Service Based Logistics Optimization
by
Gregory D. Price Jr.
B.S. Electrical Engineering, University of Texas, 2005
Submitted to the MIT Sloan School of Management and the Department of Systems Engineering in partial fulfillment of the requirements for the degrees of
Master of Business Administration
and
Master of Science in Engineering Systems
in conjunction with the Leaders for Global Operations Program at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2014
© Gregory D. Price Jr., MMXIV. All rights reserved.
The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature redacted

Author
MIT Sloan School of Management and the Department of Systems Engineering
May 9, 2014

Certified by
Deborah Nightingale, Thesis Supervisor
Professor of the Practice of Aeronautics, Astronautics, and Engineering Systems Division

Certified by
Dimitris Bertsimas, Thesis Supervisor
Boeing Leaders for Global Operation, Professor Operations Research MIT Sloan

Approved by
Richard C. Larson
Mitsui Professor of Engineering Systems
Chair, Engineering Systems Division Education Committee

Approved by
Maura Herson
Program Director
MIT Sloan School of Management MBA Program
Service Based Logistics Optimization

by

Gregory D. Price Jr.

Submitted to the MIT Sloan School of Management and the Department of Systems Engineering on May 9, 2014, in partial fulfillment of the requirements for the degrees of Master of Business Administration and Master of Science in Engineering Systems

Abstract

This thesis explores the use of a service based logistics optimization (SBLO) methodology for an inbound reverse logistics network. Currently, Quest Diagnostics solves the vehicle routing problem with time windows (VRPTW) in designing routes. The objective of the canonical VRPTW is to find a minimum cost route that visits every node once while meeting time window and capacity constraints without consideration to service levels. Since many of the nodes in Quest’s logistics network receive multiple pickups per day, have time-sensitive biological specimens, and require different service levels, the SBLO is more aligned with service objectives.

First, a spatio-temporal network model is created for every client in the logistics network. Next, a key service level metric (logistics turn-around-time) is defined. Finally, the SBLO is developed and tested on a small geographic area in Brighton, MA. The results of the two-week pilot were promising; service levels improved 25%, labor costs per requisition decreased by 10%-15%, and additional capacity was created the 2nd and 3rd shifts.

Although the effectiveness of the SBLO will be different for each route, the gains in service, reductions in cost, and increases in efficiency of the pilot warrant an investigation of the new optimization methodology applied to the entire logistics network. Quest could theoretically start processing 28% of the total New England testing volume by the 1st or 2nd shift, lowering operational costs, increasing efficiencies, and improving service levels dramatically. Additionally, this service based optimization strategy provides a value proposition that is more aligned with customer value expectations.

Thesis Supervisor: Deborah Nightingale
Title: Professor of the Practice of Aeronautics, Astronautics, and Engineering Systems Division

Thesis Supervisor: Dimitris Bertsimas
Title: Boeing Leaders for Global Operation, Professor Operations Research MIT Sloan
Acknowledgments

There are many individuals I would like to acknowledge and thank for the time and effort they dedicated in supporting me through this project and thesis.

The faculty and staff of MIT and the Leaders for Global Operations Program for the opportunity to be part of an unique and incredible learning experience.

My thesis supervisors, Dimitris Bertsimas and Deborah Nightingale for their time, guidance, and support throughout the internship and thesis writing.

I would like to thank Quest Diagnostics for their sponsorship of this project, their continued support of the LGO Program, and dedication to patients and healthcare professionals across the United States.

The Quest Cambridge management team, especially Chris Radke for her encouragement and guidance, Tim Anderson for his time and considerations throughout the project, Minoo Mahulkar for the times she listened and provided insight, Jim Bingel for his support and leadership throughout, Alpar Solymon for the insight and support of the logistics pilot, and Ellen Decaneas for making Quest a great experience for an intern. Thank you for a great seven months!

Finally, a special thanks to my lovely wife Tiffany Price for her endless support, encouragement, and love. And to my beautiful daughter Vivienne Isabella Price born November 6th, 2013. Their smiles were a guiding light throughout my time at MIT and helped make the experience unforgettable.

I leave you with a line from Walt Whitman’s O Me! O Life!: That you are here - that life exists and identity, That the powerful play goes on, and you may contribute a verse.
Contents

1 Introduction .................................................... 13
  1.1 Thesis Overview and Hypothesis ......................... 14
    1.1.1 Objectives ........................................... 14
    1.1.2 Approach at Quest Diagnostics ..................... 14
    1.1.3 Overview ............................................. 15

2 Background ..................................................... 17
  2.1 Company and Industry ...................................... 17
  2.2 Competitors ............................................... 18
  2.3 Laboratory Testing Sections ............................... 18
  2.4 Organizational Structure .................................. 19
  2.5 Quest Operations Overview ................................ 19
    2.5.1 Specimen Collection ................................ 21
    2.5.2 Logistics ............................................ 21
    2.5.3 Specimen Processing ................................. 22
    2.5.4 Specimen Testing .................................... 22
    2.5.5 Reporting, Releasing, and Calling .................. 22
    2.5.6 Service Levels ....................................... 23
    2.5.7 Focus Area ........................................... 23

3 Quest Diagnostics Logistics .................................. 27
  3.1 Overview .................................................. 27
  3.2 Background ............................................... 28
7.9 Backend Service Level Comparisons ........................................... 89
7.10 Discussion ................................................................................. 92

8 Conclusion ...................................................................................... 95
  8.1 Summary ................................................................................... 95
  8.2 Recommendations ..................................................................... 96
  8.3 Next Steps ................................................................................ 97
  8.4 Final Thoughts ........................................................................ 98
.1 Appendix I .................................................................................... 99
# List of Figures

2-1 Quest Diagnostics Cambridge organization structure ........................................ 20
2-2 Specimen collection to test result lifecycle ....................................................... 21
2-3 Turn-around-time broken down by stage in specimen lifecycle for a complete blood count test at Quest Cambridge ................................................................. 25
2-4 Daily collection times of unique accessions collected in New England ............. 26
2-5 Collection time versus hour-in-week with load time (put on analyzer) for an average 7-day week ................................................................. 26

3-1 Hypothetical service (TAT) versus cost efficient frontier for a logistics network 32

5-1 Turn-around-time record description of an example test .................................... 43
5-2 Process step to TAT record connections ............................................................. 45
5-3 Spatial-temporal network model creation process ............................................... 48

6-1 Example client location output of the spatial network model based on the service level database ................................................................. 51
6-2 Example PSC location output of the spatial network model based on the service level database ................................................................. 51
6-3 Aggregation of accessions based on collection time hour-in-day ...................... 52
6-4 Collection times for the majority of specimens collected in the New England network ................................. 54
6-5 A view of the spatio-temporal network model of Boston proper on October 21st, 2013 ................................. 55
6-6 Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 12 PM ................................. 56
6-7 Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 4 PM ........................................ 56
6-8 Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 8 PM ........................................ 57
7-1 One possible route solution of the SBLO to the maximizing service problem . 65
7-2 Another possible solution to the maximizing service problem .................. 66
7-3 SBLO framework system view ........................................................ 73
7-4 Time resolution versus min-max of the maximum TAT improvement ..... 76
7-5 Time resolution versus min-max of the average TAT improvement .......... 77
7-6 Improvement in LTAT for various time resolutions of design vector .......... 78
7-7 Illustration of how local search improves the results of the genetic algorithm 79
7-8 Max TAT: Time resolution versus cost using minimum weighted average objective ................................................................. 80
7-9 Average TAT: Time resolution versus cost using minimum weighted average objective ................................................................. 81
7-10 Illustration of how local search improves the results of the genetic algorithm for the weighted average objective ......................... 82
7-11 SBLO service level (TAT) versus cost tradeoff .................................. 88
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Logistics network parameters for Quest Cambridge</td>
<td>29</td>
</tr>
<tr>
<td>5.1</td>
<td>Data needed versus data available for various sources at Quest Diagnostics</td>
<td>42</td>
</tr>
<tr>
<td>7.1</td>
<td>Example logistics turn-around-time calculations</td>
<td>70</td>
</tr>
<tr>
<td>7.2</td>
<td>Summary of logistic turn-around-time metrics for two design vectors</td>
<td>71</td>
</tr>
<tr>
<td>7.3</td>
<td>Brighton, MA route descriptions of the original, optimized, and proposed</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>schedule</td>
<td></td>
</tr>
<tr>
<td>7.4</td>
<td>Spatio-temporal network model output for the stops on the Brighton, MA route</td>
<td>86</td>
</tr>
<tr>
<td>7.5</td>
<td>Example route cost model parameters</td>
<td>87</td>
</tr>
<tr>
<td>7.6</td>
<td>SBLO output for the baseline, optimal, and proposed route for Brighton, MA</td>
<td>88</td>
</tr>
<tr>
<td>7.7</td>
<td>Megasys accession volume counts for Brighton route before and during the</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>pilot test</td>
<td></td>
</tr>
<tr>
<td>7.8</td>
<td>Backend (Laboratory Operations) comparison for specimens transported on</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>the route in Brighton, MA</td>
<td></td>
</tr>
<tr>
<td>7.9</td>
<td>Performance gains of the quick-turn test codes for specimens transported on</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>the Brighton, MA route</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Quest Diagnostics (Quest) is the largest U.S. clinical lab operator and offers a broad range of laboratory testing services used by physicians in the detection, diagnosis, and treatment of diseases [1]. Many independent laboratories and hospital laboratories have endured decreases in the prices paid by the Medicare program and major private health insurers. Consequently, many of these organizations are experiencing flat or declining revenues and are looking to cut costs [2]. Appropriately, Quest is consolidating three principle laboratories and five satellite laboratories into an ultra-modern facility in Marlborough, MA. The move to Marlborough, MA will reduce operational overhead and improve organizational synergies for Quest. However, the move will also double the average distance a specimen travels to the performing site, which will exacerbate service levels and increase specimen transporting costs. As a result, there is uncertainty in how to pick up client specimens and drop them off in Marlborough, MA without impacting client service levels, arrival patterns, and laboratory operations.

More generally, a service-based organization is undergoing a move that stands to disrupt operations, negatively impacting clients and jeopardizing profitability. In order to mitigate this risk, it is important to understand and measure the current service levels in the network and the expected service levels during the transition and afterward.
1.1 Thesis Overview and Hypothesis

1.1.1 Objectives

The objectives of the thesis are as follows:

1. Qualitatively and quantitatively analyze the current state of the value chain of a complex service organization

2. Show how to develop a spatio-temporal network model that represents the demand and service levels of all clients serviced in a geographic area

3. Develop a service based logistics optimization (SBLO) methodology that maximizes service to clients

4. Demonstrate how the SBLO methodology will improve customer service levels, lower operational costs, and increase productivity and efficiency

Finally, a pilot project based on the SBLO framework will be executed, clearly showing the results to the service organization.

1.1.2 Approach at Quest Diagnostics

The research project that supported this thesis was completed at Quest Diagnostics. The project began by a qualitative investigation of the current state of the system through: an investigation of the value chain (specimen collection to resulting), interviews with the frontline workers who process specimens, and interviews with managers who lead the various functional departments at Quest. Next, a quantitative investigation began with a thorough data collection and analysis effort aiming to benchmark current service levels for all clients.

Following the data collection effort, statistical analyses and data mining were performed to identify the relevant information required to measure and analyze service levels. Next, this data was transformed into a model that helped to visualize Quest’s client network and provide further insight into the areas for improvement.

Throughout the project, research involving vehicle routing, operations, client service levels, and logistics was performed, drawing on advanced mathematics and concepts from
operations research and optimization. These techniques and methods were utilized to test a new optimization methodology that will substantially change the service levels of clients served by Quest Diagnostics Cambridge. This new optimization methodology was tested with a route in Brighton, MA over a period of twenty days.

1.1.3 Overview

The following work is divided into seven chapters. Chapter 2 discusses the present state of Quest Diagnostics as seen through the lens of the Quest Cambridge principal laboratory. Also, Chapter 2 provides an overview of the clinical laboratory industry, competitors, laboratory testing, and Quest’s operational structure. Chapter 3 details the logistics division at Quest and how decisions made in this department have major service level implications for clients in turning a specimen into a test result. Chapter 4 provides a brief literature review of the vehicle routing problem, reverse logistics, and optimization methods used to solve various instantiations of the vehicle routing problem. Data collection, analysis, and synthesis is developed in Chapter 5 and provides the groundwork for the spatio-temporal network model created in Chapter 6 and the subsequent service based logistics optimization model in Chapter 7. A service based logistics optimization (SBLO) methodology is developed and a pilot test of the SBLO is executed in Chapter 7 to validate and verify the model and approach. Also, we discuss the ramifications of implementing this service-based approach across the organization. Finally, we consider the next steps for Quest and the service-based optimization methodology in Chapter 8, specifically outlining recommendations for Quest Diagnostics given its present state.
Chapter 2

Background

2.1 Company and Industry

Established in 1928 and headquartered in Madison, NJ, Quest Diagnostics Incorporated (Quest) operates in the health care services industry, primarily in the clinical laboratory market where specimens are tested to detect, diagnose, and treat illnesses and diseases. Quest also operates internationally in the United Kingdom, Mexico, Brazil, Puerto Rico, and India [3]. Quest maintains a network of 35 principal laboratories throughout the U.S., 150 smaller "Rapid Response" Laboratories (RRLs), and over 2000 Patient Service Centers (PSCs) [1]. In 2012, Quest had $7.3 Billion in revenue and $555.7 Million in net income; its profit margin of 7.5% is low compared to competitors.

Quest’s clinical laboratory testing business, which boasts a test menu of more than 3000 tests including gene-based and other esoteric testing, accounts for 91% of the company’s revenue. Quest also provides testing for clinical trials and nonclinical testing, which includes risk assessment services, healthcare information technology, and diagnostics products accounting for the remaining 9% of the revenue. International growth will become more important as the specialty testing market begins to mature because 97% of the yearly revenue is generated in the United States. Quest is looking to grow its offerings in the gene-based and esoteric testing market due to the implications for personalized medicine and cancer screenings [4].

The well-established clinical laboratory market is highly fragmented. The aging U.S. population, improved health insurance coverage, and ongoing preventative care will continue
to boost revenue [4]. However, revenue growth has been in the mid single digits for most clinical labs, as Medicare and insurance reimbursements have waned. In the next five years, revenue is anticipated to grow at an annualized rate of 3.6% [4]. Additionally, healthcare insurers have taken steps to control testing pricing, because "managed-cost healthcare has become a primary method of cost control", which slows demand for healthcare services [4]. To combat this problem, clinical labs are attempting to grow organically by offering more high-margin esoteric testing options for personalized medicine. Research and development in genomics is the primary driver of the latest advancements in medical technology, and Quest has started to develop tests that would help determine a patient’s susceptibility to disease. Lastly, advances in medical technologies such as point-of-care-testing (POCT), have also started to pose a threat to clinical laboratories [4].

2.2 Competitors

According to Capital IQ, "Quest is the largest company in a mature market with limited growth opportunities," where most of the growth is derived via acquisitions of smaller regional laboratories [1]. The main competitors of Quest in order of direct competition are: Laboratory Corporation of America Holdings (Lab Corp), Hospital-based laboratories, Bio-Reference Laboratories, and physician-office laboratories [1]. Quest’s direct competitor, Lab Corp, offers an equivocal testing menu and service network and is the second largest clinical laboratory. The market for clinical testing has a low level of concentration, is highly competitive, and requires hefty fixed costs. The key factors for success are pricing, reputation, government policy, economics of scale, and proximity to key markets [1]. According to most managers at Quest, the distinguishing characteristics of Quest are service, quality, and building customer relationships.

