalized to many-to-many,
- Heuristic solution approaches are employed,
- Two node types are used: rail stations and bus stops,
- It assumes demand concentrated at nodes,
- The temporal distribution of demand not considered, and
- It considers the home-end FBNDP, that is, home to rail station.

The authors provide a formal definition of the FBNDP: "the problem of designing a set of feeder-bus routes and determining the service frequency on each route so as to minimize the sum of user and operator costs". The authors make the following assumptions in formulating and solving the FBNDP, assumptions that carrying through to this report work:
- Each bus stop serviced by only one route
- Each route is linked w/ only one rail station
- Buses have standard capacity and operating speed

Kuah and Perl provide an early effort at describing and solving the FBNDP in mathematical programming terms using a spanning tree network representation. The solutions developed by their heuristic method are compared to the average performance of existing bus networks which may be operating on much less than optimal networks. Using existing routes as the baseline for comparison and calibration may result in a method that provides solutions that resemble current transit routing practice and provides a good "sanity" check of the solution, however, this comparison imbues little confidence in the nearness of the heuristic solution to the true optimal solution.

Ceder and Israeli (1998) deal specifically with objective function construction for the TNDP taking into account both passenger and operator cost. The formulation proposed is nonlinear mixed integer. The proposed method generates all possible transit routes and pares these down to smaller subsets that maintain connectivity. The optimal subset is then selected based upon user-defined parameters.

The authors describe four basic elements of the transit planning process that are performed in sequence and iterated: network design, setting timetables, scheduling of vehicle trips, and the assignment of drivers. The authors also point out shortcomings with the current math programming methods available:
- Cannot handle large networks,
- Do not consider optional objectives and constraints,
- Determine frequency using economic parameters rather than passenger counts,
- Cannot incorporate 3 of the 4 planning components simultaneously,
- No consistent good results, and
- Cannot incorporate qualitative constraints.

Ceder and Israeli use a set covering problem to create a minimal set of routes and transfers that maintains demand coverage, a method that has influenced many of the recent approaches to the TNDP and related problems. The authors introduce innovative costs into the objective function, such as a cost of empty seats, or unused capacity. The proposed method consists of seven separate yet interrelated components that are considered fundamental parts of solving the TNDP or related problems.

Jerby and Ceder (2005) concentrate on collecting passengers and distributing them to a rail station at the home end of a rail trip. The purpose is to decrease the reliance and burden on park and ride facilities. While this addresses the opposite end of the commute trip that the CRCNDP will be designed to address in this report, many aspects of this work are applicable.

Jerby and Ceder acknowledge that in this formulation not all demand will be satisfied, a cost that should be considered in the CRCNDP. The authors use average travel time on road links, assume no congestion impacts on travel time, and assume no variance associated with travel time. This particular method does not include a feedback loop between optimal routing and demand estimation though the relationship is included in the framework the authors describe.

A seemingly obvious, but good contribution of the paper is the initial step of first creating a base network, eliminating some links if they are absolutely not going to be used for transit (radii too small, too narrow, no outlet, etc.). This step is included as a manual exercise, a useful extension may be the automation of this step.

The authors develop an average walking distance measure based on the average distance one has to walk from buildings/demand centers within an approximate ¼ mile polygon around a link. This distance does not address bus stop location and it considers all people/residents to be potential users. They also create a demand potential index that is the ratio of the population of the link polygon and the average walking distance of that polygon. They conclude that this measure adequately leads toward the goal of maximizing response to potential demand (population) and minimizing walking distances (average walking distance). This measure assumes demand potential is homogeneously distributed along links. These measures are then incorporated into their formulation.

The authors use a straightforward model formulation, maximizing the demand potential index, subject to traveling salesman problem (TSP) constraints addressing node degree and subtour elimination. They employ a simple, but interesting means to dissuade the model from choosing a route that travels back the opposite direction on the same link if that link has high demand potential. For a link (i, j) included in the optimal routing they simply multiply the link (j, i) demand potential by 1/5.

An impedance ratio, which is a ratio of travel time to demand potential is used to represent the desirability of the shuttle route. A constant is used in the numerator and denominator of the impedance ratio that can potentially be used to massage the impedance ratio in any fashion that is desired. The heuristic method developed takes a greedy approach, selecting some alpha number of links with which to construct circular routes and compare the objective function value of that route to previous routes. The authors state that this algorithm is equally applicable to distributing passengers at a terminal, which is likely a true statement, subject to some necessary modifications.

This is a straightforward algorithm that solves a version of the CRCNDP. The formulation can be made much more sophisticated with the inclusion of unserved demand costs, mid-block stops, and walking trips. However, this method may serve as an interesting basis for comparison to determine if more sophisticated results do in fact result in better solutions.

Hadas and Ceder (2007) continue work on the BNDP investigating the aspect of transit systems that is very desirable from a transit network operational perspective and very undesirable from the passenger's perspective: transfers. Much of the undesirability of the transfer from the passenger's perspective stems from the uncertainty involved with a typical transfer. Currently, timed transfers are designed based on single-point (bus stop) encounters, which exacerbate the uncertainty inherent in bus arrival times. Hadas and Ceder employ a multi-agent approach to look at the timed transfer concept from a road segment perspective, reducing uncertainty in transfer coordination and improving operational flexibility. This work is methodologically interesting, however its applicability to the CRCNDP is limited as transfers are not allowed in CRCNDP solutions and transfer coordination at the rail station is a specific constraint of the formulation, itself seeking to minimize the uncertainty of the rail-to-bus transfer.

Wei Fan (2004) presents an extension of the standard mathematical programming formulations to more sophisticated formulations that account for the walking portion of the trip and non-centroid demand centers. This work presents a discussion of the formulation development and heuristic solution methods used to solve generated instances of the problem.

Fan uses a 3-step approach that has been adopted by most work on the TRNDP and its relatives: an initial candidate route generation phase, a heuristic search procedure that iterates with a network analysis procedure, and an output phase. The route generation phase generates a candidate set of routes for a particular demand configuration, the heuristic looks for an optimal route set, and the analysis procedure assigns demand to this optimal route set. This approach is necessitated by the size of the problem created when dealing with large transit networks with many-to-many demand.

Pattnaik, Mohan and Tom (1998) seek to find optimal routes and frequencies minimizing total cost using genetic algorithms (GA) for the BTRNDP. GA is a high-level simulation of a biologically motivated adaptive system. There are two GA methods: fixed string length coding and variable string length coding. The fixed method will require an extra iteration to determine the optimal number of routes to be considered, the variable method does not require this extra iteration. The proposed method utilizes the familiar 3-step procedure that includes the candidate route generation phase.

The authors make an assumption that demand is fixed and independent of service quality. The authors also develop a fitness function that is used to measure the goodness of a solution, or nearness to an optimal solution and is used in the reproduction, crossover, and mutation phases of GA. The new population of solutions (routes) is then generated using these fitness function values which are assigned to the entire previous solution population. A common fitness function is shown below:

$$F(i) = V - \frac{O(i) \cdot P}{\sum\limits_{i=1}^{P} O(i)}$$

Where $O(i)$ is the objective function value of the ith route, P is the population size, and V is large enough to ensure nonnegative fitness values. Reproduction makes copies of the better routes from the previous population in the new population. Crossover creates different routes by combining two routes from the previous population. Mutation adds new information in a random way to the current routes to avoid getting trapped at local optima.

One aspect of the TNDP that has been discussed little up to this point is the concept of variable demand, or at the very least, service-sensitive demand. Lee and Vuchic (2005) present a recent work attempting to address the TNDP considering variable demand. This paper provides

a definitional distinction between "fixed total demand" and "variable transit demand". This distinction describes the situation in which a comprehensive regional commuter O-D matrix is known and fixed and the transit mode share, specifically the commuter rail share, varies according to the route structure and the service that it provides. The authors discuss the shortcomings of the combinatorial approach of most previous work that requires a candidate route set generation phase. This phase requires that a limit be placed on the number of routes generated and this limit is based on the knowledge of the analyst. The authors' proposed algorithm starts with the minimum in-vehicle travel time network using Dijkstra's algorithm and increases travel time while reducing wait time. They use a logit model to predict the mode share based on a straightforward utility function. They assume that network auto and transit travel times are known and fixed.

The objective function developed in this work seeks to minimize travel time, maximize profit, and maximize social benefit. The study does not incorporate all of the costs in one objective, nor does it provide a single optimal solution. It produces multiple solutions from which a planner must choose, a concept discussed later as pareto-optimality.

To calculate optimal service frequency once optimal routing is determined, demand frequency is used in the form of a ratio proposed by Vuchic in 1976 and used by Ceder in later work. It is in this ratio that a key nonconvexity arises; as the demand increases, the service actually improves because of the increased frequency of the buses. Therefore wait time is significantly reduced, resulting in better travel times for at least some portion of the riders.

Transit Equilibrium Assignment

A separate, but related topic is the transit equilibrium assignment problem (TEAP). The TEAP is not a network design problem; it assigns transit trips on a pre-existing transit network based on the concept of trip strategies. This is very obviously a relevant topic as the TEAP is seeking to better characterize the impacts on ridership of various aspects of transit routes and transit trips. The TEAP is primarily applicable to transit systems in which transit capacity is an issue.

Spiess and Florian (1989) provide the formulation that the most recent TEAP work is based upon. TEAP is addressed in this paper by allowing travelers to choose the strategy that allows reaching their destination at minimum expected cost. Later work using this TEAP strategy is seeking to develop better solution methods for the transit assignment problem formulated in this manner.

The authors note that walk links can be replaced by a transit link with zero waiting time, creating more consistency in link types and potentially reducing the formulation complexity. They also make the assumption that the traveler only knows which line is going to be served next at a stop, that is, the traveler can choose to board or not board a particular vehicle. The concept of strategy is a set of rules that enable the traveler to reach a destination from any node in the network. The authors treat their problem primarily in the simplified linear fashion; however, they do extend their work to the generalized nonlinear version as well. The techniques used by the authors are based upon earlier work by Spiess and Florian (1982) in which they provide sufficient conditions for the convergence of diagonalization algorithms.

The work by Spiess and Florian is continued by Wu, Florian, and Marcotte (1994) in their efforts to improve the TEAP formulation and the methods used to solve the TEAP. As in the previous work, the assumption is made that passengers know travel times, but the only information available while they travel is which bus line arrives first at a stop. A strategy consists of a set of deterministic rules which determine the random route taken. The authors

distinguish between zone centroid (demand generating) nodes and bus stop nodes, having the links between zone centroid nodes and bus stop nodes represent walking trips. The arcs do not represent physical structures (roadways/sidewalks) they symbolically represent the portion of the transit route between the two nodes.

The authors also introduce the concept of a hyperpath in the context of TEAP: a hyperpath is defined as a directed acyclic graph with a flow distribution rule. A hyperpath k which connects an origin p to a destination q is a subgraph $G^k = (N^k, A^k)$ of G. It is important to note that in this work the method starts with a set of feasible hyperpaths, otherwise defining the set of hyperpaths will likely lead to an exponential formulation. The algorithm is polynomial solvable in K, the set of hyperpaths, which without some preprocessing to determine a set of feasible hyperpaths will render the method exponential. Additionally, the costs utilized in this and other TEAP methods are not intended to be robust cost functions that properly capture the behavior of transit travelers. The cost functions are used such that they allow for better manipulation and examination of the TEAP formulation and solution method. This work is extended to networks with vehicle capacities in later work by Hamdouch, Marcotte, and Nguyen (2004).

Constantin and Florian (1995) continue the search for practical solution methods to the complex TEAP. They convert a nonlinear, nonconvex, MIP into a bi-level min-min nonconvex optimization problem. This conversion is undertaken in an effort to make the problem more solvable, and this tactic works under certain conditions. If the second-level problem is polynomially solvable, as in the case of this paper, then the conversion has many benefits, if not, then the conversion is not likely worth the additional effort.

SUMMARY

There is a large amount of literature existing on transit network design, transit routing, and transit equilibrium assignment. However, the challenges that initially prompted many of the researchers to expend their effort addressing these problems still exist, even though significant advances have been made over the past 30 years. Formulating these transit-related problems is still an issue. In TNDP, accounting for the walking portion of trips, the impact on demand, and the inclusion of congestion are all still significant issues that need work. The TEAP has made headway by redefining the problem in terms of hyperpaths; however, determining the initial set of these hyperpaths can be an exponential undertaking itself. Practical solution methods for both of the problems are still needed. Well established metaheuristic methods have been applied in several studies and several heuristic procedures have been developed in others to solve various problems. The simplifying and limiting assumptions of these problem formulations still leave much room for improvement in solution method, speed, and quality.

Little work has investigated the CRCNDP specifically and there is much room for further investigation into better formulations and solutions methods for this problem. Some issues associated with TNDP and TEAP are found within the CRCNDP in addition to those associated with the higher resolution needed when formulating and solving the CRCNDP.

The next chapter will look at the CRCNDP formulation for this report and discuss each of the formulation elements. Much attention will be given to proper characterization of the walking portion of the commuter rail trip, bus stop location, and aspects of the CRCNDP that can be exploited to improve solution efficiency.

CHAPTER 3: FORMULATION

The CRCNDP seeks to optimize the route design of a circulator system specifically serving commuter rail stations. Two portions of this problem are designed to be completed simultaneously: optimal bus stop selection and route optimization for this set of stops. Seamless transfer from the commuter rail portion of the trip to the circulator is enforced, requiring that a bus for each circulator route be present at the rail station when a commuter train arrives at the station. A commuter O-D trip table (all modes) is assumed known and fixed; the commuter rail share of this commuter demand is also considered fixed, while the commuter rail demand served is variable depending upon the layout of the commuter rail circulator system, that is, the accessibility of commuter rail.

The CRCNDP can be formulated in multiple fashions, each consisting of its own set of beneficial and detrimental aspects. As with any formulation there are inherent tradeoffs between modeling precision and computational effort. The ideal formulation will maximize the precision of the model while minimizing the effort required to solve the problem. Of course, this is an elusive goal and there may not be a single formulation and solution method that reaches these ideals. Following is the formulation of the CRCNDP used as the cornerstone of this report.

FORMULATION

The formulation of the CRCNDP uses a network representation similar to that shown in Figure 5. Candidate bus stops are represented by the red nodes at the mid-block and intersection locations. The demand centroids in the Figure 5 idealized grid network are located at the center of each zone. The station is identified as node 1 and is highlighted in green. Obviously, real-world networks will not be organized identically to this figure; however, it is useful for demonstration.
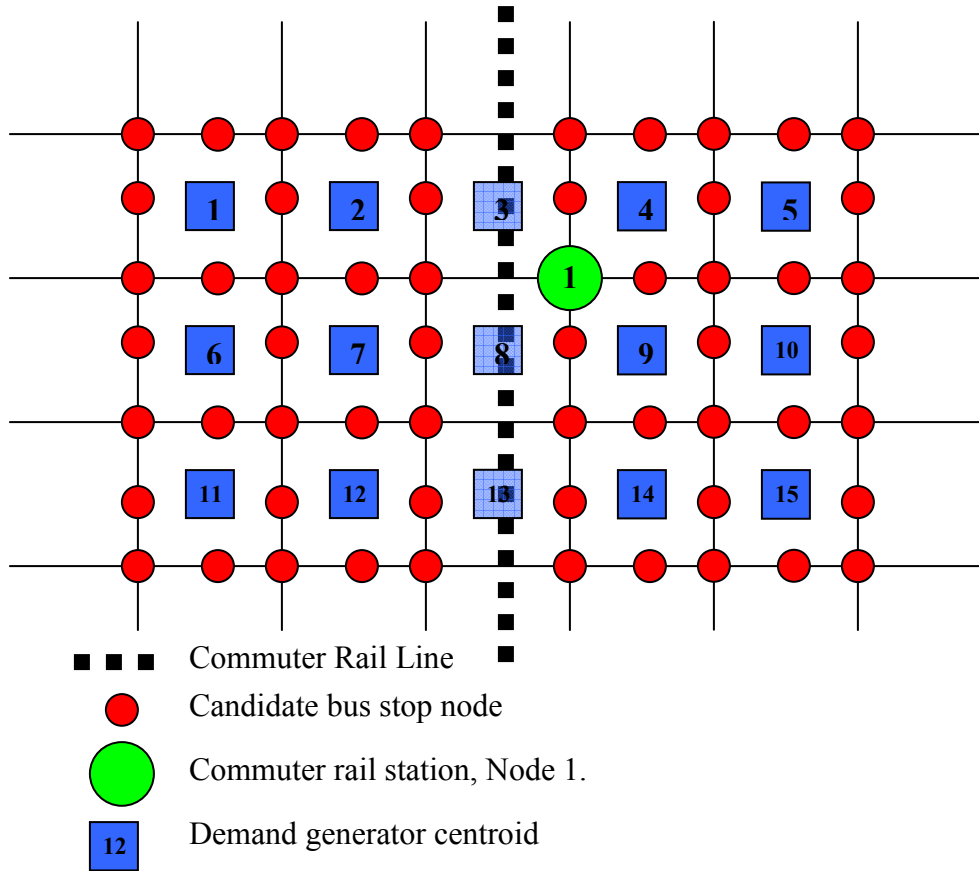
FIGURE 5: NETWORK REPRESENTATION EXAMPLE

A formalization of the CRCNDP with associated descriptions of the formulation components will now be provided. This formalization includes the appropriate sets/indices, parameters and data, assumptions, the objective function and constraint set.

Sets/Indices

$i, j, k \in I$	Candidate circulator stop locations
$g \in G$	Circulator demand centroid locations
$r \subseteq I$	Subset of demand locations/nodes to visit in circulator route

The above sets completely describe the network that will be used in representing the CRCNDP. The set I contains all candidate stop locations for the circulator route, which may be selected arbitrarily or using some method for candidate stop placement such as Murray's (2003) method. This set does not include the demand centroids, which are maintained in a separate set, G. The set G is defined as the demand activity centroids of the demand zones used in the analysis. Note that this definition is sufficiently broad to incorporate many different sized demand zones. If a high degree of resolution is desired and demand is estimated at, say, the block level, this representation is capable of this degree of resolution. However, in the likely case that reliable demand estimates at this fidelity are not available, the representation is

36

sufficient to accommodate demand estimates at the Traffic Survey (or Serial) Zone (TSZ) or Traffic Analysis Zone (TAZ), which are used interchangeably (TAZ will be used in this report from this point forward).

The subset r is the set of candidate bus stop nodes for which the CRCNDP will be solved. This subset of all possible stops is that which will be visited in a particular circulator route. In enumerative algorithms, this set will be externally generated and the CRCNDP solved for each generation of the subset.

Assumptions

- No congestion or incidents, travel time is consistently proportional to the distance between nodes.
- Demand is located at predefined zone activity centroids.
- Demand represents the general commuter trip attractiveness of a zone and is fixed for the peak period.
- Each route starts and ends at commuter rail station (Node 1).
- Circulators are not capacity constrained

The first assumption simplifies the problem somewhat by eliminating the inherent stochasticity of travel times in a transportation network. This assumption is akin to using the expected value of travel time without consideration of variance. This assumption may be relaxed in future work to provide assurance of the reliability and robustness of the circulator route design. The next two assumptions address the demand used to drive the formulation. This demand data source will likely be aggregated, meaning that one does not know individual trip trajectories, but the general attractiveness of the destination zone. Of course, the demand estimates will depend on the route configuration, as the route configuration significantly impacts the accessibility of the commuter rail line and impacts the specific travel time to a destination zone. However, the formulation is taking into consideration the prospect of not serving all demand and the demand data is being used to represent the desired trips to the destination zones. Therefore, it is reasonable in this context to assume that the desired relative number of trips (based on accessibility at the home end) will remain constant during the planning stages.

The fourth assumption is tying together the seamless transfer concept and invoking a TSP-type problem representation. Each and every route that serves the station must begin and end at the station, which is an essential component of both the seamless transfer concept of the CRCNDP and the related TSP. Finally, the assumption that circulators are not capacity constrained is a reasonable assumption in an American transit system, especially a new service. This assumption essentially states that the transit authority will provide sufficient circulator capacity to serve all commuter-rail passengers. The formulation is designed to accommodate this assumption as the fleet of buses, F, available to serve a station is an input parameter for a single route or can easily be considered for multiple routes.

Given Data

C_o cost of operating bus in \$/hr

C_{tt} traveler cost of in-vehicle travel time in \$/hr

C_d cost of unserved demand in \$/unserved passenger

C_w cost of walking in \$/hr

d_g demand for service at demand centroid g

λ_{ij} rectilinear (or shortest path) distance from node i to node j

γ_{ig} rectilinear distance from node i to demand centroid g

H commuter rail train headway

F number of buses available to the route

v_{bus} local bus operating speed

v_{walk} pedestrian walking speed

The first four given data elements are very important: the values of the various cost types used in the multi-objective formulation of the CRCNDP. The default, or base values of these parameters rely heavily upon the Transit Cooperative Research Program (TCRP) Report 78 (ECONorthwest and Parsons Brinckerhoff Quade & Douglas, Inc. 2002), which gives practical values for the cost parameters. The exception to this is the cost of unserved demand. Unserved demand is a difficult cost parameter to assign, a very wide range of values could be argued as proper. It is because of this difficulty that the unserved demand cost will be the subject of much of the sensitivity analysis later in the report. For the default/base value of the cost, the value given by Bailey (2007) of \$6,000 per household will be used to represent the cost of unserved demand. Bailey's cost is figured based on the difference between maintaining a single-vehicle and a two-vehicle household. Her assumption is that providing transit service to a household allows that household to eliminate one vehicle and reap substantial savings. Not providing service to this household therefore denies the household of these savings by forcing a two-car household.

The demand data, as discussed previously, is considered given for this formulation. The assumptions discussed above are applicable to this data. After demand data, rectilinear distances between node-node pairs and node-centroid pairs are required. These distances need to be separated as they are used in different aspects of the formulation. The node-centroid distance is used in the computation of walking costs and walking coverage of the various demand zones. The node-node distance is used in the route design portion of the solution which optimizes the order in which nodes are visited.

It should be noted that this node-node distance would be better represented by a shortest path distance between the nodes, though rectilinear distance may serve as a good proxy. Ideally, λ_{ij} will be determined beforehand for all node pairs using an all-pairs shortest path algorithm such as the Floyd-Warshall algorithm, as described by Ahuja, Magnanti, and Orlin (1993). If a set of such shortest paths were maintained, the solution would be much more precise than using rectilinear distance, and would guarantee that the final solution would map exactly to the street network. These shortest paths would need to be determined as a preprocessing step to ensure this mapping as solving the CRCNDP over the set r would allow only nodes in r to be used in the

solution. Direct paths may not exist between all nodes in *r*, in which case nodes that are not in the set *r* would need to be used to find an optimal route design. Cieslik (2001) notes that this problem arises in many network design problems in which a "minimum spanning network of all points is often shorter than the minimum spanning network of the given points alone." Cieslik is describing Stiener's Problem, which attempts to find a shortest path between a subset of nodes using all available nodes from the population. Maintaining an all-pairs shortest path distance in the CRCNDP avoids the difficulties of Steiner's Problem and ensures direct mapping of the transportation network used in the all-pairs shortest path computation.

The commuter rail headway and station fleet size will be necessary to solve the CRCNDP. Typical values for *H* range from 15 – 30 minutes and the fleet size, *F*, is a product of the capacity assumption described earlier, that the transit authority will provide sufficient circulator capacity to accommodate all commuter rail passengers. The final two values for bus speed and walking speed rely on Levinson (1983) for bus operating speeds and the TCRP Report 78 (ECONorthwest and Parsons Brinckerhoff Quade & Douglas, Inc. 2002) for walking speed. The default values for these parameters are 10 mph and 2.5 mph, respectively. The bus operating speed is estimated for city routes (as opposed to CBD or suburban routes) and the walking speed is an average value that is typically assumed in the transit planning process.

Calculated Parameters

ρ_{ig} binary parameter indicating whether demand centroid *g* is covered by walking trip from *i*

s_i demand at node *i* that is served/satisfied by route *r*

μ_g binary parameter indicating whether demand centroid *g* is unserved

δ_i dwell time at node *i*

The calculated parameters are given separate from the data simply because effort beyond simple preprocessing is required in providing these specifications. Effort will be required to create these parameters unique to the CRCNDP and their processing requirements are indeed part of the CRCNDP solution methodology. The first parameter ρ_{ig} is a parameter that signifies whether a demand centroid, *g*, is covered by a walking trip from bus stop node *i*. This parameter is developed using a walking distance threshold (usually ¼ - ½ mile) to determine first if a centroid is covered by a bus stop. The parameter is restricted to allow for walking coverage from only one bus stop (that is, passengers will always take the shortest path), so that for a given set of bus stops walking trips to demand centroids are made from the single shortest-distance bus stop. Forcing passengers to take the shortest walking path is in line with Murray's (2003) goal of reducing redundancy in transit networks, who argues that unnecessary redundancy is a great drag on transit network efficiency. The parameter takes the value "1" if a walking trip from node *i* to centroid *g* is the shortest walking trip available within the walking threshold distance and 0 otherwise.

The second parameter s_i is rather simpler to determine. In the current representation with bus stops and centroids separate, s_i adds up the demand that is served via the shortest walking trip origination at the particular bus stop *i*. The third parameter, μ_g is the set of demand centroids that are not covered by any walking trip within the walking threshold. This parameter takes the

value "1" if the demand is not covered and 0 otherwise. The parameter μ_g is then used in the computation of unserved demand cost in the CRCNDP objective function.

Dwell time is computed using a simple linear relationship between dwell time and the number of alighting passengers developed by Levinsion (1983). The particular relationship used stems from observations on an urban Boston bus route that mainly discharges passengers. This particular set of observations by Levinson best represents the general route conditions in which a commuter rail circulator route will operate. As the circulator is operating at the destination end, suburban characteristics are not as applicable. CBD characteristics are inappropriate as well, as circulation of commuter rail passengers in the CBD are more likely to use existing transit or reach their final destination on foot. The expression for dwell time is $\delta_i = 4.0 + 1.7 s_i$ and can be calculated alongside the demand served at each node, s_i.

Decision Variables

x_{ij} takes the value 1 if bus travels from i to j in route r, and 0 otherwise

w_{ij} the number of passengers that travel from i to j in route r

u_i arbitrary real number (similar to node potential) at node i used for subtour elimination

The decision variables used in this formulation are straightforward. As in traditional TSPs, the binary variable x_{ij} signifies if a trip from i to j is made during the tour. The variable w_{ij} captures the number of passengers that travel from i to j for a tour of the subset r. This variable is important to maintain so that accurate in-vehicle travel costs can be computed for each segment of the route and flow conservation is maintained at all nodes within the subset r.

The formulation of the CRCNDP utilizes the subtour elimination strategy developed by Miller, Tucker and Zemlin (1960) and introduces the unrestricted decision variable u_i. This subtour elimination constraint does not provide as intuitive a method of subtour elimination as other common methods seen today that identify and eliminate disconnected subtours, however, it performs very well practically in solving problems with the familiar TSP constraints that are discussed below. This method of subtour elimination was compared with three other subtour elimination strategies and it produced optimal solutions in less time than the other strategies. A further discussion of this analysis is reserved for later in this report.

Formulation

$$\text{minimize } z = \sum_i \sum_g C_w \frac{\gamma_{ig}}{v_{walk}} \rho_{ig} d_g + \sum_i \sum_j C_{tt} \frac{\lambda_{ij}}{v_{bus}} w_{ij} + \sum_i \sum_j C_o \frac{\lambda_{ij}}{v_{bus}} x_{ij}$$

$$+ \sum_g C_d \mu_g d_g + \sum_i \left(C_o \delta_i + C_{tt} \sum_j w_{ij} \delta_i \right) \tag{1}$$

The objective function (1) contains five separate cost components. There are three competing agents represented in these five components: user cost, operator cost, and unserved demand cost. The first two cost components are user costs, walking and in-vehicle travel time, respectively. The first component, walking cost, identifies the demand served by the walking trips and the lengths of these trips, and computes the cost accordingly. The second component, in-vehicle travel time, applies the travel cost for trip segment to each passenger onboard the vehicle during that trip segment.

The third cost component applies an hourly operational cost parameter to each trip segment in the route. This cost component can be refined to represent operator cost on a passenger-mile basis, for which cost estimates are also available. The fourth component applies the unserved demand cost to the total demand that is left unserved by the circulator system. Finally, the fifth component formalizes the dwell time cost incurred at each stop in the route. This component accounts for both the operating cost incurred while waiting and the waiting cost incurred by all passengers that remain on the circulator vehicle.

