

© 2015 by Xi Chen.

INTEGRATED PLANNING OF MULTI-TYPE RAILROAD SERVICE
FACILITIES UNDER LOCATION, ROUTING AND INVENTORY
CONSIDERATIONS

BY

XI CHEN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Adviser:

Associate Professor Yanfeng Ouyang

Abstract

In the North America railroad network, thousands of locomotives are running everyday and they are in need of receiving various types of mechanical work that can be random, periodic, or on-demand. Each type of work can be performed either at fixed facilities or via movable facilities. It is an important and difficult problem to make planning of the whole railroad system, including facilities locations, locomotives assignment, fuel inventory strategy, and movable facility routing trips.

This paper formulates a large-scale linear mixed-integer mathematical model for the railroad system that integrates decisions about (i) locations, capacities, and types of facilities, (ii) assignments of locomotive work to facilities, (iii) fuel strategy of trains, and (iv) routes of movable facilities. We also propose a framework consisting of several solution algorithms, including network consolidation, network generation, reduce candidate location number, free move, and trip merge. The solution framework is designed to solve large-scale problem efficiently. Empirical case studies using field data show that the proposed model and algorithms are capable of providing near-optimum solutions effectively.

Acknowledgments

It was wonderful experience participating in the project as well as writing thesis. I would like to acknowledge the following individuals who encouraged, inspired, supported and assisted me in my pursuit of the master of science degree in full gratitude.

Great thanks to my advisor Yanfeng Ouyang for giving me the opportunity to join the most ever wonderful program, recommending to me great projects, and working with me to overcome difficulties both in study and research.

Special thanks to Siyang Xie and Zhaodong Wang for supporting me with helpful ideas and algorithms and accompanying me to make study and life in UIUC more colorful.

Sincere thanks to Kamalesh Somani and Jing Huang from CSX. Corp for working together on the project, helping me learn more advanced techniques and providing funding for the project.

Last thanks to my family for supporting me emotionally and mentally during my two-year study far from my hometown.

Contents

List of Tables	vi
List of Figures	vii
List of Abbreviations	viii
Chapter 1 Introduction and Literature Review	1
Chapter 2 Methodology Overview	6
2.1 Assumptions and Rules	6
2.1.1 Demand generation assumptions	6
2.1.2 Demand assignment assumptions	6
2.1.3 Fueling rule assumptions	7
2.1.4 Truck assumptions	7
2.2 Methodology Framework	8
2.3 Model Decomposition Framework	9
Chapter 3 Location Inventory Subproblem	13
3.1 Objective	13
3.2 Constraints	14
3.2.1 Location and assignment	14
3.2.2 Truck	17
3.2.3 Number of truck trips estimation	18
3.2.4 Fueling strategy	20
3.3 Overview	21
3.4 Free Move	22
Chapter 4 Truck Routing Subproblem	23
4.1 Objective	23
4.2 Constraints	23
4.2.1 Serving other yards	24
4.2.2 Serving home location	26
4.3 Truck Trip Merge	26
Chapter 5 Data Processing Methods	28
5.1 Network Consolidation	28
5.2 Network Generation	29
5.3 Shortest Path Algorithm	30
5.4 Reduce Number of Candidate Locations	33
5.4.1 Gravity algorithm	33

5.4.2	P-Median scaling algorithm	34
5.4.3	Maximum throughput algorithm	35
Chapter 6	Case Study	36
6.1	Base Case	38
6.2	Shut Down LSCs	40
6.2.1	Shut down 1 LSC	40
6.2.2	Shut down 2 LSCs	41
6.2.3	Shut down 3 LSCs	43
6.3	Allow to Add Candidate Locations for LSCs	45
Chapter 7	Conclusions	47
Appendix A	Tables	48
Appendix B	Model Step 1: Notations and Variables	50
B.1	Set notation	50
B.2	Subscript notation	50
B.3	Parameter notation	50
B.4	Decision variable notation	51
Appendix C	Model Step 2: Notation and Variables	53
C.1	Parameters	53
C.2	Decision Variables	53
C.3	Extra User Constraints in step 1	54
Bibliography		56