2.3 Laboratory Testing Sections

Generally, laboratory medicine is divided into two sections: anatomic pathology, and clinical pathology. Anatomic pathology is comprised of histopathology, cytopathology, and electron
Clinical pathology includes:

- Microbiology - bacteriology, virology, parasitology, immunology, and mycology
- Chemistry - blood components, enzymology, toxicology, and endocrinology
- Hematology - coagulation and blood bank
- Genetic - cytogenetics
- Reproductive biology - semen, sperm bank, and reproductive technology

2.4 Organizational Structure

Quest maintains a fairly flat hierarchical structure divided by functional areas. Figure 2-1 illustrates Quest’s organizational structure from the CEO down to the frontline workers. Each functional department has direct reports and managers that are omitted from the figure for brevity purposes. This thesis work primarily focuses on interactions and interviews with people from the Regional VP down to the frontline workers. Route Service Representatives (RSRs), sales representatives, and client solutions are the customer facing segments of Quest.

2.5 Quest Operations Overview

Operationally, Quest is responsible for picking up specimens from client locations and Patient Service Centers, delivering those specimens to performing sites — the locations where specimens are turned into results by running on analyzers or run through a microbiology process — and finally reporting the specimen’s test results to clients. This simple description of the way a specimen becomes a test result is referred to as the specimen lifecycle. Since most of the principal components of Quest’s costs are fixed and located in operations, it will be critical to understand each component in the value chain to determine areas for improvement and to drive profitability. As Simchi-Levi points out, the smaller the profit margins, the more important it is to reduce costs to drive profitability, as those reduced costs will transfer directly into net profit. In other words, the same increase in net profit
Figure 2-1: General organizational structure of Quest Diagnostics Cambridge
due to an increase in sales is disproportionally higher than a decrease in costs (e.g. a 1% increase in net income for Quest requires an 13.3% increase in revenue, but a 1% reduction in COGS would translate directly into net income)[6]. In order to better understand the current operational state of Quest, it is important to understand the typical specimen lifecycle. Figure 2-2 illustrates the typical lifecycle of a specimen from collection to resulting. It is also worthwhile to note the functional department in charge of each step in the value chain.

![Figure 2-2: The typical specimen to test result lifecycle at Quest Diagnostics](image)

### 2.5.1 Specimen Collection

There are two primary specimen collection points for Quest Diagnostics: client locations and patient service centers (PSCs). Client locations are typically physicians' offices, hospitals, and long-term care facilities. At client locations, blood-based specimens are collected by accredited professionals, such as doctors, nurses, and phlebotomists. At every PSC, and some high-volume doctors' offices, Quest employs certified phlebotomists whom are responsible for specimen collection, preparation, and logistics readiness.

### 2.5.2 Logistics

After specimen collection, route service representatives transport specimens from various client locations and PSCs to performing sites. The types of performing sites are:

- Principle Laboratories - perform most of the testing catalog and contain most of the capital equipment and people required to process a large volume of specimens
• Rapid Response Laboratories - only perform a small subset of the testing catalog which are ordered "STAT" in order to speed up testing results for healthcare professionals providing care

• Specialty Laboratories - perform a very small subset of the testing menu, typically categorized as genetic and esoteric testing

As this thesis focuses primarily on logistics, a detailed discussion is provided in Chapter 3.

2.5.3 Specimen Processing

After delivery to a performing site, the specimen goes through the specimen processing department. Specimen processing is responsible for sorting, accessioning, labeling, and transporting specimens to their respective testing departments. Based on current staffing models, the specimen processing department is one of the largest labor department by number of people.

2.5.4 Specimen Testing

Following specimen processing, a specimen is tested and analyzed in one of the testing departments. The testing departments at Quest Cambridge are Automated Chemistry and Immunology, Cytology, Histology, Microbiology, Cytogenetics, Toxicology, and Hematology.

Depending on the testing department, the specimen will follow a certain process in order to be turned into a result (i.e. in Automated Chemistry and Immunology, blood samples are de-capped, poured off, stacked, scanned, and placed onto an analyzer). After the specimen is finished running on the analyzer, the test results are recorded into Quest’s Laboratory System (QLS) by the analyzer or a medical technologist; a discussion of QLS will follow in Chapter 5.

2.5.5 Reporting, Releasing, and Calling

Finally, after a specimen has completed the process of being turned into a test result, and the result is placed in QLS, a backend software process will look up a client’s release preference,
such as partial reporting or full reporting, and a client’s report preference: electronic, mail, fax, or telephone call. This final step is called releasing a test result. If the client has requested a call, QLS will wait for a test result to be released and will then indicate to the calling department that a notification is required.

2.5.6 Service Levels

Quest offers three service level agreements (SLAs) to clients for certain tests. The following descriptions outline the service levels:

- Routine - test will be resulted by 8 AM the following day of specimen collection if the test analytic time is less than 12 hours
- Same Day - specimens collected by 11 AM will be resulted by 5 PM depending on the test and reporting and releasing preferences of the client
- STAT - test must be resulted within 5 hours of collection time of the specimen or within 5 hours when the status is changed from routine to "STAT"

In investigating Quest’s service levels, we observed that the service level agreements are not clearly defined for every client and location, and some agreements are sold to clients by the sales team on an individual basis. Depending on the service level requested, the physical location of the specimen in the network, or type of test requested, a specimen may follow a different process from the specimen lifecycle in Figure 2-2.

2.5.7 Focus Area

After a cursory examination of the value chain from specimen collection to resulting, we determined that for most tests, the majority of time was spent in the logistics network and waiting for pickup. Figure 2-3 shows a boxplot of the specimen lifecycle for a comprehensive blood count test at the Cambridge, MA performing site, and illustrates the distribution, median, and variance of turn-around-times (TATs) at each step in the specimen lifecycle. This figure clearly indicates that the majority of a specimen’s TAT for a test is spent in the logistics network and this process step is the most variable. Moreover, the time spent in the
specimen processing department and running on an analyzer, is small in comparison to the
time spent in the logistics network. Additionally, Figure 2-4 plots the collection times for all
of the specimens collected in October 2013, normalized to obscure actual volume numbers.
A detailed discussion of the data collection technique will be discussed later in Chapter 5.

Finally, Figure 2-5 plots accession collection time versus hour in a 7-day week and overlays
load time. Load time is recorded when specimens are built onto a worklist (a collection of
specimens loaded onto the same cassette for an analyzer) and placed onto analyzers in the
performing site.

This analysis indicates that most specimens are collected before 12 pm but are loaded
onto analyzers more than 12 hours later. This spike in processing requires purchasing enough
machine capacity and employing enough labor to meet the 8 AM TAT service level agreement
promised by Quest for the majority of routine specimens. This initial analysis combined
with the fact that Quest Cambridge is moving to Marlborough, indicates that logistics is
the primary area to investigate. If Quest is able to get specimens earlier to the performing
site, most of the downstream operations become more efficient, staffing becomes easier, and
labor costs decrease.
Figure 2-3: Shows the distribution of the TAT for a complete blood count (CBC) for STAT tests routed to Quest Cambridge. Collection to Mini is an analog for the time spent in logistics, Mini to Build is an analog for Specimen Processing, Build to Release is akin to Testing, Release to Report is the time to report the test, and CallTAT is the amount of time until the healthcare professional is contacted.
Figure 2-4: Illustrates the daily collection times for every accession collected in the New England Network. Note that this data was produced by redistributing the actual collection times based on an empirical collection distribution. The cumulative % line indicates that approximately 68% of accession have a collection time before 12 PM.

Figure 2-5: Collection time versus hour-in-week overlaid with load time for an average 7-day week
Chapter 3

Quest Diagnostics Logistics

3.1 Overview

The logistics department at Quest Cambridge is one of the largest by headcount, with one director, one senior manager, three route managers, and over 90 Route Service Representatives (RSRs). A RSR provides service to clients by picking-up clinical specimens, and delivering supplies and reports within a designated route. The RSR maintains the integrity of specimens and safely transports them to performing sites for testing; assignments are typically limited to one route and one shift. Each day, the logistics department in Cambridge picks up tens of thousands of specimens from Maine to Rhode Island and delivers those specimens to various performing sites in and around New England. This requires a fleet of more than 70 vehicles and multiple vendors to successfully visit all of the required sites during a given time period.

Quest’s logistics network is unique in many ways. Specimen stabilities, client expectations, mix of service level agreements, pickup locations, health emergencies, and geographical areas each impose different constraints and priorities on the logistics network. Quest’s logistic network is considered to be a reverse logistics network as vehicles depart from a depot, follow a route to pickup specimens at multiple locations, and return to the depot or a performing site. The generally accepted definition of a reverse logistics is "the process of planning, implementing and controlling backward flows of raw materials, in process inventory, packaging and finished goods, from a manufacturing, distribution or use point, to a point of
recovery or point of proper disposal” [7]. Due to the time sensitive and perishable nature of biological samples, the logistics network must transport specimens to the laboratory in a timely manner and under environmental conditions that will not compromise the integrity of the specimens. For example, a complete blood count (CBC) test has a room temperature stability of 48 hours, a platelet poor plasma prothrombin specimen has a room temperature stability of 2 hours, and a neonatal bilirubin direct has a stability of 24 hours. If all of these specimens are picked up together, then the specimens would need to be transported to the performing site within one hour to make it through every step in the value chain. Further complicating matters, each client has a different set of expectations, time-windows, and internal processes to consider when picking-up specimens. Additionally, certain pickup locations are more difficult and constrained than others. For example, one client might have an external lockbox, which is easily accessible from the exterior of the building, but another client stores specimens inside a building which requires an RSR to park in a congested lot and navigate to a secluded locked room. Lastly, the distance from the performing site places the largest constraint on the logistics network in where to start a vehicle and when to get to the pickup location.

### 3.2 Background

Quest’s current logistic network has evolved with the testing needs of clients, the geographic locations served, and the service level agreements sold to clients by Quest’s sales team. Table 3.1 hints at the scale of the logistics operation for New England by showing how many trucks are on the road to service thousands of stops and touch points just for Quest Cambridge. A touch point is defined as a location that is scanned by an RSR when he or she checks for specimens; this is usually a lock box, tray, refrigerator, freezer, countertop, or box.

Routes vary depending on the day of the week, pickup location type, time of day, service level required, and geographic location. However, at Quest Cambridge, once routes are initially designed, they do not change substantially due to the following:

- It is a tedious and time-consuming task to find a route solution that meets all client
Table 3.1: Logistics network parameters for Quest Cambridge

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routes</td>
<td>154</td>
</tr>
<tr>
<td>Unique vehicle stop locations</td>
<td>2200</td>
</tr>
<tr>
<td>Unique touch points</td>
<td>8800</td>
</tr>
<tr>
<td>Vehicles</td>
<td>70</td>
</tr>
<tr>
<td>RSRs</td>
<td>90</td>
</tr>
<tr>
<td>Vendors</td>
<td>5</td>
</tr>
</tbody>
</table>

and testing requirements and does not result in a further disruption of service.

- Phlebotomists or clients maintain their own internal processes based on a scheduled pickup time agreed upon by logistics and sales.

- The existing client service requirements are outdated or not fully understood.

- Individual clients act individually and request the same schedules and time windows as other clients.

- Quest’s cost restrictions and full-time-equivalent (FTE) metrics incentivize logistics management to offer the lowest cost service which typically does not meet client service level agreements.

- Making changes to an earlier stop in a route will affect all downstream stops which could negatively impact clients.

- A reluctance to make any change that could affect a large client or subsidiary of a large client for fear of losing revenue.

- Same day or STAT service levels will be negatively impacted.

- Downstream operations are unable to successfully result specimens from certain clients within the service level agreement.

- Testing volume isn’t completely understood for every pickup location.
• Clients with lower volumes or with the best pickup schedules are unlikely to give up or change their schedules despite a change in volume or a shift in demand.

All of these issues create a "sticky" route environment in which clients and PSCs become accustom. Each route has a schedule, and every stop on a route has an estimated RSR arrival time. The scheduled arrival time is used heavily in the internal processes of client locations and results in a reluctance to change the arrival time once set. Additionally, phlebotomists at Quest's PSCs use the scheduled RSR arrival time as a signal to get specimens ready for transport through the logistics network. When routes change, the RSR arrival times change, and this has the potential to negatively affect clients and PSCs. Consequently, when a route gets created, its stops and estimated schedule do not change frequently enough to keep up with fluctuations in demand and to keep the route running efficiently.

There are many decisions in designing an effective, scalable, and flexible logistics network. The scope of logistics reaches across the organization and this requires logistic managers to contend with a many issues and priorities. Logistic decisions happened at three basic levels: strategic planning level, network level, and operations level [8]. The strategic level includes performance objectives, degree of vertical integration, outsourcing, and definition of custom service and associated metrics [8]. Network level decisions are concerned with the physical facilities and communication networks. Finally, operations level decisions involve demand forecasting, inventory management, productions, supply management, and transportation [8].

At Quest, the generally agreed upon logistics strategy is to create the lowest cost network that picks up all specimens within every client's service level and pickup constraints. Since the facilities are already built and the organization is moving to Marlborough, the network decisions remaining to be made are mostly concerned with the communication between clients and PSCs. Lastly, Quest Cambridge's operational decisions have created an environment were objectives are not clearly defined across PSCs, logistics, laboratory operations, and client services. This results in functional departments that optimize separately and an entire operation that runs sub-optimally.
3.3 Logistics Incentives

It is very difficult to design, implement, manage, and maintain an optimal logistics network of more than 2,000 pickup locations and more than 8,000 touch points with perishable goods. From upper management’s perspective, logistics is a necessary evil to doing business in the clinical laboratory market and not typically seen as an organization where value can be derived. Consequently, management incentivizes logistics to build the lowest cost logistics network possible. Quest is regularly looking to reduce full-time-equivalents (FTEs), in this context RSRs, from the logistics network to reduce costs. This puts undue pressure on route managers to redesign existing routes to service more stops with fewer vehicles and RSRs. This ultimately reduces service levels to many clients, providing that the network was close to optimal before the labor reduction. This is a very typical tradeoff between cost and service level provided. Figure 3-1 provides an illustrative example of the tradeoff between offering a higher level of service (lower TAT and more pickups) to clients by utilizing more vehicles and RSRs. As a logistic network adds vehicles, service improves considerably, but improvements to service reaches a lower limit as additional vehicles result in marginal improvement. This logistics TAT metric isn’t being captured at Quest Cambridge for current routes because collection times and volumes of specimens are not incorporated into the optimization methodology. Also, the metrics currently being used at Quest are meant to manage cost not to manage the service to clients.

3.4 Sub-optimization

Quest incentivizes the logistics management team to create low cost routes which satisfy most customer requirements. Whereas laboratory operations is run so that labor and analyzers are employed at levels that maximize utilization. Because these two departments are optimized separately, the whole system runs sub-optimally and ultimately reduces service to clients at a specified cost.