This multi-objective function is constrained by the following:

subject to

$$\sum_{j:j \neq i} x_{ij} = 1 \qquad \forall i \in r \tag{2}$$

$$\sum_{i:i \neq j} x_{ij} = 1 \qquad \forall j \in r \tag{3}$$

Constraints (2) and (3) are familiar TSP constraints restricting the optimal solution to contain one incoming and one outgoing trip segment for every node in the route subset r.

$$w_{ij} \leq \left(\sum_g d_g \right) x_{ij} \qquad \forall i, j \in r \tag{4}$$

$$s_i = \sum_{k \in r} w_{ki} - \sum_{j \in r} w_{ij} \qquad \forall i \in r \tag{5}$$

Constraint (4) places an upper limit on the number of passengers on a particular trip segment. Constraint (5) ensures conservation of flow at each node in the subset r. The number of passengers alighting at node i is equivalent to the number of passengers arriving at a node on the circulator vehicle less the number leaving on the vehicle.

41

$$\sum_j w_{1j} = \left(\sum_i s_i\right) - s_1 \qquad\qquad (6)$$

$$\sum_k w_{k1} = 0 \qquad\qquad (7)$$

Constraints (6) and (7) set the initial and final trip segment conditions for the route. The first segment contains all passengers that are to be served along the route, (6), and the final segment contains no passengers, (7). Note that the station is represented as node "1" in these constraints, consistent with the assumption stated previously.

$$\sum_i \sum_j \frac{\lambda_{ij}}{v_{bus}} x_{ij} + \sum_i \delta_i \leq HF \qquad\qquad (8)$$

Constraint (8) acknowledges that certain routes may have more than one vehicle available and may need to stagger vehicles to serve a larger route. This may occur when commuter rail train headways are short enough that a single circulator route is able to provide service to only a very few (or no) demand centroids and still maintain seamless transfer. In this situation, multiple vehicles could serve the route with staggered service, allowing for longer routes and better service coverage.

$$u_i - u_j + |r|x_{ij} \leq |r| - 1 \qquad\qquad 1 \leq i \neq j \leq n \qquad\qquad (9)$$

Constraint (9) presents the Miller, Tucker and Zemlin (1960) subtour elimination constraint. The nebulous, unrestricted variable u_i is used along with the cardinality of the subset r for which the route is being designed and the binary variable x_{ij}. Again, this version is being applied because of its performance in test applications.

$$x_{ij} \in \{0,1\} \qquad\qquad \forall i, j \in I \qquad\qquad (10)$$

$$w_{ij} \geq 0 \text{ and integer} \qquad\qquad \forall i, j \in I \qquad\qquad (11)$$

$$u_i \text{ unrestricted} \qquad\qquad \forall i \in I \qquad\qquad (12)$$

The final three constraints are definitional, restricting x_{ij} to a binary variable, w_{ij} to a positive integer, and formalizing the unrestricted variable u_i.

Walking Cost Component Parameter Development

The walking coverage parameter used in this report is developed based upon the rule of thumb for walking trips: travelers are willing to walk ¼-mile to access local bus service, whereas they are willing to walk approximately ½-mile to access a rail service. The data supports this rule of thumb, 80% of walking trips in the United States are less than 3000 ft, or 0.57 mile (Grava 2003). Within this 3000 foot threshold, it is generally accepted that nearly everyone will

walk ¼-mile, while at ½-mile there is a drop-off of 25-50% in the number of pedestrians willing to walk that distance (Grava 2003). Murray (2003) echoes Grava's assertions by stating that 400 meters (1/4 mile) is the standard acceptable walking distance used in both Columbus, Ohio and Brisbane, Australia.

Loukopoulos and Gärling (2005) investigate the threshold distance at which travelers indicate no preference between walking and driving to and from a particular destination in Swedish study. They estimate an average of 4.1 km (2.55 miles) as the threshold total journey distance for walking for a Swedish university town. This represents a one-way trip of nearly 1-¼ miles, much greater than the threshold given by Grava (2003). However, Loukopoulos and Gärling (2005) further state that this distance is highly dependent upon the perceived effort involved in the walking trip and how often the traveler drives an automobile (both factors that are included in their model) and are likely highly dependent upon the type of trip being made (not controlled for). Given that this report is concerned with American travelers who tend to walk less, drive more, and are less physically active than many other countries, it seems likely that the perceived effort of a walking trip will be higher than that of the average Swedish study participant. Finally, the commute trip in general, and the commuter rail trip specifically, as described in this report, will likely decrease the acceptable walking threshold for reasons given in earlier sections of this report.

Canepa (2007) highlights the fact that acceptable walking distance is not a static number, but a variable dependent upon several factors, chief among these being the density of the development surrounding the rail station. Canepa suggests that increasing density, residential and commercial, within approximately 1 ¼ miles of a rail station will increase the average distance rail patrons are willing to walk beyond the 2000-ft threshold suggested by Calthorpe (1993), which is accepted as a good representation of the distance a traveler is willing to walk. However, in the context of this report, dense development around the commuter rail stations is assumed to be negligible. Therefore, a reasonable walking distance threshold of ¼-mile to 2000-ft will likely better capture the walking habits of commuter rail passengers. Ewing (1999) found that the 2000-ft walking threshold captures approximately 80% of bus and light rail transit passengers.

In this report the specific service being accessed is a feeder or circulator bus, which is most closely associated with local bus service. However, as this feeder service is designed specifically to optimally serve commuter rail passengers, it is conceivable that the feeder service would attract passengers willing to walk a further distance than for typical local bus service. This report will therefore use the walking thresholds suggested by Grava (2003) of ¼ - ½ mile, keeping in mind that optimistic estimates suggest that the range may be as high as 1 mile or more.

The walking coverage parameter, which is defined between a candidate bus stop node and a demand centroid, will be developed using the rectilinear, or Manhattan, distance between the potential stop and the demand centroid. An example of rectilinear distance is given in Figure 6. The rectilinear distance is simply the sum of the difference in x and y coordinates between two points. More formally, $g((x_1, x_2),(y_1, y_2)) = |x_1 - x_2| + |y_1 - y_2|$ for all points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ (Barila 2007). This distance represents (and was named for) the distance a Manhattan taxicab would have to traverse in order to travel between two points in Manhattan's city grid. This distance is a much better representation of the distance that must be traversed by a pedestrian traveling between two points in an urban environment than the straight-line distance.

It is assumed that coordinates are available for each of the candidate bus stop nodes and the demand zone centroids. This assumption is reasonable as the prevalence of geographic information systems (GIS) and global positioning systems (GPS) in metropolitan planning

authorities increases. If coordinates are not readily available, creating a coordinate system in the traditional fashion of using a map and scale is a reasonable alternative.

The walking cost component is a means of accurately representing the walking distance between a bus stop and a traveler's final destination. This distance requires several assumptions be made about the pedestrian network and the pedestrian:

- Pedestrian walkways are arranged in a rectilinear fashion
- Pedestrian walkways exist along all rectilinear paths
- Pedestrians are aware of and will use the shortest rectilinear path
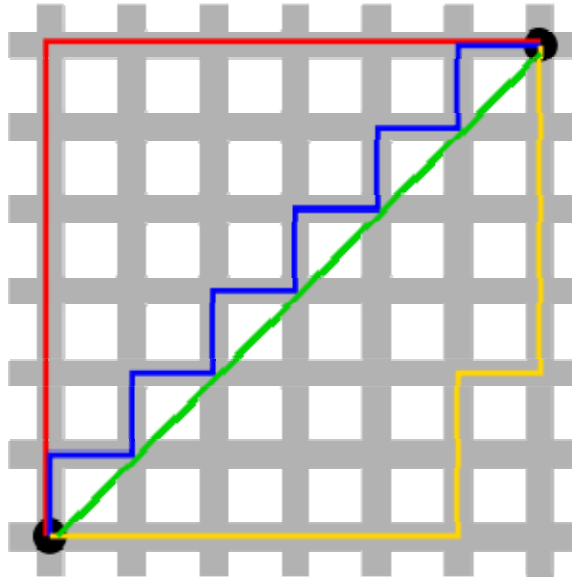
FIGURE 6: RECTILINEAR DISTANCE EXAMPLE: GREEN LINE = EUCLIDEAN DISTANCE, RED, BLUE, AND YELLOW LINES = RECTILINEAR DISTANCE.
Source: http://en.wikipedia.org/wiki/Taxicab_geometry

Using the rectilinear distance as the measure of walking distance has several benefits, chiefly that it represents the typical walking trip in an urban area better than straight-line distance. The first two assumptions are reasonable given that most urban street networks have a large portion of their roads organized in a rectangular fashion and most of these roads will have some sort of pedestrian facility. The final assumption is reasonable given that most pedestrians will in fact work to find the shortest path and once known, will use that path.

Ideally, the first two assumptions would be rendered unnecessary through the application of a shortest path algorithm between a bus stop and demand centroid. Using a shortest path algorithm would not require a rectilinear pedestrian network and would allow for direct paths. However, to improve upon the simple rectilinear metric used in this report one would have to maintain an accurate pedestrian network in addition to the roadway network. This accurate pedestrian network would be needed to find the shortest paths not constrained to be rectilinear or to find paths that avoid inadequate pedestrian facilities. The improvement to the solution provided by using the shortest path on a pedestrian network as opposed to the simple rectilinear distance is unlikely to compensate for the increased effort required to maintain the pedestrian network.

Comparison of Subtour Elimination Strategies

The initial implementations of the CRCNDP solver utilized the subtour elimination strategy developed by Miller, Tucker and Zemlin (1960) (MTZ). This subtour elimination constraint does not provide as intuitive a method of subtour elimination as the more common methods seen today, however, as shown below it performs very well in solving TSP instances.

The Generic Algebraic Modeling System (GAMS) is an independent optimization IDE that can be called upon to model and solve a variety of optimization problems. There are four subtour elimination strategies included in the GAMS model library, all of which can be found in Kalvelagen (forthcoming) labeled TSP1, TSP2, TSP3, and TSP4. TSP1, TSP3, and TSP4 utilize cut generation algorithms to eliminate subtours and resolve the problem. TSP2 is the MTZ strategy, which utilizes unrestricted variables akin to node potentials as a means of eliminating subtours.

The example problem used in this comparison uses nodes i1 through i12 of the GAMS model library included in file br17.inc. This is the example problem designed for the three subtour elimination strategies, used in its original form. The purpose of this comparison is to evaluate the running time and the goodness of the solutions provided by each of the subtour elimination strategies. As solving the CRCNDP will be time intensive and dependent upon the ability of the method to eliminate subtours efficiently, this is of critical importance. Table 5 displays the approximate running time of each method for the 12-node instance of br17.inc.

TABLE 5: SUBTOUR ELIMINATION RUNNING TIME AND SOLUTIONS

Strategy	Running Time (s)	Best Solution
TSP1	84	6*
TSP2	3	39
TSP3	5	32*
TSP4	24	39

It is obvious that the two fastest strategies are TSP2 and TSP3, in that order. However, TSP3 has a * associated with its optimal solution. This asterisk indicates that there are actually subtours present in the optimal solution given by this strategy, meaning that the optimal solution given by TSP3 is infeasible as it includes subtours.

TSP1 also has an infeasible solution at optimality and this is due to the design of the subtour elimination strategy. The maximum number of cuts generated by the algorithm is a user input, with the default value at 100 cuts. This value was increased to 1000 allowed cuts to determine if TSP1 would eventually arrive at a feasible optimal solution. TSP1 ran for a long time (> 30 min for this single problem instance) and eventually drew nearer a feasible optimal solution (it was gradually eliminating subtours). However, the computation time is excessive, which eliminates TSP1 for consideration in this report work. It should also be noted that TSP4 is designed to generate "smarter" cuts than TSP1 using the same basic algorithmic structure. As TSP4 does in fact solve the problem in a reasonable amount of time and produces a feasible optimal solution it is preferable to TSP1 by all measures.

TSP3 generates subsets, with the default value of 500 subsets generated. With $2^{12} = 4096$ subsets of nodes possible, 500 will not be nearly enough to eliminate subtours effectively and consistently. In TSP3 the default value of subsets was increased to 10000 to investigate whether the algorithm drew closer to a feasible optimal solution and the time required doing such. The

algorithm did draw closer to a feasible optimal solution, however, subtours were still present in the final solution and the solve time was greater than that for TSP2. It should be noted that that subtour elimination strategy represents the more common subtour elimination constraints found today and those originally formulated in the CRCNDP.

TSP2 is the fastest and its optimal solution includes no subtours, meaning that the solution is valid and is obtained quickly. TSP1 takes the longest time to run and produces the lowest-valued optimal solution, however, this solution has a number of subtours and is infeasible. In fact, the number of subtours allowed by this strategy suggests that it is a poor strategy for problems of this size. TSP4 arrives at an optimal solution of 39, the same as TSP2. However, the final tour arrived at by TSP4 differs from the TSP2 tour. Below are the tours for each of the strategies, with TSP1 and TSP3 stopping when a subtour is encountered. Note the difference between TSP2 and TSP4.

TSP1: $1 - 12 - 9 - 8 - 1$
TSP2: $1 - 12 - 7 - 6 - 4 - 5 - 8 - 9 - 2 - 10 - 11 - 3 - 1$
TSP3: $1 - 12 - 1$
TSP4: $1 - 3 - 2 - 10 - 11 - 6 - 7 - 5 - 4 - 9 - 8 - 12 - 1$

TSP2 and TSP4 provide feasible solutions with an equal objective function value. If one inspects the problem file, br17.inc, it is apparent that both strategies sought to take advantage of the zero-cost links in the network and both were successful in doing this. TSP4 output states that its solution is "proven optimal", which implies that the TSP2 solution can also be proven optimal.

It should also be noted that TSP1, TSP3, and TSP4 are much more extensive subtour elimination strategies in terms of algorithmic development and coding effort. TSP2 requires a couple of lines of code. TSP4 requires approximately 130.

An additional comparison was conducted between TSP2 and TSP4 on a larger network from the br17.inc file which utilized 15 of the network nodes included in br17.inc. This was done to ensure that TSP2 maintained its run-time superiority in larger networks as well. In fact, TSP2 did maintain its superiority, and even exaggerated it. TSP2 solved the 15-node problem in 15 seconds arriving at an optimal solution value of 39. TSP4 took 179 seconds to arrive at the same objective function value. This suggests that TSP2 is an order of magnitude faster than TSP4 at arriving at an optimal tour for the given problem. Both solutions were checked for feasibility, the tours are shown below.

TSP2: $1 - 12 - 9 - 8 - 4 - 5 - 6 - 7 - 15 - 10 - 13 - 11 - 2 - 14 - 3 - 1$
TSP4: $1 - 12 - 15 - 6 - 7 - 4 - 5 - 9 - 8 - 11 - 13 - 2 - 10 - 3 - 14 - 1$

It is therefore logical that one would select TSP2 as a subtour elimination strategy. TSP2 allows the fastest running time by an order of magnitude, is easiest to implement, and arrives at a solution just as good as the proven optimal solution of the slower strategy, TSP4. Therefore, unless evidence is presented to contradict this comparison or a superior subtour elimination strategy is discovered, the TSP2 MTZ subtour elimination strategy will be used in this report to solve the CRCNDP.

SUMMARY AND CONCLUSION

This chapter concentrated on formulating the CRCNDP in a precise, yet solvable manner. The final formulation has features of both a set covering problem (SCP) and a traveling salesman problem (TSP), modified to allow for multiple routes and incorporate the walking portion of the typical commuter rail trip. The walking portion of the trip is incorporated using rectilinear distance from the bus stop node to the final destination. This distance serves as a much more manageable means of incorporating the walking portion of the commuter rail trip as opposed to using exact shortest path between candidate bus stops and demand centroids on the pedestrian network.

The SCP and TSP are inherently difficult to solve for realistic-sized networks. Incorporating multiple routes and the walking trip in the formulation add to the complexity and may require that heuristic methods be employed in the solution of the problem. However, a benefit of using features of SCP and TSP formulations is that they are well-studied problems and many different solution methods have been applied and tested. It is likely that one of these (or more than one) will be applicable to the CRCNDP, be it an exact or heuristic method.

A more in-depth discussion of potential solution methods for the CRCNDP is put forth in Chapter 4. The formulation will be implemented for a test network and implications of these preliminary results discussed in detail. Specifically, Chapter 4 will discuss the development of the set r that represents the circulator route stops in the formulations. It becomes apparent that defining this set will require additional consideration in the solution development process. At the conclusion of Chapter 4 a clear vision of the process of solving and testing the CRCNDP will be presented.

CHAPTER 4: ALGORITHMIC DEVELOPMENT

The goal of this report is to develop and implement an innovative, efficient method for the optimization of circulator bus routes exclusively serving commuter rail stations. This process is developed using network analysis and optimization tools the application of these tools in an innovative manner. The report effort is to be expended towards development of the general solution methodology for the CRCNDP and implementing this methodology. The focus is not the improvement of optimization solution techniques or algorithms, but the unique application of these methods to the CRCNDP.

Figure 7 provides a flow chart of the proposed solution process to be implemented and the data requirements and tasks required in each step. Details of the flow chart are given in subsequent discussion. Following the discussion of flow chart elements is a discussion of the development of the first solution methodology applied to the CRCNDP, complete enumeration. This solution method will be discussed in some detail, with pseudocode provided for clarity. An application of this technique is presented and the limitations and implications of complete enumeration highlighted. Following the enumerative discussion, methods of selecting smarter subsets of candidate stop locations are discussed. These methods seek to employ a form of implicit enumeration as a preprocessing step, eliminating infeasible branches ahead of time, limiting the processing efforts to those subsets that may potentially yield good solutions.

| Data Collection/Generation | ← | • Travel demand estimates
• Roadway travel times
• Operator cost estimates
• Candidate bus stop coordinates |

| Preprocessing | ← | • Determine location of demand centroids from digital network representation.
• Compute rectilinear distance between all candidate stop and centroid node pairs. |

| Solution | ← | • Complete Enumeration
• Smarter candidate stop selection procedures.
• Heuristics |

| Sensitivity Analysis | ← | • Analyze range of parameter and cost values *ceteris paribus*
• Seeks to verify default parameter values |

| Case Study | ← | • Proposed Austin commuter rail MLK station
• CAMPO data
• Pareto-optimal solutions |

FIGURE 7: SOLUTION METHODOLOGY DEVELOPMENT FLOW CHART

Data Collection/Generation

The flow chart in Figure 7 is intended to represent the general flow of activity in this report. There will be preliminary runs of the formulation to determine the validity of the formulation (it does what it is expected to do) and any computational issues that are implicit in the formulation. Actual data acquisition will be more laborious, especially the acquisition of travel demand estimates. The Capital Area Metropolitan Planning Organization (CAMPO) is the source for the demand (O-D) information. The Capital Area Council of Governments (CAPCOG) and the city of Austin will provide good geospatial data.

Preprocessing

Ceder and Wilson (1986) suggest that an initial paring of the potential bus network be performed to exclude extraneous links from the bus route optimization problem. This network preprocessing will not play a significant role in the current network and cost representation. The current CRCNDP formulation does not include physical roadway links. The trip segment between nodes is represented either by the shortest path (determined and maintained offline) or the rectilinear distance between nodes (given data).

Once the basic network structure has been determined and demand zones have been selected (TAZs), demand centroids are assigned based on the activity centroid of that TAZ. For example, in a demand zone that includes several office buildings in one zone quadrant and green space, the demand centroid will be placed in the quadrant with the office buildings as they are much more likely to attract commuter travelers than green space. These centroids will be used to complete the next preprocessing step, determining rectilinear distances between candidate bus stops and demand centroids.

Defining the walking distances between candidate bus stops and reachable demand centroids will require that some form of coordinate data for the candidate stops and the demand centroids be readily available. Using the coordinate data, it will be a relatively simple process to compute the rectilinear distance between candidate stop and centroid and create the set of reachable centroids for each candidate stop based upon the desired threshold value for walking distance. These walking costs will then be input as parameters in the optimization of the bus circulator route(s).

Solution

Preliminary solution efforts will utilize the General Algebraic Modeling System (GAMS) to model the problem and use its integrated solvers (Cplex, namely) to solve the problem in a complete enumeration effort. As will be discussed shortly, the MIP formulation of the CRCNDP includes the set of candidate stops to be visited in a route, r, a given set. Therefore, solution efforts will need to identify this subset of candidate stop nodes for which the CRCNDP will be solved. Some characteristics of the CRCNDP may help reduce the solutions space, in this way serving as a sort of preprocessed implicit enumeration technique. These efforts will focus on eliminating infeasible subsets from consideration, eliminating the wasted solution time spent on these infeasible subsets.

In the event that exact methods to solve the CRCNDP require computational effort that is impractical (several hours to solve a single problem to optimality), heuristic methods will be explored. Popular metaheuristic solution methods include Genetic algorithm (GA), Tabu Search, and Simulated Annealing. These solution methods are well studied and have been implemented

in many fields of study, most recently related to the CRCNDP by Fan (2004). As metaheuristic solution methods are generic by design, enabling them to be applied to many fields, tailoring the metaheuristic methods to the problem of interest in a clever and correct fashion would be of the utmost importance and would strongly influence the success of the implementation.

Heuristic methods may be employed that exploit certain aspects of the problem, and break the problem into more manageable and solvable pieces. These problem pieces can be recombined to provide a solution, though no guarantee of optimality is implied. Heuristic solution methods can be very useful if employed in an intelligent way and can provide near-optimal or optimal solutions in many cases. The CRCNDP may be one particular problem in which near-optimal solutions may be acceptable, as the problem is inherently limited by the uncertainty surrounding the cost parameters.

The cost parameter variance leads to the need to define optimality for this formulation of the CRCNDP. Cost parameters (operating, walking, in-vehicle traveling, unserved demand) will vary from city to city, and location to location. Therefore, any optimal solution will only be optimal for the specific set of cost parameters used to obtain that solution. In practice it will be useful to give several equally (or nearly) good solutions that correspond to different cost parameter sets. This type of solution is referred to as Pareto-optimal, a concept that has been applied in the area of construction management extensively in recent years by Hyari and El-Rayes (2006) and Kandil and El-Rayes (2006) in developing the best construction plan for a given set of project requirements. Zitzler and Thiele (1999) provide the following definition of Pareto-Optimality:

Many real-world problems involve simultaneous optimization of several incommensurable and often competing objectives. Often, there is no single optimal solution, but rather a set of alternative solutions. These solutions are optimal in the wider sense that no other solutions in the search space are superior to them when *all* objectives are considered. They are known as *Pareto-optimal* solutions.

These alternative cost parameter sets could represent a variety of weighting schemes that a user of this methodology might employ. For instance, a particular metropolitan area may feel that the most important cost to consider is unserved demand, followed by in-vehicle traveling, operating, and walking costs, in that order. A cost parameter set could then be defined representing these priority levels and a solution obtained accordingly. If the user wanted to compare this with other schemes, redefining cost parameters and solving the problem with the new weighting scheme to compare results would be a straightforward process. Pareto-optimal solutions will often be the most useful solution type in practice, and a heuristic that significantly reduces computational effort (allows for quick solution to different cost weighting schemes) and provides very good solutions (near-optimal) may be an excellent option.

Sensitivity Analysis

The sensitivity analysis step will seek to determine the best set of cost parameters to provide non-trivial, good solutions. This process will necessarily entail a sensitivity analysis in which a wide range of cost parameters are evaluated relative to ranges of the other cost parameters to determine the best ranges for each cost parameter. These ranges can then be used to help define the boundaries for the cost parameters and the alternative cost parameter weighting schemes to be evaluated.

Case Study

Another contribution of this report is the application of the proposed method to a case study in Austin, Texas. Austin is currently implementing a commuter rail system on existing rail ROW owned by the local transit authority, Capital Metropolitan Transportation Authority (Capital Metro). This case study will provide valuable insight into the method, the effort required for implementation, data needs, solution techniques, and the justification for complex formulations incorporating the walking portion of the commuter rail trip. Pareto-optimal solutions will be pursued, each corresponding to the best solution for a particular set of cost parameters. The generated solutions will be discussed in light of other Austin circulator/feeder plans, the goals of a transportation system, and other expectations discussed in this report.

COMPLETE ENUMERATION

Complete enumeration represents the most "brute force" approach to solving this problem. This approach solves the CRCNDP for every possible combination of candidate bus stops in a network. This objective function includes costs incorporating both in-vehicle and out-of-vehicle costs. The first step in an enumerative method simply selects the combination of candidate stops to be served and the second solves the CRCNDP for this particular subset, r. This process is repeated for every possible combination on the subset r. One then knows the objective function value, or the total cost of the solution, for all possible candidate bus stop combinations and one can then simply select the globally optimal value.

A complete enumeration procedure as outlined above has several benefits. Firstly, the chosen solution is easily proven to be optimal, as all possible solutions have been considered. Secondly, complete enumeration will shed light into the relationship between the stop selection and route design portions of the problem, providing insight for future, more efficient solution method application. Lastly, complete enumeration may be a viable solution method for the CRCNDP because of the problem size reductions inherent in the problem.

Algorithm

The algorithm in Figure 8 finds the optimal set of bus stops as well as the optimal sequence of stops for a particular network by evaluating all possible sets, $r \subset I$, where $|I| = N$ and each combination must include the station. A parameter, Δ_g is introduced to compare walking distances from a demand centroid to a bus stop. Another parameter, U, represents the walking threshold as defined earlier in this report, to ensure that walking coverage is provided only to those demand centroids that are within this threshold. This check ensures that bus passengers will take the shortest walking trip to their final destination. At the end of the algorithm, the optimal solution for a particular subset, r, is compared with the current global optimal solution. The global optimal solution, z^*, is updated if the current solution is better. The subroutines defining s_i, μ_g, ρ_{ig}, and δ_i are utilized in all solutions techniques developed in this report.

algorithm CRCNDP Complete Enumeration
begin

       $r := \{\emptyset\}$;
       $q := 1$;
       $z^* := \infty$;

      while $q \leq N$ **do**

      for all $r \subseteq I$ such that $1 \in r$ and $|r| = q$

          for all $g \in G$

              $\Delta_g := \infty$;

              for all $i \in r$

                  if $\gamma_{ig} < \Delta_g$ and $\gamma_{ig} < U$

                      $\Delta_g := \gamma_{ig}$;

                      for all $j \in r$

                          $\rho_{jg} := 0$;

                    end;

                  $\rho_{ig} := 1$;

                end if;

                $s_i := \sum_g \rho_{ig} d_g$;

                $\delta_i := 4 + 1.7 s_i$;

              end;

              if $\sum\limits_{i \in r} \rho_{ig} = 0$

                $\mu_g := 1$;

              end if;

          end;

          solve CRCNDP(r, s_i, μ_g, ρ_{ig}, δ_i,) $\rightarrow z$;

          if $z \leq z^*$

             $z^* = z$;

          end if;

        end;
        $q := q + 1$;
        end;
      end;

FIGURE 8: CRCNDP COMPLETE ENUMERATION PSEUDOCODE

The complete enumeration algorithm consists of several components: initialization, combination generation, parameter calculation, local solution determination, and global solution update. The initialization phase assigns the intial values of the subset r, the cardinality of r, and the global optimal solution. The subset r is initially set to contain no elements, after which the cardinality of r, q, is set to one. This ensures that the first iteration will look at only one node, and subsequent checks will ensure that this node is the station. The initial global optimal solution is set to infinity, guaranteeing that the first feasible subset's solution will be the new global optimal solution.

The combination generation phase uses the cardinality of I, $|I| = N$ and the current "sample" size, q, to generate a subset combination, r for which the CRCNDP will be solved. Once all sets of size q are evaluated, q is incremented by one and the process continues until the $q = N$ phase is complete. It should be noted that because the station, node 1, must be included in every combination, there will be a maximum of 2^{N-1} combinations evaluated.

The parameter calculation phase is concerned with determining the parameters s_i, μ_g, ρ_{ig}, and δ_i. The walking coverage parameter ρ_{ig} uses an intermediate parameter, Δ_g to ensure that the walking trip to a particular centroid is minimal. The parameter U is used to ensure that this walking coverage is also within the predefined walking limit (usually ¼ - ½ mile). The binary parameter ρ_{ig} takes the value "1" if a centroid g is covered by a minimal walking trip from a circulator stop at node i within the walking trip threshold. It takes the value "0" for all other pairs.

The parameter ρ_{ig} is then used to determine the amount of demand served at a particular stop location. A summation over all demand centroids for the product of walking coverage from i to g and demand at g, gives the total demand served at stop i. Restated, all of the walking trips that begin at stop i are multiplied by the number of people using that particular walk trip. Walking coverage is also used to determine those zones that are completely unserved. A summation of ρ_{ig} over all bus stops for each demand centroid will yield either a "1" or a zero "0", as each centroid is served by at most one stop. For those centroids whose summation is zero, the binary parameter μ_g takes the value "1", as that zone is now considered unserved. This parameter will be used in determining the unserved demand cost in the objective function. The dwell time computation uses a simple linear relationship developed by Levinson (1983) that incorporate the number of passengers alighting at a stop, s_i.

All parameters have been computed prior to calling the CRCNDP solver. The CRCNDP uses the current subset r and the four parameters as input to the formulation given in Chapter 3. These five inputs will produce a unique, locally optimal solution for the CRCNDP. In the final algorithm stage, this locally optimal solution is checked against the current global optimal solution. If the local solution is better (lower cost), then the global optimal solution is updated accordingly. After this final test, the next combination r is generated, its parameters determined, and the problem solved for this new combination of parameters. In this enumerative form, the algorithm runs for each of the 2^{N-1} combinations.