List of Tables

2.1	Locomotive work type and facility type mapping. “Y” indicates that the corresponding type of facility is capable of conduct the corresponding type of work.	7
6.1	Facility cost/benefit	36
6.2	Work cost at different facilities	37
6.3	Work time (hr) at different facilities	37
6.4	Base result: LSC location and capacity	39
6.5	Base result: Shop location and capacity	39
6.6	Base Result: LTHOME locations, LT types and serving cities	40
6.7	Cost comparison between base case and current case (Shut down 1 LSC) . . .	42
6.8	Cost comparison between base case and current case (Shut down 2 LSCs) . .	43
6.9	Cost comparison between base case and current case (Shut down 3 LSCs) . .	45
6.10	Cost comparison between base case and current case	46
A.1	SHOP capability and work type mapping	48
A.2	Truck types, changed serving cities in the case of allowing to add candidates .	48
A.3	Locations and capacities of LSCs in the case of allowing to add candidates .	49

List of Figures

1.1	CSX railroad network map with various facilities	2
2.1	A framework of solution methods.	8
2.2	The layers for point-specific and path-specific demands.	10
2.3	A framework of model part, consisting step 1 and step 2.	11
3.1	A framework of constrains relations in step 1.	21
4.1	Merge truck trips into one trip	27
5.1	Create the network (Xie et al., 2014)	29
6.1	Location base result illustration	38
6.2	Location result illustration when 1 LSC is shut down	41
6.3	Location result illustration when 2 LSCs are shut down	42
6.4	Location result illustration when 3 LSCs are shut down	44
6.5	Location result when adding candidate locations for LSCs	46

List of Abbreviations

LSC	Locomotive service center
SHOP	Locomotive shop
MFS	Mainline fueling station
LTHOME	Locomotive truck home
LT	Locomotive truck
LST	Locomotive service truck
LFT	Locomotive fueling truck
CFLP	Capacitated facility location problem
VRP	Vehicle routing problem
MICP	Multi-period inventory control problem
LRP	Location routing problem
LRICP	Location routing with inventory control problem

Chapter 1

Introduction and Literature Review

Thousands of locomotives are running every day in the North America railroad network, and must receive various types of mechanical work that can be random (e.g., repair), periodic (e.g., routine maintenance and service), or on-demand (e.g., fueling). Each type of work can be performed either at fixed facilities (e.g., locomotive shops (SHOP), locomotive service centers (LSC), and mainline fueling stations (MFS)) or via movable facilities (e.g., locomotive service trucks (LST) and locomotive fuel trucks (LFT) deployed from locomotive truck homes (LTHOME), and Vendor).

Each of those facilities is characterized by the types of locomotives it can handle and the level of work it can conduct. For example, a LSC is able to perform routine service, maintenance, and provide locomotives with fuel. A SHOP can only provide locomotives with random repairs and routine maintenance. MFS is only constructed for fast fueling of pass-by locomotives. LST/LFT will stay at LTHOME most of the time and occasionally go out to offer routine service or fueling to locomotives at other yards if needed. Trucks are categorized as different types of LST or LFT (e.g., LST60/120/168, LFT60/120/168) based on the number of available work hours per week. Vendor is a backup virtual facility for emergency fueling if a locomotive accidentally runs out of fuel or there are no other normal fueling facilities on the train route. Figure 1.1 shows the current CSX railroad network with various facilities. These facilities are able to do different combinations of various work tasks for locomotives.