Most, if not all, logistic route designs or redesigns start with an initial route structure and the route manager changes that route based on a new set of requirements. As an
Figure 3-1: Illustrates the tradeoff between cost and service for a hypothetical example. If the vehicle network is over-utilized, adding additional vehicles to the network will increase service (lower TAT) significantly. However, further additions will only improve service marginally.
example, if a new client is added, removed, or a service change is requested, then the route manager will look at the current routes through that geographic location. If there is a current route through that location and close in proximity and time, the manager will either remap the route by hand, or use a route optimization package provided by ORTEC. ORTEC is a software company providing software solutions for supply chain optimization. The route optimization package essentially solves what is called a Vehicle Routing Problem (VRP) with time windows (VRPTW). A detailed literature review of route optimization and vehicle routing can be found in Chapter 4. If adding a client, the route manager will attempt to add that client to the route and still meet all of the previously mentioned requirements. This is typically done in small geographic areas without regard to the entire logistics network or with regard to downstream operations. This myopic optimization is one cause of a system-wide sub-optimality in logistics and operations. Additionally, route decisions are made and changed by sales representatives and dispatchers, whom are not informed of the downstream implications for a change in a route. These seemingly innocuous changes can radically change the service levels of clients on a route.

3.5 Sales

The sales team wields tremendous influence and power at Quest Diagnostics. Sales is a client facing function and is the primary communication channel to the client when selling the capabilities of Quest. This is problematic for operations because some sales representatives promise service levels which may not be feasible given current operational capabilities. This can lead to unnecessary route changes, missed pickups, and can worsen service for other clients on a route if logistics tries to shoehorn the client into an existing route. This problem exists due to a misalignment of objectives and incentives in the sales organization and operations and the lack of a unifying strategy between all functional units.
3.6 Implications of the Move to Marlborough

Quest is combining its two largest laboratories - Wallingford and Cambridge - and five other satellite laboratories. Cambridge and Wallingford currently process 91% of the total testing volume on average in New England. Since Quest is effectively doubling the average distance a specimen will travel to the performing site, service levels are likely to be negatively impacted for those customers furthest away from the new performing site. In order to mitigate this issue, Quest needs to modify its route optimization methodology and overall logistics strategy from lowest cost routes to maximizing service to clients (minimizing logistics lead time - the amount of time waiting for pickup and in the logistics network).
Chapter 4

Literature Review

4.1 The Vehicle Routing Problem

The classic vehicle routing problem (VRP) is a well-researched combinatorial optimization problem and has been solved sufficiently with exact, approximate, and heuristic methods for many scenarios. The constrained VRP (CVRP), is currently classified as NP-hard (in the strong sense), and generalizes the well-known traveling salesman problem [9]. The canonical vehicle routing problem typically involves a fleet of vehicles, delivery and pickup locations, routes, and depots. The problem is to create minimum cost routes with capacitated vehicles that service every location (nodes), at least once, and return to the depot. However, there are often several contrasting objectives in solving a typical VRP:

- minimization of the global transportation cost, dependent on the global distance traveled (or on the global travel time) and on the costs associated with the vehicles utilized and with the corresponding drivers
- minimization of the number of vehicles required to serve all customers
- balancing of the routes for travel time and vehicle load
- minimization of the penalties associated with partial service of the customers

When the number of vertices (nodes) increases to real-world numbers, exact and approximate solutions become intractable and heuristics are generally employed [10]. There
are many extensions to this basic problem such as: VRP with Time Windows (VRPTW), VRP with Backhauls, VRP with Pickup and Delivery, and the Capacitated or Distance-Constrained VRP (CVRP). There are more than 18 different common variants of the Vehicle Routing Problem [11]. Although, many of these variants have parts applicable to Quest’s logistics structure, none of them satisfy the service objective that seeks to maximize service or minimize lead time. However, the VRPTW is the variant used to initially generate routes. In order to borrow from concepts used in this optimization, a complete formulation of the VRPTW is developed in Section 4.2.

Quest Diagnostic’s main objective is building and maintaining a minimum cost logistics network based on the VRPTW problem. However, this thesis suggests that Quest should change its optimization objective from lowest cost routes to routes that maximize service (minimize logistics turn-around-time). This new optimization methodology is called the service based logistics optimization. Service to clients in the context of logistics is defined as minimizing the overall logistics turn-around-time (LTAT) for specimens in the logistics network. The LTAT is defined as the difference in time from a specimen’s collection to when the specimen is delivered to the performing site.

It is important to understand how Quest’s vehicle routing problem and logistics network differs from traditional VRP variants. First, Quest’s VRP is a reverse logistics problem where a vehicle leaves from a depot and picks up specimens from client locations and PSCs and returns to the depot. Second, the goods being transported are biological specimens that have stability, temporal, and temperature requirements, requiring the vehicle to return to the depot within a short time frame. Next, some pickup locations require multiple stops per day in order to speed up the TAT from collection to test resulting. Fourth, the heterogeneity of pickup locations requiring one pickup, multiple pickups, and strict time windows within a small geographic area is large. Fifth, many of the pickup locations are in an urban environment and most of the routes through these locations are time-dependent due to traffic, weather, construction, and other road hazards. Sixth, Quest’s vehicles are essentially non-capacitated due to the size and number of specimens picked up at each supply node during the day. Seventh, the number of specimens ready for pickup is not known before the RSR visits a location and the demand can vary considerably depending on the time of day.
and weekday. Eighth, the mix of service levels for specimens on each route an in geographic locations varies considerably and isn’t know ahead of time.

Many of Quest routes contain PSCs and client locations that receive multiple stops per day, which are also on the same route. Despite many attempts at creating minimum cost routes, logistics creates routes that are suboptimal due to manner in which they are constructed and subsequently changed by logistic managers, dispatchers, and sales representatives. These changes effectively result in increased costs across the entire laboratory operation. Additionally, many stops get more than one pickup per day which essentially changes the objective from minimum cost routes to ones that minimize the TAT of a test. The background of this hypothesis is discussed in more depth in Chapter 3.

4.2 Canonical VRPTW Formulation

Typically, the VRPTW is defined on a graph $G = (V, A)$, where the set of vertices (nodes) $V$ is defined as $V = S^d \cup S^c$ where $S^d$ is the depot node, the set of client nodes is $S^c = (1, ..., n)$, and the set of arcs $A$ is $A = (i, j) i, j \in V, i \neq j$. The fleet of vehicles is defined by vector $K = (1, ..., k)$. The following set of variables are used to define the problem.

$$x_{ijk} = 1 \quad \text{if use arc (i,j) \forall k \in K} \quad (4.1)$$
$$d_{ij} \quad \text{distance between nodes i and j} \quad (4.2)$$
$$a_i, b_i \quad \text{time windows of service for node i} \quad (4.3)$$
$$w_i \quad \text{the start time of service node i} \quad (4.4)$$
$$t_{i,j} \quad \text{travel time to node i from node j} \quad (4.5)$$
$$c_{i,j} \quad \text{cost associated with arc (i,j)} \quad (4.6)$$
$$q_i \quad \text{number of units picked up at node i} \quad (4.7)$$
$$s_i \quad \text{associated service time for node i} \quad (4.8)$$

Equation 4.9 is the canonical minimization of the global transportation cost formulation
for the VRPTW [10].

\[
\min \left( \sum_{(i,j) \in A} c_{ij} \sum_{k \in K} x_{ij} \right) \tag{4.9}
\]

Subject to the following constraints:

\[
\sum_{j \in V} x_{0jk} = 1 \quad \forall k \in K \tag{4.10}
\]

\[
\sum_{i \in V} \sum_{k \in K} x_{ijk} = 1 \quad \forall j \in V^c, \forall k \in K \tag{4.11}
\]

\[
\sum_{j \in V} x_{0jk} = \sum_{j \in V} x_{ak} \quad \forall a \in V^c \tag{4.12}
\]

\[
a_i \leq w_i \leq b_i \quad \forall i \in V, k \in K \tag{4.13}
\]

\[
w_{ik} + s_i + t_{i,j} \leq w_{jk} + (1 - x_{ijk})M \quad \forall (i, j) \in A, k \in K \tag{4.14}
\]

\[
\sum_{(i,j) \in A} t_{ij} \leq 8 \quad \forall (i, j) \in A, k \in K \tag{4.15}
\]

\[
\sum_{(i,j) \in V} d_{i,j} = D \quad \forall (i, j) \in V^c \tag{4.16}
\]

\[
x_{ijk} \in \{0, 1\} \quad \forall (i, j) \in A, k \in K \tag{4.17}
\]

\[
d_{ik} \geq 0 \quad \forall i \in V, k \in K \tag{4.18}
\]

\[
d_{ik} \leq C \quad \forall i \in V, k \in K \tag{4.19}
\]

The following section discusses the constraints given previously. Equation 4.10 ensures that all vehicles must start and end at the depot. Vehicles must abide by conservation of flow for every node by Equation 4.12. Each node is visited only once on this particular route by Equation 4.11. Time windows are respected in Equations 4.13 and 4.14. A driver can’t be on road longer than 8 hours or overtime will come into effect is show in Equation 4.15. Integrality and non-negativity constraints are captured in Equations 4.17 and 4.18. Finally, Equation 4.19 ensures that vehicle capacities are not violated.
4.3 Solution Methods: Exact, Approximate, Heuristics

Most exact methods for the VRP involve a capacitated fleet of homogeneous vehicles and one of three formulations: vehicle flow, two-commodity flow, and set partitioning. The vehicle flow formulations use binary decision variables to indicate if a vehicle traverses an arc. The two-commodity flow formulation requires an extended graph, and uses two flow variables to represent an edge of a feasible CVRP solution, and the load of the vehicle. The set partitioning formulation uses a binary variable for each feasible route through the network $G$. Unfortunately, even for moderately sized VRPs, the resulting LP programs are unable to be solved directly as the number of decisions variables is exponential [9]. This requires the use of cutting plane and column generation techniques described in the Large Scale Optimization Chapter of Introduction to Linear Optimization [12].

Exact algorithms can only solve relatively small instances of the VRP and their computational times are highly variable. Heuristics are generally employed when the number of vertices becomes even moderate in number (20 - 100). These strategies are used during the initial search process and effectively traverse the search space to obtain optimum or near-optimum results [13]. Metaheuristics are not problem-specific, and can be generally applied for all combinatorial problems, but they are not generally deterministic [13]. Some popular heuristics used to solve VRPs are: Genetic Algorithms, Simulated-Annealing, Ant Colony Optimization, and Particle Swarm Optimizations (PSO).

4.4 VRP with Service Based Optimization

A service based optimization strategy is considered due to the nature of Quest’s business and the alignment with client expectations. Since many client locations and PSCs get multiple stops per day, the optimization objective in this subset of routes should be to minimize logistic turn-around-time. The literature supporting a vehicle routing network based on overall service time is somewhat sparse. Typically, the optimization objective of the VRP with deterministic demands, is to minimize cost, which amounts to using a finite
number of vehicles and routes, where a node is only visited once. These solutions offer little to incorporate temporal demand into the formulation or to maximize service. Service, in this context, is defined as what the customer wants. For Quest, customers want quick, correct results for tests. This type of optimization is fundamentally different than the vehicle routing problem with stochastic demand, where demand is probabilistically defined [14]. The service based approach aligns a service level metric used throughout an organization and communicated to clients. The crux of the problem amounts to carefully defining the service level metric and incorporating the necessary information to minimize lead time (maximize service) to clients.
Chapter 5

Data Collection and Analysis

5.1 Data Source Overview

A detailed, accurate data source is required in order to successfully measure the service levels at every step in the specimen lifecycle for all accessions. In order to measure the service levels for each node in the network, a statistically relevant sample size, for each test and in each node needs to be available. We can draw reasonable conclusions about the service levels of clients if we make simplifying assumptions at each pickup location, but in order to understand service levels at a sufficient level of detail in the specimen lifecycle for every segment besides logistics, an enormous amount of data is required. Since this puts a large strain on the IT infrastructure, the data reports would be pulled in the off hours as to lower the impact to the system. At a macro level, there are approximately 2000 pickup locations, over 3000 tests, and over 20 processing nodes in the New England network. Since many of the tests are not ordered frequently, larger sampling time scales are required in order to draw reasonable TAT metrics. Additionally, some pickup locations process many specimens and thus the estimation of the service levels will be easier than those locations with low volume. Those pickup and drop off locations with less than five accessions per day were excluded from the network model in order to simplify the analysis. Lastly, in order to provide a time dimension to the network model, the data collection process needed to be detailed and robust enough to measure past and present performance, and needed to be updatable as more information became available. Figure 5.1 itemizes the data required to analyze the
specimen lifecycle for every accession. Once accessions are geotagged to their pickup and drop off locations, a directed graph can be constructed showing the direction of the flow of specimens to each of the respective performing sites.

Table 5.1: A comparison of the three potential data sources at Quest. Megasys is primarily used solely by the logistics organization. The Quest Laboratory System (QLS) is used by Lab Ops, Management, and Client Services and contains the highest fidelity dataset. Cognos is an IBM business intelligence software package with reporting, analysis, and dashboards, but its information is essentially summary data pulled from QLS and other sources.

<table>
<thead>
<tr>
<th>Data Required</th>
<th>Megasys</th>
<th>QLS</th>
<th>Cognos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accession Number</td>
<td>Limited</td>
<td>X</td>
<td>Limited</td>
</tr>
<tr>
<td>Pickup Location</td>
<td>Limited</td>
<td>Limited</td>
<td>Limited</td>
</tr>
<tr>
<td>Dropoff location</td>
<td>Limited</td>
<td>Limited</td>
<td>Limited</td>
</tr>
<tr>
<td>Service Level</td>
<td>Same-days and STATs*</td>
<td>STATs and Routine</td>
<td></td>
</tr>
<tr>
<td>Test/Specimen</td>
<td>Limited</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Pickup time</td>
<td>Limited</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Dropoff time</td>
<td>Limited</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Client</td>
<td>Limited</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Collection Time</td>
<td>Limited</td>
<td>X (75%)</td>
<td></td>
</tr>
<tr>
<td>Ready for pickup Time</td>
<td>Limited</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Processing Time</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Testing Time</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Resulting Time</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Reporting Time</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Calling Time</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>Limited (Inaccurate)</td>
<td>Limited</td>
<td>Limited</td>
</tr>
</tbody>
</table>

Unfortunately, there is not one data source to pull all of the required information. Therefore, a data fusion approach was utilized to incorporate disparate datasources. However, we decided to use QLS as our primary data source, and utilize an internally built data report in QLS called a turn-around-time (TAT) report. According to programmers and analysts at Quest, the TAT report records information for every test performed in New England. The detailed information contained in the TAT report is given in Table 5-1.