Enumerative Example Application

The case study to be pursued in this report is the Martin Luther King Jr. Boulevard (MLK) station of the Austin, TX MetroRail system. This station is approximately 2 miles east of the University of Texas campus, the Texas State Capital complex, and the Austin CBD. The MLK station is located in a primarily single-family residential neighborhood with low residence and commercial density.

Figure 9 presents a snapshot of the MLK station area and the TAZs that are within a 2-mile "service boundary" of the station. This 2-mile boundary is selected based on the net operating speed of buses on an urban route, which Levinson (1983) suggested is approximately 10 mph. At 10 mph, an urban bus could reach the 2-mile boundary and return within 24 minutes. However, since dropping off passengers requires a nonzero amount of time, the total 4-mile round trip will likely approach 30 minutes, which is the headway of commuter rail trains on the MetroRail system. This service boundary could be expanded if desired, in fact, it does not have to be a circular boundary at all. However, in an effort to keep the problem size manageable and as objective as possible, a 2-mile circular boundary is used for this example. It should be noted that a TAZ does not have to be completely covered by the boundary to be included in the analysis, so in fact, centroids > 2 miles from the station are considered.
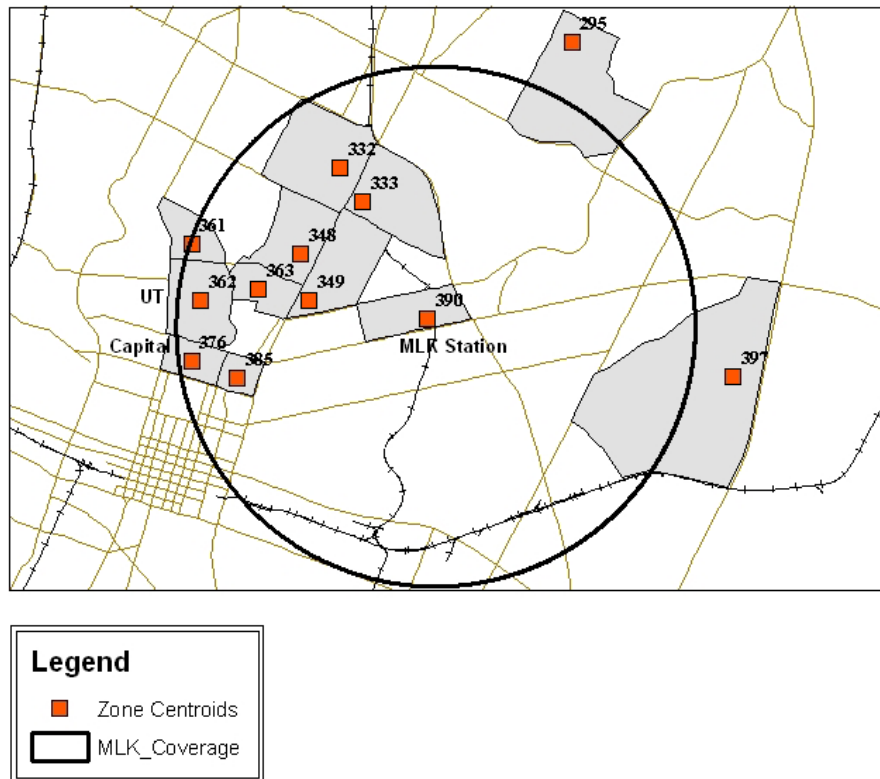
FIGURE 9: TAZS USED IN MLK ENUMERATIVE APPLICATION

In the enumerative example, computational resource issues played a role in the implementation scheme. As discussed previously, the enumerative algorithm detailed in Figure 8 will solve the CRCNDP for a total of 2^{N-1} combinations, where N is the total number of candidate bus stops to be considered. From preliminary investigation, it was found that the enumerative algorithm's solution time exceeded 8 hours for greater than 12-node examples. Therefore, in the enumerative example, only the 12 highest demand TAZs within the MLK service area are considered part of the set *I*. These 12 TAZs are highlighted in Figure 9 along with their assigned demand centroids. These centroids were assigned offline based upon the activity concentration apparent from aerial photography of the area.

56

The 12 nodes' demand was assigned using 2-hour peak vehicle counts obtained from the Capital Area Metropolitan Planning Organization (CAMPO). The count data provided origin-destination (O-D) estimate between each of the Austin Metro area's 1074 TAZs during the AM peak period. The first step in determining the demand for the TAZs covered by the 2-mile service boundary shown in Figure 9 was to determine which TAZs are covered at the home end of the commuter rail trip.
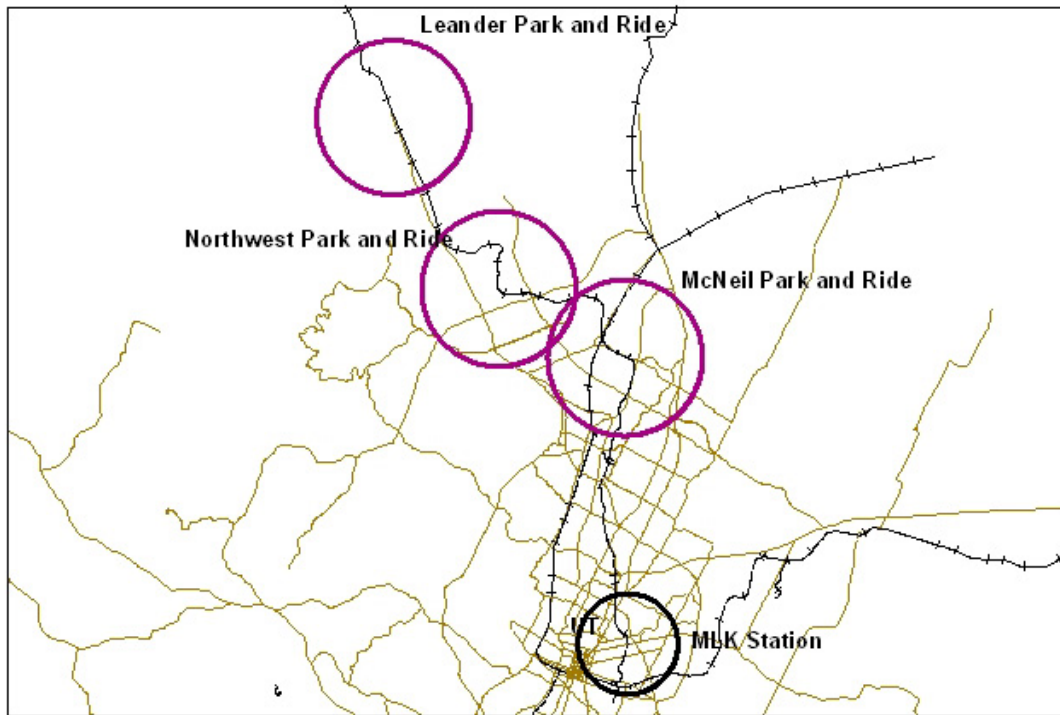
FIGURE 10: REGIONAL VIEW OF METRORAIL STATION COVERAGE ON

AUSTIN'S ARTERIAL AND RAIL NETWORK

As has been discussed previously, the commuter rail trip in new U.S. commuter rail systems will likely begin at a park and ride facility in the suburban areas surrounding a metropolitan area. The primary means of accessing a new park and ride facility is the personal auto, to which most households in a suburban area will have access. In this application, a coverage radius of three miles around each station was used to signify commuter rail accessibility to an origin TAZ.

This corresponds to a 10-minute trip from home to park and ride at a net operating speed of 15 mph. Acknowledging that those furthest away from the station may be less likely to utilize commuter rail, this application will use a binary definition of accessibility. The TAZ either has access to commuter rail, or it does not. Figure 10 depicts the three park and ride locations that are used as origins for the demand data in this application.

It is assumed in this application that the relative O-D trips within the commuter rail accessible regions will be the same after commencing commuter rail service as prior to commuter rail service. That is, commuter rail will not have a significant immediate impact on

the majority of commuters' mode choice. This result of this assumption is that this 2-hour peak hour O-D table, which was developed without MetroRail service in place, will provide a demand distribution for commuter rail trips as well. This demand distribution is obtained by capturing all trips in the CAMPO dataset that originate in one of the three park and ride coverage zones and terminate in the MLK station coverage zone. The trips from all three origin zones can be aggregated as all commuter rail trips to the MLK coverage area will go through the MLK station, so they will all share a common origin at the station. Hence, the problem becomes a one-to-many, with all demand originating at the MLK station. The demand figures do not represent a precise number of trips as the current analysis does not attempt to estimate the commuter rail mode share. However, the assumption that the relative O-D distribution is applicable to the system prior to and after commuter rail service commencement makes a precise mode share unnecessary for this analysis. Using the relative demand data obtained from the aggregation of covered origin zones, one can use this relative demand to estimate the general commuter rail attractiveness of a particular TAZ.

For example, this general attractiveness of a TAZ could be used to distribute passengers from a train. If a full train is assumed to arrive at the station each half-hour, then the aggregate demand percentages can be applied (assuming that the relative demand is constant throughout the 2-hour period). Table 6 shows the application of this relative demand assumption to a full trainload of passengers arriving at MLK station. The percentage of relative demand is simply applied to the train capacity.

TABLE 6: TAZ DEMAND ASSUMING A FULL, TWO-CAR TRAIN AT 108 PASS/CAR (* STATION)

TAZ	Aggregated Demand Destined for TAZ	Demand %	Estimated Trips to TAZ
390*	10	1.7%	4
295	19	3.2%	7
332	17	2.9%	6
333	17	2.9%	6
348	58	9.8%	21
349	18	3.0%	7
361	121	20.5%	44
362	185	31.3%	68
363	19	3.2%	7
376	75	12.7%	27
385	21	3.6%	8
397	30	5.1%	11
TOTAL	589	100%	216

This estimation could be performed with the circulator capacity as the limiting factor instead of the train capacity. In this case, one would simply use the relative demand percentages

and apply them to the capacity of the circulator buses. In this application, since train capacity is used to estimate trips to each TAZ, one must also assume that sufficient circulator capacity will be made available for the final route configuration.

It was previously noted that the enumerative technique requires that a limited number of nodes be used in solving the CRCNDP for this application. Therefore, the top 12 highest aggregate demand TAZs (shown in Table 6) were selected for analysis. Because representing every possible bus stop that could serve each of these TAZs would exceed the number of allowable nodes in this application, the nodes used will be the demand centroid of each TAZ. This will imply a zero-distance walk trip from the bus stop serving each zone to that zone centroid. The final solution will therefore provide a sequence of demand centroids to visit and not provide input into bus stop placement decisions. The 12-node restriction does limit the applicability of the results and is shown here to demonstrate the CRCNDP enumerative solution method, highlight aspects of the problem and solution in a simple setting, and provide motive for future efforts to improve the solution method.

Table 7 shows the data used in this application of the CRCNDP Complete Enumeration algorithm. The first three cost data and the walking speed are estimated from TCRP Report 78 (ECONorthwest and Parsons Brinckerhoff Quade & Douglas, Inc. 2002), the unserved demand cost from Bailey (2007) and the bus operating speed from Levinson (1983). Note that the rectilinear distances between nodes is not given in this table. These rectilinear distances were computed for every node pair using *x-y* coordinate data of the demand centroids.

TABLE 7: MLK ENUMERATIVE APPLICATION DATA

Data	Definition	Value
C_o	cost of operating bus	$82/hour
C_{tt}	traveler cost of in-vehicle travel time	$13/hour
C_w	cost of walking	$25/hour
C_d	cost of unserved demand	$6000/unserved passenger
d_g	demand for service at demand centroid g	See Table 6
H	commuter rail train headway	30 minutes
F	number of buses available to the route	2
v_{bus}	local bus operating speed	10 mph
v_{walk}	pedestrian walking speed	2.5 mph

Applying the complete enumeration algorithm yielded an optimal objective function value of $393,359 in 3983 seconds (1.10 hours) of run time. The optimal set of bus stops is r = {390, 362, 376, 385, 349, 333} with a demand centroid sequence of: 390 – 362 – 376 – 385 – 349 – 333 – 390. This optimal route has a total length of approximately 43 minutes.

The parameter ρ_{ig}, which indicates walking coverage between node-centroid pairs is shown in Table 8 using a walking threshold of ½-mile. Using ρ_{ig} the other parameters for this particular set are $s_i = \{0, 0, 0, 12, 0, 35, 0, 112, 0, 35, 0, 0\}$, $\mu_g = \{0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1\}$, and $\delta_i = \{0, 0, 0, 24.4, 0, 63.5, 0, 194.4, 0, 63.5, 0, 0\}$. Figure 11 provides a graphical depiction of this final solution.
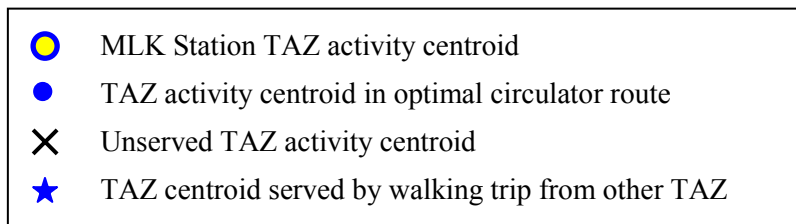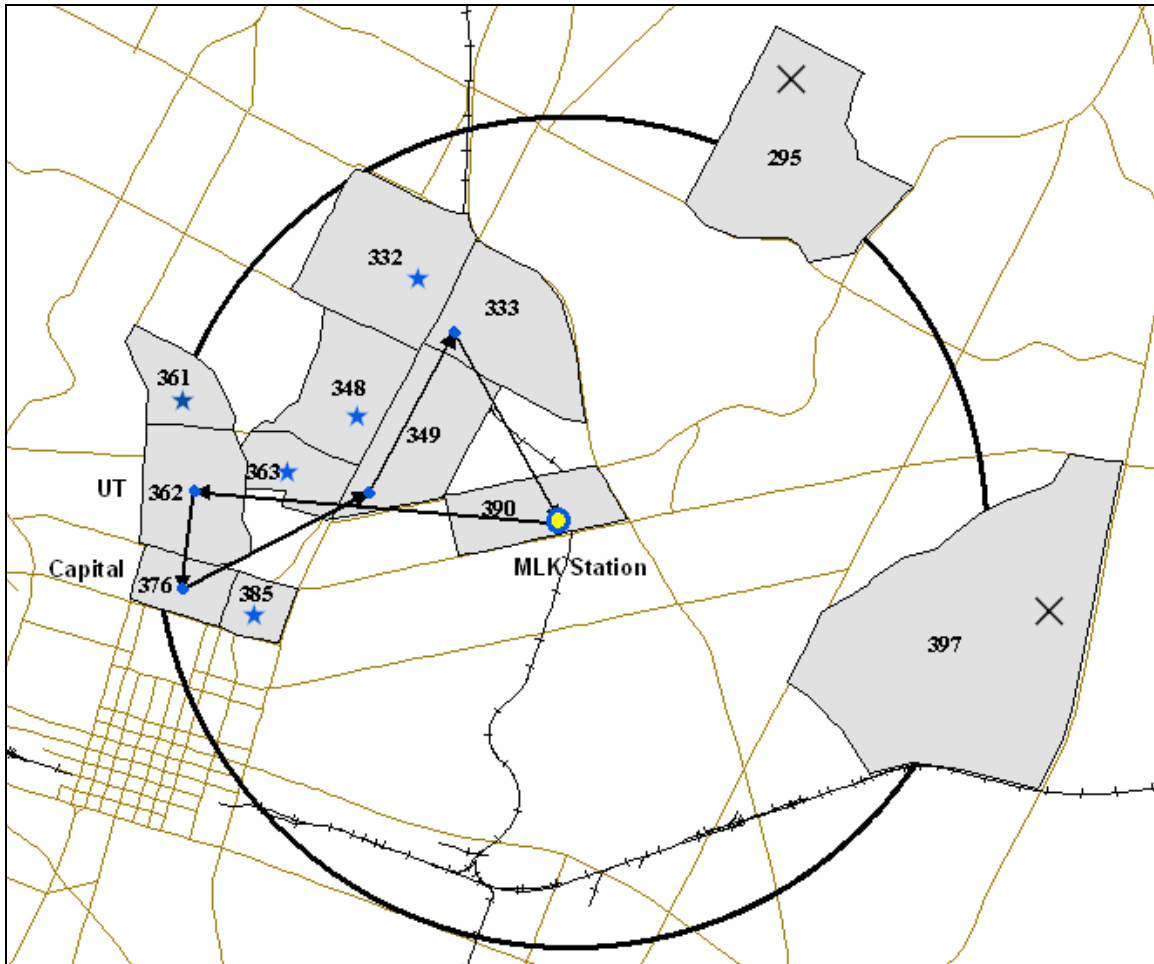
FIGURE 11: EXAMPLE APPLICATION SOLUTION

Several interesting observations can be made of this particular solution. First, the optimal route does not use all available headway in the final solution. Two buses were available to serve this particular route, and staggering for 30 minute train headways would allow for a maximum 60-minute route. However, the problem formulation allowed for the tradeoff between the increased costs from longer walking trips and unserved demand and the increased costs of passengers enduring a longer in-vehicle travel time.

In the end, the optimal route is 43 minutes, with the last 8 minutes used for the empty bus to return from node 333 to the station. This is significant because it suggests that the default cost parameters here allowed for this tradeoff to occur. A poor set of cost parameters could result in trivial solutions in which the maximum number of stops is made within the given headway regardless of the in-vehicle time.

TABLE 8: BINARY WALKING COVERAGE PARAMETER
FOR R = {390, 362, 376, 385, 349, 333}

		To				
		3	5	7	9	11
From	4	1				
	6		1		1	
	8			1		
	10					1

This solution shows that the cost parameter set is reasonable (though this does not negate the need for sensitivity analysis) and that a 60-minute headway may not be a good use of circulator resources. A better use may be to allow for two routes, each with a 30-minute maximum route length, especially since a circulator route greater than 30 minutes is unlikely to attract many commuter rail riders. Not surprisingly, the route visits the highest demand centroids first at UT and the state capital complex. This is a good result; a solution that did not place priority on these nodes would have been suspect. However, the solution does ignore nodes 349 and 363 en route to UT at the beginning of the route (it would have required the bus to divert approximately 500 feet). This is reasonable, as the waiting time incurred by diverting to and dwelling at node 349 or 363 by the UT and State Capital passengers is not technically justified. Though technically accurate, this may not be the most practical solution and may be a situation in which external engineering judgment is used in the final route design.

It is also a positive result that walking trips are relied upon to serve demand in a cost-efficient manner. It is hypothesized that by including physical bus stop locations, the use of walking trips will be further enhanced allowing greater confidence in the results and a direct interpretation. In summary, it is evident that the CRCNDP formulation and complete enumeration algorithm work as intended. However, the run time of 1.1 hours for just twelve nodes is a cause for concern. To truly reap the benefits of formulating the CRCNDP as it has been in this report, one needs to be able to evaluate a larger number of nodes. The following section discusses what improvements can be made to the enumerative algorithm and what the limitations of such improvements may be.

SUMMARY

This chapter discussed the enumerative algorithm used to solve the CRCNDP and its performance in a simplified network. A significant positive result of these efforts is the confirmation that the formulation works as intended. However, the limitations of enumerating NP-hard problems present significant computational challenges that must be overcome if the CRCNDP is to be utilized in practical applications.

The following chapter presents the first of two methods to overcome the inherent difficulties of the CRCNDP by presenting an improved enumerative algorithm. This algorithm seeks to eliminate infeasible solutions prior to solving the problem using 1-Trees and stop spacing requirements. Additionally, a stopping criterion is presented that should perform well in the average-case. The performance of these improvements is presented and absolute and relative limitations discussed.

CHAPTER 5: IMPROVEMENTS TO THE ENUMERATION ALGORITHM

The above example application of the enumerative algorithm required over an hour to solve a 12-node problem. Every combination generated a call to GAMS, the independent development environment for mathematical programming. Each call to GAMS initiates a CRCNDP solution, which takes 1-2 seconds for four or fewer nodes and up to approximately 15 seconds for 10 nodes. The application solution included a large number of infeasible solutions. In fact, over 70% (out of 2048) of the calls to GAMS resulted in infeasible solutions. The time spent determining the infeasibility is often on the order of 1-2 seconds. If one could eliminate these infeasible solutions prior to calling GAMS, the 3983 second run time could potentially be reduced by up to 2800 seconds. Several preprocessing options exist to potentially trim these infeasible solutions from the solution space prior to running a solution algorithm. The three methods investigated are discussed below.

IMPROVING THE ENUMERATIVE ALGORITHM

Three additions to the enumerative algorithm are considered to improve the running time by reducing the solution space. The first two methods of reducing this solution space incorporate preprocessing to eliminate infeasible stop combinations from consideration by the CRCNDP. The third method establishes a stopping criterion that for practical implementations of this method should result in the algorithm terminating with a provably optimal solution without having to evaluate all possible stop combinations.

Preprocessing Step: Use 1-Trees to Determine Feasibility

As mentioned previously, the CRCNDP is closely related to the TSP. The TSP is a well-studied problem that over the years has had many algorithms and solution methods designed to solve it in an efficient manner. The goal of this portion of the research is not to develop another TSP solution method, as the computational issues with the CRCNDP lie not within the TSP solution method, but the necessity of having to apply it 2^{n-1} times, where n is the number of candidate stops in a network. Therefore, a method of better selecting which stop combinations for which to solve the CRCNDP is truly useful.

As noted, a large percentage (70%) of the calls to solve the CRCNDP for various stop combinations results in infeasible solutions due to long tour lengths. These infeasible calls can be reduced using a preprocessing step that can recognize a good portion of these infeasible sets and eliminate them prior to solving the CRCNDP.
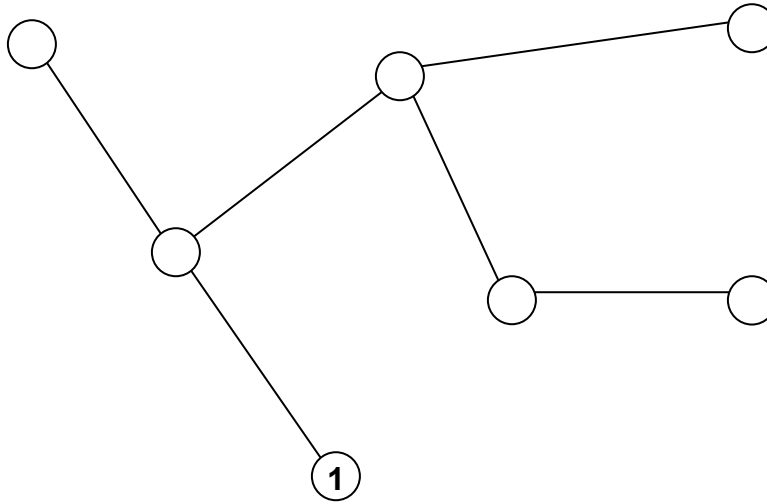
FIGURE 12: EXAMPLE OF MINIMUM SPANNING TREE SOLUTION

A good means of eliminating these infeasible sets utilizes the lower bounds to the TSP. A well-documented method of establishing a lower bound for the TSP uses the 1-tree, which is a minimum spanning tree (MST) over a set of nodes along with the second-lowest cost arc emanating from the station node. Figures 12, 13 and 14 display the difference between an MST, a 1-Tree, and a TSP solution.

Figure 12 shows an MST solution for a set of 7 arbitrary nodes. Note that each node is connected but the degree of each node is not necessarily equal to 2. An MST can provide both a lower and upper bound for the TSP, though in some cases the bounds are not very good. Figure 14 displays a TSP solution for the same node set, note that if one arc incident to node "1" is removed, the solution is an MST. So, by removing this arc, we arrive at an MST, suggesting that any MST provides a lower bound for a TSP solution over the same set of nodes. Conceptually, this makes sense, as a TSP solution cannot do better than the minimum path to connect all nodes as it must add at least one arc onto this solution to satisfy TSP constraints. An MST can also provide an upper bound to the TSP. The goal of a TSP is to visit every node and return to the "home" node in minimal cost fashion. A naïve method of accomplishing this task would be to visit each node along the MST and return to the home node along that same MST, resulting in a tour cost of 2*MST. In the worst case, this solution would visit each node directly from the home node, return to the home node, and then visit the next node directly again, and so on. Obviously, this is a very inefficient means to accomplish the goals of the TSP, but ignoring the node degree constraints, it accomplished the task of visiting each node. We have therefore established bounds on the optimal solution of the TSP, shown in (13).

$$z^*_{MST} \leq z^*_{TSP} \leq 2\left[z^*_{MST}\right] \qquad (13)$$

As mentioned, these bounds are not very tight. Examples can be easily constructed in which the bounds provided by the MST solution are significantly different from the true TSP

64

solution. A common method of tightening the lower bound is the use of 1-trees, an example of which is in Figure 13.
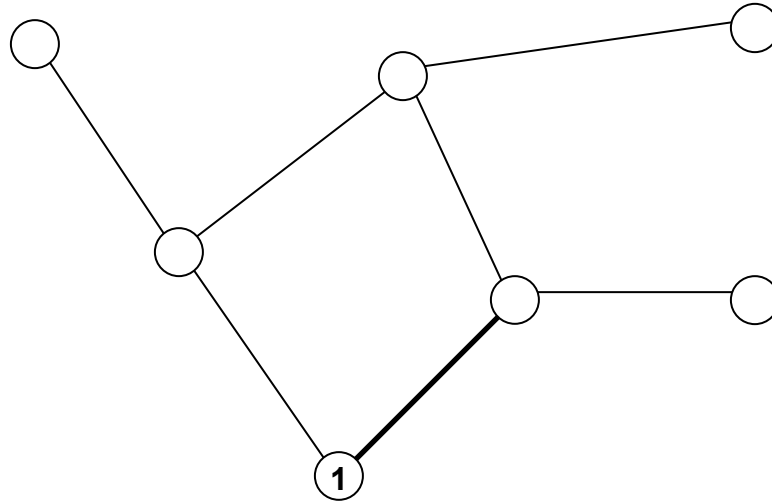
FIGURE 13: 1-TREE SOLUTION FOR EXAMPLE NODE SET

The 1-Tree solution is exactly the MST solution with the addition of the minimum cost arc not already in the MST solution from the home node, or node "1", to a node in the MST. This is exactly the situation described in which the MST converges to the TSP solution, in which one arc connecting the home node is removed from the TSP solution, resulting in an MST. This removed arc is the minimum cost remaining arc that would be added to the MST to give a 1-Tree solution. All TSP solutions are therefore 1-Trees, but not all 1-Tree solutions are TSP solutions, such as the 1-Tree in Figure 13. This 1-Tree violates the node degree constraint on several occasions. However, the 1-Tree does provide a tighter lower bound than the MST solution, as it does have this extra arc that creates a cycle. This tighter bound is formalized in (14).

$$z^*_{MST} \leq z^*_{1-Tree} \leq z^*_{TSP} \tag{14}$$

The bound provided by the 1-Tree solution may only be marginally better than the solution provided by the MST, however, it is an improvement and can certainly be useful in the context of the CRCNDP enumerative algorithm. For comparison, Figure 14 displays a possible TSP solution for the same example node set. Note the substantial difference between the MST and 1-Tree solutions and this TSP solution, one can easily imagine scenarios in which the optimal 1-Tree solution is far from the optimal TSP solution. It is important to remember though that the bounds gained by MST and 1-Tree solutions can be found in polynomial time.
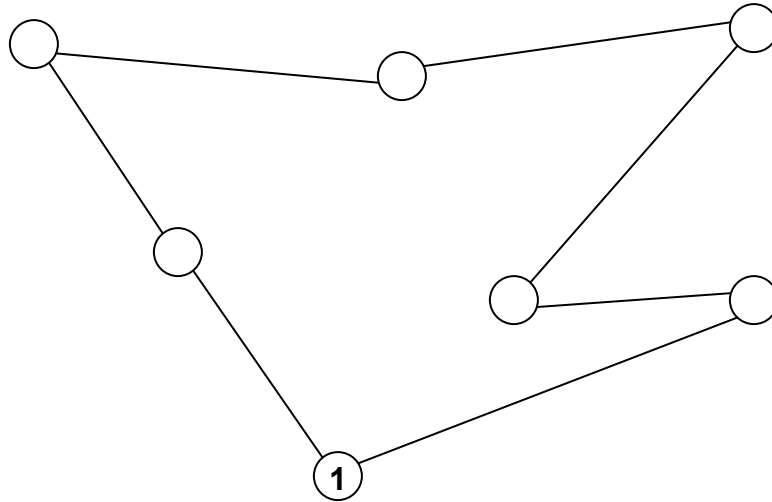
FIGURE 14: TSP SOLUTION TO EXAMPLE NODE SET

While in any application tighter bounds are going to be useful (especially if they can be maintained in polynomial time), less tight bounds are still very useful in the context of this research. Above it is noted that 70% of the calls to GAMS resulted in infeasible solutions, wasting considerable computational effort. It is stated that a means of eliminating these infeasible solutions (or a significant portion thereof) would be of considerable benefit. The lower bounds provided by the 1-Tree present such an opportunity.