There are about 4000 locomotives in the CSX railroad network. They are the power engine for trains to perform daily tasks. Each locomotive can be road or nonroad type. Based on different types, locomotives have various attributes such as tank capacity, failure rate, service interval, fuel consumption, etc. For example, the burn rate is 4.5 gallons per



Figure 1.1: CSX railroad network map with various facilities

mile per locomotive for road locomotives and is twenty gallons per hour per locomotive for nonroad locomotives. There are two main types of train: normal trains and light trains. Normal trains are used for transporting goods in cars. A normal train may have twenty cars but only two locomotives. Based on the type of locomotive and practical operation properties of normal train, trains can be further divided into nonroad trains, road trains with high priority and road trains with low priority. Each type of normal train has different attributes like number of working locomotives, fueling constraints, etc. For example, a nonroad train is allowed to get fuel from LSC by being transported to other yards after arriving at its destination, but a road train is not allowed to be transported. Light trains are quite different from normal trains, they are used to transport locomotives instead of goods. If a locomotive runs out of fuel or is broken, it can only be transported by a light

train. This is called a light move.

The railroad network is not only multi-type in facilities, work, locomotives and trains, but also capacitated in some of them. For SHOP, LSC and MFS, capacity is defined as the number of locomotives that can be worked on at any given point of time based on space, manpower and productivity. Capacity increase is allowed in a step-wise function (e.g. 3-unit capacity increase if there is one-step increase or 6-unit capacity increase if there is two-step increase). The capacity of a specific type of LT is defined as the number of available working hours per week multiplied by the number of that type of LT. Besides that, LFT has tank capacity which indicates how much fuel it can carry each time. For locomotive, tank capacity varies with the type. For example, it can be 4000 gallons per road locomotive and 3250 gallons for nonroad locomotive.

Having the complexity of the railroad network in mind, our work on optimizing all types of related cost demonstrates its difficulty and value. There are several main modules of costs in the network, including facility cost, locomotive transportation cost, work cost and fuel material cost. Facility cost includes construction cost, shutdown benefit, capacity expansion cost, and capacity reduction benefit. To construct a new facility, there is a construction cost. Besides one time investment for construction, the existing facility has maintenance cost every year. When the existing facility expands capacity, there is capacity expansion cost, including labor cost and physical cost. On the other hand, there is a shutdown benefit when an existing facility is shut down and possible savings when a facility reduces its capacity. Locomotive running cost is calculated yearly to be comparable with facility cost. A locomotive will cost some money per hour to indicate life time reduction when it's running. Besides, it has a fixed cost every time there is a pick up or drop off along the route, indicating the delay effect caused by that. An extra locomotive cost, crew cost, and fuel cost is applied to light train movement expense. When a locomotive receives work at a facility, there is work cost, including service cost, repair cost, maintenance cost, and fuel fixed cost. Work costs vary according to different work type and facility capability. Fuel material cost is a relatively larger part than other cost introduced above because fuel consumption is huge for all locomotives per year. Locomotive needs to buy fuel along its route in case that it runs out of fuel so locomotive pays fuel material cost, which is decided by fueling amount and

fuel price at a specific location or facility. Those are all the costs needed to be optimized.