Figure 5-2 illustrates how the TAT report maps back to the specimen lifecycle illustrated in Figure 2-2. Although the TAT report has quite a bit of information, there is a large data gap from when specimens are collected to when they get entered into QLS. This due to the lack of IT infrastructure that has the ability to track at the accession level. This process
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Example</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
<td>Overall Performing Business Region</td>
<td>&quot;New England&quot;</td>
<td>String</td>
</tr>
<tr>
<td>Client ID</td>
<td>Unique identifier for a client</td>
<td>1234567890</td>
<td>Number</td>
</tr>
<tr>
<td>Accession</td>
<td>Unique identifier for a Requisition</td>
<td>XX12345X</td>
<td>String</td>
</tr>
<tr>
<td>Mini Date</td>
<td>Date that basic patient/test information is first entered into QLS</td>
<td>7/25/13</td>
<td>String</td>
</tr>
<tr>
<td>Mini Time</td>
<td>Time that basic patient/test information is first entered into QLS</td>
<td>02:17</td>
<td>String</td>
</tr>
<tr>
<td>Maxi Date</td>
<td>Date that full patient/test information is entered into QLS by another person</td>
<td>7/26/13</td>
<td>String</td>
</tr>
<tr>
<td>Maxi Time</td>
<td>Time that full patient/test information is entered into QLS by another person</td>
<td>01:13</td>
<td>String</td>
</tr>
<tr>
<td>QLS Date</td>
<td>The date the record was completed</td>
<td>7/25/13</td>
<td>String</td>
</tr>
<tr>
<td>QLS Time</td>
<td>The time the record was completed</td>
<td>20:17</td>
<td>String</td>
</tr>
<tr>
<td>Specimen Receive Date</td>
<td>Unknown</td>
<td></td>
<td>String</td>
</tr>
<tr>
<td>Specimen Receive Time</td>
<td>Unknown</td>
<td></td>
<td>String</td>
</tr>
<tr>
<td>Auto-Specimen Receive</td>
<td>Unknown</td>
<td></td>
<td>String</td>
</tr>
<tr>
<td>Load Build Date</td>
<td>Date when the specimen is loaded onto the analyzer</td>
<td>7/25/13</td>
<td>String</td>
</tr>
<tr>
<td>Load Build Time</td>
<td>Time when the specimen is loaded onto the analyzer</td>
<td>21:39</td>
<td>String</td>
</tr>
<tr>
<td>Collection Date</td>
<td>Date that the specimen was collected at the PSC</td>
<td>7/25/13</td>
<td>String</td>
</tr>
<tr>
<td>Collection Time</td>
<td>Time that the specimen was collected at the PSC</td>
<td>09:50</td>
<td>String</td>
</tr>
<tr>
<td>Full Log</td>
<td>If a requisition is input in the field, and all of the data is present, it is fully logged</td>
<td>N</td>
<td>String</td>
</tr>
<tr>
<td>STAT</td>
<td>Flag which indicates if specimen management flagged the test as a STAT</td>
<td>Y</td>
<td>String</td>
</tr>
<tr>
<td>ASAP</td>
<td>Flag which indicates if specimen management flagges the test as ASAP</td>
<td>N</td>
<td>String</td>
</tr>
<tr>
<td>Order Entry Site</td>
<td>Where the test was ordered</td>
<td>QWA</td>
<td>String</td>
</tr>
<tr>
<td>Performing Site</td>
<td>Where the test was performed</td>
<td>QWA</td>
<td>String</td>
</tr>
<tr>
<td>Order Code</td>
<td>What test is actually being performed</td>
<td>6399</td>
<td>Number</td>
</tr>
<tr>
<td>Worklist</td>
<td>Worklist number that the test was run with</td>
<td>SBCBC</td>
<td>String</td>
</tr>
<tr>
<td>Release Date</td>
<td>Date when the analyzer finished processing the test</td>
<td>7/25/13</td>
<td>String</td>
</tr>
<tr>
<td>Release Time</td>
<td>Time when the analyzer finished processing the test</td>
<td>23:13</td>
<td>String</td>
</tr>
<tr>
<td>TAT (hH:MM)</td>
<td>Relative time difference between collection to release</td>
<td>2:56</td>
<td>String</td>
</tr>
<tr>
<td>Report Date</td>
<td>Date when QLS reported out the information to clients</td>
<td>7/26/13</td>
<td>String</td>
</tr>
<tr>
<td>Report Time</td>
<td>Time when QLS reported out the information to clients</td>
<td>08:17</td>
<td>String</td>
</tr>
<tr>
<td>Call Date</td>
<td>Date when call center called the asking physician</td>
<td>7/26/13</td>
<td>String</td>
</tr>
<tr>
<td>Call Time</td>
<td>Time when call center called the asking physician</td>
<td>08:17</td>
<td>String</td>
</tr>
<tr>
<td>Call TAT (hH:MM)</td>
<td>Relative time between when a call was succesfully made to when test was released</td>
<td>1:30</td>
<td>String</td>
</tr>
<tr>
<td>Coll to Release TAT</td>
<td>Collection of specimen to test release</td>
<td>13:23</td>
<td>String</td>
</tr>
<tr>
<td>Coll to Report TAT</td>
<td>Collection of specimen to report</td>
<td>22:28</td>
<td>String</td>
</tr>
<tr>
<td>Coll to Call TAT</td>
<td>Collection of specimen to call</td>
<td>10:00</td>
<td>String</td>
</tr>
<tr>
<td>Logistics (Coll) to Mini'</td>
<td>Collection of specimen to mini timestamp</td>
<td>10:27</td>
<td>String</td>
</tr>
<tr>
<td>Mini to Maxi TAT</td>
<td>Time spent double checking mini data and putting in complete information</td>
<td>03:56</td>
<td>String</td>
</tr>
<tr>
<td>Mini to Build TAT</td>
<td>Time spent in specimen management</td>
<td>1:23</td>
<td>String</td>
</tr>
<tr>
<td>Build to Release TAT</td>
<td>Relative time between load build time and release time</td>
<td>1:33</td>
<td>String</td>
</tr>
<tr>
<td>Release to Call TAT</td>
<td>Relative time between test release to a successful call.</td>
<td>1:23</td>
<td>String</td>
</tr>
</tbody>
</table>

Figure 5-1: The table provides fields and associated descriptors for a turn-around-time record. Every test performed in New England with have a turn-around-time record. Most of the information in this report is calculated based on the timestamps of certain events. The pickup and delivery location is not in the report.
is typically called a specimen audit and provides where and when specimens are collected, picked up, and delivered to performing sites.

5.2 Assumptions

Since the logistic information is missing from the report, we need to make some assumptions about where Quest picked up the specimens, where it dropped them off, and when the specimens were ready for pickup. If we make six key assumptions and fuse data from different sources, then we can obtain most of the data required to recreate the specimen lifecycle and the logistics chain of events. The following six key assumptions are made for every TAT record:

1. **Performing Site** - the performing site code is a unique identifier of where the test was actually performed. Although a specimen can potentially be routed through multiple locations, the performing site is typically the final endpoint.

2. **Order Entry Site** - the order entry site code is a unique identifier of where the requisition was recorded into QLS. Many times this occurs electronically through Care360 (an IT system created by Quest and used by physicians), but can also be manually entered into IDAA by a technician (another IT system which handles accessioning), or electronically by scanning the data into IDAA. We can use this field to assume where logistics picked up the specimen. If the order code is from a recognized PSC, then Quest Logistics probably picked up the specimen from a PSC, if it is not from a PSC then most likely logistics picked up the specimen from a client location.

3. **Collection Time** - is a proxy for when specimens are ready to be picked up by logistics. Typically specimens go through a preparation process to ready them for transport. This typically takes 40 minutes if done sequentially.

4. **Unique accessions** - in order to obtain volume information, we need to count the number of unique accessions, and then count the number of tests on each accession. This will be our analog for testing and accession volumes. There is the complications
Figure 5-2: Describes the connections between the typical "STAT" specimen lifecycle events, illustrated earlier in Figure 2-2, and the TAT report described in Figure 5-1
that some records in the TAT report can bias the data. These records are filtered out of the dataset.

5. **Sum number of TAT records** - since we are dealing with tens of millions of records, and some of those records are not well-formed, we could potentially overestimate the actual number of tests. Also, some of the tests could not be in the TAT report, due to a coding error, and thus we will underestimate other tests.

6. **Zero or Null Collection Time** - up to 20% of the tests in the TAT report have a "null" or "12:00:00" entered for the specimen collection time. This problem is due to clients not recording and entering a valid collection time for particular specimens. We will need to statistically redistribute these specimens according to the 80% that have a collection time.

### 5.3 Building Spatio-Temporal Model of Network

Designing an optimal reverse logistics network requires a detailed understanding of the operations of an organization, a well-developed company strategy, and as much information about the demands of clients as possible. In order to determine the service levels in the New England network, we needed to answer the following questions:

1. Where are specimens entering the network?
2. What volume of specimens are at a pickup location?
3. When are the specimens ready for logistic pickup?
4. What is the TAT for every accession in the network?
5. What is the TAT for every test, client, and accession in the network through each stage in the specimen lifecycle?
6. What are the TATs for every test in the network?
7. Who are the largest clients and most requested tests?
In order to sufficiently answer these questions, we performed data fusion with the TAT report and ancillary information to determine client and PSC locations, test order code information, site code information, and service level information. Figure 5-3 presents a high-level graphical overview of the data fusion procedure that led to a general network model. We pulled TAT reports for every test performed in New England from June 2013 to November 2013; this amounted to processing approximately 10 Million records. In order to process this amount of data, an extensive code base was built in Matlab. The Matlab code handled data cleaning, filtering, removing garbage records, and reducing dimensionality of spurious data. Patient results and biographical information was redacted from the dataset. The data were input into a singular service level database (SLD). This became the primary source for the spatio-temporal network model.
Figure 5-3: Illustrates how the data model was created and the inputs to the model creations process. The turn-around-time (TAT) report and test-not-performed (TNP) reports are pulled from QLS for a time range. The report contains timestamps for most steps in the specimen lifecycle as listed in Figure 5-2.
Chapter 6

Spatio-Temporal Network Model

6.1 Spatial Formulation

Using the service level database developed in Chapter 5, we first build a spatial model of the network, each point representing a supply node (pickup location), and the size of the point representing the number of accessions (amount of demand). Essentially this first step is determining the nodes in a graph [15]. In order to produce this model, we geotagged patient service centers (PSC), rapid response labs (RRL), in-office phlebotomists (IOP), mobile phlebotomists (MP), long-term care businesses (LTC), and client location; this was approximately 10,000 points. Next, we used the following rules to link every TAT record in the database to a physical pickup location and delivery location.

- If the order entry code is either a PSC, RRL, IOP, or LTC, then the pickup location is the physical address for the associated order entry code. This rule mapped approximately 40% of the records to a physical location.

- If the order entry code is none of the above, then we use a client number to lookup physical business address for the associated client. This physical address is then used as the pickup location. This rule mapped approximately 55% of the records to a physical location.

- If the business address was located outside the physical region of New England, or a client number wasn’t found, then we geotagged the TAT record with a latitude and
longitudes of 0.0, 0.0.

Using this methodology, we created Figure 6-1, which presents the data for October 2013. Notice the density of demands are currently close in proximity to the two principle laboratories in Cambridge, MA and Wallingford, CT. This graphical network model also provided insight into the top clients by volume, test, location, and performing site. This information can be used strategically in determining which logistic routes will have the largest impact to arrival volumes into the performing sites.

There are some limitations to this network model which need to be recognized. First, the model makes assumptions based on the record in the turn-around-time report. A specimen could potentially take a multi-path route through the network, stopping at multiple locations in the logistics network, but only the start and endpoint locations are recorded in the QLS system. Secondly, many specimens are not entered into the QLS system through Quest’s electronic medical records system and only enter into the system at the performing site. Consequently, the assumption is that this specimen originated from the doctor’s location, but this could potentially come from another location if the doctor practices at multiple locations but only records one location in the QLS system. Thirdly, the pickup locations used for geotagging the TAT records are determined based on a client information database which could be out of date when the geotagging process is completed. This could result in erroneous pickup locations or underestimated volume numbers for some known locations. The number of unrecognized locations was below 3%. Fourth, some of the records in the TAT report are not well-formed and could contain spurious information that will inflate the volume numbers for various clients. Lastly, the TAT report was pulled every month with the same input parameters, but there is the possibility that not every record in the New England network will be pulled due to when the QLS records are finalized.

The following section incorporates the time dimension to the graphical model to provide the final piece to the optimization model.
Figure 6-1: The results of the spatial network model based on the service level database developed in Chapter 5. Each blue bubble represents a physical pickup location for logistics. The size of the bubble represents the number of unique accessions for the month of Oct 2013. The figure shows the highest density areas of the complete model.

Figure 6-2: The results of the spatial network model based on the service level database developed in Chapter 5. Each bubble represents a physical pickup location for logistics and the size each bubble is scaled relative to the location with the most unique accessions in the month of October 2013. This figure demonstrates the distribution of Client Locations, PSCs, and other miscellaneous pickup locations.
6.2 Temporal Formulation

We had two options to build out the time dimension of the model. We could either use the collection time dimension or the mini-time dimension to aggregate accessions by hour in the day. We decided to use the collection time dimension because 80% accessions include a valid collection time and this timestamp represents the start of the specimen lifecycle at the pickup location. Using the spatial model create above, we binned all the TAT records by hour-in-day listed in the collection time dimension. Next, we determined the number of unique accessions to determine a volume collected per hour-in-day. The results of this analysis is shown in Figure 6-3. A key decision here is the size of the time units chosen for the temporal aggregation of the accessions. The time discretization was chosen so that a computational advantage was found in the calculations, but small enough so that improvements could be found.

![Figure 6-3](image_url)

Figure 6-3: The results of aggregating accessions based on collection time hour-in-day. 20% of the collection times either did not have a collection time, or a collection time of 12:00:00, which probably indicates an incorrect collection time as it is the default value.

Since 20% of the accessions didn’t have a collection time, we redistributed these accessions and built a non-parametric estimated model based on the 80% of accessions that have a collection time. As inferred by the binned demand model, the arrival of patients to a patient
service center or client location has vastly different average hourly rates throughout the day. Figure 6-3 does not lend itself to any conventional parametric probabilistic function, so we utilized a histogram to determine the estimated probability density function (PDF). The histogram estimate of the pdf is then given by Equation 6.1. The origin \( x_0 \) was selected to be 12 AM and the resolution \((h)\) of the histogram was chosen to be one hour as this represents a balance between resolution, complexity, and smoothing. There are many other PDF estimation techniques which have nice properties, such as Kernel Density Estimation, and Gaussian Mixture Models but a histogram approach was used due to the simplicity and fidelity required of the estimation technique.

\[
\hat{H}_n(x) = \phi(x) = \frac{1}{nh} \sum_{i=1}^{n} 1_{x_0,x_0+h}(x_i) \text{ number of } x_i \text{ in the same bin as } x \quad (6.1)
\]

After removing the zero and null collection time records, and calculating the average for bin 0, we use the equations in 6.6 to create the estimate of the PDF for all supply nodes in the network. Next, we resample the TAT records, which were initially removed, and give them an estimated collection time based on the estimated PDF. We then take these resampled records and add them back to what is now an estimated demand model \( \hat{D}(x) \) which is shown in Figure 6-4. The following variables are used to estimated the collection time model.

\[
\begin{align*}
    d_i & \quad \text{the number of unique accessions with a collection time hour of } i \quad (6.2) \\
    d_{null} & \quad \text{the number of unique accessions with no collection time} \quad (6.3) \\
    d_0 & \quad \text{the number of unique accessions with a collection time of 12:00} \quad (6.4) \\
    \hat{d}_i & \quad \text{estimate of the number of unique accessions with a collection time hour of } i \quad (6.5)
\end{align*}
\]
23

\[ \sum_{i=0}^{23} d_i = D(x) \]

\[ d_{null} + d_0 = d_{no \text{ collection time}} \]

\[ \hat{d}_0 = \frac{d_1 + d_{23}}{2} \]

\[ d_{no \text{ collection time}} = d_{null} + d_0 - \hat{d}_0 \]  

(6.6)

\[ n = \sum_{i=0}^{23} \hat{d}_i \]

\[ h = 1 \]

\[ \phi(x) = \frac{\hat{D}(x)}{\sum_{i=0}^{23} \hat{d}_i} \]

Figure 6-4: Illustrates the daily collection times for every accession collected in the New England Network. Note that this data was produced by redistributing the actual collection times based on an empirical collection distribution. The cumulative % line indicates that approximately 68% of accession have a collection time before 12 PM.