We can guarantee that if the 1-Tree solution for a particular set of nodes is infeasible (that is, the route length is too long), the TSP solution (and therefore the CRCNDP solution) will be infeasible as well. Since we can find the 1-Tree solution in polynomial time, a preprocessing step finding the 1-Tree solution for all combinations of nodes will take much less time than running the CRCNDP for the same combinations of nodes and will produce a list of node sets that should be considered for CRCNDP solution. In this manner a provable optimal solution can be guaranteed, and while more time is spent on preprocessing, considerable computational gains will be made in solving the CRCNDP.

Preprocessing Step: Use Minimum Stop Spacing to Determine Feasibility

An additional means of trimming the solution space prior to solving the CRCNDP once again utilizes the nature of the CRCNDP and the characteristics of bus circulator systems in general. Grava (2003) suggests that for typical urban bus systems stops should not be placed within 1000 feet of each other in typical corridors and not within 1500 feet ideally. This restriction on stop spacing presents another preprocessing opportunity for the CRCNDP. Eliminating sets of candidate stops in which at least one pair of stops is within 1000 (or 1500) feet of another will certainly reduce the number of stop sets that the CRCNDP will have to evaluate. The sets eliminated may not be infeasible sets, however, the resulting CRCNDP solution will still be provably optimal with the condition that the stops be some minimum distance apart.

Stopping Criterion: Cardinality of Stop Set

A large source of inefficiency in any enumerative algorithm is continuing to search for an optimal solution when the optimal solution has already been found and any subsequent investigations have no chance of providing an improvement to the global optimal solution. Developing a stopping criterion that will help eliminate wasteful computational effort can potentially provide vast improvement to the solution time.

Consider the CRCNDP. In practice, a common restriction on the size of the circulator route will be 30 minutes, as any circulator trip exceeding this value (or even approaching it) will likely be very unattractive to rail passengers after already traveling to the home-end rail station and making the rail trip. If one considers both operational speed of urban buses (typically 10 – 12 mph) and dwell times at stops, it becomes apparent that a fairly small number of bus stops will utilize all 30 minutes of the allowed route length in the CRCNDP solution. In a network with 50 candidate stop locations, it is highly unlikely that the optimal solution is going to have more than perhaps 8 – 10 stops, and maybe fewer. All effort to evaluate stop sets with greater than 8 – 10 stops would likely be wasteful efforts and could be eliminated. However, if a provable optimal solution is desired, greater rigor is needed than this intuitive look at a stopping criterion.

Included in the improved enumerative algorithm is a stopping criterion that does not set a particular set cardinality as the maximum possible, the stopping criterion simply looks at the improvement of a network's current optimal solution. If the current optimal solution never changes past a certain point, then it would be useful to identify a stopping point earlier than complete enumeration after which one could say that there is no possibility of improving the current optimal solution, thus proving its global optimality.

The improved enumerative algorithm employs a stopping criterion that looks at the cardinality of the current optimal solution, which is the cardinality of the set of stops for which the CRCNDP is being solved (remember, the CRCNDP is solved for all possible combinations of r in the complete enumeration case). Let C be the cardinality of the set of stops, $C = |r|$. The stopping criterion will allow the algorithm to evaluate all possible combinations of size $C = 1, 2, 3, \ldots, n$ so long as each size, C results in at least one improvement to the current optimal solution. Consider a case in which the current optimal solution has cardinality C. If no improvement to the current optimal solution is made during the evaluations of all sets of size $C + 1$, then the algorithm terminates as there is no possibility that an improvement will be made for $|r| > C + 1$. A formal proof of the correctness of this stopping criterion is provided later. Figure 15 displays the stopping criterion employed in this improved enumerative algorithm and seek to give an intuitive feel for the subsequent proof.

Consider the set of five nodes shown in Figure 15. If this were a shortest path or TSP problem, it would be easy to show that if no improvement were made to the solution with the addition of one node, then no improvement can be made with two nodes. In fact, in these problem types it is simple to show that in Euclidean space, where the shortest distance between two points is a straight line, a feasible solution cannot be improved by even the addition of *one* node. However, the CRCNDP is a multi-objective problem and the addition of a node may lower some costs while increasing others. For example, adding an additional bus stop in a route may provide service to previously unserved customers thereby reducing the unserved demand cost while simultaneously increasing the travel time and operating cost. If the unserved demand cost reduction outweighs the increasing in travel time and operating cost, then there would be a net improvement to the solution.

Figure 15 shows five nodes with the three nodes on the right-hand side comprising the current optimal solution z^*. The two nodes on the left, i and j are nodes that are evaluated at later points in the enumerative algorithm when the cardinality of the sets being evaluated increases to $C + 1$ and then $C + 2$, where $C = |r_{z^*}|$.

$z^* + \pi_i$

$z^* + \pi_i + \pi_j$

$z^* + \pi_j$

r_{z^*}
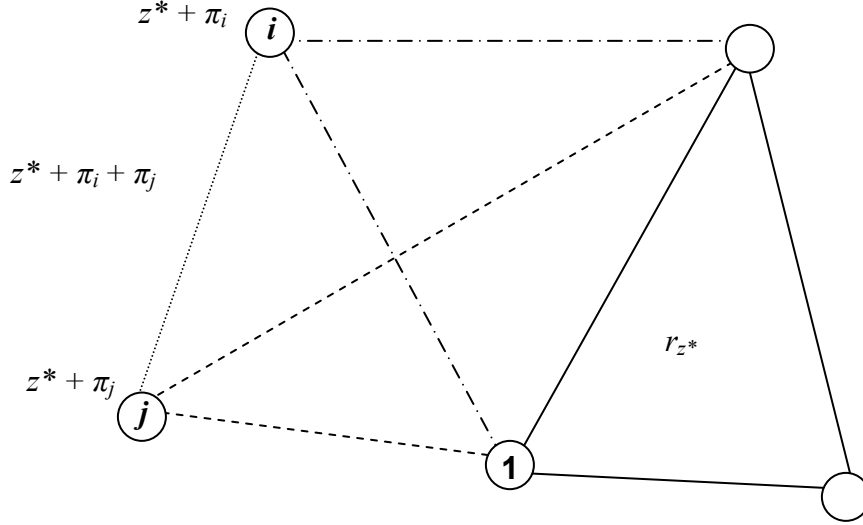
FIGURE 15: VISUAL DESCRIPTION OF STOPPING CRITERION

Node potentials, or marginal cost reductions are given for the two nodes not in r_{z^*} as π_i and π_j. These node potentials represent the improvement to the optimal solution should the node i or j enter the solution in addition to the three nodes comprising r_{z^*}. Using the node potential concept allows a good introduction to the more formal proof of stopping criterion correctness in the next section. In Figure 15, if adding nodes i and j to the original set of three improves the solution but adding either i or j alone does not, then $[z^* + \pi_i + \pi_j] < z^* < [z^* + \pi_i]$ and $[z^* + \pi_i + \pi_j] < z^* < [z^* + \pi_j]$ This requires both that $[\pi_i + \pi_j] < 0 < \pi_i$ and $[\pi_i + \pi_j] < 0 < \pi_j$. Here we find an inconsistency, the sum of the two node potentials is given in both statements as less than zero, however, to maintain either of the two statements one of the node potentials must be less than zero. If this is the case, then we would see an improvement by adding one node into the solution by itself. Either one of these requirements can be true independently. However, they cannot both be true.

Stopping Criterion Proof

Following is a formal proof of the correctness of the previously described stopping criterion for the enumerative algorithm. Simply stated, this criterion stops the enumerative algorithm if no improvements are made to the optimal solution for all feasible sets of a particular cardinality. Let $z^*(C)$, represent the current optimal solution of the CRCNDP, which covers the candidate stops in set r, where $C = |r|$, the cardinality of set r. We want to show that if there is no improvement to the optimal solution after examining all candidate stop sets of size $C + 1$, then the algorithm should terminate with $z^*(C)$ as the optimal solution.

If we are given that the current optimal solution has cardinality C and no improvement is made to the optimal solution after all feasible sets of size $C+1$ are examined, then we have:
$$z*(C) \leq z*(C+1) \tag{15}$$
Suppose that after all sets of size $C+1$ are examined with no improvement to the optimal solution an improvement is made to the optimal solution upon examination of sets of size $C+2$, then we have:
$$z*(C+2) \leq z*(C)$$
If the previous two conditions hold then the following must also hold:
$$z*(C+2) \leq z*(C+1) \tag{16}$$
If (16) holds then for some set of candidate stops q where $|q| = C+2$ the following must be true:
$$z(q) \leq z*(C)$$
Let S represent a subset of cardinality $C+1$ of the candidate stops in set q, $S \subseteq q$. Then the following must hold:
$$z(q) \leq z(S) \qquad \forall S \subseteq q, |S| = C+1$$
This implies that there is a node i that can be added to the subset S such that the following holds:
$$z(S + \{i\}) \leq z(S) \qquad \text{and} \qquad z(S + \{i\}) \leq z*(C)$$
Note that the condition in (15) requires that:
$$z*(C) \leq z(S) \qquad \forall S \subseteq q, |S| = C+1 \tag{17}$$
Provided that all feasible combinations are evaluated for each size of candidate stop set:
$$z(S + \{i\}) \leq z*(C + \{i\})$$
therefore,
$$z(S) \leq z*(C)$$
which contradicts (17). ∎

IMPROVED ENUMERATIVE ALGORITHM FLOW CHART

Figures 16a and 16b display the flow chart for improvements to the enumerative algorithm. These flow charts depict the checks in the algorithm used as stopping criteria within the algorithm. Figure 16a shows the preprocessing steps used to improve the efficiency of the enumerative algorithm. The preprocessing begins with given data: candidate stops, demand centroids, distances between stop-stop pairs and stop-centroid pairs, and the demand at each zone. The stopping check *ftest* is initialized so that a set size of 1, which will produce no feasible sets, does not stop the algorithm. The algorithm proceeds to the computation of walking coverage, demand served at stops, unserved demand and dwell time.
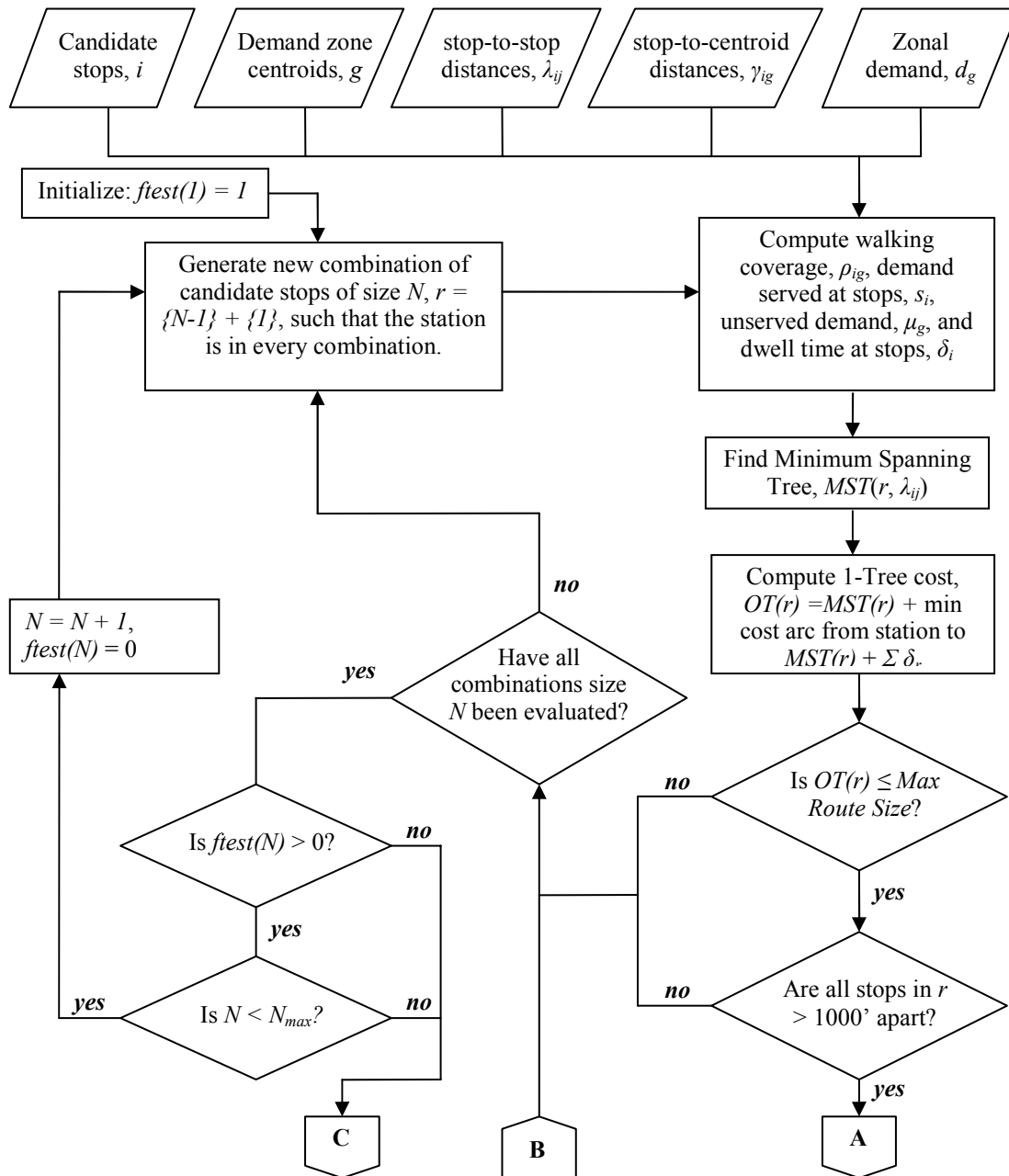
FIGURE 16A: IMPROVED ENUMERATIVE ALGORITHM FLOWCHART – PREPROCESSING PORTION
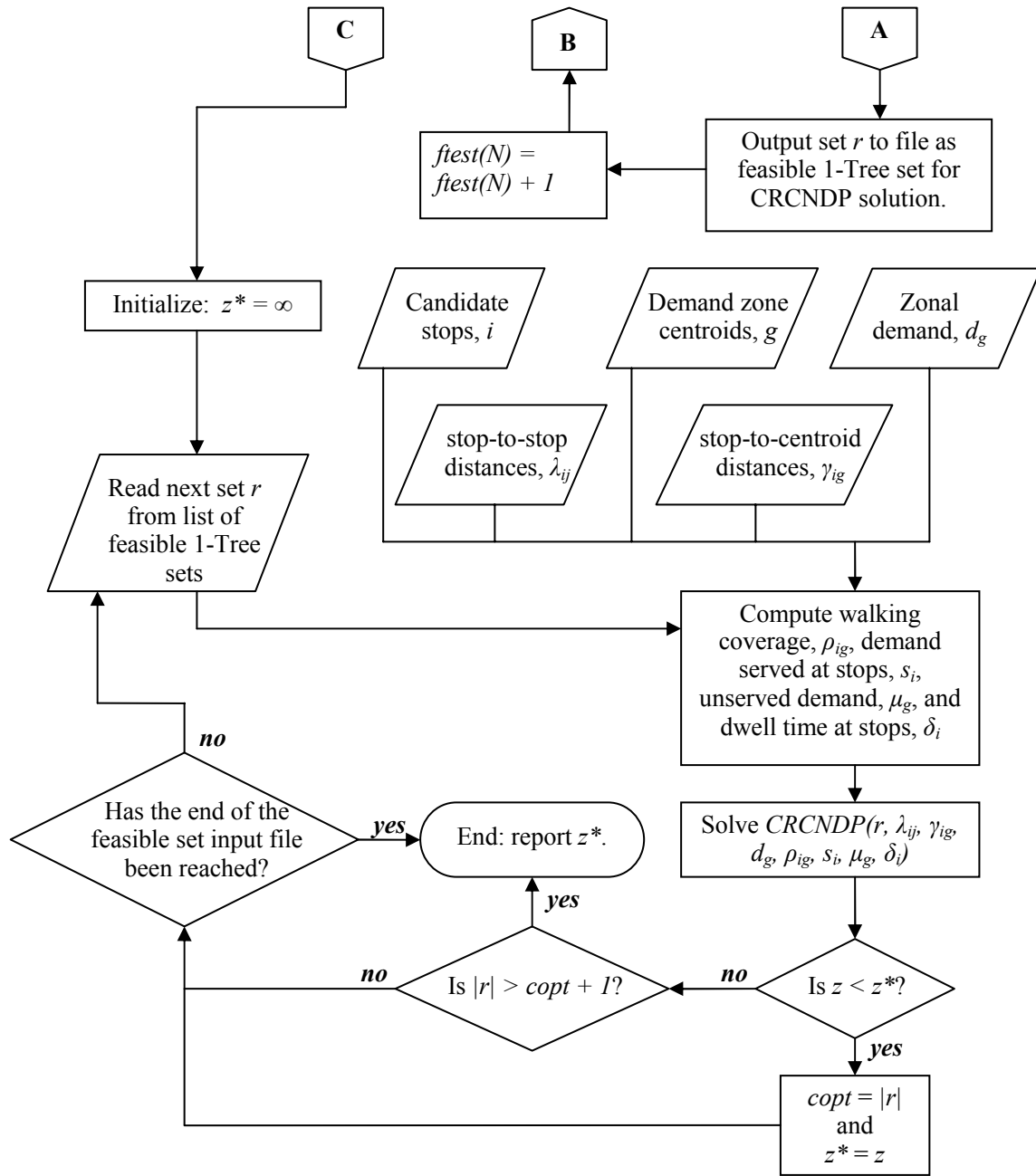
FIGURE 16B: IMPROVED ENUMERATIVE ALGORITHM FLOWCHART – CRCNDP PORTION

This initial information is used to find the minimum spanning tree (MST) for each set of candidate stops and the 1-Tree. If the 1-Tree cost is less than the allowable route size and all stops in the set are greater than 1000 feet apart, then the candidate node set is added to the list of feasible sets and *ftest* is updated. If these conditions are not met, then the next combination of candidate stops is evaluated, provided one exists and the *ftest* conditions are satisfied.

The check *ftest(N)* is zero if no sets of cardinality size *N* have had feasible 1-Trees. A feasible 1-Tree of size *N* adds 1 to *ftest(N)*. If all sets of size *N* have been evaluated and none are

feasible, the algorithm terminates as there will not be any sets of size *N+1* that will be feasible either. The basic concepts of the previous proof for the CRCNDP stopping criterion are applicable in this case.

The preprocessing steps are undertaken for all possible combinations or until there are no feasible sets of a particular size. All feasible sets will have been written to file for use as input to the CRCNDP. Figure 16b details the CRCNDP solution process, starting with much of the same information as the preprocessing stage. The CRCNDP solution process begins by initializing the global optimal solution to infinity. The algorithm then selects a feasible set from the list created by the preprocessing stage and solves the CRCNDP. If an improvement to the global optimal solution results, then set the test *copt* equal to the cardinality of the set and update the global optimal solution. If no improvement results, then check *copt*, the CRCNDP stopping criterion. If the cardinality of the current set is greater than *copt + 1* then the algorithm terminates and $z*$ is reported. Otherwise, the next feasible set in the list is evaluated if one exists, if not, the algorithm is terminated and $z*$ is reported.

PERFORMANCE

The improved enumerative method's performance with the 1-Tree preprocessing and practically advantageous stopping criterion engaged is depicted in Table 9. The two methods are compared for the 12-centroid geometry used in the enumerative algorithm application and a generated 20-centroid network. Each of the networks employed a 30 minute commuter train headway (maximum route length for seamless transfer), and unserved demand costs of 6000 dollars/unit unserved demand. They both incorporate a maximum walking distance of ½-mile in determining the number and nature of walking trips to centroids.

TABLE 9: ENUMERATIVE AND IMPROVED ENUMERATIVE ALGORITHMIC PERFORMANCE

Network Description	Enumerative Optimal Solution	Enumerative Solution Time (s)	Improved Algorithm Solution	Improved Algorithm Solution Time (s)
12 Centroids	$z* = 500356$, $r_{z*} = \{1, 6, 8\}$	178	$z* = 500356$, $r_{z*} = \{1, 6, 8\}$	11*
20 Centroids	$z* = 1037340$, $r_{z*} = \{1, 5, 12, 18\}$	50180	$z* = 1037340$, $r_{z*} = \{1, 5, 12, 18\}$	366**

*Required 39 Seconds of preprocessing time
**Required 7140 Seconds of Preprocessing time

Neither of these two scenarios distinguish between bus stops and centroids, as the limits of the enumerative technique are practically reached (12-hour running time on a 16 parallel Xeon 3.0 GHz processor configuration) with the 20-centroid example. The remaining cost parameters are unchanged from the earlier enumerative example.

It is readily apparent that the improved enumerative technique vastly outperforms the complete enumeration method. This is not surprising as complete enumeration should represent about the worst (computational effort-wise) method of solving the problem. It should be noted that the solutions arrived at by both methods are exactly the same. This supports the earlier proof that the stopping criterion and the implemented 1-Tree preprocessing scheme maintain the guarantee of solution optimality.

The difference in computational effort required by the two methods is striking as the improved algorithm solved the 12-centroid problem in 11 seconds compared to 178 and the 20-centroid problem in just over 6 minutes compared to just under 14 hours for the enumerative algorithm. The sources of these improvements are 1) The reduction in infeasible calls to GAMS, and 2) The reduction in the overall number of candidate sets considered. In the 12-centroid problem, for example, the enumerative method solves the CRCNDP 2048 times, with 1462 of these resulting in infeasibility. By contrast, the improved algorithm, after 39 seconds of preprocessing, solved the CRCNDP only 70 times, with a mere 36 of these resulting in infeasibility. For the 20-Centroid problem, enumeration required the CRCNDP to be solved 524,288 times with only 445 of these resulting in feasible solutions, meaning that almost 524,000 galls to GAMS were for infeasible sets. The improved algorithm reduced this inefficiency to 1069 calls to GAMS (to solve the CRCNDP), of which 718 were for infeasible sets. A more detailed comparison of the two sets of results shows that the 1-Tree preprocessor is responsible for eliminating 3971 infeasible sets before the stopping criterion eliminates the rest.

The contrast between the algorithms is reduced when the preprocessing time of the improved algorithm is taken into account. For a one-time solution of the problem, the improved algorithm only outperforms enumeration by a factor of three for the 12-centroid problem and a factor of about 6 for the 20-centroid problem.

This reduced improvement for a one-time solution does not negate its utility in the context of problems of this size. While the improvements may be less profitable if one desires to solve the CRCNDP only once, they are compounded if one desires to solve the CRCNDP for a variety of cost parameter sets. The only things that would need to remain constant from implementation to implementation are the walking threshold, lambda and gamma parameters, and the dwell time computation. This allows the analyst to perform a variety of scenario analyses in a very short amount of time using the improved algorithm. The same variety of analyses would require a much more substantial time investment should complete enumeration be used.

SUMMARY

There are limits to the improved method and these are discussed in more detail in Chapter 6, in which a metaheuristic approach to solving the CRCNDP is presented. For many purposes the improved technique of this chapter will be sufficient for a network design analysis undertaken by a transit authority seeking to optimize the circulator network of a new commuter rail line. Some networks may exceed the limits of this improved method or the processing capabilities available to the analyst. In these cases, one may be willing to sacrifice a guarantee of optimality for computational expediency. In these cases a metaheuristic method of solving the CRCNDP would be most welcome as they can often provide very good, near-optimal solutions at a fraction of the computation cost.

The next chapter details the development of a tabu search metaheuristic designed specifically for the CRCNDP. This chapter includes a broad discussion of tabu search and the specific means of tailoring this generic method to the CRCNDP. Along with the details of the development of the method a comparison of the performance of the tabu search metaheuristic to the enumerative technique and the improved exact method. A new example problem is also described for the MLK station in Austin, Texas that incorporates distinct demand centroids and candidate bus stops.

CHAPTER 6: METAHEURISTIC APPROACH

Improvements to the naïve enumerative method of solving the CRCNDP have been discussed; however, even these have practical limitations. The improved technique solves the CRCNDP well for small- to medium-sized networks provided that devoting resources to a preprocessing effort is not overly undesirable. The improved technique solves a CRCNDP problem of 20 centroids in roughly 6 minutes, provided that 2 hours of preprocessing are undertaken.

There are several scenarios in which this preprocessing will be undesirable. First, if an analyst seeks to obtain a solution for planning purposes in a very small amount of time, then the improved enumerative method may not provide the needed utility. If online routing decisions are needed, for instance, in a dial-a-ride program or for an ITS-enabled bus system with real-time passenger destination data, two hours of preprocessing time will not be available. In addition, if larger networks are to be considered when solving the CRCNDP then one must expect that the computational effort in solving a larger problem will be greater than that for the example problems given. In this chapter a larger network will be considered, one in which 10 demand centroids are served by 20 candidate bus stops, and the improved enumerative algorithm does indeed require substantially more time to arrive at a provably optimal solution.

In each of these cases it would be desirable to have an option for solving the CRCNDP that provides a good solution quickly. Of course this is always desirable, and in some cases can be remedied through the application of increasingly sophisticated hardware. If one does not have the resources to address computational issues with hardware (and even if one does), then a clever technique for solving the problem could prove very useful.

Of course there is a tradeoff for improved computational speed. In most cases of constrained multi-objective optimization, this tradeoff is in the guarantee of optimality. For every unit of efficiency gained through non-exact solution techniques, there is at least some loss in confidence that one can have in the optimality of the obtained solution, or the confidence one has that the solution method will provide good solutions to *all* problems. The goal is then to develop a robust solution method that not only solves difficult problems well consistently, but does it in an efficient manner.

This is an elusive goal, and as such, has attracted much interest over the years as optimization methods have become more and more common in a variety of fields. There is a class of solution methods, metaheuristics, which are general heuristic solution methods that can be tailored to solve a specific problem. Common examples include tabu search, genetic algorithms, simulated annealing, and ant colony optimization. These metaheuristics, in their general form, cannot be guaranteed to perform any better on average than any other heuristic algorithm. But, one can tailor these algorithms for specific problems, using the general form as an outline, that do in practice perform very well. Tabu search is a very general strategy that has been applied successfully in several applications that are very relevant to the CRCNDP, a recent example of which is Fan and Machemehl (2004).

Tabu Search

There is no single way to implement a tabu search strategy to solve a particular problem. The general tabu search framework can be applied to a large number of combinatorial optimization problems, the versions that are tailored to individual problems are such that they capitalize on specific aspects of the problem of interest to arrive at good solutions with minimal

effort (compared to enumerative strategies). Figure 17 provides the generic version of tabu search, as can also be found in Glover and Laguna (1997).

Step 1: (Initialization)

(A) Select a starting solution $x^{now} \in X$

(B) Record the current best known solution by setting $x^{best} = x^{now}$ and define $best\_\cos t = c(x^{best})$

Step 2: (Choice and termination)

Determine $Candidate\_N(x^{now})$ as a subset of $N(H, x^{now})$. Select x^{next} from $Candidate\_N(x^{now})$ to minimize $c(H, x^{now})$ over this set. Terminate by a chosen iteration cut-off rule.

Step 3: (Update)

Re-set $x^{now} = x^{next}$ and if $c(x^{now}) < best\_\cos t$, perform Step 1(B). Then return to Step 2. Update the history record H.

FIGURE 17: GENERIC TABU SEARCH METHOD

The first step selects a starting solution and performs the appropriate accounting procedures. This is a critical step in the tabu search strategy, as defining a good starting point for the method can significantly influence the goodness of the solutions evaluated as the method is running. The specific method of identifying the starting solution for the CRCNDP will be presented later.

The second step of the method uses the starting solution to define a neighborhood of solutions that are near the starting point, $N(H, x^{now})$. These neighborhoods solutions have a history associated with them, H. This history is what determines whether or not a neighborhood solution can be evaluated. The neighborhood solutions that may be visited, $Candidate\_N(x^{now})$, are then evaluated to determine the best solution to evaluate next, x^{next}. This step is iterated a predefined number of times, which is why one can guarantee good computational performance. The third step then updates the current solution and checks to see if the best neighborhood solution from Step 2 is the best global solution uncovered. If so, then the global best solution is updated and the method continues.

Glover (1989) discusses the performance of tabu search in a TSP context, a problem that is of considerable relevance to the CRCNDP. Glover describes an "easy" 42 city TSP problem and experience in applying tabu search. The specifics of the 42 city problem are not of interest, however, the results of this analysis are illustrative of tabu search in general. Figures 18 and 19 are borrowed from Glover (1989) and display the nature of tabu search heuristics well.