Due to its complexity, planning of these facilities to satisfy all work demand within the railroad network is currently done manually based on expert knowledge and experience. An integrated optimization model is strongly needed to enhance operational efficiency of the locomotives and trains and minimize total cost involved. Based on the introduction of different properties of facilities, the problem described generally involves three sets of fundamental problems for logistics systems design: multi-type capacitated facility location problem (CFLP) with demand allocation, vehicle routing problem (VRP) and multi-period inventory control problem (MICP). Traditionally, these problems are studied separately. CFLP itself is a mixed-integer problem and some typical related applications and methods were discussed in Daskin (1995) and Drezner (1995). They introduced key classical location problems (covering, center, median, and fixed charge) in modeling practical applications (e.g., production and distribution facilities, interacting services and facilities, etc). Geoffrion and Bride (1978) proposed a Lagrangian Relaxation (LR) solution to CFLP. A combination of Benders decomposition and LR solution methods was later studied to solving CFLP (Van Roy and Erlenkotter, 1986). Some other approximate algorithms were implemented by Levi and Shmoys (2012) and Ahuja et al. (2004). Xie et al. (2014) further proposed multi-type CFLP with facility upgrading and co-location. VRP has been researched for decades, Raff (1983) summarized the state of the art in solving basic problems of routing and scheduling of vehicles and crews. Furthermore, Desrochers et al. (1992) proposed an algorithm to solving variant of VRP with time windows. Wagner and Whitin (1958) discussed a dynamic version of the economic lot size model to addressing MICP, which allows demands, inventory holding cost, setup cost to vary over periods. Various efforts have also been made to solve combinations of these decisions. For example, Perl and Daskin (1985) presented a “location routing” problem (LRP) which was studied in both deterministic (e.g., Laporte and Nobert, 1981; Belenguer et al., 2011) and stochastic settings (e.g., Laporte et al., 1989; Kleywegt et al., 2004; Ahmadi-Javid and Seddighi, 2013). In Laporte and Nobert (1981), only a single depot was considered to be located among many points with determinations about associated optimal delivery routes simultaneously. A Branch-and-Cut algorithm was developed in Belenguer et al. (2011) for solving small-size LRP (only 10 depots are tested)

with capacity constraints. Later, Laporte et al. (1989) published models and solutions to stochastic LRPs with unknown random customer supplies. Meanwhile, affects in changing production capacity caused by random disruptions were considered in VRP in Ahmadi-Javid and Seddighi (2013). A heuristic solution to multi-depot LRP is proposed in Tai-Hsi et al. (2002). In the other direction, Campbell et al. (1998); Campbell and Savelsbergh (2004) combined the VRP and inventory considerations into an “inventory routing” problem. Inventory routing problem was firstly presented in Campbell et al. (1998) and the methods to solving it was later improved in Campbell and Savelsbergh (2004) by decomposing the problem into two phases. Nourbakhsh and Ouyang (2010) integrated a capacitated fueling station location problem with fueling schedule decisions. However, no published work was done to integrate all three sets of problems (CFLP, MICP, VRP).

Therefore, a large-scale linear mixed-integer mathematical model to solve location routing with inventory control problem (LRICP) is proposed to optimize (i) locations, capacities, and types of facilities, (ii) assignments of work to facilities, (iii) fuel strategy of trains, and (iv) routes of movable facilities. To solve the model efficiently for large-scale railroad networks and reach a near-optimum solution, we propose sequential solution algorithms (e.g., network consolidation, network generation, reduce candidate location number, and etc).

In chapter 2, we first state essential model assumptions and rules to make the model more general. Then a methodology framework is shown to explain the solving technique flow. A model decomposition framework is proposed later to reduce the complexity of the model, including two subproblems which are introduced in chapter 3 and chapter 4 respectively. After solving the model, a series of data processing methods are introduced in chapter 5 to solve large-scale network efficiently. Several case studies are conducted in chapter 6 to show the correctness and efficiency of the model. At the end, we make a conclusion of this paper in chapter 7.

Chapter 2

Methodology Overview

2.1 Assumptions and Rules

In this paper, we build a mixed-integer linear model to address a complicated large-scale real world problem. There are many associated realistic requirements, rules, and issues which bring difficulty in the model development. For example, how to transform real-world raw data into the form which model can utilize, how to incorporate the work and fuel demand allocations into the model, etc. In this section, we present our assumptions and rules, based on which we build our model in later chapters.

2.1.1 Demand generation assumptions

In order to generate demand input in the network, first we assume that all the repair, service, and maintenance demand occur randomly and point-wisely, and happen only at destinations of the trains (If a locomotive breaks down en-route, it will be first moved to destination). Then those demands will be calculated using given data (e.g., locomotives frequency, demand occurrence ratio) as exogenous input to the optimization model. Next, for fueling demand, which is train based, is assumed to happen at locations along the train route. Different trains correspond to different fueling strategies, including where to stop, which facility to purchase fuel, and how much fuel purchased at each facility. The fueling strategy will imply the fueling demand, however, fueling strategies are model decisions instead of inputs.