After developing the spatio-temporal model for the entire network, this same estimation technique was applied to the individual supply nodes in the network. For each client location and PSC, we determined volume and collection times of accessions collected from June 2013
to October 2013. The data were pulled from QLS in the same manner for each month. We then aggregated by location and performed the same calculations shown in Equation 6.6 for each physical location. Some of the locations have very low volumes and highly variable collection times which makes forecasting difficult. Consequently, all nodes that have an average daily volume of less than five accessions is discarded from the model. This simplifying assumption is made because logistics is unlikely to offer scheduled pickup for supply nodes that have less than five accessions per day. Figure 6-5 illustrates the results of the model for the network on October 21st, 2013 at 8 AM, Figure 6-6 12 PM, Figure 6-7 4 PM, and Figure 6-8 8 PM for a small geographic area in Brighton, MA.

Figure 6-5: A view of the spatio-temporal network model of Boston proper on October 21st, 2013. This view omits the accessions that have zero or null collection times. Most PSC locations open at 8 AM. This figure illustrates the output of the model of the demand seen at 8 AM.

### 6.3 Specimen Readiness

A specimen must go through a preparation and manifesting processes in order for logistics to received and transport the specimen to a performing site. Typically, a specimen will go
Figure 6-6: Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 12 PM.

Figure 6-7: Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 4 PM.
Figure 6-8: Illustrates the output of the spatio-temporal model for a small geographic area in Boston at 8 PM. At this hour most PSCs and client locations are closed which is why there are so few dots in this area.

through the following process if collected at a client location or PSC:

1. Labeling - an identification barcode is printed and placed on the specimen collection tube

2. Clotting - most, but not all, specimens will have to wait 20-30 minutes for the blood to clot in the serum separation tube

3. Centrifugation - specimens will have to spin for 10-15 minutes in a serum separation tube

4. Pour Off - some specimens will have to be decanted into a smaller tube

5. Re-labeling Pour Off Tube - a label will be placed on the pour-off tube

If a specimen is collected at a PSC, then it will go through an additional manifesting processes in order to keep track of the specimen during transportation and delivery to specimen processing at the performing site. The entire specimen readiness process, from collection to
manifesting, takes an average of one hour. Therefore, the first chance for logistics to pickup a specimen is one hour from collection time if the phlebotomist starts the specimen readiness process immediately. By neglecting the internal complex queueing, process prioritization, staffing, and service rates at each supply node, we can shift the spatio-temporal demand model and put a bound on when specimens are first ready to be picked up by logistics. We make this simplifying assumption for the optimization model developed in Chapter 7.

6.4 Distinguishing Service Levels

Lastly, we wanted to be able to distinguish the level of service requested for each accession or test in accession (service levels were defined in Section 2.5.6). This was difficult to determine based on the data provided in the TAT report. There were two fields in the dataset which could help distinguish each test, but these flags are not used consistently throughout Quest. For example, a specimen processor inputs a "Y" into the STAT field if a STAT tests comes into specimen processing at the principle laboratory, but if the specimen is delivered to some of the RRLs, the processor doesn’t record the test as a STAT because Quest uses this information to call doctors with test results. Consequently, we made assumptions based on the performing site locations, as to the volume of STAT specimens coming from various locations. The TAT reports did not adequately record the same-day designation and therefore we were unable to determine the volume of same-days at each supply node. Consequently, most test were considered to be routine tests.

6.5 Inputs to the Optimization Model

Now that we have a spatio-temporal model that describes when specimens are ready for pickup for all locations in the network, we can now feed this information into an optimization model which can help us determine the optimal way to design routes in the logistics network. For a given geographic area, the spatio-temporal model can be queried to pull the demands for every pickup location. Next, we can use this specimen collection data to determine the optimal time to schedule pickup for every location in that geographic area. Additionally,
this database will help to inform the route managers as to how many specimens are typically collected on every route and when those specimens are delivered to the processing site. The following chapter outlines the optimization model and the new logistics objective in combination with the spatio-temporal model developed in this chapter.
Chapter 7

Optimization Model

In this Chapter, we develop a Service Based Logistics Optimization (SBLO). The SBLO incorporates the demand profile for all of the supply nodes in the logistics network, and tunes routes to minimize the maximum logistics turn-around-time (LTAT). This new optimization methodology: (1) improves specimen flow to the processing sites, (2) lowers overall operational costs, (3) provides a service level bounds (LTAT) on the routes in a given geographic area, and (4) guards against batching specimens (units) in the logistics network. The following sections discuss route generation, route tuning, incorporation of demand into the optimization model, and the LTAT key performance indicator. Next, the SBLO is mathematically formulated and discussed. The optimization model is then implemented in Matlab and tested with demand data for a small geographic area. Finally, the SBLO methodology is applied to a small geographic area in Brighton, MA, and the results are discussed.

7.1 Service Based Optimization

As discussed in Chapter 4, the VRPTW problem has been solved sufficiently with many different techniques. Quest uses an optimization tool provided by ORTEC to solve the VRPTW for a specific region, with time time windows set according to morning, afternoon, and night schedules. ORTEC does not take into account specimen TAT (the client focused metric) or when specimens are collected and ready for pickup, a key piece of information for logistics managers. Moreover, ORTEC does not optimize with respect to the level of demand
(number of specimens ready) at any time of day, and simply ensures every location is visited at least once. Additionally, the route manager is not aware of the average demand for specific pickup locations or when these specimens are available. Quest initially asks clients when they would like a pickup and attempts to accommodate everyone on a specific route. The client makes these decisions without regard to those clients in the same geographic region that will be placed on the same route. This creates a sub-optimal route structure and ultimately leads to longer specimen TATs. Notwithstanding, Quest Routes are initially VRPTW optimal based on the client data and constraints entered into the tool, but the routes become sub-optimal due to the decisions of dispatchers, poor input data quality, and route manager changes. The data and constraints input into the tool are based on information from Quest’s sales organization and the client, and in many instances these data are incorrect or assumed to be the default value. This results in generated routes that are sub-optimal or routes that can be improved to offer better service to select clients. The following SBLO is different from the canonical VRPTW based on the following facts:

- Many client locations and PSCs receive more than one pickup per day: morning, afternoon, and evening

- The morning and afternoon time windows have typically been chosen based on: (1) how long it takes to get test results on specimens collected earlier in the day, (2) the distance to the performing site, and (3) how quickly the client would like the specimens resulted

- The time windows for the morning and afternoon roots are typically “soft” constraints and picked without regard to when specimens are ready for pickup and without regard to reducing overall specimen TAT

- The vehicles transport specimens that have varying stability times.

- The heterogeneity, and size of the specimens at every pickup location does not impact vehicle capacity

Essentially, clients want some specimens picked up and resulted quickly, and other specimens resulted at a lower service level. As evidenced by looking at the current Cambridge morning
and after route structures, many of the morning and afternoon time windows are initially selected based on the severity of not meeting service levels and on the lead time through backend operations. In fact many location receive multiple pickups due to singular instances where specimens get left behind. This negatively impacts the client and the route manager adjusts the entire route to compensate. An optimization objective that is more aligned with every client is to minimize the total logistics turn-around-time (LTAT) for specimens on a set of client nodes currently served by a route. LTAT is defined as the amount of time from when a specimen is collected to when that specimen is delivered to the performing site. Since most of the morning and afternoon routes are not time window constrained why use a VRPTW to select a route? A more robust service objective would be to incorporate the hourly demands of supply nodes in the network, and to minimize the amount of time between when a specimen is collected to when it is transported to a performing site. This allows laboratory operations to get specimens in sooner, which allows machines to be stocked earlier, and test results to be delivered faster to clients. Additionally, this optimization methodology allows the logistics manager to track service level metrics for every client on the route, which will better inform the client services team and route manager of service level interruptions.

7.2 The SBLO Formulation

Given a set of routes which visits a set of supply nodes (client locations and PSCs), decisions need to be made as to how many pickups a site requires based on service level, demand, volume, and service level mix. Likewise, a decision needs to be made as to when these routes should begin if not truly constrained by a time window. Using each pickup node’s associated collection time demand profiles \( \hat{D}_{t,d} \), developed in Chapter 6, we can now determine which routes best minimize the maximum LTAT. The problem is formulated by Equation 7.11. The current formulation is classified as a mixed-integer non-linear program (MINLP). The SBLO optimization is solved in two phases. The first phase uses a genetic optimization to find routes and start times that minimize the maximum aggregate LTAT for all the nodes in a given geographic region. There are many tradeoffs to consider when deciding the algorithm to use in solving a MINLP [16]. A genetic optimization approach was utilized based on the
following:

- The concept is easy to understand and explain to stakeholders
- Supports multi-objective optimization
- Is inherently parallel and easily distributed so that more of the search space can be explored
- Performs well when the objective space is highly non-linear and discontinuous
- Does not require analytical gradients for the objective function

Genetic Algorithms are defined as probabilistic search algorithms that iteratively transforms a set (population) of mathematical objects (design vectors), each with an associated fitness value, into a new population of offspring objects, using the Darwinian principle of natural selection, where the operations are patterned after naturally occurring genetic operations such as crossover and mutation [17].

After the genetic search algorithm was applied, a local search algorithm was then employed to refine and further optimize the genetic algorithm results. The use of local search algorithms for binary optimization is well established in operations research and produces good results [18]. The local search algorithm is an iterative search algorithm that terminates when no further improvement can be found. Essentially, the local search algorithm proceeds in the following manner:

1. Given a starting point \( x_0 \), set the current best design point, \( x_{best} \), to \( x_0 \)
2. Set a test variable \( x_{test} \) to \( x_{best} \)
3. Remove one of the selected binary variables in \( x_{test} \) and set to 0
4. Step once in each dimension of the binary variable, other than those already selected, and calculate the objective \( J_{test} \)
5. Find the best improvement across every dimension and every selected dimension for variable \( x_{test} \) that still passes the inequality and equality constraints \( g(x_{test}) \) and \( h(x_{test}) \)
6. If the best objective $J_{best}$ is better than $J_{test}$ set $x_{best}$ to $x_{test}$ and go back to step 2

7. If there is no improvement in the objective within tolerance, terminate

After the local search was run, a trade study was completed to understand the tradeoff between time resolution in the optimization model and how much the objective can be improved.

Figure 7-1: This figure illustrates an example of one possible solution to the SBLO. A route is started at time t and services all of the nodes before returning to the depot.
Figure 7-2: This figure illustrates another example of a possible solution to the SBLO. One route is selected to only service nodes 1 and 6 before returning to the depot, whereas another route is started at a different time to service nodes 2, 3, 4, and 5 before returning to the depot.
Decision Variables

Given a set of predetermined routes, each servicing some or all of the nodes in a small geographic area, in which time windows are relaxed, one must decide which routes to use and at what time to start those routes to minimize the maximum logistic turn-around-time for all specimens in the geographic area. An illustrative example is presented with Figures 7-1 and 7-2, optimization can choose from a set of three routes to service all six nodes in a given geographic area. There are many different routes that could service all of the nodes in this geographic area but only a subset of them are used. Due to the combinatorial nature of which routes to run, when to start them, and the time resolution, the number of design variables grows substantially.

Equation 7.1 defines a binary decision variable that is one if route i starts at time t, and zero otherwise.

\[ x_{it} = \begin{cases} 1 & \text{if route } i \text{ starts at time } t, \\ 0 & \text{otherwise} \end{cases} \quad (7.1) \]

Data and Derived Data

The following data is required to run the SBLO optimization. Note that the total travel times are deterministic and do not change throughout the day.

\[ \hat{D}_{jt} \] The estimated demand generated at node j at time t. \hspace{1cm} (7.2)
\[ T_{i,t} \] The total travel time for route i started at time t for a depot D \hspace{1cm} (7.3)
\[ t_{j,k} \] The travel time to node k from node j \hspace{1cm} (7.4)
\[ C_i \] The marginal cost of running route i \hspace{1cm} (7.5)
\[ a_{j,i} \] Arrival time of route i at node j if vehicle leaves at time \( S_{i,t} \) \hspace{1cm} (7.6)
\[ a_{j,i}^p \] Previous arrival time of route i at node j if vehicle leaves at time \( S_{i,t} \) \hspace{1cm} (7.7)

Intermediate Variable Definitions

Equation 7.8 presents the intermediate variable \( Z_j \), which is defined as the aggregate logistics turn-around-time (LTAT) for node j based on a design vector \( x \). LTAT is defined as the
relative time difference from when a specimen is collected to when it is delivered to the performing site.

\[ Z_j = f(x) \] logistics turn-around-time for node j based on a design vector \( x \) (7.8)

Since there are multiple specimens collected at different times, an aggregation metric is required for each node \( j \). The maximum aggregate LTAT for node \( j \) is defined as the maximum of the weighted average of the LTAT for node \( j \) across all visits of node \( j \) throughout the day. Since the optimization algorithm uses a single scalar variable for all of the nodes, scalarization was explored with two options:

1. Option 1 - a weighted average based on the number of specimens transported from each node (weighted average LTAT)

2. Option 2 - maximum LTAT across all nodes (maximum LTAT)

It is important to understand the aggregation method in order to understand the metric being optimized. Table 7.1 provides an example of how to calculate the aggregate LTAT for a node based on a design vector \( x \) and assuming only one type of route \( (r_1) \). Table 7.1 shows the calculations based on two different design vectors \( x_1 \) and \( x_2 \). The first design vector indicates that a route of type \( r_1 \) will start at 10:00 and another will start at 19:00. Design vector two indicates that a route of type \( r_1 \) starts at 13:00, 17:00, and 19:00. Note that the design vector \( x_1 \) has dimensionality that is set by the number of routes \( r \) multiplied by the time resolution. For example, if there are seven routes and 30 minute resolution is desired, then the \( x \) vector will have dimensionality (number of rows by number of columns) of 48 by 7, for a total of 336 different dimensions.

Table 7.1 assumes a time resolution of 60 minutes (1 hour). The weighted average LTAT for design vector 1 is shown in the Node1*TAT Option1 row. Table 7.2 provides a summary of the LTAT calculations for the example presented in Table 7.1. For \( x_1 \), a route of type \( r_1 \) is able to pickup all the demand from time bin 7 to time bin 11. Since the route returns to the depot at 12:00, the specimens in time bin 7 wait 5 hours until they are delivered to the performing site. Additionally, the two specimens in time bin 8 wait 4 hours and so forth.
for each specimen in other time bins. Since the first route arrived at Node 1 at 11:00, it picks up all the previous demand from last visit to time bin 11. Next another route of type \( r_1 \) arrives at Node 1 at 20:00 and picks up all the demand from time bin 20 to time bin 12, since all the other demand was picked up by the previous route. The weights for the LTAT are calculated by multiplying the number of specimens times the difference between the depot time and the collection time for a specific time bin. Equation 7.9 is the weighted average LTAT for the node \( j \) based on the design vector \( x \). \( a_{j,i}^p \) is the time of the previous visit of node \( j \) by route \( i \) and \( a_{j,i} \) is when route \( i \) arrives at node \( j \). \( T_{depot,i} \) is the time when route \( i \) returns to the depot. Finally, \( k \) is the index for time used in Equations 7.9 and 7.10. Equation 7.10 is the total demand picked up from node \( j \) by all \( r \in R_{routes} \) that visit node \( j \) based on design vector \( x \).

\[
f_j(x) = \frac{\sum_{i \in R_j} \sum_{k=a_{j,i}^p}^{a_{j,i}} \hat{D}_{j,k}(T_{depot,i} - k)}{\sum_{i \in R_j} \sum_{k=a_{j,i}^p}^{a_{j,i}} \hat{D}_k}
\]

\[
M_j = \sum_{i \in R_j} \sum_{k=a_{j,i}^p}^{a_{j,i}} \hat{D}_{j,k}
\] (7.10)

The total weighted average LTAT for Option 1 in Table 7.1 is found by summing the Node1*TAT Option 1 row and dividing by the sum of the Node 1 Demand row. Note that is calculation follows Equation 7.9. The first summation in the equation would go from 1 to 2 since there are two route run. The LTAT column in Table 7.2 shows the LTAT of Node 1 as a result of running a route of type \( r_1 \) starting at 10:00 and again at 19:00. The maximum LTAT for this design vector is 6.56 decimal hour whereas the weighted average LTAT is 4.88 decimal hours. Option 2 indicates a lower maximum LTAT and lower weighted average LTAT. This example presents the two different ways of producing a scalar LTAT for a node given different design vectors \( x_1 \) and \( x_2 \). Returning to Equation 7.8, we now understand that the LTAT for each node depends on the design vector \( x \) so that \( f(x) \) represents the LTAT for node \( j \).
Table 7.1: Provides an example of LTAT calculations for node 1 given two different design vectors. Design vector 1 indicates that route 1 starts at 10:00 and 19:00. The second design vector indicates that route 1 starts at time 13:00, 17:00, and 19:00.