TABU SEARCH — 42 CITY TSP
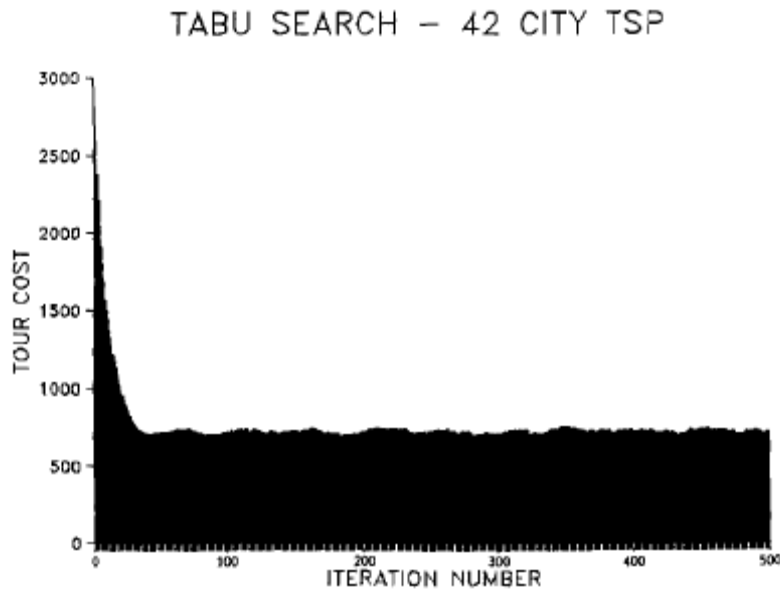
FIGURE 18: FIRST 500 ITERATIONS OF GLOVER (1989) 42 CITY TSP SOLUTION
source: Glover(1989)

Figure 18 provides a visualization of the first 500 iterations of Glover's solution. One can see that starting from a relatively poor starting solution, the algorithm quickly targets a good solution range and after the first 30 iteration or so the solutions oscillate about this good solution range. From this perspective, it appears that tabu search is very efficient at identifying a good solution and that the subsequent 470 iterations are really wasted effort. Figure 19 provides a different perspective on this data, viewing iterations 30 to 130 at higher resolution.

TABU SEARCH — 42 CITY TSP
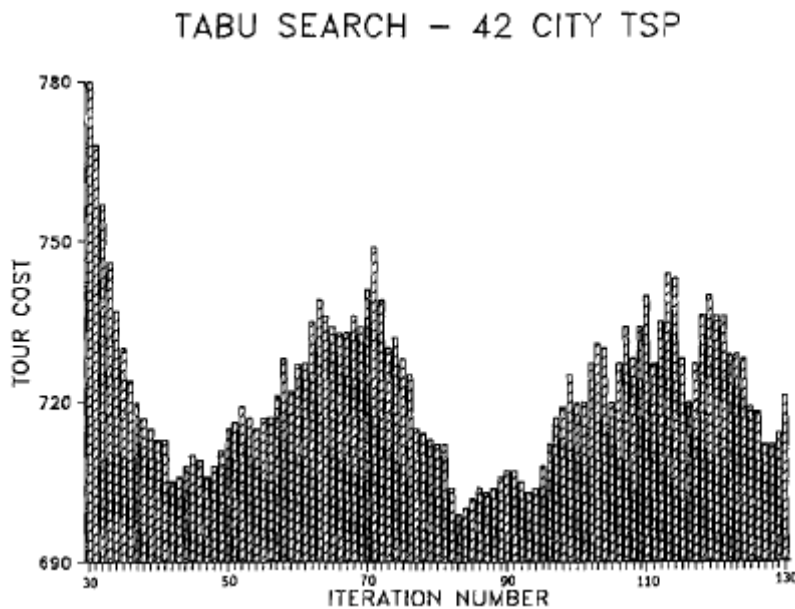
FIGURE 19: ITERATIONS 30 TO 130 OF GLOVER (1989) 42 CITY TSP SOLUTION
source: Glover(1989)

77

Viewing the range of good solutions at this resolution highlights the variability of solutions and the relatively small difference between one of the many good solutions and the optimal solution (which was found in iteration 83). This aspect of tabu search, identifying many solutions within a good range, can be very useful in a multi-objective optimization context. As discussed previously, much of the problem data, including costs, has a fair amount of variability or uncertainty associated with it. If this uncertainty is not accounted for specifically in the formulation, then the definition of "optimal" is of a more subjective nature. One must qualify any solution obtained (by any method) with the particular restrictions imposed by the uncertainty of the data. Multi-objective problems used in a decision support context, such as the CRCNDP, may then benefit equally from a series of good solutions for presentation to a decision-maker and from a globally optimal solution subject to a multitude of qualifications.

This consideration is important in the decision-making context of the CRCNDP and deserves some thought. As has been mentioned, the CRCNDP data in practice will always have some degree of associated uncertainty. Cost data, especially unserved demand, will always be subject to debate and will be impossible to pin down to an exact value or function. Even an elaborate function seeking to precisely define unserved demand based upon other system characteristics will necessitate an error term to account for the additional uncertainty that will certainly be present. Therefore, if any optimal solution must have a qualification for the uncertainty inherent in the problem, then perhaps the extra computational effort required to arrive at a provably optimal solution is not always justified. In fact, obtaining several good solutions will be necessary in most cases because one would want to display the effects of varying cost parameters on the solution, that is, demonstrate the robustness of the optimal solution.

Tabu search will not aid in showing theoretical robustness, as one needs an optimal solution as a benchmark by which to evaluate subsequent solutions. Detailed planning processes, in which online routing is not of interest, would be the sort of application in which solution robustness is of interest and would benefit from an exact solution method. Sketch planning or online route design, by contrast, would benefit more from a good solution quickly. It is this second case in which a tabu search method will perform well.

TABU FOR THE CRCNDP

Defining the history that will be used to determine whether a particular neighborhood solution is tabu can be as simple or complex as desired. For implementation with the CRCNDP, the strategy was to develop as simple a tabu search heuristic as could produce good solutions consistently. As will be shown and discussed subsequently, the CRCNDP tabu search is a relatively simple implementation, using two recency-based restrictions to avoid repetitions (getting stuck in local optima) and a restriction to avoid "bad", or infeasible solutions. These restrictions are used in combination with a simple attribute of each solution to guide the solution process. Following is a detailed description of the CRCNDP tabu process.

Initialization

Tabu search, as with many heuristics, benefits from defining a good starting point. Since tabu search is restricted to a particular number of iterations, getting the most from these limited iterations is of the utmost importance. Because the CRCNDP is driven in a significant fashion by the unserved demand cost, it is logical that one of the most important goals of the formulation is to serve as many people as possible with the circulator system. The tabu search algorithm for

the CRCNDP uses this aspect of the problem in selecting its starting point in the following two steps:

1. For a set *r*, select the |*r*|-1 highest demand serving locations (stops or centroids depending upon the application) and set these locations plus the station as the initial set *r*.
2. If this set produces an infeasible solution, replace members of the set until a feasible set *r* is encountered. This feasible set will then be the new starting point.

In practice this system of finding an initial solution works well. For very small sets, |*r*| < *4*, the initialization will require only one or two iterations to arrive at an initial feasible solution. For larger sets, the time taken to find a feasible set is longer, but not prohibitively so. Additionally, as discussed later, the initial solutions have been found to lead to very good (or optimal) solutions with the prescribed iteration limit.

Neighborhood Generation (Choice and Termination)

Tabu search for the CRCNDP operates on the outer level of the problem. The heuristic is searching for good sets of bus stops to include in the route design optimization routine. When defining the neighborhood solutions that tabu search is to evaluate the algorithm will attempt to cleverly select candidate stop sets for which the CRCNDP will be solved. This heuristic employs a method of selecting a stop to leave the current set *r* and several options to fill the empty slot in the set *r*, these stop options will define the various neighborhoods for consideration. The following steps summarize the neighborhood generation phase of the CRCNDP tabu search algorithm.

1. Select leaving stop randomly, provided the stop is not tabu as a recent addition to the set. Increment the leaving tabu parameter for the leaving stop accordingly.
2. Select entering nodes using one of two strategies: based upon the following attribute, a_i, of the candidate entering stops, or random selection. The attribute a_i is defined as:

$$a_i = \frac{d_i}{\sum_{j \in r} \lambda_{ij}}$$

Conceptually, this attribute is a ratio of the demand served by the candidate stop to the total distance from that stop to the remainder of the set *r*. By selecting the non-tabu stops with the highest value of a_i, the algorithm is seeking to include stops that serve a high amount of demand and are a small distance from the other stops in the set *r*. It can be argued that the nature of the CRCNDP tends to favor clusters of stops, which this attribute promotes in the entering stops it selects. This tendency towards clustering has been acknowledged and exploited in the dial-a-ride problem, which is related to the CRCNDP, by Cordeau (2006).

The second step, selecting entering stops, is performed randomly at a predefined interval. This is done to push the algorithm to look for non-cluster solutions alongside the ones maximizing the attribute, a_i. Some analysts may prefer only using, a_i and the algorithm can be

tailored as such. However, in this report, every other iteration will select a neighborhood randomly in an attempt to possibly capture a far superior route that may not be well represented by a_i. In practice within the CRCNDP, this has been an effective means of choosing entering stops.

Updating

The final step shown in the general tabu method in Figure 18 is the updating phase. This can be a trivial step in which mostly accounting processes are undertaken, but the CRNCDP tabu also uses this step to force a broader search scheme. This broader search scheme seeks to ensure that the search does not get stuck in neighborhoods surrounding infeasible sets or neighborhoods of local optima that are far from the global optimum value. The following steps are part of the update tabu search phase:

1. If a neighborhood solution results in an improvement to the global best solution, update the global best solution and set the neighborhood solution as the baseline solution for the subsequent iteration.
2. If no improvement is made to the global best solution, maintain the same baseline solution and investigate new neighborhoods in the next iteration.
3. After a set number of iterations, if no improvement to the best solution is made, randomly select the next baseline solution.
4. Update tabu durations.
    a. If a neighborhood resulted in an infeasible solution, increment the infeasible tabu parameter for the entering stop accordingly.
    b. If a neighborhood solution is selected as the next baseline solution, increment the entering tabu parameter of the entering stop accordingly.
    c. At each iteration, decrement the tabu parameters (that are greater than zero) for every stop by one.

The first two steps direct the algorithm to find and investigate good solutions. If a neighborhood has made an improvement to the global best solution, then it is selected to be the next baseline solution and new neighborhoods are developed around it. The third step addresses the situation in which no improvements are made within a set number of iterations. In this situation, the algorithm is instructed to choose one of the previous neighborhoods at random for the next baseline solution. This is done to investigate a wide variety of sets and works well in practice for the CRCNDP.

The final step of the updating phase specifically addresses the tabu parameters of each candidate stop. Each time a stop leaves the set r, enters the set r, or produces an infeasible solution, its associated tabu parameter is incremented. In this manner, the algorithm is instructed to "stay away" from these tabu stops so that it does not get stuck in unproductive solutions and neighborhoods. The duration of the tabu is a parameter set by the analyst and is discussed in some detail later. It is sufficient at this point to know that the three tabu parameters are incremented when their conditions are met and that all tabu parameters are decremented by one each iteration so that the stops can be members of r in later iterations when they might be in more productive company.

The tabu search algorithm for CRCNDP is relatively simple to state, which is the beauty of this particular metaheuristic. The algorithm identifies good solutions, digs deeper to see if improvements can be made to these good solutions, and if not, moves on to see if other productive solutions can be found. This algorithm is especially effective with the CRCNDP, as

the size of the problem is restricted (by the very characteristics that were used in the improved enumerative algorithm) in a practical sense and a good implementation of tabu search should be able to identify the best neighborhoods for investigation rather efficiently. The performance of the algorithm and sensitivity to the various components of the algorithm are discussed next.

TABU PERFORMANCE ON CRCNDP

The gains in computational effort offered by heuristic methods will be best seen in larger applications. Table 10 contrasts the performance of tabu search with the previous two methods used: enumeration and improved enumeration. Tabu was implemented in both of the previous networks along with a third generation of the network surrounding the MLK station in Austin, Texas. For current purposes, it is sufficient to know that the third generation network includes 10 demand centroids and 20 candidate bus stops serving these centroids. This 20 – 10 network is significantly more complex and requires a significant increase in computational effort, as is shown in Table 10.

Table 10: Tabu Search Performance Comparison for Three Sample Networks

Network Description	Enumerative Optimal Solution	Enumerative Solution Time (s)	Improved Algorithm Solution	Improved Algorithm Solution Time (s)	Tabu Solution†	Tabu Solution Time (s)
12 Centroids	$z^* = 500356$, $r_{z^*} = \{1, 6, 8\}$	178	$z^* = 500356$, $r_{z^*} = \{1, 6, 8\}$	11*	$z^* = 500356$, $r_{z^*} = \{1, 6, 8\}$	56
20 Centroids	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	50180	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	366**	$z^* = 1038780$, $r_{z^*} = \{1, 5, 18\}$	266
10 Centroids / 20 Bus Stops	--	> 12 hours	$z^* = 90105$, $r_{z^*} = \{1, 10, 13\}$	1731***	$z^* = 90105$, $r_{z^*} = \{1, 10, 13\}$	344

*Required 39 Seconds of preprocessing time
**Required 7140 Seconds of preprocessing time
***Required 11986 Seconds of preprocessing time
†Tabu Search performed 100 iterations at each size of set r

Note: The difference in tabu solution for the 20 centroid network stems from nodes 11, and 14 not being served (12 is served by walking trip in tabu solution) nodes 11 and 14 only account for about 4% of demand, which explains why the tabu solution, though not optimal, is very good (within 4% of optimal).

Note that the 20 − 10 network required approximately an additional hour of preprocessing for the improved algorithm beyond to the 20 centroid network.

The computational effort benefits are obvious. Tabu search requires no preprocessing and yet still provides good solutions to the CRCNDP in less time than the improved enumerative algorithm. The computational benefits are negated in the smallest network, as completing the 100 iterations for each size of r required more time than the preprocessing phase of the improved algorithm. The benefits are obvious for the 20-centroid and 20-10 networks, especially when the preprocessing time is considered.

This brings one back to the point made earlier that the application will drive the method used. If a baseline optimal solution is needed, for say, investigating the robustness of an optimal solution or performing sensitivity analysis, then the improved algorithm will be the better choice. Because the preprocessing is completed only once, the improved algorithm will perform well in these circumstances. If one desires a quick, good solution, e.g., for sketch planning or real-time applications, the tabu search method will be superior.

The performance of tabu search will also vary with the parameters of the algorithm. The number of iterations, number of neighborhoods evaluated and tabu durations may have considerable impact on the goodness the solutions. Tables 11 - 14 contain the results of a sensitivity analysis of the tabu algorithm to these particular parameters. Tables 11 and 12 show investigations into the impact of the number of iterations on the solution goodness and the solution time. This investigation is undertaken for the 20-centroid and 20-10 networks, looking at the solutions obtained from 10, 20, 50, 100, and 200 iterations. Each iteration count is for each size of the set r, so an application that is looking at all set sizes from 1 to 5 (which is the cutoff for these examples), would actually perform 50 total iterations at 10 iterations per size of r.

Tables 13 and 14 present the sensitivity of tabu solutions to the number of neighborhoods and tabu durations employed, respectively. Each scenario will run for 100 iterations on the 20-Centroid network.

TABLE 11: TABU SENSITIVITY TO NUMBER OF ITERATIONS, 20 CENTROID NETWORK

20 Centroid Network				
Iterations per each set size $\|r\|$	Tabu Solution	% Difference from Optimal	Iteration Best Solution Obtained	Running Time (s)
10	$z^* = 1082160$, $r_{z^*} = \{1, 5, 12\}$	4.3%	16	35
20	$z^* = 1070190$, $r_{z^*} = \{1, 4, 5, 18\}$	3.2%	58	53
50	$z^* = 1070190$, $r_{z^*} = \{1, 4, 5, 18\}$	3.2%	102	133
100	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	206	260
200	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	546	506

The 20 centroid network, which has a more evenly distributed demand, reached the optimal solution using 100 iterations for each size of the set, r. At iteration 206 (the 6th set of size $|r| = 4$), the optimal solution of $r = \{1, 5, 12, 18\}$ was found. Note that at the lower iteration counts, the solutions obtained were still very good. Using only 10 iterations, the best solution

83

found was within 4.3% of optimality. This improved to 3.2% for the 20 and 50 iteration implementations. Considering that the 100 iteration implementation required less than five minutes to arrive at the optimal solution, one could certainly argue that the improved solution is worth the investment of effort.

For the 20 – 10 network, it appears that the tabu search method is even more efficient at finding the optimal solution. The demand distribution of this network is less evenly distributed so using the highest demand serving stops as the initial solution serves as an excellent starting point for this type of network. Additionally, since the starting point of a particular set size, $|r|$ is based on the previous best solution, this particular network lends itself well to this solution method. Note that in the 50 iteration and 200 iteration examples the optimal solution was found very shortly after the 3-stop set r evaluations were started.

TABLE 12: TABU SENSITIVITY TO NUMBER OF ITERATIONS, 20 - 10 NETWORK

20 Candidate Bus Stop – 10 Centroid Network				
Iterations per each set size $|r|$	Tabu Solution	% Difference from Optimal	Iteration Best Solution Obtained	Running Time (s)
10	$z^* = 90435,$ $r_{z*} = \{1, 10, 12, 13\}$	0.4%	22	52
20	$z^* = 90105,$ $r_{z*} = \{1, 10, 13\}$	0.0%	48	87
50	$z^* = 90105,$ $r_{z*} = \{1, 10, 13\}$	0.0%	103	188
100	$z^* = 90105,$ $r_{z*} = \{1, 10, 13\}$	0.0%	131	454
200	$z^* = 90105,$ $r_{z*} = \{1, 10, 13\}$	0.0%	201	834

Table 13 displays the sensitivity of the tabu search to the various tabu durations. "Added" is a tabu parameter for stops recently added to r, "Removed" is for those recently removed, and "Infeasibility" is for those whose addition resulted in an infeasible solution. The added and removed durations are given as functions of the set size and infeasibility as a fixed value. This is done primarily based on observation of performance during the development of the tabu search. It was found that single-value durations for added and removed tabu parameters tended to result in infinite loops frequently as compared to basing the durations on set size. The default value of these tabu durations is given in the first row, with Added = $|r|$, Removed = $|r|$, and Infeasible = 2 iterations.

TABLE 13: TABU SENSITIVITY TO TABU DURATIONS: 20 – CENTROID NETWORK

20 Centroid Network -- 100 iteration per set size $|r|$					
Added Tabu Duration (iterations)	**Removed Tabu Duration (iterations)**	**Infeasibility Tabu Duration (iterations)**	**Tabu Solution**	**% Difference from Optimal**	**Iteration Best Solution Obtained**
$\|r\|$	$\|r\|$	2	$z^* = 1073860$, $r_{z^*} = \{1, 4, 5, 12, 18\}$	3.5%	394
$\|r\| - 1$	$\|r\| - 1$	2	$z^* = 1038780$, $r_{z^*} = \{1, 5, 18\}$	0.1%	142
$\|r\| - 1$	$\|r\| - 1$	1	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	206
$\|r\|$	$\|r\|$	1	$z^* = 1070190$, $r_{z^*} = \{1, 4, 5, 18\}$	3.2%	242
$\|r\| + 1$	$\|r\| + 1$	2	$z^* = 1087230$, $r_{z^*} = \{1, 11, 15\}$	4.8%	136
$\|r\| + 1$	$\|r\| + 1$	1	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	224
0	0	0	$z^* = 1038780$, $r_{z^*} = \{1, 5, 18\}$	0.1%	156

An important observation is that the tabu durations do not appear to limit the ability of the algorithm to find good solutions. While certain duration combinations appear to work better, this is based on a single run of each duration combination. All produced good solutions, two produced optimal solutions. Note that the default settings, which produced an optimal solution in Table 11 in 100 iterations did not produce one here in Table 13. This is the nature of a heuristic solution method. The ability of the method to produce good solutions consistently is a product of finding a good starting point and directing the heuristic to good candidate solutions using the attribute a_i. This is supported by the fact that durations of zero for each tabu parameter also yielded a good solution. This suggests that the initial solutions and attributes are directing the method toward a good solution and that tabu parameters may come into play for problems involving longer routes and larger networks.

Table 14 shows the sensitivity of the algorithm to neighborhood size. This is again presented relative to the size of the set r. It is readily apparent that all neighborhood sizes produced good solutions and that the method is rather insensitive to neighborhood size, provided that the number of neighborhood solutions investigated is at least equal to the current cardinality of the set r.

TABLE 14: TABU SENSITIVITY TO NUMBER OF NEIGHBORHOODS:
20 – CENTROID NETWORK

20 Centroid Network			
Neighborhood Size (# evaluated each iteration)	**Tabu Solution**	**% Difference from Optimal**	**Iteration Best Solution Obtained**
$\|r\| + 1$	$z^* = 1038780$, $r_{z^*} = \{1, 5, 18\}$	0.1%	184
$\|r\|$	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	228
$\|r\| + 2$	$z^* = 1037340$, $r_{z^*} = \{1, 5, 12, 18\}$	0.0%	236

SUMMARY

In certain contexts the tabu search method presented in this chapter outperforms the enumerative method presented in Chapter 4 and the improved enumerative method discussed in Chapter 5. When quick, good solutions are desirable, tabu search performs admirably an produces very good solutions under a wide variety of conditions. The tabu search method implemented for the CRCNDP has also shown itself to be robust, providing good solutions for a range of iteration counts, tabu durations, and neighborhood sizes. It also gives good solutions in an example where the demand is somewhat evenly distributed throughout a network and an example with more clustered demand.

The following two chapters will deviate from the previous three and discuss the CRCNDP in general rather than addressing specific solution methods. Chapter 7 will look at the sensitivity of various CRCNDP formulation parameters, most notably unserved demand and walking distance. Chapter 8 will apply what has been learned to the Austin MLK case study and present the results and policy implications in an applicable manner. The final chapter will seek to summarize the work in this report and provide closing thoughts and future research directions.

CHAPTER 7: CRCNDP SENSITIVITY ANALYSIS

To this point, it has been shown that the CRCNDP formulation and the solution methods developed for the CRCNDP work well in producing good, non-trivial solutions. However reasonable these solutions are, the correctness of the solutions is always subject to debate because of the uncertainty inherent in many of the parameters used in acquiring the solutions. There are two prominent examples of this uncertainty in the CRCNDP formulation, and these examples will be addressed in much detail in this chapter: unserved demand cost and walking distance threshold.

Unserved demand cost is a difficult cost to place an exact value upon. Conceptually, it is attempting to account for the cost of potential passengers not having access to the commuter rail system. This requires one to define both "potential passengers" and "access", which are interrelated terms. Potential passengers, in theory, could be everyone in a metropolitan area. This is not reasonable. Commuter rail systems, as developed in the U.S., do not operate on infrastructure that is placed to provide access to everyone in a metropolitan area. By design, commuter rail systems will serve only a limited proportion of the population, at least in the system's infancy. For the purposes of this report, the potential passengers of a commute rail system are considered those with a trip end within a 2-mile radius of a destination station. This represents a circulator trip of roughly 10 - 15 minutes, considered reasonable given that the circulator trip will likely be the end of a multi-modal, considerably longer (30 - 60 minute) overall journey. Arguments could be made for larger or smaller service areas, and the method developed here would still be applicable.

Accessibility is a concept that is frustratingly difficult to define with certainty as everyone's idea of what an accessible system is will differ somewhat. There are those who consider a ½-mile walk a perfectly reasonable definition of accessibility, and there are those who are unwilling or unable to walk even a few hundred feet. Here, unserved demand ties in with the other parameter under investigation in this chapter: walking distance threshold. This threshold is considered a measure of the accessibility of the circulator and hence, the commuter rail system. Does it represent an accurate definition of accessibility for everyone? Not likely. Grava (2003) does suggest that 80% of people consider ½-mile the upper limit of a reasonable walking distance to access public transportation.

In this report, unserved demand cost and the walking distance threshold have been selected as a measure of the accessibility of commuter rail. These two variables do not provide a comprehensive picture of how travelers determine accessibility, though they do serve as a good proxy, as has been argued here and in previous chapters.

In any case, these two variables and their impact on the CRCNDP's ability to provide good, correct solutions must be understood to justify any confidence in the methods proposed here as decision making and design tools. This chapter will seek to scrutinize these two variables in a variety of situations and better understand the goodness, correctness, and robustness of these measures.

Again, the 20 - 10 and 20-centroid networks will be used to test the characteristics of these parameters. Each analysis will be undertaken for both networks, as the 20 - 10 demand distribution is concentrated in one region whereas the 20-centroid demand distribution is more distributed. Because this type of analysis needs to compare the optimal solutions from a variety of parameter combinations the improved enumerative algorithm will be used for each portion of the analysis. The default parameters in both the 20 - 10 and 20 - centroid scenarios are given in

Table 15, not including the parameters that are the subject of the sensitivity analysis in this chapter.

TABLE 15: DEFAULT SENSITIVITY ANALYSIS DATA AND PARAMETER VALUES

Data	Definition	Value
C_o	cost of operating bus	$82/hour
C_{tt}	traveler cost of in-vehicle travel time	$13/hour
C_w	cost of walking	$25/hour
H	commuter rail train headway	30 minutes
F	number of buses available to the route	1
v_{bus}	local bus operating speed	10 mph
v_{walk}	pedestrian walking speed	2.5 mph

The data above are gathered from the same sources as the values depicted in Table 7. The values of these costs can also play a significant role in direction the solution of the CRCNDP, however, their role is more predictable. The sensitivity analysis pursued in this chapter will seek to identify performance characteristics of the CRCNDP under different combinations of walking threshold and unserved demand cost. Three walking thresholds will be investigated: ¼-mile, 3/8-mile, and ½-mile. These represent the endpoints and midpoint of the range suggested by Grava (2003) and other related research (see Chapter 3). Unserved demand will be scrutinized for a range of values within each of the three walking thresholds, from $1000 - $10000/unserved demand unit. This includes the previous default value of $6000/unserved demand unit. Results are presented in sections by walking threshold, as the preprocessing phase of the improved algorithm must be completed for each threshold value. Any sensitivity analysis performed within each walking threshold will not require additional preprocessing, which again highlights the benefit of the improved enumerative method.

20 CENTROID NETWORK

The first portion of the sensitivity analysis will deal specifically with implementation on the 20-Centroid network. The details of this network are given in Appendix A. This is a generated network borrowing lambda values from the 20-10 network but with generated demand figures. The 20-centroid network is intended to represent a larger network than the previous 12-centroid network with demand distributed over a wider range of centroids compared to the 20-10 network in which demand is concentrated in a fairly small area.

Walking Threshold: ¼-mile (1320 ft)

In Table 16 one can see that the CRCNDP is initially rather sensitive to unserved demand cost and later not sensitive at all. One would expect significant differences in the solution to the CRCNDP if unserved demand cost varied between $2,000 and $4,000. Outside of this range, however, the solution is relatively insensitive to the change in unserved demand cost. It is noteworthy also that each solution builds upon the previous. What appears to be happening in this case is the most straightforward situation – stops are added to serve a segment of the demand only once the unserved demand cost reaches a certain threshold. That is, the route adds stops but does not remove stops as the unserved demand cost goes higher.

TABLE 16: CRCNDP SENSITIVITY TO UNSERVED DEMAND:
THRESHOLD ¼-MILE, 20-CENTROID NETWORK

Walking Threshold - 1/4 mile (1320 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal z^* improvements
1000	$r_{z^*} = \{1, 8\}$	277141	5
2000	$r_{z^*} = \{1, 8\}$	542141	5
3000	$r_{z^*} = \{1, 7, 8\}$	764737	11
4000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	976776	13
5000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	1184780	13
6000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	1392780	13
7000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	1600780	14
8000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	1808780	14
9000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	2016780	14
10000	$r_{z^*} = \{1, 7, 8, 12, 18\}$	2224780	14

An explanation for this result lies in the walking threshold distance. With such a small walking threshold, the probability of serving multiple demand centroids goes down significantly as the walking coverage radius (walking threshold) goes down. Therefore, there is less overlap in walking coverage throughout the network and the decision is a one-to-one decision (See Table 17). That is, if a stop is added, it serves one demand centroid and that centroid's demand must be such that the reduction in unserved demand cost outweighs the additional travel and operation costs. What is happening in Table 16 as unserved demand cost goes up is:

- When unserved demand cost reaches about $3,000, serving stop 7 is justified.
- When unserved demand cost reaches about $4,000, serving stops 12 and 18 is justified.
- The network is laid out such that any additional stops simply do not justify the increased travel and operation cost.

The last column of Table 16 shows the number of suboptimal improvements made to z^* throughout the duration of the improved algorithm. Naturally this number rises as the number of stops in the set r increases as an improvement to z^* is a requirement for the algorithm to investigate the next greater set size. The number of improvements appears to level off once the largest set is uncovered suggesting that the increased unserved demand cost is not inducing the

algorithm to explore new regions of the solution space. At this walking threshold, it appears that past a certain point the solution is not very sensitive to the unserved demand cost.