2.1.2 Demand assignment assumptions

Since different facility types can do different types of work, there exists a facility type and work type mapping as illustrated in Table 2.1, in which "Y" implies that the facility in that

column can do the corresponding work in that row. Observing from Table 2.1, multiple types of facilities can serve the same type of work demand. For example, a locomotive can be fueled by three types of facilities: (i) by a MFS, (ii) by a LSC, or (iii) by a truck. It shall be emphasized that certain categories of facilities in this table actually contain many subcategories (e.g., various types of shops), depending on facility capability (i.e., the type of permitted work) and capacity (i.e., the amount of permitted work per unit time). More types of shops and related work types can be referred in appendix Table A.1.

Table 2.1: Locomotive work type and facility type mapping. “Y” indicates that the corresponding type of facility is capable of conduct the corresponding type of work.

Work Type	SHOP (9 types)	LSC (3 types)	LTHOME (9 types)	MFS
Repair (12 types)	Y	Y		
Maintenance (4 types)	Y	Y		
Service (2 types)		Y	Y	
Fueling		Y	Y	Y

2.1.3 Fueling rule assumptions

As stated before, fueling strategy implies where to stop, which facility to purchase fuel, and how much fuel purchased at each facility for each locomotive, three constraints are enforced: (i) Locomotives cannot run out of fuel; (ii) Locomotives cannot fuel to beyond the tank capacity; (iii) flow conservation, i.e., locomotives origin from or arrive at the same location should bring the same amount of fuel. There are several additional fueling rules for locomotives attached to different types of trains, for example, locomotives on high priority road trains should purchase fuel only from dedicated truck (waiting at station for trains) or MFS, while locomotives on non-priority road trains could fuel at any possible fueling facility type, even from trucks elsewhere. In addition, in order to avoid unexpected running out of fuel, locomotives can get vendor fuel from any locations they pass by.

2.1.4 Truck assumptions

As movable facilities, locomotive trucks have limitations in doing service or fueling work due to limited resources or specific requirements, for example, a truck can only serve 3-4

locomotives or fuel 7000-10000 gallons before returning home to replenish labor, fuel or other materials (which we call reservice or refueling), and it takes some time for reservice or refueling. Also, truck can serve more than one yard before coming back home but each truck trip has an upper bound.

2.2 Methodology Framework

The LRICP aims at applying optimization theory to the real-world railroad network planning problem, which is very complicated and requires careful methodology development in order to be solved efficiently. In this study, we build an integrated sequential methodology framework as shown in Fig. 2.1.

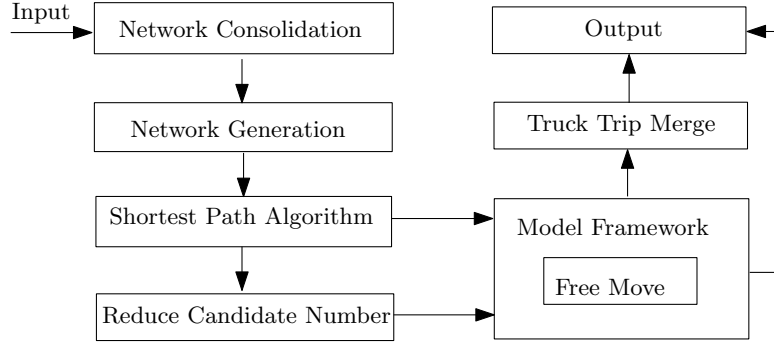


Figure 2.1: A framework of solution methods.