<table>
<thead>
<tr>
<th></th>
<th>Time Bins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23</td>
</tr>
<tr>
<td>Node 1 Demand</td>
<td>0 0 0 0 0 0 0 1 2 2 2 1 2 2 1 1 1 1 1 0 0 0 0 0 0</td>
</tr>
<tr>
<td>Design Vector 1</td>
<td>0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0</td>
</tr>
<tr>
<td>TAT Option 1</td>
<td>12 11 10 9 8 7 6 5 4 3 2 1 9 8 7 6 5 4 3 2 1 0 0 0 0</td>
</tr>
<tr>
<td>Node1*TAT Option 1</td>
<td>0 0 0 0 0 0 0 5 8 6 4 1 18 16 7 6 5 4 3 0 0 0 0 0</td>
</tr>
<tr>
<td>Design Vector 2</td>
<td>0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0</td>
</tr>
<tr>
<td>TAT Option 2</td>
<td>15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 4 3 2 1 2 1 0 0 0</td>
</tr>
<tr>
<td>Stop1*TAT Option 2</td>
<td>0 0 0 0 0 0 0 8 14 12 10 4 6 4 1 4 3 2 1 0 0 0 0 0</td>
</tr>
</tbody>
</table>
Table 7.2: Summarizes the calculations for the example given in Table 7.1. The two different options correspond to design vectors 1 and 2 in the previous table.

<table>
<thead>
<tr>
<th></th>
<th>Route Start (h)</th>
<th>Node1 Arrival (h)</th>
<th>Depot Time (h)</th>
<th>Units</th>
<th>LTAT (decimal h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option 1</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>9</td>
<td>6.56</td>
</tr>
<tr>
<td>Option 2</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>13</td>
<td>4.54</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>4</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>0</td>
<td>NaN</td>
</tr>
</tbody>
</table>

Objective

The following objective function seeks to minimize across all routes that reach a set of nodes in the geographic region of interest. This presumes that a set of VRP optimal routes is known ahead of this optimization step. Since the complete set of nodes $N_{nodes}$ may be serviced by a number of these VRP routes, each route is considered in the service maximization problem. These routes were generated using the approach outlined in Section 4.2. Note that this objective is using a min-max optimization method in order to provide a solution that bounds the worst case LTAT. This thesis also explores the minimization of the weighted average LTAT for all of the nodes in the set $N_{nodes}$.

$$\min \{ \max_{j \in N_{nodes}} \{Z_j\} \} \tag{7.11}$$

Constraints

Most of the constraints are relaxed for this optimization, but a few additional constraints are required. First, Equation 7.12 ensures that all of the specimens need to be picked up from all of the nodes. $M(x)$ is defined by Equation 7.10. Next, Equation 7.13 ensures that all specimens are delivered to the performing site within 12 hours of collection due to specimen stability, temperature, and quality requirements. Equation 7.14 controls the number of times that a specific route i can run per day. Equation 7.15 puts a bound on how many total routes can run in one day. A cost constraint could be added to the model, but we decided to relax
this requirement in order to build a cost versus LTAT (service) tradespace.

\[
\sum_j M_j = \sum_{j,t} D_{j,t} \quad \text{Routes must pickup all the demand in every node for the day}
\]  
(7.12)

\[
Z_j \leq 12, \forall j \in N_{\text{nodes}} \quad \text{The maximum logistic turn-around-time for every node } j
\]  
(7.13)

\[
\sum_t x_{it} \leq 4, \forall i \in R_{\text{routes}} \quad \text{The max number of times a single route } i \text{ can run in a day}
\]  
(7.14)

\[
\sum_{i,t} x_{it} \leq 4 \quad \text{The total number of routes that can run in a day}
\]  
(7.15)

7.3 Implementation

An optimization simulation was designed and implemented in Matlab. Figure 7-3 illustrates the optimization simulation at a system level. A logistics module was developed in order to simulate a logistics network in a small geographic area. The module runs a route plan based on a given design vector \( x \) and returns the associated maximum aggregate logistic turn-around-times, route costs, average turn-around-time, demand picked up at every node, and total demand.

A module was also implemented to incorporate the non-linear constraints listed in the previous section. This is indicated as the constraints and options module in Figure 7-3. The genetic algorithm in Matlab uses a penalty function to incorporate the non-linear constraints. Consequently, each constraint listed above should be scaled to avoid over/under penalizing each new design variable. If this scaling is not done, the genetic algorithm could produce a result that doesn’t violate the greatest in magnitude constraint, but does violate the smallest constraint. In the example above, the max number of routes constraints, Equations 7.14 and 7.15 would be violated before Equation 7.12 due to the difference in scale.

After successfully validating the logistics module, wrapper code was developed to pull in
Figure 7-3: Illustrates the optimization framework that was implemented in order to complete the time resolution, max routes, and max versus average LTAT trade study. All modules were designed and implemented in Matlab. Matlab's genetic algorithm (ga) was used for the initial global search.
the demand data, route encodings, and constraint constants (parameters) discussed in the
previous section. These are shown at the top of Figure 7-3. The optimization trade study was
run and tested with six different nodes, each with demands pulled from the spatio-temporal
model, and seven different route options servicing those six nodes. Matlab’s internal genetic
algorithm function (ga) was used for the initial global search, and a local search algorithm
was developed from scratch and validated with various know test runs. The local search
algorithm, logistics model, and cost model functions can be found in the appendix of this
document.

7.4 Performance

We ran the optimization model for time intervals of 240, 120, 60, 30, 15 and 10 minutes.
For each time interval, the optimization was run and varied the maximum number of routes
run per day constraint from 1 to 4. Figure 7-4 and 7-5 demonstrate the results of the trade
study. The maximum LTAT and average LTAT improved when the maximum number of
routes constraint was relaxed from 1 to 4. This parameterization of the constraint allowed the
optimization model to select a route plan that drastically reduces the LTAT by almost 61%
from 1 route per day to 4 routes run per day. The optimization model improves by an average
of 27% when reducing the time resolution from 240 minutes to 10. Figure 7-6 demonstrates
the incremental improvements in the model when reducing the time interval. The percentages
are relative to the previous time interval. For example, the percent improvement from moving
to 60 from 120 minutes is 11.7% for maximum LTAT and 11% for the weighted average LTAT.
The largest improvement was observed in going from 60 minute to 30 minute resolution.

Figure 7-7 illustrates the improvement using the local search heuristic on results of the
genetic algorithm for time resolutions of 240, 120, 60, 30, 15, and 10 minutes. Each subplot
shows a different time interval and sweeps over the max number of routes constrain show
in Equation 7.15. When the time interval (time resolution) is large (small), the local search
algorithm is unable to improve the genetic algorithm’s solution. However, once the time in-
terval decreases, the local search solution improves the genetic algorithm substantially. This
is likely to the increase in dimensionality of the design variable as the resolution increases.
Next, the same trade study was run with a different objective shown in Equation 7.16. This objective uses a weighted average aggregation of all of the nodes and the demands from those nodes. This was explored to investigate the possibility of one node having demand that is much different than any other node and potentially skewing the results.

\[ \min \sum_j M_j * Z_j \quad \text{weighted average of the logistic turn-around-time for nodes in } N \] (7.16)

The trade study indicates that the optimal time resolution for the model is the 30 minute interval. This provides improvement to both the maximum LTAT and weighted average LTAT objectives shown in Equations 7.16 and 7.11. The highest percentage improvement in maximum and weighted average LTAT is in going from 60 to 30 minutes. This was illustrated with Figure 7-6. The local search algorithm improves upon the genetic algorithm at time resolutions greater than 120 minutes. This was confirmed with both the maximum LTAT and weighted average LTAT objectives presented in Figures 7-7 and 7-10. Interestingly, at time resolutions of 240 minutes, the genetic algorithm and local search algorithms produced the same result. This is likely due to the dimensionality of the design space and the minimum population size option passed to the genetic algorithm, which was set at 500. This adequately covered the design space and both the genetic and local search algorithm settled on the same local optimum. Furthermore, the genetic algorithm coupled with the local search algorithm produces substantially better results when the time resolutions are in the ranges of 15 to 10 minutes. The local search algorithm executes faster than the higher resolution genetic algorithm runs. This is due to the size and complexity of the design space. Finally, the weighted average LTAT objective produces results that are monotonically decreasing in LTAT for increases to the maximum number of route constraints. This was demonstrated with Figures 7-8 and 7-9. Lastly, this trade study demonstrated the benefits of increasing the time resolution and using a local search algorithm on top of genetic algorithm outputs.
Figure 7-4: This graph provides a family of curves exploring the tradeoff between time resolution and objective function (Maximum TAT) improvement. As the time intervals are decreased (increased in time resolution), the optimization solution improves substantially across the cost dimension.
Figure 7-5: This graph presents the family of curves exploring the tradeoff between time resolution and objective function (Average TAT) improvement. As time resolution increases, the objective improves by almost 30% across the cost dimension.
Figure 7-6: This graph shows the improvement (reduction) in LTAT for the validation runs of time resolutions 240, 120, 60, 30, 15, and 10 minutes. For each interval, the optimization was run with 4 different maximum number of routes run per day constraints. Interestingly, the improvement from 240 to 120, and 120 to 60, is not as large as the improvement observed in going from 60 to 30 minutes.
Figure 7-7: This figure shows a collection of plots that illustrate the improvement in maximum LTAT across a given time interval by using the local search algorithm. Each plot illustrates the improvement in object by using the local search algorithm seeded with the genetic algorithm solution for that time interval.
Figure 7-8: This graph provides a family of curves exploring the tradeoff between time resolution and maximum TAT based on the weighted average of the TAT objective. As the time intervals are decreased (increased in time resolution), the optimization solution improves substantially across the cost dimension. This chart is very similar to Figure 7-4 but provides results that are monotonically decreasing in maximum TAT for increases to the number of maximum routes per day.
Figure 7-9: This graph presents the family of curves exploring the tradeoff between time resolution and cost based on the weighted average of the TAT objective. As time resolution increases, the objective improves by almost 30% across the cost dimension. This chart is very similar to Figure 7-5 but provides results that are monotonically decreasing in average TAT for increases to the number of maximum routes per day.
Figure 7-10: This figure shows a collection of plots that illustrate the improvement in maximum LTAT across a given time interval by using the local search algorithm. Each plot illustrates the improvement in the objective by using the local search algorithm seeded with the genetic algorithm solution for that time interval. This plot is similar to Figure 7-7 but shows the improvement based on the weighted average objective.
7.5 Pilot Test of the Optimization Model

Since the current client and PSC network served by Quest Cambridge is large, we down-selected the set of logistic locations in order to investigate, implement, and test the optimization model. The objective was to change a small geographic area to improve client service without impacting outside routes. However, this was difficult because many client locations and PSCs are serviced by multiple routes per day, and we had clearance to change the morning route. Nevertheless, we selected a small geographic area located in Brighton, MA satisfying the following criteria:

- the geographic area contains routes that run everyday
- the area contains client locations and patient service centers
- the area is close in proximity to the depot
- the three types of service levels are represented in the geographic area (routine, same-day, STAT)

The route and associated service area selected contains three PSCs and nine client locations. Table 7.3 summarizes the route information - order of stop in route, and scheduled pickup time - for each stop in the geographic area. This area will be referred to as Brighton in the following sections. The optimized schedule was determined using the SBLO outlined previously. However, additional constraints were added to the optimization model to ensure: only one route type was selected, and the second and third route start times did not change from the baseline schedule. The SBLO framework produced the "optimal" schedule shown in the Optimized column of Table 7.3. The Proposed column of Table 7.3 presents the schedule as it was run during the pilot test. This was called "proposed" because the sales representative, logistics director, route manager, and all the clients needed to approve the schedule change for the route and only approved the morning route change. Table 7.3 only shows the morning route information, but the afternoon and night routes are similar in sequence and in drive times.
Table 7.3: Outlines the Brighton route schedule under three scenarios: original, optimized, and proposed. The original schedule is how the route is currently scheduled. The optimized schedule was determined using the optimization methodology outline previously. The proposed schedule was run during the pilot test of the new optimization methodology. This was called "proposed" because the sales representative, the logistics director, the route manager, and all the clients needed to approve the changes.

<table>
<thead>
<tr>
<th>Number</th>
<th>Type</th>
<th>Order</th>
<th>Time</th>
<th>Order</th>
<th>Time</th>
<th>Order</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop 0</td>
<td>Depot</td>
<td>0</td>
<td>10:45</td>
<td>0</td>
<td>11:15</td>
<td>0</td>
<td>11:15</td>
</tr>
<tr>
<td>Stop 1</td>
<td>Client</td>
<td>1</td>
<td>10:55</td>
<td>1</td>
<td>11:30</td>
<td>1</td>
<td>11:32</td>
</tr>
<tr>
<td>Stop 2</td>
<td>Client</td>
<td>2</td>
<td>11:05</td>
<td>7</td>
<td>12:20</td>
<td>8</td>
<td>12:20</td>
</tr>
<tr>
<td>Stop 3</td>
<td>Client</td>
<td>3</td>
<td>11:05</td>
<td>8</td>
<td>12:22</td>
<td>9</td>
<td>12:22</td>
</tr>
<tr>
<td>Stop 4</td>
<td>PSC</td>
<td>4</td>
<td>11:10</td>
<td>9</td>
<td>12:24</td>
<td>7</td>
<td>12:13</td>
</tr>
<tr>
<td>Stop 5</td>
<td>Client</td>
<td>5</td>
<td>11:20</td>
<td>2</td>
<td>11:39</td>
<td>2</td>
<td>11:40</td>
</tr>
<tr>
<td>Stop 6</td>
<td>Client</td>
<td>6</td>
<td>11:35</td>
<td>4</td>
<td>12:09</td>
<td>4</td>
<td>12:08</td>
</tr>
<tr>
<td>Stop 7</td>
<td>Client</td>
<td>7</td>
<td>11:35</td>
<td>5</td>
<td>12:11</td>
<td>5</td>
<td>12:10</td>
</tr>
<tr>
<td>Stop 8</td>
<td>Client</td>
<td>8</td>
<td>11:35</td>
<td>6</td>
<td>12:13</td>
<td>6</td>
<td>12:11</td>
</tr>
<tr>
<td>Stop 9</td>
<td>PSC</td>
<td>9</td>
<td>11:40</td>
<td>10</td>
<td>12:15</td>
<td>10</td>
<td>12:13</td>
</tr>
<tr>
<td>Stop 10</td>
<td>Client</td>
<td>10</td>
<td>11:50</td>
<td>Removed</td>
<td>Removed</td>
<td>Removed</td>
<td>Removed</td>
</tr>
<tr>
<td>Stop 11</td>
<td>PSC</td>
<td>11</td>
<td>11:55</td>
<td>3</td>
<td>11:59</td>
<td>3</td>
<td>11:58</td>
</tr>
<tr>
<td>Stop 12</td>
<td>Client</td>
<td>12</td>
<td>12:05</td>
<td>11</td>
<td>12:37</td>
<td>11</td>
<td>12:24</td>
</tr>
<tr>
<td>Stop 13</td>
<td>Depot</td>
<td>13</td>
<td>12:15</td>
<td>12</td>
<td>12:50</td>
<td>12</td>
<td>12:47</td>
</tr>
</tbody>
</table>
7.6 Pilot Route Optimization

For each stop on the Brighton route, we pulled the data from the spatio-temporal model discussed in Chapter 6. The volume of accessions (demand) for each of the stops is listed in Table 7.4. Since Stop 10 had less than one accession per day on the route, it was successfully removed from the optimization model and the pilot test. If this unproductive stop is removed from daily service, while volume stays low, Quest will enjoy substantial cost savings on the route over time. Unfortunately, two other stops on this route also have low volume numbers, but remain on the route due to client obligations and the inability of the sales team to convince the client to remove the scheduled stop. This has detrimental effects to the operation and productivity of a route, and causes substantial downstream costs for Quest.