There are two important observations from this first sensitivity analysis. First, the tradeoffs are straightforward when dealing with networks where the walking threshold allows for a one-stop, one-centroid coverage as shown in Table 17. Nearly every stop in this 20-centroid network (in which stops are at the centroids) has only itself within the walking threshold. Second, the critical values of unserved demand for this arrangement appear to be about $3,000 and $4,000. At approximately $3,000/unserved demand unit, the solution moves away from simply serving a single, nearby stop with high demand, to a multi-stop route. This suggests that at near this value the unserved demand cost begins to force non-trivial solutions. At greater than $4,000/unserved demand unit the solution fixes on a higher cardinality route set, suggesting that the solution obtained at $4,000 is the largest route set desirable. There are certainly route sets with more stops that are feasible, however, no improvements were made in this configuration.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	1	0																			
	2		0	1013	2298	1773	959	2352													
	3		1013	0	1791	1882	1502														
	4		2298	1791	0	959	1339	2443				2624									
	5		1773	1882	959	0	814	1484	2280												
	6		959	1502	1339	814	0	1864													
From	7		2352		2443	1484	1864	0													
	8					2280			0	1122		1049									
	9								1122	0	1809	2171			1809						
	10									1809	0										
	11			2624					1049	2171		0	1791	2425	1773						
	12											1791	0	1936	2298				1791		
	13											2425	1936	0	1701	2135			1846		
	14									1809		1773	2298	1701	0						
	15													2135		0	2497	2370	2533		
	16															2497	0	1574		2070	1538
	17															2370	1574	0			
	18												1791	1846		2533			0		
	19																2070			0	2106
	20																1538			2106	0

TABLE 17: WALKING COVERAGE TABLE, VALUES GIVEN IN RECTILINEAR FEET

Values in **Red**: Walking coverage with ½-mile threshold (in addition to 3/8 and ¼ mile coverage)
Values in **Blue**: Walking coverage with 3/8-mile threshold (in addition to ¼ mile coverage)
Values in **Black**: Walking Coverage with ¼-mile threshold

Walking Threshold: 3/8-mile (1980 ft)

Table 17 also contains information regarding the walking coverage of a 1980-foot walking threshold. In contrast to the results of the 1320-foot threshold, there is more coverage if a 3/8-mile walking distance is used as the maximum distance that a commuter rail circulator passenger will endure. Both the black and the blue highlighted numbers represent potential walking trips, and because of this increased coverage, the tradeoffs between stop location become more interesting and less simple to predict outright. This decrease in predictability is illustrated well in Table 18, which shows the complete results of the sensitivity analysis at this threshold.

Interestingly, for unserved demand cost less than $2,000, the solution is to serve only stop 9. However, stop 9 is not included in any other solution directly. In fact, until unserved demand cost reaches $9,000 per unserved demand unit, the demand associated with stop 9 is not served at all. The solution containing stop 9 serves two high demand stops, 8 and 10, via walking trips, which is why it is considered optimal even at this low level of unserved demand cost. Unserved demand cost of $3,000 and $4,000 produces the same solution, which includes stops 5 and 11, the stops upon which all subsequent solutions build upon. After unserved demand cost reaches $3,000, the solution behaves much as it did in Table 16, adding stops to the set once serving the demand justifies the travel and operating cost and not removing a stop from the route set.

For easier comparison, the critical values of unserved demand appear to be in the range of $4,000 – 5,000 and $8,000 – 9,000. Comparing to the shorter walking threshold, one notices that there are two fairly distinct critical regions. This is likely due to the additional coverage provided by each stop. In the 1320-foot scenario the coverage was limited to one stop (see Table 17). In this example, there is greater walking coverage and the tradeoffs through the addition of a bus stop are more complex. In this example, the second critical value represents the point at which stop 8 enters the solution. Stop 8 is a high demand stop and is covered by a walking trip in all previous solutions. At $8,000 – $9,000 it is a better solution to not only serve zone 8 with the shortest walking distance possible, but also serve the demand at zone 9, which has a small associated demand.

TABLE 18: CRCNDP SENSITIVITY TO UNSERVED DEMAND: THRESHOLD 3/8-MILE, 20-CENTROID NETWORK

Walking Threshold - 3/8 mile (1980 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal z^* improvements
1000	$r_{z^*} = \{1, 9\}$	255087	3
2000	$r_{z^*} = \{1, 9\}$	494087	3
3000	$r_{z^*} = \{1, 5, 11\}$	710072	5
4000	$r_{z^*} = \{1, 5, 11\}$	873072	6
5000	$r_{z^*} = \{1, 5, 11, 18\}$	1036070	7
6000	$r_{z^*} = \{1, 5, 11, 18\}$	1187980	8
7000	$r_{z^*} = \{1, 5, 11, 18\}$	1343980	8
8000	$r_{z^*} = \{1, 5, 11, 18\}$	1499980	9
9000	$r_{z^*} = \{1, 5, 8, 11, 18\}$	1654510	11
10000	$r_{z^*} = \{1, 5, 8, 11, 18\}$	1803510	11

At $8,000 – $9,000, there is a transition from simply serving high demand zones to serving the low demand zones as well. For later analysis these numbers may be useful as thresholds representing route design priorities: $4,000 as a cutoff between minimal routes and the larger, non-trivial routes and $8,000 as a cutoff between ensuring service to high and low demand zones.

The final column of Table 18 shows that nearly each increase in unserved demand cost results in more sub-optimal improvements to z^* prior to finding the optimal solution. This suggests that at 1980 feet, the CRCNDP is more sensitive to changes in unserved demand and higher unserved demand opens up new regions of the solution space for exploration. Notice the difference between $5,000 and $8,000 per unserved demand unit. Each of these costs resulted in the same solution, but at $8,000 two additional sub-optimal improvements were made to z^*. If one investigates the output from the sensitivity analysis, one finds that the higher unserved demand cost found several suboptimal improvements to the solution, and hence, potential good solutions that the smaller unserved cost ignored. If one desires to investigate a large number of solutions during the algorithm, then selecting a value somewhere within the range of the two previously discussed cutoff points may provide a good value for this type of analysis.

Walking Threshold: 1/2-mile (2640 ft)

Lastly, Table 17 contains the reachable centroids for a walking threshold of 1/2-mile (2640 ft) on the 20-centroid network. Expectedly, there is an increase again in coverage with this increased threshold and all of the values in the table are within the ½-mile threshold. This increase in coverage implies that zonal demand has a higher probability of being served by multiple stop locations than either of the two previous threshold values. Previous reasoning suggests that this will lead to more complex tradeoffs and likely more sensitivity to unserved demand. However, this is not supported by the results in Table 19. Table 19 shows that once the $5,000 – 6,000 range is met, the solution remains the same and until $10,000 is reached, there

are no additional sub-optimal improvements made while the algorithm is being applied. At 2640 feet, it appears that the CRCNDP is somewhat less sensitive to unserved demand cost.

Why would this be? If one removes the '0' distances from Table 17 and averages the walking distance between all stops within the three ranges, one finds that the average distance is 1830 feet with an associated standard deviation of 492 feet. 1830 feet is very near the 3/8-mile threshold used in the previous section, yet is greater than one standard deviation away from either of the two other walking thresholds.

TABLE 19: CRCNDP SENSITIVITY TO UNSERVED DEMAND:
THRESHOLD 1/2-MILE, 20-CENTROID NETWORK

Walking Threshold - 1/2 mile (2640 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal z^* improvements
1000	$r_{z^*} = \{1, 5\}$	270610	2
2000	$r_{z^*} = \{1, 5\}$	401365	2
3000	$r_{z^*} = \{1, 5\}$	590365	2
4000	$r_{z^*} = \{1, 5, 18\}$	764776	4
5000	$r_{z^*} = \{1, 5, 18\}$	901776	4
6000	$r_{z^*} = \{1, 5, 12, 18\}$	1037340	7
7000	$r_{z^*} = \{1, 5, 12, 18\}$	1159340	7
8000	$r_{z^*} = \{1, 5, 12, 18\}$	1281340	7
9000	$r_{z^*} = \{1, 5, 12, 18\}$	1403340	7
10000	$r_{z^*} = \{1, 5, 12, 18\}$	1525340	9

This observation suggests that the magnitude of the walking threshold is not the most relevant factor in sensitivity to unserved demand and the complexity of the tradeoffs, but it is the walking threshold relative to the average distance (or some aggregate measure) between stops. If the threshold is near the average distance the tradeoffs are more subtle and can yield more sub-optimal improvements on the path to optimality. This theory will be tested in the next network, the 10-centroid, 20-stop network in which the demand and spatial configuration is different.

CRCNDP Sensitivity to Walking Distance

The previous sections have dealt primarily with the CRCNDP sensitivity to unserved demand and how walking threshold influences this sensitivity. In this section, walking threshold will be looked at independently. Table 20 displays the sensitivity of the CRCNDP to walking threshold. For the 20-centroid network, the solutions to the problem are shown for each walking threshold at unserved demand values of $3,000, $6,000, and $9,000 per unserved demand unit.

94

TABLE 20: CRCNDP SENSITIVITY TO WALKING THRESHOLD VALUES: 20-CENTROID NETWORK

Unserved Demand Cost	Walking Threshold		
	1/4-mile	3/8-mile	1/2-mile
$3,000	$r_{z*} = \{1, 7, 8\}$	$r_{z*} = \{1, 5, 11\}$	$r_{z*} = \{1, 5\}$
$6,000	$r_{z*} = \{1, 7, 8, 12, 18\}$	$r_{z*} = \{1, 5, 11, 18\}$	$r_{z*} = \{1, 5, 12, 18\}$
$9,000	$r_{z*} = \{1, 7, 8, 12, 18\}$	$r_{z*} = \{1, 5, 8, 11, 18\}$	$r_{z*} = \{1, 5, 12, 18\}$

Table 20 provides three interesting insights. First, as one would expect, the optimal solutions include fewer stops (regardless of unserved demand cost) as the walking threshold increases since a greater amount of demand can be covered with longer walking trips and fewer stops are needed. Second, though the combinations change, the optimal solutions are all subsets of the relatively small superset, {1, 5, 7, 8, 11, 12, 18}. This suggests that any good route would also be a subset of these nodes. Third, it is useful to note that though in certain cases the smallest routes may be optimal, the larger routes are indeed feasible (able to return to station within the allotted train headway). Therefore, since one does not know the demand distribution precisely, one could safely serve stops at all of the stops in the larger sets (unserved cost = $9,000) knowing that if there is demand at each stop it can all be served with the seamless transfer concept intact.

10 CENTROID / 20 STOP NETWORK

The 10 centroid / 20 stop network (20-10) is a portion of the network surrounding the MLK Station in Austin, Texas. The demand is concentrated near the University of Texas and the State Capitol areas. The data used in this network is shown in Appendix B.

Walking Threshold: 1/4-mile (1320 ft)

The 1/4 –mile threshold yielded results consistent with the performance of the CRCNDP in the 20-centroid network. Once again, at about $3,000 - $4,000 the solution goes through a transition from small routes to large route sets that serve maximum demand.

TABLE 21: CRCNDP SENSITIVITY TO UNSERVED DEMAND: THRESHOLD 1/4-MILE, 20-10 NETWORK

Walking Threshold - 1/4 mile (1320 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal z^* improvements
1000	$r_{z*} = \{1, 8\}$	155880	3
2000	$r_{z*} = \{1, 8, 13\}$	295318	7
3000	$r_{z*} = \{1, 8, 11, 13\}$	405569	9
4000	$r_{z*} = \{1, 4, 8, 11, 13\}$	502234	10
5000	$r_{z*} = \{1, 4, 8, 11, 13\}$	598234	11
6000	$r_{z*} = \{1, 4, 8, 11, 13\}$	694234	11
7000	$r_{z*} = \{1, 4, 8, 11, 13\}$	790234	11
8000	$r_{z*} = \{1, 4, 8, 11, 13\}$	886234	11
9000	$r_{z*} = \{1, 4, 8, 11, 13\}$	982234	11
10000	$r_{z*} = \{1, 4, 8, 11, 13\}$	1078230	11

The large set of stops noted for the $4,000 and greater unserved demand cost are, like the 20-centroid network, a function of the walking threshold. Because each centroid is served by fewer stops with a ¼-mile threshold, the rising cost of unserved demand prompts the CRCNDP to try and include as many stops as possible so that as much demand can be served as possible.

Walking Threshold: 3/8-mile (1980 ft)

Table 22 displays the results of the unserved demand cost sensitivity analysis being applied to the 20 − 10 network with a 3/8-mile walking threshold. In the 20 − 10 network, the average distance from stop to centroid (those stops that are within the threshold distance) is 1550 feet with an associated standard deviation of 603 feet.

Walking Threshold - 3/8 mile (1980 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal z^* improvements
1000	$r_{z^*} = \{1, 7\}$	130275	3
2000	$r_{z^*} = \{1, 7\}$	245275	3
3000	$r_{z^*} = \{1, 4, 7, 11\}$	353702	8
4000	$r_{z^*} = \{1, 5, 7, 11, 12\}$	441820	10
5000	$r_{z^*} = \{1, 5, 8, 13\}$	514104	9
6000	$r_{z^*} = \{1, 5, 8, 13\}$	580104	9
7000	$r_{z^*} = \{1, 5, 8, 13\}$	646104	10
8000	$r_{z^*} = \{1, 5, 8, 13\}$	712104	10
9000	$r_{z^*} = \{1, 5, 8, 13\}$	778104	10
10000	$r_{z^*} = \{1, 5, 8, 13\}$	844104	10

The 1980-foot threshold is then within one standard deviation of the mean, and like the $20 -$ centroid network, there appears to be a fair amount of complexity in the tradeoffs being made within the CRCNDP. Table 22 shows that between $2,000 and $4,000, the optimal solution changes radically from a trivial one-stop route to a four-stop route. Additionally, in this span, stop 4 enters and leaves the optimal solution. Once again, it appears that $4,000 is a good approximation of the critical value at which the solution converges to a particular route and does not change.

A somewhat surprising result is the increase in cardinality to 5 when unserved demand is at $4,000 per unit with a subsequent decrease when unserved demand cost increases to $5,000 per unit. What is happening here is that the 4-stop set, $r = \{1, 5, 8, 13\}$ actually serves more demand than the 5-stop set. When unserved demand cost is less than $5,000 per unit, the optimal solution is as shown in Table 22, $r = \{1, 5, 7, 11, 12\}$. At the $5,000 threshold, not serving the demand at centroid 4 becomes too costly and subsequent solutions maintain the 4-stop route. This is an interesting result, suggesting that greater route size does not always mean greater demand coverage.

Walking Threshold: 1/2-mile (2640 ft)

The sensitivity analysis undertaken for a walking threshold of ½-mile produced unsurprising results (shown in Table 23). Whereas the ¼-mile threshold yielded larger route stop sets, the ½-mile threshold produced routes that place much of the burden of serving demand upon the walking portion of the trip. This burden results in fewer stops in an optimal route and less sensitivity overall to the unserved demand. These results are consistent with those obtained using the 20-centroid network. Again, the critical value of unserved demand appears to be at approximately $4,000 / unserved demand unit.

Walking Threshold - 1/2 mile (2640 ft)			
Unserved Demand Cost ($/unit demand unserved)	Optimal Route	Objective Function Value ($)	Sub-optimal $z*$ improvements
1000	$r_{z*} = \{1, 9\}$	107372	5
2000	$r_{z*} = \{1, 9\}$	195372	5
3000	$r_{z*} = \{1, 9\}$	283372	5
4000	$r_{z*} = \{1, 9, 13\}$	342868	7
5000	$r_{z*} = \{1, 9, 13\}$	399868	7
6000	$r_{z*} = \{1, 9, 13\}$	456868	8
7000	$r_{z*} = \{1, 9, 13\}$	513868	8
8000	$r_{z*} = \{1, 9, 13\}$	570868	8
9000	$r_{z*} = \{1, 9, 13\}$	627868	8
10000	$r_{z*} = \{1, 9, 13\}$	684868	8

The next section will take a more independent look at the relationship between optimal route sets and walking threshold. After this brief investigation, an overall summary of what has been learned about CRCNDP sensitivity to unserved demand cost and walking threshold will conclude the chapter.

Sensitivity to Walking Distance

The same observations that were relevant to the 20-Centroid network can be made for the 20-10 network in Table 24. The larger the walking threshold, the smaller the optimal set of stops is. In general, there is a relatively (to the total number available) small superset of stops, {1, 4, 5, 7, 8, 9, 11, 13, 14} which contributes to the stops in the various optimal solutions. Therefore, one can be safe in assuming that any good solution to the CRCNDP for the 20-10 network should be a subset of these nodes.

TABLE 24 CRCNDP SENSITIVITY TO WALKING THRESHOLD: 20-10 NETWORK

	Walking Threshold		
Unserved Demand Cost	1/4-mile	3/8-mile	1/2-mile
$3,000	$r_{z*} = \{1, 8, 11, 13\}$	$r_{z*} = \{1, 4, 7, 11\}$	$r_{z*} = \{1, 9\}$
$6,000	$r_{z*} = \{1, 4, 8, 13, 14\}$	$r_{z*} = \{1, 5, 8, 13\}$	$r_{z*} = \{1, 9, 13\}$
$9,000	$r_{z*} = \{1, 4, 8, 13, 14\}$	$r_{z*} = \{1, 5, 8, 13\}$	$r_{z*} = \{1, 9, 13\}$

There is much that can be learned from the analysis of the previous sections, though as with most research, more questions are raised than answered. Following are four useful observations that are true of the CRCNDP and can be used to guide future applications of this problem and solution method.

Observation #1: The CRCNDP behaves as expected

Though the sensitivity analysis did produce some unexpected happenings, the relationship between the optimal solution and walking threshold or unserved demand is in agreement with logic. Smaller walking thresholds resulted in routes with more stops and vice versa. Increasing unserved demand cost resulted in optimal routes with more stops and vice versa. The solutions found were in agreement with logic, yet were not trivial: a promising result.

Observation #2: Sensitivity to Walking Threshold is Relative

Rather than magnitude alone, it appears that the behavior of the CRCNDP is dependent somewhat on the walking threshold value relative to the distance between stop and centroid. If the threshold is near the average distance from stop to centroid, then the tradeoffs being evaluated by the CRCNDP are a bit more complex and result in a higher sensitivity to changing unserved demand costs. If the threshold is far from the average stop to centroid distance (greater than one standard deviation), then the tradeoffs are less complex and the solution tends to converge quickly toward a particular route that is more stable.

Observation #3: A good default value for Walking Threshold is 3/8-mile

This observation stems from the predictability of the solutions generated by the threshold and its relation to walking threshold values suggested in the literature. Three-eighths of a mile produced good solutions that were not wholly predictable. Smaller threshold values tend to visit as many stops as feasible, larger as few stops as possible. Additionally, the literature suggests a walking threshold of ¼ - ½ mile, and 3/8 is the midpoint of this range. Validating this distance with data collected from commuter rail circulator passengers would be of much use.

Observation #4: A good default value for Unserved Demand Cost is $6,000

Bailey (2007) can be used as a baseline for unserved demand cost, which suggests that $6,000 is the annual cost of not providing transit to a household. This cost is based on the cost difference of maintaining a one-car and a two-car household. This value has worked well in the analysis performed earlier in this work, and can certainly be employed without fear of prompting bad solutions. This value represents a very good starting point for any application of the CRCNDP; its use is supported by the analysis in this chapter. $6,000 per unserved demand unit sets serving demand as a high priority and using a cost any higher does not impact the results positively or negatively. However, depending on the importance of reducing unserved demand, different values of its cost should be considered (See Table 25).

TABLE 25: PRIORITY-BASED UNSERVED DEMAND COSTS

Importance of Reducing Unserved Demand	Unserved Demand Cost
Low	$2,000
Medium	$4,000
High	$6,000

It is easy to say that $4,000 or $6,000 per unserved demand unit are good values to induce non-trivial solutions to the CRCNDP. It is another thing altogether intuitively grasping what this value means. Bailey (2007) calculated this value based on the cost difference between a one- and two-car household (assuming that transit access allows one to eliminate a personal auto from the household).

This figure, in the context of the CRCNDP, may be accounting for more than auto ownership on an annual cost basis. It is certain that the unserved demand cost includes costs spread over a longer time frame than one trip on a circulator service. This cost could include environmental costs, social impacts of land use, congestion costs, and a variety of other factors that influence or are influenced by one's daily commute. The values shown in Table 25 work well in the CRCNDP; and Bailey (2007) provides a good explanation of where these costs originate. However, the issue is far from settled. It makes logical sense that access to transit affects many aspects of a household and a community and therefore any cost attempting to represent this access (or lack of) should account for this variety of factors. This issue will be a significant source of interest for the transit community and the research community in the coming years.

SUMMARY

The sensitivity analysis undertaken in this chapter has confirmed that the CRCNDP performs as expected and provides optimal routing schemes for a given set of input costs and parameters. Of course, other costs included in the CRCNDP objective function will impact the solutions, but these tend to be better understood than walking threshold and unserved demand. Operator costs, in-vehicle travel time, and walking cost should all impact the solutions in predictable manners. The next chapter, in investigating pareto-optimal solutions for the Austin case study will reveal some of these impacts of costs not dealt with in this chapter.

This chapter also revealed that the CRCNDP is indeed sensitive to unserved demand cost and walking threshold. Encouragingly, this sensitivity was displayed in a manner that is consistent with logic. Lower unserved demand cost and higher walking thresholds resulted in smaller routes sets. Greater unserved demand cost and smaller walking thresholds resulted in larger route sets.

Finally, this chapter confirmed that the estimate assembled by Bailey (2007) is a good estimate of unserved demand cost and that 3/8-mile is a good default value for walking threshold. However, these parameters are serving as a proxy for accessibility and the cost of not providing this accessibility to commuter rail. Public transportation accessibility is a more

complex issue and will be the subject of future work in this area. For the present, unserved demand cost and walking threshold will serve as good proxies.

Chapter 8 will apply the CRCNDP and developed solution methods to a full-scale case study in Austin, Texas. The interplay between cost parameters and accessibility measures will be further illustrated. Results from this case study will be presented in a manner suitable for contrasting route design priorities and aiding decision-makers in designing the best circulator route system for the Austin MLK station.

The final chapter of this report will bring together the goals and objectives of this work and the results of the efforts of this research. As with most research, more questions are raised than have been answered; the questions raised as part of this report will be noted and future research directions in support of these identified.

CHAPTER 8: CASE STUDY

In previous illustrations of the development and performance of the CRCNDP and the solution methods developed in this report example have been drawn from the MLK station of the new MetroRail system in Austin, Texas. These examples have been somewhat limited because of the size of the problem and its impact on the feasibility of using enumerative methods. The lessons learned through the development of these enumerative methods led to the development of a tabu search solution method for the CRCNDP, allowing for larger problems to be solved in an efficient manner. During the developmental stages it was found that the tabu search method performed well, providing good (and in some cases, optimal) solutions in a very short amount of time compared to the enumerative and improved enumerative algorithms. This performance generates confidence in the tabu search's ability to provide good solutions, an ability largely due to picking good starting solutions and its breadth-based search.

The case study undertaken required that the tabu search method be used, as the case study network is large enough to make even the improved enumerative algorithm prohibitively computationally expensive. The case study will be presented in the context of developing good solutions for several different implementation strategies. The three strategies are: 1) Cultivating Ridership, 2) Minimal Operation cost, and 3) Balanced Approach. As each of these three strategies is investigated, the details of each will be presented. First, the network used in this case study will be briefly presented. Because of the size of the network, the input data is not shown in its entirety.

NETWORK DESCRIPTION

Figure 20 provides an aerial overview of the MLK Station area. The station, UT, the state capitol complex, and the CBD are all identified. These are overlaid upon aerial photography and the arterial street network.
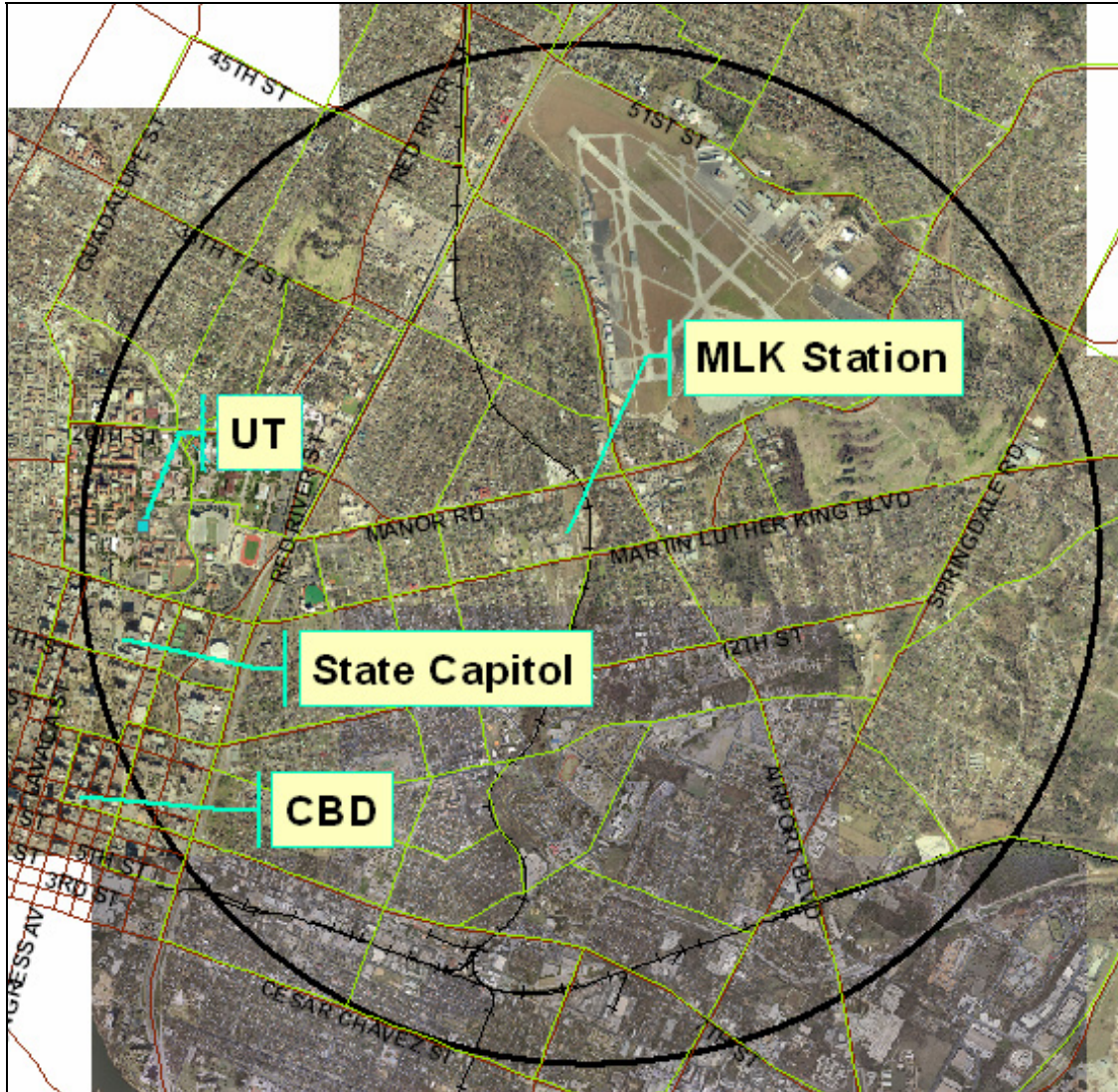
FIGURE 20: OVERVIEW OF MLK STATION COVERAGE REGION