We first process the raw data (including information about stations, train routes, locomotives, demands, facilities, etc.) into the desired format as part of the input. In this stage, multiple types of point-wise work demand (e.g., service, repair, etc.), each with unique properties, are estimated from historical data. For example, service and repair demands are estimated based on locomotive shipment tasks. When an en-route locomotive breaks down, it is always first hauled to the shipment destination (to avoid shipment delay), then either hauled to a nearby shop facility, or repaired on-site by passing service trucks. Hence, we assume that these types of demand happen only at shipment destinations and generate related input data accordingly. The fueling demand, on the other hand, may occur anywhere during the shipment of a locomotive. It is jointly determined by typical locomotive running schedules, fuel consumption rates (depending on locomotive type), availability of fuel sta-

tions on shipment routes, and location-dependent fuel sales prices, and therefore cannot be exogenously calculated.

Next, we generate the railroad network as a graph. Because the network is quite large with more than 25 thousand stations (nodes) and even more links, it is inefficient to both calculate point-to-point distance and add variables about all those stations and links. Therefore, network consolidation is performed before generating network. Network consolidation eliminates nodes and links by masking nodes and links that are irrelevant to model decisions and make the network into a more compact one with necessary yards and railroad links. Masking nodes and links will not affect the accuracy of distance calculation.

Several algorithms (e.g., shortest path algorithm) are then developed and implemented to compute and extract useful information and data (e.g., all pair shortest travel distances). In addition, since the numbers of candidate facility locations are also extremely large and obviously some locations are predominant than others, we further reduce the number of candidate facility locations by introducing a set of heuristic methods, which helps reduce the number of variables and constraints

Finally, taking all those processed information and data as input, we formulate and solve the optimization model to obtain solutions, and conduct post-processing procedures involving various heuristic algorithms to improve the results before the ultimate results are generated for use.

2.3 Model Decomposition Framework

The core of the methodology framework is the optimization model. In this optimization model, the objective is to minimize the total cost which includes the facility construction cost or shut down benefit, capacity cost, work cost, transportation cost, truck travel cost, and fueling cost. The constraints can be divided into several basic categories: (i) those related to facility location, capacity setup, and demand assignment, which basically follow the formulation for standard multi-type capacitated facility location problems (Xie et al., 2014), (ii) those related to mobile facility routing and demand fulfillment, which follows the formulation for standard capacitated vehicle routing problems (Toth and Vigo, 2001), and (iii) those related to fueling locations and purchase quantities, which follows those in

the inventory model developed by Wagner-Whitin model (Wagner and Whitin, 1958). In addition, a set of demand fulfillment constraints are added to ensure that demand of each type is satisfied. The main trade-off of this problem is to balance between facility cost (e.g., construction), work cost (e.g., repair) as well as locomotive-related cost (e.g., shipping, fueling, and delay).

There are several layers of model decisions, shown in 2.2. First, the model selects a subset of candidate locations for each type of fixed facilities (i.e., LSC, SHOP, MFS, LTHOME), their capabilities and capacities, as well as demand allocations. Then for movable facilities, the decisions include the number of trucks of each type, the locations each truck served, and the typical routing plans of each truck. Finally the fueling strategies will output the locations of fueling stops, the facilities to get fuel, and the amount of fuel to purchase at each stop.

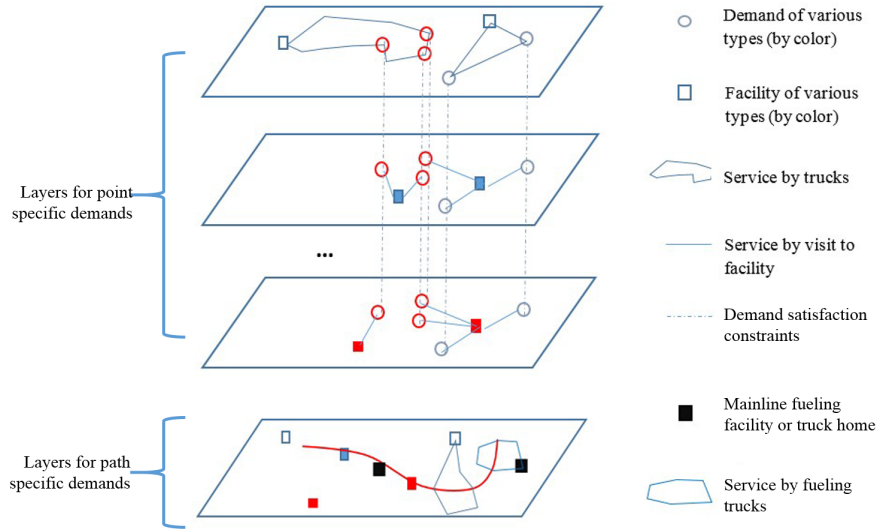


Figure 2.2: The layers for point-specific and path-specific demands.