The demands were input into excel and the service based optimization model was run on the 12 Stops. Table 7.3 lists the new ordering and scheduled pickup times that result in a minimized global turn-around-time.

7.7 Estimated Model Performance

Most of the stops listed in Table 7.3 receive three stops per day. The optimization model was run assuming the same route was performed in the same order each tour, and the drive times and service times are identical throughout the day. The optimization model estimates significant gains in reducing logistics TAT and improving the number of specimens delivered to the first and second shift. The optimization model also computes a bound on the maximum aggregate LTAT and computes this metric for each stop and then finds the maximum aggregate LTAT on the route. Table 7.6 illustrates the metrics used to compare each route structure.

The optimization model was run multiple times, parameterizing over the max number of routes constraint. For the Brighton area, we explored one to three pickups per node per day. Based on representative costs from Quest, a cost model was calculated for each route based on the route length, time on route, and number of pickups per node per day. This provides an efficient frontier, or pareto front diagram, showing the relationship between cost
Table 7.4: The spatio-temporal network demand model for the stops on the Brighton, MA route. Each column represents the average number of accessions collected during an hour in a day of the week.

| Stop   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|--------|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Stop 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 2 | 0 | 1 | 1 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 11| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 3 | 4 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 7 | 8 | 8 | 5 | 4 | 4 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 5 | 6 | 6 | 4 | 2 | 4 | 4 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stop 12| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
and service as discussed earlier in Chapter 3 and illustrated in Figure 3-1. The inputs to the cost model are show in Table 7.5 and the service versus cost trade space is illustrated in Figure 7-11. Figure 7-11 illustrates that logistics TAT can be improved by 48% by adding an additional pickup per day, and another 18% for 3 stops per day. The brighton route currently has three stops per day, but the scheduled times are different than those based on the optimization model. If we manually put in the currently scheduled times in the optimization model, which pulls demand information, the resultant LTATs worsen, a move upward on the efficient frontier mapped out in Figure 7-11. Intuitively, the optimization model provides the logistics manager with a way to make a informed decision about when to schedule a stop to pickup specimens and an associated service level based on the number of stops per day. This is balanced against the cost of providing that service.

Table 7.5: Provides contrived cost model data for calculating the route costs. These parameters were used in the optimization model and show in the calculateCosts function in the appendix.

<table>
<thead>
<tr>
<th>Cost Model Inputs</th>
<th>Parameter Description</th>
<th>Parameter</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per truck per year</td>
<td>50,000</td>
<td>USD</td>
<td></td>
</tr>
<tr>
<td>Business Days per year</td>
<td>251</td>
<td>Days</td>
<td></td>
</tr>
<tr>
<td>RSR SWB per hour</td>
<td>20</td>
<td>USD</td>
<td></td>
</tr>
<tr>
<td>Variable vehicle costs per mile</td>
<td>0.5</td>
<td>USD</td>
<td></td>
</tr>
<tr>
<td>Time on route</td>
<td>1.5</td>
<td>Decimal hours</td>
<td></td>
</tr>
<tr>
<td>Route distance (miles)</td>
<td>10</td>
<td>Miles</td>
<td></td>
</tr>
</tbody>
</table>

Since many stakeholders needed to sign off on changing the route schedule, the proposal for the pilot test was modified from the optimal schedule. The model estimates a logistics TAT improvement of 28% with the optimal scenario and approximately 15% with the proposed scenario. Additionally, the morning route will produce 14% more specimens on the early route, providing lab operations with 14% more specimens sooner on this route.
Figure 7-11: Illustrates the tradeoff between cost and service based on outputs from the new optimization model. As indicated earlier, adding additional pickups per day to the Brighton route lowers logistics TAT significantly. After two pickups per day, each additional pickup will only give a marginal improvement to logistics TAT.

Table 7.6: Illustrates the optimization model’s output given three different scenarios: baseline, optimal, and proposed. The baseline is how the route is currently scheduled and run. The optimal is the best the routes can perform with three stops, and the proposed is the schedule for the pilot test. KPIs are recorded for the morning route, because the pilot test altered the morning route. Additionally, the table shows how the global TAT is minimized across the three scenarios.

<table>
<thead>
<tr>
<th>KPIs</th>
<th>Brighton - Morning Route</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>Route 1</td>
<td></td>
</tr>
<tr>
<td>Average TAT (Hours)</td>
<td>3.23</td>
</tr>
<tr>
<td>Max - TAT (Hours)</td>
<td>3.97</td>
</tr>
<tr>
<td>Avg Demand (Reqs)</td>
<td>75</td>
</tr>
<tr>
<td>All Routes</td>
<td></td>
</tr>
<tr>
<td>Weighted Avg (Hours)</td>
<td>3.38</td>
</tr>
<tr>
<td>Max TAT (Hours)</td>
<td>5.36</td>
</tr>
<tr>
<td>Weighted Flow (Reqs)</td>
<td>72</td>
</tr>
</tbody>
</table>
7.8 Actual Model Performance

Quest does not have a system that can track and audit a specimen throughout collection, logistics, and lab operations. Therefore, we used the collection to release statistic to measure the end-to-end performance of operations, and to measure the performance of the optimization model. To compare the volume of accessions for each stop on the route, we used Megasys counts which are data input by route service representatives when transporting specimens to the performing site. This is an error prone measurement but the closest proxy to accession volume from logistics. Table 7.7 shows the accession volumes for the Brighton morning route in November 2013 compared to the optimized route which ran from December 9th to the 20th, 2013. Since the variance and number of samples for the two routes are likely unequal, a t-test statistic was used to test the null hypothesis that the means are the same. The data suggests that the null hypothesis can be rejected at less than 1% significance level. This hypothesis test was done mechanically only to show that the by changing the route, a different number of accessions will be collected. The fact that the route was scheduled later in time almost certainly guarantees more accessions will be collected by that route.

Table 7.7: Megasys accession volume counts for Brighton route before and during the pilot test.

<table>
<thead>
<tr>
<th>Category</th>
<th>metric</th>
<th>Current Nov</th>
<th>Pilot 5-day</th>
<th>Pilot 8-day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route</td>
<td>average standard dev.</td>
<td>68.9</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.5</td>
<td>11</td>
<td>16.2</td>
</tr>
<tr>
<td>Stops</td>
<td>Reqs/stop avg.</td>
<td>5.7</td>
<td>9.1</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>PSC avg. (req)</td>
<td>17.4</td>
<td>25.3</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>T-test p-value</td>
<td>0.0016</td>
<td>0.0073</td>
<td></td>
</tr>
</tbody>
</table>

7.9 Backend Service Level Comparisons

Table 7.8 compares the end-to-end specimen lifecycle for all stops serviced on the Brighton morning route before and during the pilot test. Stop 10 was removed from the morning route and placed on another route that arrived at the stop, two hours later. Consequently,
all specimen TATs increased by an average of two hours for Stop 10. Nevertheless, all other stops enjoyed an increase in service of at least 10.9%. All specimens picked up from the Brighton route were considered routine because there was no way to use QLS to definitively determine service levels. However, Quest labels all of the specimens that come from the PSCs on the Brighton route as same-day, and thus these specimens will go through a different prioritization at the performing site which could speed up their TATs.

Table 7.8: QLS backend operations comparison for the specimen lifecycle of all the specimens picked up on the Brighton, MA route. Stop 10 was remove from the Brighton route but was added to another route, pending removal due to a low volume of requisitions per day.

<table>
<thead>
<tr>
<th>Stop Number</th>
<th>Avg Collection To Release Gain (Loss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop 1</td>
<td>21.70%</td>
</tr>
<tr>
<td>Stop 2</td>
<td>26.70%</td>
</tr>
<tr>
<td>Stop 3</td>
<td>12.10%</td>
</tr>
<tr>
<td>Stop 4</td>
<td>20.60%</td>
</tr>
<tr>
<td>Stop 5</td>
<td>10.90%</td>
</tr>
<tr>
<td>Stop 6</td>
<td>22.40%</td>
</tr>
<tr>
<td>Stop 7</td>
<td>19.00%</td>
</tr>
<tr>
<td>Stop 8</td>
<td>20.40%</td>
</tr>
<tr>
<td>Stop 9</td>
<td>21.70%</td>
</tr>
<tr>
<td>Stop 10</td>
<td>-30.10%</td>
</tr>
<tr>
<td>Stop 11</td>
<td>13.40%</td>
</tr>
<tr>
<td>Stop 12</td>
<td>49.70%</td>
</tr>
</tbody>
</table>

Lastly, we compared individual tests on the Brighton routes because there are quick-turn tests that need to be resulted quickly and we wanted to verify that these tests were not negatively affected by the new schedule. Table 7.9 shows the improvement to the end-to-end TAT for certain tests that are based on specimens collected in locations served by the Brighton route. We observed enough requisitions on these routes to perform statistically relevant t-tests except for Lipid Panel tests collected at client locations. For all of these tests, service levels improved substantially.
Table 7.9: Lists the performance gains (losses) for quick-turn test codes run on specimens picked up on the Brighton route during the pilot test from Dec. 9th - Dec. 20th. The metric used for comparison was average collection to build and average collection to release. The table looks at specimens that were collected at client locations and at PSCs on the Brighton Route. The number of samples for each test code is based on number of accessions listed.

<table>
<thead>
<tr>
<th>Test</th>
<th>Clients</th>
<th></th>
<th>PSCs</th>
<th></th>
<th>Accessions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg Coll - Build</td>
<td>Avg Coll-Rel</td>
<td>Reqs</td>
<td>Avg Coll-Build</td>
<td>Avg Coll-Rel</td>
</tr>
<tr>
<td>Complete Blood Count (6399)</td>
<td>19.70%</td>
<td>16.20%</td>
<td>116</td>
<td>16.80%</td>
<td>20.40%</td>
</tr>
<tr>
<td>Prothrombin (8847)</td>
<td>16.60%</td>
<td>15.30%</td>
<td>31</td>
<td>7.30%</td>
<td>10.90%</td>
</tr>
<tr>
<td>Comprehensive Metabolic Panel (10231)</td>
<td>21.60%</td>
<td>18.90%</td>
<td>112</td>
<td>19.90%</td>
<td>25.70%</td>
</tr>
<tr>
<td>Lipid Panel (7600)</td>
<td>8.90%</td>
<td>10.60%</td>
<td>13</td>
<td>20.50%</td>
<td>19.40%</td>
</tr>
</tbody>
</table>
7.10 Discussion

The pilot test of the Brighton route reduced costs, improved service, and increased RSR efficiency. Because approximately 30% more specimens were obtained on average four hours sooner, these specimens were processed on the 1st shift where labor is 15% cheaper than the 3rd shift. This means that the labor costs to sort, accession, and place on an analyzer are reduced by 15%. By incorporating the demand model, we were able to statistically determine which stops were productive and which should be removed. This saves labor, increases productivity, and allows the RSR to become more efficient. Additionally, we improved service for all clients on all test on the Brighton route by improving end-to-end turn-around-time by an average of 25%. This was achieved by maximizing service versus minimizing vehicle routing costs.

In addition to the immediate operational gains, the change in service levels has long-term, strategic implications for a service-based organization. First, by incorporating demand, tracking service metrics, and optimizing scheduling, the logistics team can provide a bounds on service levels (logistics turn-around-times in this context) to the sales team, which can better inform clients, managers, and other operations. For the Brighton route, the model predicted a maximum LTAT of 8 hours. By having two stops, that maximum LTAT is reduced to just over 4 hours; a significant improvement in service. This LTAT service level metric provides a clear indicator to logistics managers and provides a upper bound for clients in that geographic location. Moreover, this service level metric can be used to incentivize managers in a logistics organization, and to provide business decisions about what level of service to offer versus the cost of providing that service. For example, this optimization methodology allowed logistic managers to make a quantitative tradeoff between cost of service and the service level provided to clients on a particular route.

Logistical improvements, however, are not the only considerations when optimizing the end-to-end specimen lifecycle. The scheduling of the Brighton route impacted RSRs, clients, managers, and specimen processors. Since the optimal solution was to increase the start time of the morning route by approximately one hour, the RSR’s personal schedule was affected and this caused some initial doubt to the reason for changing the scheduling. If the RSR
isn’t fully informed of their impact on downstream operations, then the route will continue to run inefficiently. Additionally, the sales professional assigned to the clients affected by the route was initially against any change related to the route for reasons discussed in Chapter 3. In order to successfully implement a schedule change, the sales representative and logistic managers must seek consensus. After logistics, specimen processing obtained these specimens sooner, but in some instances, the specimens remained in a holding bin until a specimen processor became free. This was due to the prioritizations of clients and the work that was first in the specimen processing queue.

The SBLO methodology demonstrated that maximizing service begins by having the proper objective. Next, scheduling routes according to this new objective improves service and increase efficiencies. However, some specimens just sat in specimen processing despite getting in sooner. The end goal is overall specimen TAT. Therefore, in order to fully optimize the end-to-end lead time of the specimen lifecycle, each step in the process needs to be optimized. Furthermore, stakeholders in each functional unit of the service based organization need to be in consensus and agree to the necessity for end-to-end optimization.
Chapter 8

Conclusion

8.1 Summary

Increased competition, lower medicare reimbursements, higher costs, and increased customer constraints have all contributed to low profit margins seen in companies in the clinical laboratory market. Since service is the most important distinguishing characteristic between Quest and its competitors, a more aligned service operation is required. Turn-around-time is one of the service metrics used across the organization and communicated to clients. In order to improve TATs across the organization, Quest needs to optimize the entire operation, from logistics to laboratory operations, instead of designing, managing, and optimizing each functional unit separately.

In Chapter 7 we demonstrated how by changing Quest’s current logistic’s objective from lowest cost routes to minimizing the maximum logistics turn-around-time, Quest can offer much better service to clients. Essentially this new objective seeks to globally minimize the aggregate time specimens spend in the logistics network and waiting for pickup. Technically, we defined logistics turn-around-time as the relative difference between a specimen’s collection time to when that specimen arrives to the performing site. We demonstrated this empirically with the service base logistics optimization methodology, and use the spatio-temporal network model developed in Chapter 6 to find where and when specimens are available. This new optimization methodology tunes the routes serving a geographic area, to arrive when the specimens are ready for pickup. This is fundamentally different that the
current method of scheduling, which is essentially finding the lowest cost routes that solve the vehicle routing problem with time windows. This was discussed in Chapters 3 and 4.