Figure 22 depicts the TAZs that will define the demand zones used to determine demand distribution for the case study application. The black boundary centered on the station represents a 2-mile boundary about the station used to limit the number of demand zones that are considered in the analysis and represents those passengers that are theoretically considered to have access to commuter rail at the destination end of the commuter rail trip.
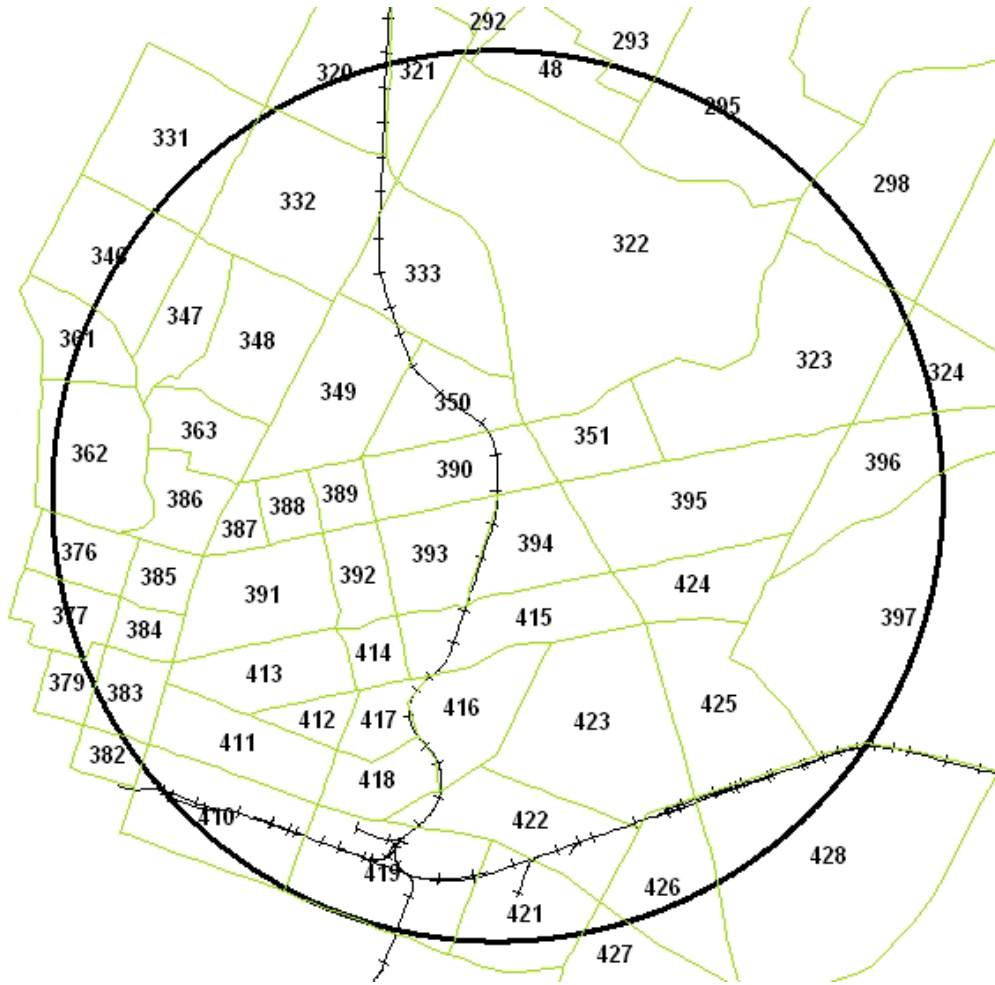
FIGURE 21: TAZ (DEMAND ZONES) LOCATION FOR MLK STATION COVERAGE REGION

Using the TAZs of Figure 22 and the aerial photography of Figure 21 to determine activity intensity, the demand centroids were located as shown in Figure 23 as blue squares. There are 48 demand centroids within the 2-mile boundary.
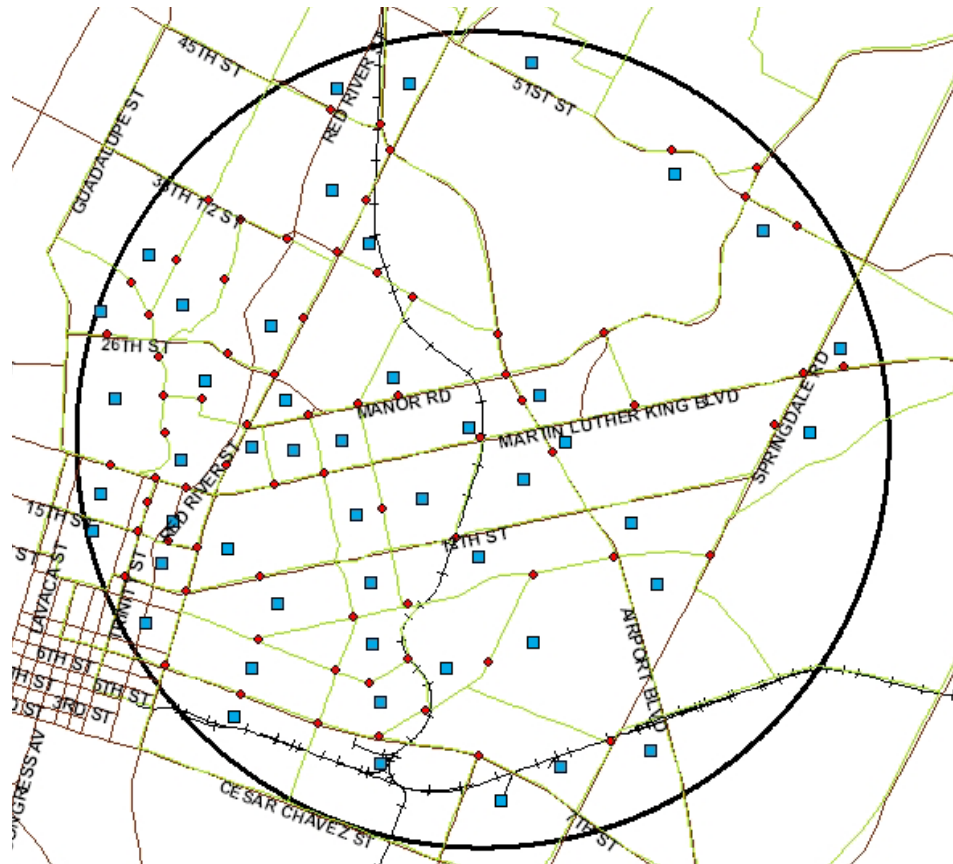
FIGURE 22: DEMAND CENTROID AND CANDIDATE STOP LOCATIONS WITHIN MLK STATION COVERAGE REGION

The demand of each of the 48 centroids is given in Table 26. Also shown in Figure 23 are the 72 candidate stop locations (red dots) from which the route to serve MLK Station will be constructed.

TABLE 26: CASE STUDY NETWORK ZONAL DEMAND

Centroid	Demand	Centroid	Demand	Centroid	Demand	Centroid	Demand
1	1	13	1	25	1	37	0
2	3	14	2	26	1	38	1
3	4	15	30	27	2	39	1
4	3	16	46	28	2	40	1
5	3	17	5	29	1	41	1
6	0	18	19	30	1	42	3
7	4	19	9	31	1	43	2
8	4	20	7	32	1	44	0
9	4	21	6	33	1	45	3
10	2	22	5	34	5	46	0
11	14	23	4	35	1	47	1
12	4	24	1	36	2	48	1

STRATEGY 1 – CULTIVATE RIDERSHIP

This strategy seeks to cultivate, or serve, as much of the potential demand as possible. This may represent a good strategy for a new rail system that is seeking to build a ridership base upon which to build. In general, the priorities of this strategy are to provide best service to as many potential riders as possible. This is implemented in the following manner:

- Doubling walking and travel cost to $50 and $25 per hour, respectively.
- Setting Unserved demand cost at the high priority level of $6,000/unserved demand unit.
- Reducing the walking threshold to ¼-mile

These modifications to the cost and walking threshold parameters place a high value on rider costs and rider convenience. By doubling the walking and travel costs, this particular strategy intends to minimize the impact of operator costs and seek to limit their impact on route design. Maintaining a high level of unserved demand cost seeks to serve the largest number of potential demand possible in a convenient manner, via the ¼-mile walking threshold. In essence, this strategy seeks to serve as much demand as possible as conveniently as possible, placing little emphasis on operator cost.
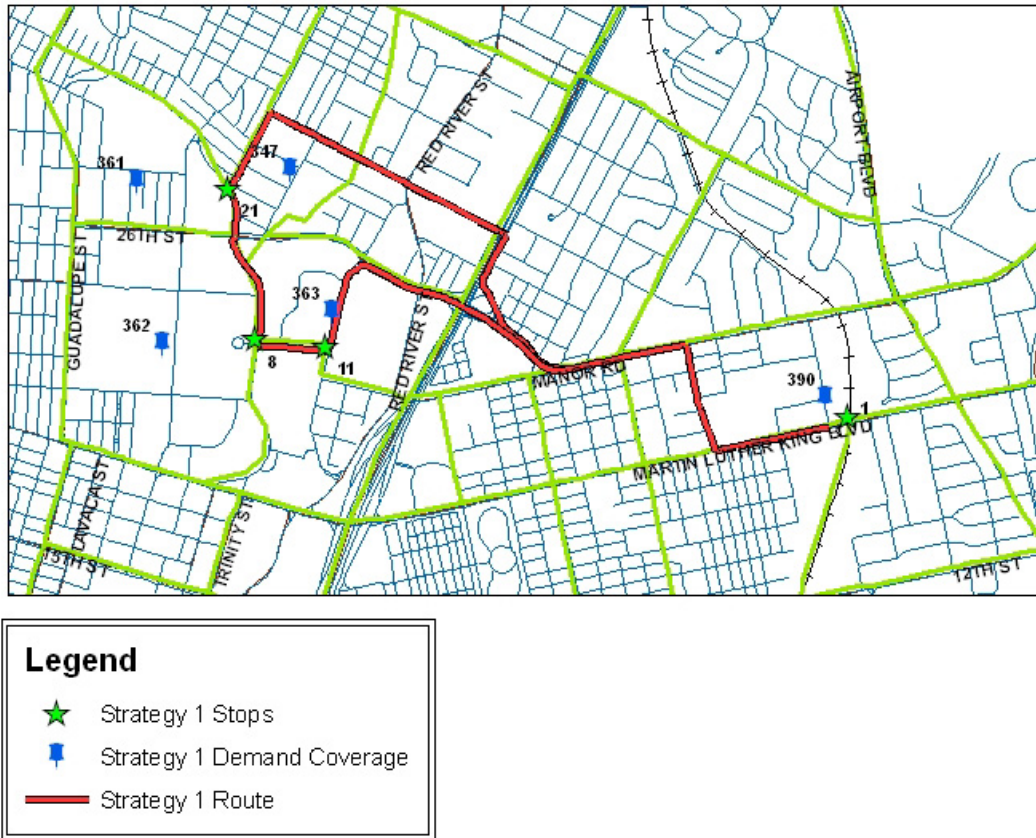
FIGURE 23: STRATEGY 1 ROUTE DEPICTION: BEST ROUTE 1 – 21 – 8 – 11 – 1

Figure 23 visually depicts the optimal route while Table 27 shows the evolution of the solution using the tabu search methodology. It is evident early on that the good starting point positively influenced the performance of the algorithm. The first solution, $r_{z*} = \{1, 8\}$ provides the foundation for all other solutions. Somewhat surprisingly, this first strategy, which intends to best serve passengers, yields a best solution with only three stops. There is no guarantee that this is an optimal solution, as the solution space for a network with 48 centroids and 72 candidate stops is enormous. This surprising result, in practice, may imply that additional runs be made for this particular strategy and that investigating other strategies is a wise investment of resources.

The optimal route, $1 – 21 – 8 – 11 – 1$, not surprisingly, provides service primarily to the University area. The particular path depicted above is an approximation, as the CRCNDP in the form used in this report is based on rectilinear distances, which do not necessarily map directly to shortest paths. Table 27 illustrates that there are several good solutions with objective function values near that of the best solution found by the tabu search method. Interestingly, the 7-stop solution is the second-best solution, suggesting that if a larger number of stops is desired on this route, the 7-stop solution is a good choice.

TABLE 27: STRATEGY 1 ROUTE EVOLUTION

Route Size (# Stops)	Route Set	Objective Function Value
2	{1, 8}	1003200
3	{1, 8, 15}	950401
4	*{1, 8, 11, 21}*	*872233*
5	{1, 8, 11, 15, 18}	924854
6	{1, 8, 11, 12, 15, 70}	930706
7	{1, 4, 8, 11, 13, 30, 70}	906698

The cost and walking threshold parameters used in this strategy appear to have a double-edged effect. They place a premium on passenger convenience; however, by assuming only a ¼-mile walking threshold, a large amount of demand remains unserved in the final solution because this strategy "refuses" to serve demand if it cannot serve it in a convenient fashion. The demand that is served is served in a convenient fashion, but the solution may be missing some of subtle benefits of employing the CRCNDP.

STRATEGY 2 – MINIMAL OPERATION COST

Strategy 2 represents the opposite end of the spectrum in route design, seeking only to provide a "bare bones" network from which to provide accessibility to commuter rail. This type of strategy may very well come into play if there are severe budgetary and resource restrictions on the transit authority. This strategy is implemented using the following devices:

- Doubling operating cost to $160 per hour.
- Setting Unserved demand cost at the low priority level of $2,000/unserved demand unit.
- Increasing the walking threshold to ½ - mile

The intent of these modifications is obvious: a focus on operating cost at the expense of convenience to the rider. Note that when the cost parameters are "doubled", they are actually being doubled relative to the other costs. So, in absolute terms, the values may not be entirely accurate. However, the relative weight of the costs is conveyed and will impact the solution accordingly. A more correct way of stating the modification to operating cost would be "making operating a bus twice as onerous" as the default situation.

The second strategy indeed yielded a route that requires minimal operation, as the route only visits a single stop on the UT campus and then returns to the station. An interesting aspect of this solution is that the station also serves several zones with walking trips because of the ½-mile walking threshold employed in this strategy. The ½-mile threshold also allows for serving a large portion of UT demand from a single stop. From a practical standpoint, a single stop along the route may not be desirable as there is likely much demand that could potentially be served along the route that is ignored in this strategy's pursuit of minimal operational cost.
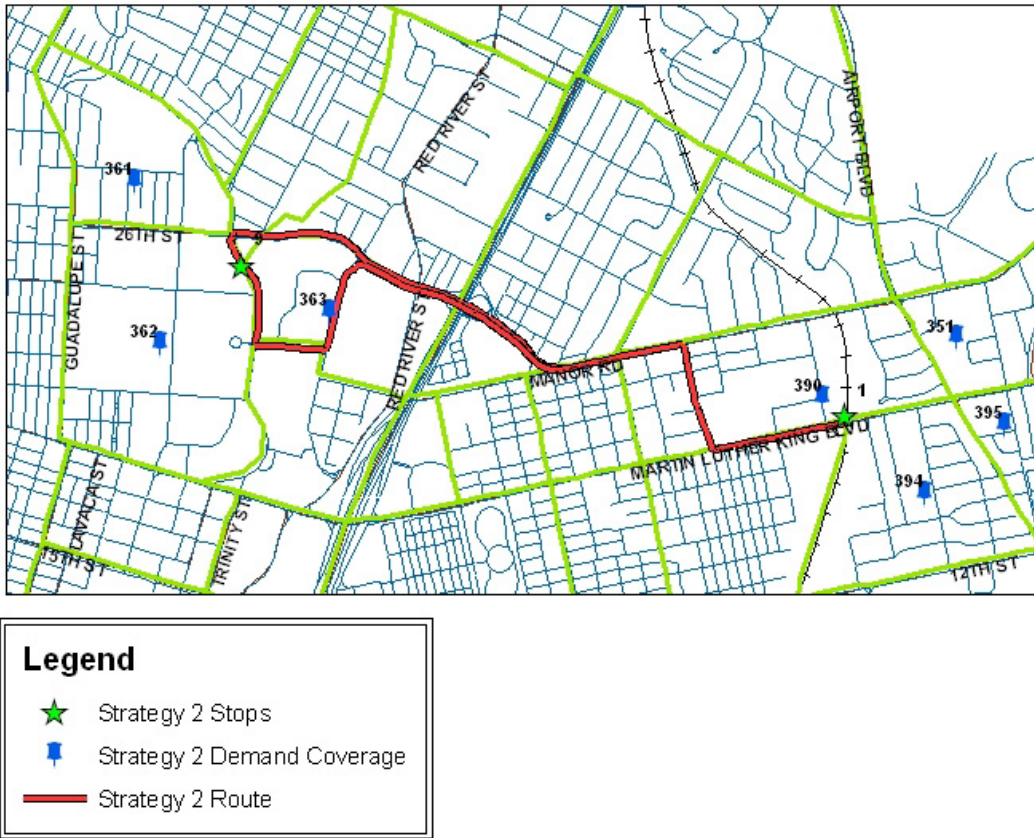
FIGURE 24: STRATEGY 2 ROUTE DEPICTION: BEST ROUTE 1 – 9 – 1

The changes undergone by the best route throughout the tabu search solution process are shown in Table 28, and display six good solutions. The objective function values are all on the same order of magnitude, suggesting that even for this strategy, which does not place much value on serving demand, there is a general tendency for the CRCNDP to seek solutions that provide good service to potential passengers.

TABLE 28: STRATEGY 2 ROUTE EVOLUTION

Route Size (# Stops)	Route Set	Objective Function Value
2	*{1, 9}*	*274432*
3	{1, 7, 18}	295344
4	{1, 9, 18, 69}	289966
5	{1, 11, 12, 31, 44}	333272
6	{1, 11, 13, 16, 17, 20}	354554
7	{1, 11, 12, 26, 27, 28, 36}	333606

The shortcomings of this strategy are evident, especially the fact that additional stops can be added and demand served within the allotted headway of 30 minutes. Compared to Strategies 1 and 3, which use routes of approximately 26 minutes, this strategy uses a route of only 23 minutes. Granted, this allows for the circulator bus to maintain seamless transfer with greater reliability and certainty, but the extra 3 minutes spent waiting at the rail station could be better spent serving passengers.

STRATEGY 3 – BALANCED APPROACH

The deficiencies of the first two strategies highlight the need for another option for applying the CRCNDP. Because the previous two strategies represent somewhat extreme situations, the third strategy attempts to find a middle ground. In fact, this strategy is what would be recommended as the default settings for the CRCNDP, should a single solution be desired. Following is a list of the balanced approach parameters:

- Operation, travel, and walking cost at their default values of $82, $13, and $25 per hour, respectively.
- Unserved demand cost at the high priority level of $6,000/unserved demand unit.
- A midpoint value for walking threshold of 3/8-mile.

The goal of this strategy is to strike a balance between the two extremes of the first two strategies and best represent the parameters being used to evaluate the tradeoffs inherent in the CRCNDP. Of course, this strategy is the strategy that is best supported by data. The operating, walking, and traveling costs are all gathered from ECONorthwest and Parsons Brinckerhoff Quade & Douglas, Inc (2002) and brought into 2007 dollars. Unserved demand cost is supported by Bailey (2007) and the walking threshold of 3/8-mile is within the suggested range for transit service and places the burden upon both operator and passenger more evenly (in regards to providing service).
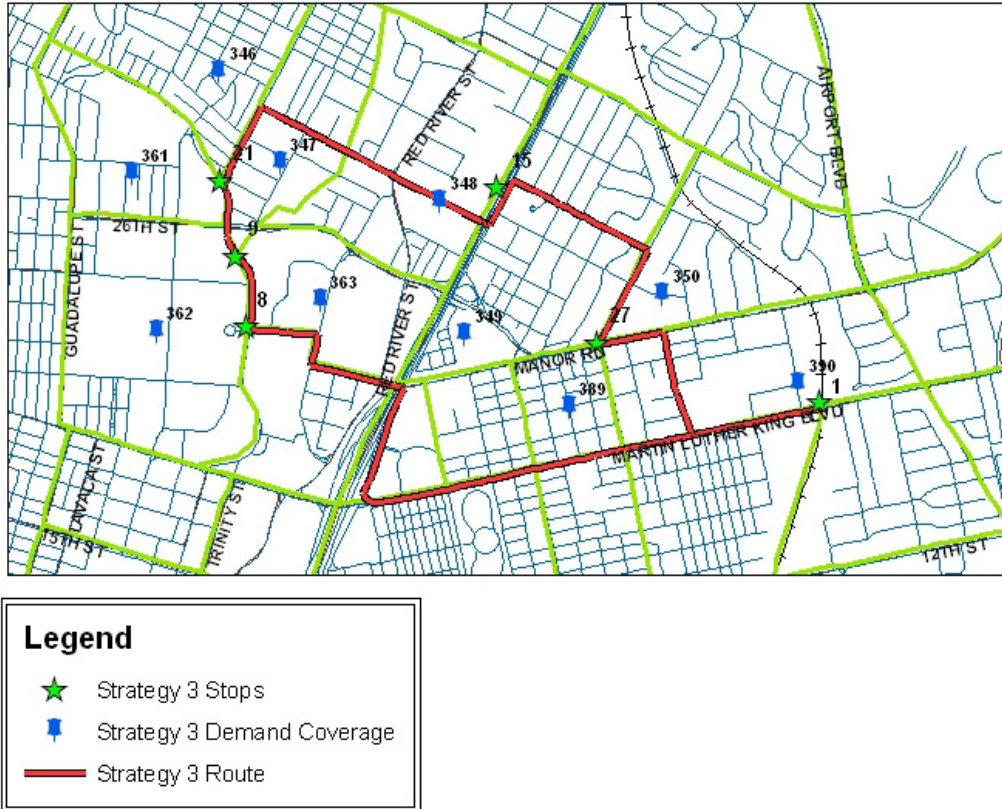
FIGURE 25: STRATEGY 3 ROUTE DEPICTION: BEST ROUTE 1 – 8 – 9 – 21 – 15 – 27 – 1

The best route given in Figure 25 is much more elaborate than the previous two strategies' solutions. There are five stops in the route, serving 10 different demand centroids. The route, 1 – 8 – 9 – 21 – 15 – 27 – 1, first serves the university and then three other small business areas on the way back to the station. The demand served at stop 27 will have to endure quite a long circulator trip to reach their destination, and in
the current CRCNDP configuration this does not reduce the likelihood that these passengers will take commuter rail. Because demand is figured as the general attractiveness of a particular TAZ based on 2-hr peak commuter trips, there is no built-in relationship between in-vehicle (circulator) time and propensity to use commuter rail.

A glance at Table 29 again shows the number of good solutions that the tabu search method found, indeed, the best solution that is depicted above very narrowly bettered the solution that utilizes only four non-station stops. This similarity in good solutions drives home the fact that the cost parameters chosen for a particular implementation of the CRCNDP can have profound impacts on the route chosen for implementation. One can easily construct a cost parameter set that would choose a different best route.

112

TABLE 29: STRATEGY 3 ROUTE EVOLUTION

Route Size (# Stops)	Route Set	Objective Function Value
2	{1, 7}	868353
3	{1, 10, 11}	807090
4	{1, 15, 18, 21}	964257
5	{1, 8, 15, 18, 21}	735846
6	*{1, 8, 9, 15, 21, 27}*	*735214*
7	{1, 8, 9, 12, 15, 18, 21}	740220

The balanced strategy did serve its purpose. It produced a route that serves the primary demand zones in a convenient fashion, while not simply attempting to maximize the cardinality of the route set, *r*. It would appear that the balanced strategy provided a good set of cost parameters by which to evaluate the tradeoffs included in the multi-objective formulation of the CRCNDP.

SUMMARY

The aim of any decision-aiding tool should not be a substitute for engineering judgment and experience. It is a supplementary tool that can be very useful in proposing new solutions, supporting engineering judgment, and providing a quantitatively supported solution to the problem at hand. The CRCNDP certainly falls within the realm of decision-aiding tools, especially when it is implemented using the tabu search solution method. The tabu search cannot guarantee optimal solutions, only best solutions from those investigated by the algorithm. In such a case, the engineer must interpret the solution carefully, and as presented in this chapter, perhaps identify other good solutions and look at the problem thoroughly.

This chapter provides a brief example of the application of the CRCNDP to a new commuter rail station in Austin, Texas. Three strategies were investigated, one that favors passengers, one that favors operators, and a balanced approach. The balanced approach yielded the best solution, serving a large number (10) of demand centroids in a route nearly the same duration as the passenger-oriented strategy (which only served 5 demand centroids). The increased coverage of the balanced approach allows the engineer to add stops along the route if desired and, over time, may induce additional demand along the coverage area of the circulator route. As commuter rail is implemented with long-term goals in mind, providing circulator coverage in line with the long term goals of shaping urban form is an added benefit of using the balanced approach.

Chapter 9 will conclude this report. The original objectives will be revisited and the accomplishments in support of these objectives highlighted. Many questions have been raised throughout this work and these will be noted along with potential avenues of research that could be pursued to answer these questions. Concluding thoughts and remarks will close this treatment of the CRCNDP.

CHAPTER 9: SUMMARY

Five objectives have been the foci of this report. These five objectives sought to define, understand, and improve the accessibility to commuter rail and express transit services through the commuter rail circulator network design problem (CRCNDP). The five stated objectives of this report are:
1. Investigate Current State of Knowledge
2. Solve the Commuter Rail Circulator Network Design Problem
3. Account for Walking Portion of Trip
4. Investigate Formulation Performance
5. Case Study Application

The first two chapters address the first objective of this report. The first chapter is primarily concerned with defining the CRCNDP and its relevance in today's transportation system. A summary of the existing and planned commuter rail systems and their degree of reliance on circulators is presented. This summary shows that while not all commuter rail systems will require a dedicated circulator, a significant portion will need at least some form of circulator and some will rely exclusively on circulators for access.

The first two chapters also seek to define the limits of this report and part of this is a precise definition of commuter rail itself. Commuter rail is a term that is often used to describe a fairly wide variety of services in practice. For the purposes of this work this definition was narrowed to include only peak-hour, peak-direction service on existing rail right-of-way. This precise definition was used to define the problem mathematically, an important step in being able to solve the CRCNDP.

The first two chapters also describe the existing CRCNDP-related work and its relevance. There are three primary branches of relevant literature: mathematical programming treatments of transit route optimization problems, analytic treatments of route optimization problems, and transit equilibrium assignment problems. The mathematical programming treatments are the most relevant, as this report would fall into that category. This report makes a contribution to this body of work by modeling the problem more directly including bus stop locations and walking trips and solving it using exact and heuristic methods.

Analytic methods represent much of the early work on the CRCNDP and related areas. Because of limitations in computational power, the problem is simplified through various assumptions to maintain a convex objective function. Second-order equations are then used to find the optimal expressions for these convex functions and the various relationships between problem parameters used to suggest values of the parameters relative to other values. The assumptions used to maintain convexity limit the direct applicability of the results, these limitations have led to greater use of mathematical programming methods in recent years to tackle the CRCNDP.

The transit equilibrium assignment problem (TEAP) provides innovative methods of addressing the CRCNDP from a mathematical perspective and creative ways of modeling the transit trip decision process. However, the TEAP is relevant primarily in capacity-constrained environments as the problem models the decision between two competing transit options. While this situation is certainly relevant in particular situations, a new commuter rail system is not likely to be one of these in which capacity is of primary concern. In most commuter rail systems the choice is between rail and personal auto, a decision process that the TEAP is not designed to address.

The introductory material to the CRCNDP includes several examples of commuter rail systems in the United States that demonstrate the transitional nature of commuter rail. When commuter rail service is initiated it is in the long-term that the primary contributions of the system are reflected. It is because of these long-term goals that commuter rail tends to transition to higher level-of-service forms of rail transportation. Because commuter rail eventually seeks to reduce congestion and promote smart land use practices it is necessary that the service eventually provide more than peak hour and peak direction service. In the short term, however, it is essential that the commuter rail cultivate a ridership base upon which to later expand service. Cultivating this ridership will require optimal service and optimal access to the commuter rail system. Hence, the CRCNDP and the pressing need for circulators systems in regions in which the final destinations of commuters do not lie within walking distance of rail stations being built on existing right-of-way designed to serve 19th century industrial needs.

Chapter 3 sets about formalizing the CRNCDP mathematically and settles upon a multi-objective, mixed-integer program. This formulation allows for the precise modeling of the circulator trip considering user and operator costs, the exact location of bus stops, the inclusion of walking distance, and the incorporation of the "seamless transfer" concept. Seamless transfer requires that a circulator vehicle be available at the rail station every time a commuter train arrives. This allows for minimal passenger wait and transfer time and requires that circulator routes strictly adhere to routes maintaining this transfer policy. It is argued that without this policy in place, the commuter rail system as a whole becomes much less attractive and cultivating ridership becomes that much more difficult. This limit on route size provides computational benefits as well, contributing to the development of efficient exact solution methods for the CRCNDP.

The formulation of the CRNCDP seeks specifically to avoid any nonlinearities and the computational complexity inherent in them. Avoiding these nonlinearities requires that preprocessing be performed in any solution of the CRCNDP to calculate the appropriate parameters. These preprocessed parameters include walking coverage, passengers served at a stop, and unrelated to the nonlinearities, dwell time. It is through this mathematical program that an optimal solution to the CRCNDP can be found. It is important to note that the optimal solution to any one implementation of the CRCNDP is optimal relative to the cost parameters used in its formulation. Different route structures will result from different objective function cost sets. This pareto-optimality is demonstrated well in chapter 8 with the case study application of the CRNCDP.

Chapter 4 presents the first efforts directed at solving the CRCNDP rather than describing it. In any optimization problem, complete enumeration represents the most naïve solution method and also the most complete description of the solution space. The algorithm presented in this chapter is designed simply to evaluate every possible solution to the CRCNDP and pick the minimal solution. The CRCNDP, as formulated, has a two-level structure that first seeks to select the optimal set of bus stops in a route and then design the best route to serve these stops. The algorithm in Chapter 4 divides the problem along this natural split and evaluates every possible combination of bus stops and optimizes the route to serve each of these combinations. It is easy to see the limitations of such a method as the number of combinations grows exponentially with the addition of every candidate stop to a network. The chapter concludes with an example application of the enumerative method that displays that the CRCNDP produces good solutions and appears to solve the desired problem, highlights the limitations due to the exponential explosion of stop combinations, and provides insight into improvements that could be made to an exact solution method. These improvements are dealt with in greater detail in Chapter 5.