By using the SBLO, more specimens arrive at the performing site sooner, which efficiently utilizes Quest’s current machine and labor capacity. This decreases labor costs per accession due to shift differential and specimen processor productivity. This lowers the cost of operations and analyzers can be stocked sooner with efficient batch sizes. This also results in a decrease in aggregate specimen TAT for clients on the affected routes and provides management with a key performance indicator of service levels per client and per geographic region. Moreover, we demonstrated how by incorporating demand into the optimization strategy, Quest is more informed of the volume of accessions coming from individual logistic stops, and can report and manage the logistics turn-around-time service level metric during and after the move to Marlborough, MA. This will help mitigate service interruptions and will provide a qualitative metric to management.

8.2 Recommendations

Based on the analyses performed throughout the project and thesis, we recommend the following actions to be taken by Quest:

1. Organizationally combine Logistics and Lab Operations in Marlborough to create a unified strategic vision, associated metrics and KPIs, and to achieve the operational excellence goals set out in Quest’s global strategy

2. Better align the value proposition of Quest to clients by changing the primary logistics optimization objective from minimizing logistics costs to maximizing service levels (minimizing logistics TAT or reducing lead time for test results)

3. Change the misalignment between the sales teams and lab operations by informing the sales organization of the current capabilities of operations at Quest and changing the incentive structure

4. Increase IT infrastructures to clients, PSCs, and principal laboratories to completely remove unproductive stops from the logistics network
5. Overhaul internal PSC processes by creating single piece specimen flow, removing process bottlenecks, removing/modify manifesting, and improving the communication between phlebotomists and route service representatives.

6. Standardize internal layouts and processes of PSCs to improve patient wait times and to reduce specimen collection to logistic readiness lead time

7. Incentivize productivity improvements and lead time reductions over strict Full-Time-Equivalent (FTE) reduction and cost cutting across the organization

8. Investigate and analyze the impact of holdovers, one-offs, and client specific service constraints in logistics, specimen processing, and laboratory operations

9. Investigate and analyze the cost of serving certain clients (logistics and laboratory operations) with the revenue those clients generate

10. Investigate removing the clotting process from PSC/client internal processes by using plasma validated tests (green-top serum separation tubes); plasma validate tests account for 22% of volume in New England.

8.3 Next Steps

Although the actual performance of service based optimization model was proven for a small geographic region in Brighton, MA, Quest should investigate this methodology across the entire logistics network in New England. This will require a significant effort to create a general model that can be applied across the organization and integrated into the logistics, lab operations, and client services departments. Next, Quest needs to invest in distinguishing service between levels in its QLS database in order to uniformly integrate this information into the logistics network. The optimization and demand model will need to be refined to take into account service levels. This could potentially change the optimization strategies of some of the locations if the distance from the performing site is very large. However, the service based approach helps to bound the maximum logistics time on a route such that a specimen that is collected at any time of the day could be resulted within a time period
indicated by the logistics TAT and laboratory TAT. However, this is also a sizable effort as there is inconsistent use and definition of the service levels across the organization. Next, additional data is required in order to audit an accession number all the way to the beginning of the specimen lifecycle. This will help in providing service level information about each test and every client in the network.

8.4 Final Thoughts

Quest Diagnostics is an amazing clinical laboratory operator, offering a comprehensive testing catalog to many healthcare professional across the U.S. and abroad. Quest inherently is a health services organization and seeks to do what is right for the client and patient at every step in the value chain. This thesis married demand, service levels, and optimization to help improve a very complex operation. We observed substantial improvements to service in a small geographic area by using the service based logistics optimization methodology rather that minimal-cost optimization. This tool can continue to help improve operations at Quest, as it begins to further its growth and dominance in the clinical laboratory market. However, every solution that is output from a model should be validated, tested, and well-understood before a complete implementation. This thesis demonstrated that myopic cost optimization in separate departments can have detrimental impacts to customer service, and that by changing a simple logistics objective, an organization can optimize its end-to-end operations.
.1 Appendix I

The following Matlab code runs the local search algorithm previously discussed.

```matlab
function \([x,fval] = runLocalSearch(X0, data, configObj, opts)\)

\(x = [];\)
\(fval = 0;\)

TolFun = opts.TolFun;
maxIters = opts.maxIters;

curTol = inf();
notDone = 0;
curSol = X0;

% add remove var
nVars = size(X0,2);  % dimensionality of the design vector
nSelected = sum(X0);  % number variables that are 1
nAdjacent = nVars-nSelected;  % number of adjacencies
curBest = X0;  % current best solution
iters = 0;  % number of iterations
xHist = {};  % keep a history of design vectors
fHist = [];  % keep a history objective vlues
J.min = sbo.fitness.function.discretized(curBest, data, configObj);
xHist(end+1) = {curBest};
fHist(end+1) = J.min;
precisionZero = eps('single');

while (curTol > TolFun && iters < maxIters && ~notDone)
    iters = iters + 1;
    posIdxs = find(curBest>0);
    zeroIdxs = find(curBest==0);
    % create space for all adjacent search directions
    J.test = zeros(nSelected, nVars-nSelected);
    V.constr = zeros(nSelected, nVars-nSelected);
```
X.test = cell(nSelected, nVars-nSelected);

% search all adjacent dimensions
for i_selected = 1:nSelected
    removedIdx = posIdxs(i.selected);
    for i.nvars = 1:nAdjacent
        curIdx = zeroIdxs(i.nvars);
        X = curBest;
        X(removedIdx) = 0;
        X(curIdx) = 1; % add a point

        % check to see if we have already done this one
        [found, fidx] = findDesign(xHist,X);
        if ~found
            % run new point
            J.test(i.selected, i.nvars) = ...;
            sbo.fitness.function.discretized(X, data, configObj);
            X.test(i.selected, i.nvars) = {X};
            c = sbo.nonlcon.discretized(X, data, configObj);
            if (any (c>0))
                % violates constraints set to inf
                V.constr(i.selected, i.nvars) = 1;
            end
        else
            J.test(i.selected, i.nvars) = fHist(fidx);
            X.test(i.selected, i.nvars) = xHist(fidx);
        end
    end
end

% find best solution that passed constraints
% set all the violated constraints to inf() [could do a barrier search]
J.test(logical(V.constr)) = inf();
[min.val, min.idx] = min(J.test(:));
[row, col] = ind2sub(size(J.test),min.idx);
J.local = J.test(row,col);
X.local = X.test{row,col};
% if best solutions is better store it and redo search
del = J.min - J.local;
if del > precisionZero && del > TolFun
    xHist(end+1) = {X.local};
fHist(end+1) = J.local;
curTol = abs(J.min - J.local)/J.min;
    J.min = J.local;
    curBest = X.local;
    continue;
else
    % couldn't improve local solution given constraints
    notDone = 1;
end
end

if isequal(curBest,curSol)
    x = curSol;
else
    x = curBest;
end
fval = sbo_fitness_function_discretized(x,data,configObj);
end

The following Matlab code runs the logistics route simulation and returns parameters that will be fed into the optimization engine.

function [ maxTAT, avgTAT, pickedUpDemands, totDemand, marginal_costs, ...
    avgCostPerSpecimen, routePlan, startTimes, nVisitsPerNode ] = ...
    runLogisticsPlan.discretized( X, data, configObj )
pickedUpDemands = [];
const service.time.min = configObj.service.time.min;
logistics.readiness.offset.min = configObj.logistics.readiness.offset.min;
hours.in.day = 0:1:23;

node.idxs = data.node.idxs;
dMat = data.distance.matrix;
tMat = data.time.matrix;
demand = data.demand;
routes = data.routes;
numStopsOnRoutes = data.num.stops.on.routes;
totRouteRunningTimes.min = data.route.run.times.min;
nRoutes = size(routes,2);
nNodes = numel(node.idxs);
route.idxs = 1:1:nRoutes;

% unpack structure
time.interval = configObj.time.interval;
nIntervals = 24*60/time.interval;
route.indicators = reshape(X,[nIntervals nRoutes]);
a = 0:time.interval/60:(24-(time.interval/60));
A = repmat(a,nRoutes,1)';
start.times = A.*route.indicators;

% get the route indexes to run
route.indicators = logical(route.indicators);
pos.route.idxs = sum(route.indicators,1) > 0;
routes.to.run = route.idxs(pos.route.idxs);
start.times.per.route.rs = start.times;
start.times.per.route.rs(~route.indicators) = nan();
start.times.per.route.rs = start.times.per.route.rs(:,pos.route.idxs);

% get time to start each route
route.starts = start.times.per.route.rs;
% build route plan
routeRunIdxs = repmat(routes_to_run,[nIntervals 1]);
[sortedStartTimes, sortIdxs] = sort(route_starts(:,1),1);
notNan = ~isnan(sortedStartTimes);
sortedStartTimes = sortedStartTimes(notNan);
sortIdxs = sortIdxs(notNan);
sortedEndTimes = sortedStartTimes + ...
    totRouteRunningTimes_min(routeRunIdxs(sortIdxs))/60;
sortedRouteIdxs = routeRunIdxs(sortIdxs);

% get time to each node for each route
relativeTimeToNodes = {}; demandIdxs = {}; nodeTimes = {}; nodeVisitTimes = cell(nNodes,1);
nodeVisitsRouteIdx = cell(nNodes,1);
routeInfo = cell(nNodes,1);
for i-route = 1:numel(sortedStartTimes)
    routeIdx = sortedRouteIdxs(i_route);
    route = routes(:,routeIdx);
    routeStartTime = sortedStartTimes(i_route);
    routeEndTime = sortedEndTimes(i_route);
    numStopsOnRoute = numStopsOnRoutes(routeIdx);
    relativeTimeToNodes{i_route} = zeros(numStopsOnRoute-1,1);
    nodeTimes{i-route} = zeros(numStopsOnRoute,1);
    nodeTimes{i-route}(1) = routeStartTime;
    for i-stop = 2:numStopsOnRoute
        curNode = route(i.stop);
        relativeTimeToNodes{i.route}(i.stop-1) = ...
            tMat(route(i.stop-1)+1,route(i.stop)+1)+service.time.min;
        nodeTimes{i.route}(i.stop) = ...
            nodeTimes{i.route}(i.stop-1)+...
            (relativeTimeToNodes{i.route}(i.stop-1))/60;
        if curNode
            nodeVisitTimes{curNode}{end+1} = nodeTimes{i.route}(i.stop);
nodeVisitsRouteIdx{curNode}{end+1} = routeIdx;
routeInfo{curNode}{end+1} = routeEndTime;
end
end
end

% allocated specimens based on FIFO
node.tats = cell(nNodes,1);
node.demands = cell(nNodes,1);
total.specimens.collected = 0;
node.tats = cell(nNodes,1);
nVisitsPerNode = zeros(nNodes,1);
for i-node = 1:nNodes
    visitTimes = cell2mat(nodeVisitTimes{Kinode});
nVisits = size(visitTimes,2);
    if nVisits == 0
        node.tats{i.node} = [];
        node.demands{i.node} = [];
    else
        nVisitsPerNode(i.node) = nVisits;
        routeIdxs = cell2mat(nodeVisitsRouteIdx{Kinode});
        routeDepotTimes = cell2mat(routeInfo{Kinode});
        endIdxs = round(visitTimes-(logistics.readiness.offset.min/60));
        lessThanZero = endIdxs < 0;
        endIdxs(lessThanZero) = 0;
        startIdxs = [0 endIdxs(1:end-1)+1];
        sameTimeSlotIdxs = startIdxs >= endIdxs;
        if any(lessThanZero)
            asdf = 10;
        end
        demandsPickedUp = arrayfun(@(startIdx,endIdx)...
sum(demand(i.node, startIdx:endIdx)), startIdxs+1, endIdxs+1);

if ~sum(demandsPickedUp)
    TAT = nan();
else
    TAT = arrayfun(@(startIdx,endIdx,idx)...
        sum((routeDepotTimes(idx)-...
        hours.in.day(startIdx:endIdx)).* ...
        demand(i.node,startIdx:endIdx))/...
        sum(demand(i.node,startIdx:endIdx)),...
        startIdxs+1, endIdxs+1,1:nVisits);
end

if ~isempty(sameTimeSlotIdxs)
    demandsPickedUp(sameTimeSlotIdxs) = 0; % zero out race condition
end

zeroDemand = demandsPickedUp == 0;
if ~isempty(zeroDemand)
    TAT(zeroDemand) = nan();
end

node.tats{i.node} = TAT;
node.demands{i.node} = demandsPickedUp;
total.specimens.collected = ...
    total.specimens.collected + sum(demandsPickedUp);
end
end

[costsPerYear, avgCostPerSpecimen, ~] = ...
    calculateCosts(routes, sortedRouteIdxs, nodeTimes, tMat, dMat,...
    total.specimens.collected);

marginal.costs = costsPerYear;
totDemand = total.specimens.collected;
maxTAT = zeros(nNodes,1);
avgTAT = zeros(nNodes,1);
pickedUpDemands = zeros(nNodes,1);
for i_node = 1:nNodes

tats = node_tats{i_node};
demands = node_demands{i_node};

if isempty(tats)
    maxTAT(i_node) = inf();
    avgTAT(i_node) = inf();
else

    metrics = zeros(2,size(tats,2));
    metrics(1,:) = tats;
    metrics(2,:) = demands;
    nnanIdxs = ~isnan(metrics(1,:));

    maxTAT(i_node) = max(maxTAT(i_node), max(metrics(1,:)));
    avgTAT(i_node) = sum(metrics(1,nnanIdxs).*...
        metrics(2,nnanIdxs))/sum(metrics(2,nnanIdxs));
    pickedUpDemands(i_node) = sum(demands);
end
end

routePlan = sortedRouteIdxs;
startTimes = sortedStartTimes;
end

The following Matlab code calculates the costs for a route plan output from runLogisticsPlan_discretized.m listed above.

function [costsPerYear, avgCostPerSpecimen] = ...
    calculateCosts(routes, routePlan, node_times, tMat, dMat, ...
numSpecimensCollected

businessDaysPerYear = 251;
RSR_SWB_per_hour = 20;
vehicle_costs_per_mile = 0.5;
km_to_mile = 0.621371;

total_distance = zeros(numel(routePlan),1);
total_time = zeros(numel(routePlan),1);
for i_route = 1:numel(routePlan)
    route = routes(:,routePlan(i_route));
    numStopsOnRoute = sum(~isnan(route));
    
    for i_stop = 2:numStopsOnRoute
        total_distance(i_route) = ...
        total_distance(i_route) + ...
        dMat(route(i_stop-1)+1,route(i_stop)+1);
        total_time(i_route) = ...
        total_time(i_route) + tMat(route(i_stop-1)+1,route(i_stop)+1);
    end
end

mileage_costs = sum(total_distance)*km_to_mile*vehicle_costs_per_mile;
time_costs = (sum(total_time)/60)*RSR_SWB_per_hour;

costsPerYear = (mileage_costs+time_costs)*businessDaysPerYear;

if numSpecimensCollected == 0
    avgCostPerSpecimen = inf();
else
    avgCostPerSpecimen = (mileage_costs+time_costs)/numSpecimensCollected;
end

end
Bibliography