The fifth chapter takes a more intelligent approach to exact solution methods. Two techniques are used to improve the solution time of the CRCNDP: preprocessing using a polynomial algorithm and implementing an intelligent stopping criterion. The first technique uses the fact that the CRCNDP can be reduced to a traveling salesman problem (TSP) for any set of stops. The TSP is an NP-hard problem, however, a 1-Tree can be found for any set of nodes in polynomial time and serves as a lower bound to any TSP problem. Therefore, if a set of bus stops produces a 1-Tree solution that is greater than the route length allowed by seamless transfer, then the TSP solution is guaranteed to be greater as well. Using this preprocessing methodology is shown to be very effective in reducing the number of candidate stop sets over which the CRCNDP is solved. It is particularly effective in combination with the stopping criterion that is employed. It is shown that for any set of bus stops of size C, if no improvement is made to the optimal solution, then one need not investigate sets of bus stops of size $C+1$ as there is no possibility of improving the solution with this larger set of stops. These techniques allow for vast improvements in solution time, however, even they eventually reach their limits.

Chapter 6 describes a solution technique that is useful when the limits of exact methods have been reached. A tabu search method is employed, which is a generic metaheuristic algorithm that can be tailored to a wide variety of applications. In the case of the CRCNDP, it is a rather straightforward implementation of the technique. Chapter 6 describes an implementation that seeks primarily to reduce the possibility of being stuck in a locally optimal solution that is far from globally optimal by promoting a wide search for good solutions and then investigating these good solutions further. If the locally good solution fails to produce further improvements, the algorithm moves on. The success of this algorithm depends largely upon finding a good initial solution and this breadth-focused searching. Comparing the tabu search implementation to the exact methods produced exciting results. The tabu search found optimal solutions in many cases and found them in a fraction of the time of the exact methods. This lends confidence in the application of the tabu search method to larger networks in which it is impossible to apply even the intelligent exact solution procedure.

Chapter 7 sought to test the performance of the CRCNDP by examining the effects of walking threshold and unserved demand cost on route design. This examination continued in Chapter 8 with the case study, though in a less rigorous fashion. The sensitivity analysis of chapter 7 suggested that 3/8-mile is a good default value of walking threshold and that the walking threshold sensitivity may be relative to the average distance between stop and centroid. The default value of $6,000 per unserved unit of demand was found to be a good value, which corroborates the findings of Bailey (2007).

In Chapter 8 the techniques that were developed over the course of the report are applied to a case study network in Austin, Texas. The network under consideration is that which surrounds the future MLK station as part of the Austin MetroRail commuter rail line. The case study network consisted of 48 demand centroids (TAZs) and 72 candidate stops to serve these centroids. These numbers are far outside the practical range for the application of the exact techniques developed in this report. Therefore, the tabu search method was applied to the MLK station in three separate scenarios. The three scenarios attempted to capture three different priority schemes that could be under consideration by a commuter rail operation authority. The three strategies were 1) User-focused, 2) Operator-focused, and 3) Balanced. Happily, all strategies performed as expected and not surprisingly, the balanced approach yielded the best results. While the balanced approach provides good results and will likely perform well in many applications, it is recommended that any analysis using the techniques from this work investigate several different strategies and cost combinations just to illuminate the best and most robust strategy for a particular application.

CONTRIBUTIONS

The primary contributions of this report fulfill those that were originally proposed in the early pages of this research, summarized below:

- Develop a formulation of the CRCNDP that accounts for unserved demand, walking trip variables, and seamless transfer.
- Provide an efficient and systematic solution procedure for the CRCNDP.
- A better understanding of the relationship between optimal commuter rail circulator design and commuter rail accessibility.
- A demonstration of the utility of the CRCNDP solution method through application to the Austin, Texas proposed commuter rail system.

These contributions are self-explanatory and do not necessarily account for all of the contributions of this report. In addition to these points, this work provides a synthesis of CRCNDP-related literature throughout the past 50 years. It provides not a single solution method, but three, each building upon its predecessor. The metaheuristic method finally applied to the case study can be compared to optimal solutions and exact solution techniques because of these efforts.

There is a practical aspect to this work as well. Commuter rail is defined precisely and this definition applied rigorously to the CRCNDP. The current commuter rail systems' reliance on circulator systems is displayed and finally, the CRCNDP is applied to a system that will be operational within the next 18 months. This research is intended for application, to improve the design and management of our public transportation system.

FUTURE RESEARCH DIRECTIONS

There is no silver bullet for making a public transportation system work well. A public transportation system can flourish or perish with or without a robust circulator system. However, a public transportation system will certainly fail if it is unable to attract passengers. The CRCNDP is one tool to enable transit authorities to be able to better cultivate and retain public transportation passengers.

An obvious extension of this work is the accommodation of multiple routes directly in the formulation and solution methods. The current methods allow for multiple routes only by choosing, for example, the three "best" unique routes that the solution procedure produces. This method has several flaws, not the least of which is the lack of accounting for the impacts on unserved demand that a second or third unique route would present. Configuring the algorithms in this research to account for multiple routes will require significant effort, as the inclusion of multiple routes will certainly increase the complexity, performance and computational resource requirements of any solution method.

The CRCNDP is seeking to better model the accessibility of public transportation, and this improved accessibility is intended to have a positive impact on commuter rail ridership. This report assumes a fixed commuter O-D trip table, though acknowledges that there will naturally be some reciprocal relationship between transit accessibility and ridership. Quantifying this relationship is a topic of research that is being intensely pursued and will be for some time. There is difficulty in accounting for all of the variables that impact any traveler's decision to ride or not to ride public transportation. There is even greater difficulty in gathering data to calibrate and validate any model intending to estimate the relationship. An ideal CRCNDP formulation

would include a model that accounts for the relationship between ridership and accessibility and is certainly a worthy extension of this work.

Finally, a future direction for CRCNDP work would be the development of an interface that would allow the CRCNDP to be solved without the aid of an expert researcher. Such a tool would have great utility as a GIS tool for which an analyst could select a particular network surrounding a station, identify candidate stop locations and TAZ centroids, and the CRCNDP tool would provide an optimal (or very good) route design.

CONCLUSION

The CRCNDP is a relevant problem for today's public transportation authorities and analysts. American cities have not developed in the dense, transit-friendly manner of many other cities throughout the world where transit flourishes. Because of American prosperity and reliance upon the automobile, a development pattern has been propagated over the past 50 years that is not ideally suited for public transportation.

A perfect storm of socio-economic and political factors over the past few years has prompted a renaissance of sorts for public transportation in this country. Chief among these is the dramatic rise in gasoline prices over the past 5 years. The threefold rise in prices has had wide-ranging effects, not the least of which are a larger number of public transportation passengers, an increased demand for public transportation accessibility, and increased political feasibility for a variety of transit endeavors. This storm has led to previously auto-dominated regions looking to introduce public transportation service to a larger portion of its population. As rail is viewed as the most desirable of public transportation options, it is rail that is looked to as an expedient means of increasing the transit capacity of a region. Commuter rail service in these regions is often chosen because service during peak hours will provide the largest number of potential passengers in a cost-effective manner. Commuter rail is implemented on existing rail right-of-way and infrastructure; improving the cost effectiveness yet hurting the overall accessibility of the system. It is in improving access to commuter rail that the needs for circulators and the CRCNDP are most evident. Conditions are ripe for improving public transportation in America and the investment in public transportation made today needs to be an intelligent one; developing an accessible, sustainable system that will perform well and achieve long-term goals even if the current transit-acquiescent climate sours. It is the duty of a transportation engineer to help direct the investments in our transportation system and infrastructure and it is the goal of this report to help provide this direction.

APPENDIX A: 20 – CENTROID NETWORK DATA

TABLE A-1: 20 CENTROID NETWORK DEMAND

Centroid	Demand Units
1	0
2	7
3	6
4	6
5	21
6	7
7	44
8	68
9	7
10	27
11	8
12	11
13	5
14	7
15	34
16	54
17	14
18	2
19	9
20	11

Table A-2: 20 Centroid Network Inter-Centroid Rectilinear Distance (λ_{ij})

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	10974	10467	8676	9201	10015	9997	9020	10142	11951	7970	6179	6776	8477	7464	8260	9834	4931	8828	6722
2	10974	0	1013	2298	1773	959	2352	4053	4885	5681	4922	5410	7346	6695	9481	11978	11091	7201	14048	12488
3	10467	1013	0	1791	1882	1502	3365	3764	4885	6695	4415	4903	6839	6188	8975	11471	10585	6695	13541	11981
4	8676	2298	1791	0	959	1339	2443	2841	3963	5772	2624	3112	5048	4397	7183	9680	8794	4903	11750	10190
5	9201	1773	1882	959	0	814	1484	2280	3112	4813	3148	3637	5573	4922	7708	10205	9318	5428	12275	10715
6	10015	959	1502	1339	814	0	1864	3094	3926	5193	3963	4451	6387	5736	8522	11019	10133	6242	13089	11529
7	9997	2352	3365	2443	1484	1864	0	3076	3908	3329	3944	4433	6369	5718	8504	11001	10114	6224	13071	11511
8	9020	4053	3764	2841	2280	3094	3076	0	1122	2931	1049	2841	3293	2642	5428	7925	7038	4089	9995	8435
9	10142	4885	4885	3963	3112	3926	3908	1122	0	1809	2171	3963	3365	1809	4596	7093	6206	5211	9163	7602
10	11951	5681	6695	5772	4813	5193	3329	2931	1809	0	3981	5772	5175	3474	5320	7817	6930	7020	9886	8326
11	7970	4922	4415	2624	3148	3963	3944	1049	2171	3981	0	1791	2425	1773	4560	7057	6170	3040	9127	7566
12	6179	5410	4903	3112	3637	4451	4433	2841	3963	5772	1791	0	1936	2298	4071	6568	5681	1791	8638	7078
13	6776	7346	6839	5048	5573	6387	6369	3293	3365	5175	2425	1936	0	1701	2135	4632	3745	1846	6702	5142
14	8477	6695	6188	4397	4922	5736	5718	2642	1809	3474	1773	2298	1701	0	2786	5283	4397	3546	7353	5793
15	7464	9481	8975	7183	7708	8522	8504	5428	4596	5320	4560	4071	2135	2786	0	2497	2370	2533	4567	3007
16	8260	11978	11471	9680	10205	11019	11001	7925	7093	7817	7057	6568	4632	5283	2497	0	1574	4777	2070	1538
17	9834	11091	10585	8794	9318	10133	10114	7038	6206	6930	6170	5681	3745	4397	2370	1574	0	4903	2957	3112
18	4931	7201	6695	4903	5428	6242	6224	4089	5211	7020	3040	1791	1846	3546	2533	4777	4903	0	6847	5286
19	8828	14048	13541	11750	12275	13089	13071	9995	9163	9886	9127	8638	6702	7353	4567	2070	2957	6847	0	2106
20	6722	12488	11981	10190	10715	11529	11511	8435	7602	8326	7566	7078	5142	5793	3007	1538	3112	5286	2106	0

APPENDIX B: 20 – 10 NETWORK DATA

TABLE B-1: 20 - 10 NETWORK DEMAND

Centroid (TAZ)	Demand Units
1 (390)	1
2 (332)	2
3 (333)	2
4 (348)	5
5 (349)	2
6 (361)	11
7 (362)	17
8 (363)	2
9 (376)	7
10 (385)	2

Table B-2: 20 - 10 Stop-to-Centroid Rectilinear Distance (γ_{ig})

	To Centroid									
	390	332	333	348	349	361	362	363	376	385
	1	2	3	4	5	6	7	8	9	10
1	**544**	9893	7663	8049	5862	12724	10194	8349	10977	9862
2	10934	13421	12992	8489	6988	6455	3924	5444	**1865**	**1112**
3	10427	12914	12486	7982	6481	7468	4937	4937	2879	**605**
4	8636	11123	10694	6191	4690	6545	4015	3146	2301	**1186**
5	9160	11648	11219	6716	5215	5586	3056	3671	**1776**	**1491**
6	9975	12462	12033	7530	6029	5966	3436	4485	**1377**	**1111**
7	9956	12444	12015	7512	6011	4102	**1765**	4467	**980**	2974
8	8476	9368	8939	4436	3158	3704	**1311**	**1391**	4056	3372
9	9598	8536	8107	3604	4280	2583	2144	**1792**	4888	4494
10	11407	9260	8831	4327	6089	**773**	**1757**	3602	4165	6304
11	7427	8500	8071	3567	2108	4754	2223	**522**	4925	3810
12	5635	8011	7582	3079	**1578**	6545	4015	2170	5413	4298
13	6232	6075	5646	**1273**	**914**	5948	4604	**1902**	7349	6234
14	7933	6726	6297	**1794**	2615	4247	3953	**1251**	6698	5583
15	6920	3940	3511	**992**	2493	5260	6739	4037	9484	8369
16	7716	**1634**	**1014**	3489	4990	7392	9236	6534	11981	10866
17	9290	2330	2171	2603	4104	6506	8350	5648	11095	9980
18	4387	6220	5791	3119	**931**	7794	5263	3419	7204	6089
19	8284	**1065**	**1165**	5559	7060	9462	11306	8604	14051	12936
20	6179	3171	**941**	3999	5500	7902	9746	7044	12491	11376

Table B-3: 20 - 10 Inter-Centroid Rectilinear Distance (λ_{ij})

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	10974	10467	8676	9201	10015	9997	9020	10142	11951	7970	6179	6776	8477	7464	8260	9834	4931	8828	6722
2	10974	0	1013	2298	1773	959	2352	4053	4885	5681	4922	5410	7346	6695	9481	11978	11091	7201	14048	12488
3	10467	1013	0	1791	1882	1502	3365	3764	4885	6695	4415	4903	6839	6188	8975	11471	10585	6695	13541	11981
4	8676	2298	1791	0	959	1339	2443	2841	3963	5772	2624	3112	5048	4397	7183	9680	8794	4903	11750	10190
5	9201	1773	1882	959	0	814	1484	2280	3112	4813	3148	3637	5573	4922	7708	10205	9318	5428	12275	10715
6	10015	959	1502	1339	814	0	1864	3094	3926	5193	3963	4451	6387	5736	8522	11019	10133	6242	13089	11529
7	9997	2352	3365	2443	1484	1864	0	3076	3908	3329	3944	4433	6369	5718	8504	11001	10114	6224	13071	11511
8	9020	4053	3764	2841	2280	3094	3076	0	1122	2931	1049	2841	3293	2642	5428	7925	7038	4089	9995	8435
9	10142	4885	4885	3963	3112	3926	3908	1122	0	1809	2171	3963	3365	1809	4596	7093	6206	5211	9163	7602
10	11951	5681	6695	5772	4813	5193	3329	2931	1809	0	3981	5772	5175	3474	5320	7817	6930	7020	9886	8326
11	7970	4922	4415	2624	3148	3963	3944	1049	2171	3981	0	1791	2425	1773	4560	7057	6170	3040	9127	7566
12	6179	5410	4903	3112	3637	4451	4433	2841	3963	5772	1791	0	1936	2298	4071	6568	5681	1791	8638	7078
13	6776	7346	6839	5048	5573	6387	6369	3293	3365	5175	2425	1936	0	1701	2135	4632	3745	1846	6702	5142
14	8477	6695	6188	4397	4922	5736	5718	2642	1809	3474	1773	2298	1701	0	2786	5283	4397	3546	7353	5793
15	7464	9481	8975	7183	7708	8522	8504	5428	4596	5320	4560	4071	2135	2786	0	2497	2370	2533	4567	3007
16	8260	11978	11471	9680	10205	11019	11001	7925	7093	7817	7057	6568	4632	5283	2497	0	1574	4777	2070	1538
17	9834	11091	10585	8794	9318	10133	10114	7038	6206	6930	6170	5681	3745	4397	2370	1574	0	4903	2957	3112
18	4931	7201	6695	4903	5428	6242	6224	4089	5211	7020	3040	1791	1846	3546	2533	4777	4903	0	6847	5286
19	8828	14048	13541	11750	12275	13089	13071	9995	9163	9886	9127	8638	6702	7353	4567	2070	2957	6847	0	2106
20	6722	12488	11981	10190	10715	11529	11511	8435	7602	8326	7566	7078	5142	5793	3007	1538	3112	5286	2106	0

REFERENCES

Acken, L. Dallas Area Rapid Transit: Miles ahead with miles to go. Mass Transit, September/October (2005).

Ahuja, R., Magnanti, T. and J. Orlin. *Network Flows: Theory, Algorithms, and Applications.* Pretice-Hall, Inc., New Jersey, 1993.

American Public Transportation Association (APTA). U.S. Commuter Rail Transportation System Links, available online: http://www.apta.com/links/transit_by_mode/commrail.cfm, Accessed July 18, 2006.

Bailey, L. Public Transportation and Petroleum Savings in the U.S.: Reducing Dependence on Oil. ICF International, American Public Transportation Association, Washington, D.C., 35 pp, (2007).

Bard, J. *Practical Bilevel Optimization: Algorithms and Applications.* Kluwer Academic Publishers, Dordrecht, 1998.

Barile, Margherita. "Taxicab Metric." From MathWorld--A Wolfram Web Resource, created by Eric W Weisstein. http://mathworld.wolfram.com/TaxicabMetric.html, Accessed January 9, 2007.

Bhat, C., Sivakumar, A., Sen, S., Guo, J., and Copperman, R. Austin Commuter Survey: Findings and Recommendations. Southwest Region University Transportation Center Report 167240-1, September 2005.

Calthorpe, P. *The Next American Metropolis: Ecology, Community, and the American Dream,* New York: Princeton Architectural Press, 1993.

Canepa, B. Bursting the Bubble: Determining Transit-Oriented Development's Walkable Limits. Presented at the 86th Annual meeting of the Transportation Research Board, Washington, D.C., January 2007.

Capital Metro. *All Systems Go!,* Available online: http://allsystemsgo.capmetro.org/, Accessed July 2006.

Ceder, A. and Israeli, Y. User and Operator Perspectives in Transit Network Design. *Transportation Research Record: Journal of the Transportation Research Board, No.1623,* Transportation Research Board of the National Academies, Washington, D.C, pp. 3 – 7 (1998).

Ceder, A. and Wilson, N.H.M. Bus Network Design. *Transportation Research Part B*, Vol. 20B, No. 4, pp. 331-344 (1986).

Chang, S. K. and P. Schonfeld. Multiple Period Optimization of Bus Transit Systems. *Transportation Research Part B*, Vol. 25B, No. 6, pp. 453-478 (1991a).

Chang, S. K. and P. Schonfeld. Optimization models for Comparing Conventional and Subscription Bus Feeder Services. *Transportation Science*, Vol. 25, No. 4, pp. 281 – 298 (1991b).

Chien, S. and P. Schonfeld. Joint Optimization of a Rail Transit Line and its Feeder Bus System. *Journal of Advanced Transportation*, Vol. 32, No. 3, pp. 253 – 284 (1998).

Chien, S., Spasovic, L., Elefsiniotis, S., and R. Chhonkar. Evaluation of Feeder Bus Systems with Probabilistic Time-Varying Demands and Nonadditive Time Costs. *Transportation Research Record: Journal of the Transportation Research Board, No. 1760,* Transportation Research Board of the National Academies, Washington, D.C, pp. 47 – 55 (2001).

Cieslik, D. *The Steiner Ratio.* Kluwer Academic Publishers, Norwell, MA, 2001.

Constantin, I. and Florian, M. Optimizing Frequencies in a Transit Network: a Nonlinear Bi-level Programming Approach. *International Transactions in Operational Research*, Vol. 2, No. 2, pp. 149-164, (1995).

Cordeau, J-F. A Branch-and-cut Algorithm for the Dial-a-ride Problem. *Operations Research.* Vol. 54, No. 3, pp. 573 – 586 (2006).

Dial, R.B. Bicriterion Traffic Assignment: Basic Theory and Elementary Algorithms. *Transportation Science*, Vol. 16, No. 2, (1996).

Dubois, D; Bel, G. and Llibre, M. A Set of Methods in Transportation Synthesis and Analysis. *The Journal of the Operational Research Society,* Vol. 30, No. 9, pp. 797-808 (1979).

Dunphy, R., Cervero, R., Dock, F., McAvey, M., Porter, D. and Swenson, C. *Developing Around Transit: Strategies and Solutions that Work.* ULI-Urban Land Institute, Washington, D.C., 2004.

ECONorthwest and Parsons Brinckerhoff Quade & Douglas, Inc. Estimating the Benefits and Costs of Public Transit Projects: A Guidebook for Practitioners. Washington, D.C.: Transportation Research Board, National Research Council. TCRP Report 78, 2002.

Edwards, J.D., Jr. (ed.): *Transportation Planning Handbook*, Institute of Transportation Engineers/Prentice-Hall, 1992, 525 pp.

Ewing, R. *Pedestrian- and Transit-Friendly Design: A Primer for Smart Growth* Washington, DC, Smart Growth Network, 1999.

Fan, W. Optimal Transit Route Network Design Problem: Algorithms, Implementations, and Numerical Results. Doctoral Report, Department of Civil Engineering, The University of Texas at Austin, Austin, Texas, May 2004.

Fan, W. and R. Machemehl. A Tabu Search Based Heuristic Method for the Transit Route Network Design Problem. Presented at the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), August 9, 2004.

Florian, M and H. Spiess. The Convergence of Diagonalization Algorithms for Asymmetric Network Equilibrium Problems. *Transportation Research Part B*, Vol. 16B, No. 6, pp. 477-483 (1982).

Gabriel, S. and D. Bernstein. Nonadditive Shortest Paths: Subproblems in Multi-Agent Competitive Network Models. *Computational and Mathematical Organization Theory*. Vol. 6, 29 – 45 (2000).

Glover, F. and M. Laguna. *Tabu Search*. Kluwer Academic Publishers, Boston, July 1997.

Glover, F. Tabu Search - Part I. *ORSA Journal on Computing*, Vol. 1, No. 3, 190-206, (1989).

Grava, S. *Urban Transportation Systems: Choices for Communities*, McGraw-Hill, 840 pp, 2003.

Gray, George E. and Lester A. Hoel (eds.): *Public Transportation* (2nd ed.), Prentice-Hall, 750 pp, 1992.

Hadas, Y. and A. Ceder. Multi-agent Approach for Public Transit System Based on Flexible Routes. Presented at the 86th Annual meeting of the Transportation Research Board, Washington, D.C., January 2007.

Hamdouch, Y., Marcotte, P. and Nguyen, S. Capacitated transit assignment with loading priorities. *Mathematical Programming B*, 101, pp. 205-230 (2004).

Hyari, K. and K. El-Rayes. Optimal Planning and Scheduling for Repetitive Construction Projects. *Journal of Management in Engineering*, Vol. 22 No. 1, pp. 11 – 19 (2006).

Jerby, S. and Cedar, A. Optimal Routing Design for Shuttle Bus Service. *Transportation Research Record: Journal of the Transportation Research Board, No. 1971*, TRB, National Research Council, Washington, D.C., pp. 14 – 22 (2006).

Kalvelagen, E, *Model Building with GAMS*. forthcoming

Kandil, A. and K. El-Rayes. MACROS: Multiobjective Automated Construction Resource Optimization System. *Journal of Management in Engineering*, Vol. 22 No. 3, pp. 126 – 134 (2006).

Kocur, G. and Hendrickson, C. Design of Local Bus Service with Demand Equilibration, *Transportation Science*, Vol. 16, No. 2, (1982).

Kuah, G. K. and Perl, J. A Methodology for Feeder-Bus Network Design. *Transportation Research Record: Journal of the Transportation Research Board, No. 1120*, TRB, National Research Council, Washington, D.C., pp. 40 – 51 (1987).

Kuah, G. K. and Perl, J. Optimization of Feeder Bus Routes and Bus-stop Spacing. *Journal of Transportation Engineering*, Vol. 114, No. 3 (1988).

Kuby, M., Barranda, A. and Upchurch, C. Factors influencing light-rail station boardings in the United States. *Transportation Research Part A*, Vol. 38, pp. 223 – 247 (2004).

Lee, Y-J and V. Vuchic. Transit Network Design with Variable Demand. *Journal of Transportation Engineering*, ASCE, Vol. 131, No. 1, (2005).

Levinson, H.S. Analyzing Transit Travel Time Performance. *Transportation Research Record: Journal of the Transportation Research Board, No. 915,* Transportation Research Board of the National Academies, Washington, D.C, pp. 1-6 (1983).

Litman, Todd. Impacts of Rail Transit on the Performance of a Transportation System. *Transportation Research Record: Journal of the Transportation Research Board, No.1930,* Transportation Research Board of the National Academies, Washington, D.C, pp. 23-29 (2005).

Loukopoulos, P., and Gärling, T. Are Car Users Too Lazy to Walk?. *Transportation Research Record: Journal of the Transportation Research Board, No.1926,* Transportation Research Board of the National Academies, Washington, D.C, pp. 206 – 211 (2005).

Miller, C. E., Tucker, A. W., and R. A. Zemlin. Integer Programming Formulation of Traveling Salesman Problems, *Journal of the Association for Computing Machinery (JACM)*, Vol. 7, No. 4 , pp. 326-329 (1960).

Moellering, H., Gauthier, H. and J. P. Osleeb. An interactive Graphic Transit Planning System Based on Individuals. *Urban Systems*, Vol. 2, pp. 93-103 (1977).

Murray, A. T. A Coverage Model for Improving Public Transit System Accessibility and Expanding Access. *Annals of Operations Research*, Vol. 123, pp. 143-156 (2003).

Nelson, D., Duse-Anthony, Y., and Friemann, S. Improving Inner-City Mobility with Commuter Rail Service: Case study of Fairmount Line in Boston, Massachusetts. *Transportation Research Record: Journal of the Transportation Research Board, No.1930,* Transportation Research Board of the National Academies, Washington, D.C, pp. 44-53 (2005).

Newell, G. F. Some Issues Relating to the Optimal Design of Bus Routes. *Transportation Science*, Vol. 13, No. 1, (1979).

Pattnaik, S.B., Mohan, S., and Tom, V.M. Urban Bus Transit Route Network Design Using Genetic Algorithm. *Journal of Transportation Engineering,* Vol. 124, No. 4, (1998).

Rails-to-Trails. Rail-Trail Statistics. Available online: http://www.railtrails.org/whatwedo/railtrailinfo/trailstats.html, Accessed February 1, 2007.

Regional Transportation Authority (RTA). Music City Star Notes. Volume 1, Summer 2004. Available online: www.musiccitystar.org/starnotes/ Starnotes_V1_S04.pdf, Accessed April 3, 2007.

Schrank, D. and Lomax, T. The 2005 Urban Mobility Report. Texas Transportation Institute, Texas A&M University, College Station, Texas, 2005.

Spasovic, L. N. and Schonfeld, P. M. Method for Optimizing Transit Service Coverage, *Transportation Research Record: Journal of the Transportation Research Board, No. 1402*, TRB, National Research Council, Washington, D.C., pp. 28 – 39 (1993).

Spiess H. and Florian, M. Optimal Strategies: A New Assignment Model for Transit Networks. *Transportation Research Part B*, Vol. 23B, No. 2, pp. 83-102 (1989).

Stein, D. M. An Asymptotic, Probabilistic Analysis of a Routing Problem. *Mathematics of Operations Research*. Vol. 3, No. 2, pp. 89 – 101 (1978).

Transportation Research Board (TRB). *Highway Capacity Manual 2000.* TRB, National Research Council, Washington, D.C., 2000.

U.S. Census Bureau. U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin.: Projected Population of the United States, by Age and Sex: 2000 to 2050. <http://www.census.gov/ipc/www/usinterimproj/> Internet Release Date: March 18, 2004.

U.S. Census Bureau. U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin.: Population in Metropolitan and Micropolitan Statistical Areas in Alphabetical Order and Numerical and Percent Change for the United States and Puerto Rico: 1990 and 2000. http://www.census.gov/ipc/www/usinterimproj/ Internet Release Date: December 30, 2003.

Vuchic, V. R. *Urban Public Transportation: Systems and Technology*, Prentice-Hall, 673 pp, 1981.

Vuchic, V.R. *Urban Transit: Operations, Planning and Economics*, John Wiley & Sons, 644 pp, 2005.

Wardrop, J.G. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers. Part II; Engineering Divisions; Airport, maritime, railway, road.* Institution of Civil Engineers, Vol. 1, pp. 325-378 (1952).

Wolsey, L. *Integer Programming*. John Wiley & Sons, Inc., New York, 1998.

Wu, J., Florian, M. and P. Marcotte. Transit Equilibrium Assignment: A Model and Solution Alogirthms. *Transportation Science*, Vol. 28, No. 3, (1994).

Zitzler, Eckhart and L. Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE Transactions on Evolutionary Computing,* Vol. 3, No. 4, pp. 257 – 271 (1999).