This optimization problem combines several difficult decisions (location, routing and inventory), involving interdependent facility (LSC, SHOP, MFS, and LTHOME) and demand types (repair, maintenance, service, and fueling). And since we are applying the model to real-world applications which are generally of large scale, off-the-shelf solvers such as CPLEX and GurobiTM are found to experience excessive computation burden. Therefore, a model

decomposition framework is proposed to solve LRICP efficiently.

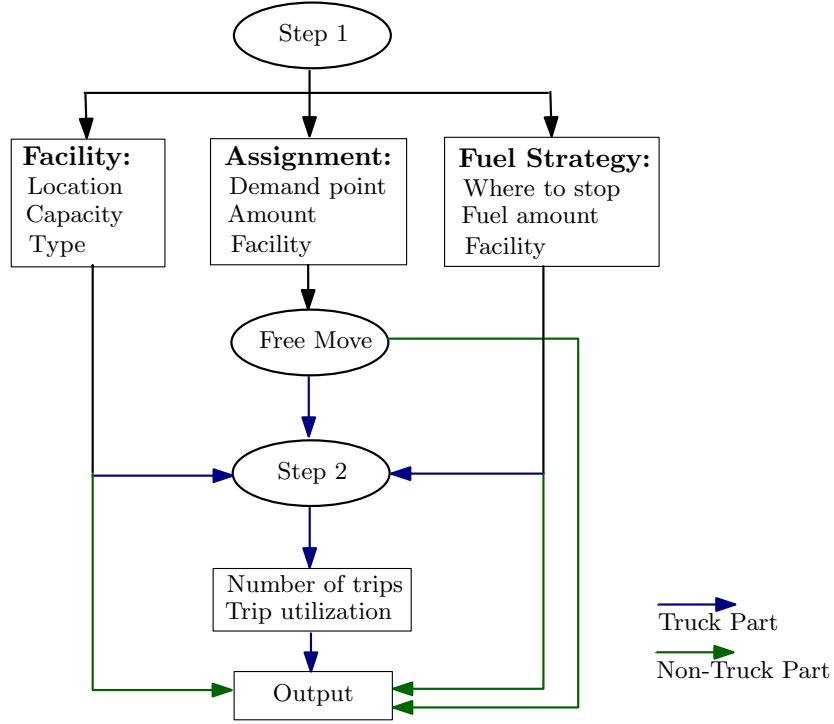


Figure 2.3: A framework of model part, consisting step 1 and step 2.

We propose a model decomposition framework consisting of two sequential steps to solve the LRICP as shown in Fig. 2.3. Modules with ellipse shape are main steps in the procedure and those with rectangle shape are output results of corresponding steps. In location inventory subproblem (step 1), the complete model is decomposed to a smaller model which integrates facility (e.g., locations, capacities, types), demand assignment (e.g., demand locations to which facilities, amount of each assignment), and inventory (fueling) strategy (e.g., where to stop, how much to fuel, and from which facility it receives fuel) decisions but leaves the routing decisions for subsequent steps. Note that truck homes are treated as fixed facilities and follow the same demand assignment rules. After finishing step 1, we fix the locations and fueling strategy, but implement a free move heuristic method to utilize as much free transportation as possible (which is a critical realistic part), which will move demand assignment between different facilities. After this free move modification, the new assignment decision, as well as the truck home and fueling strategy given by step 1

solution, are used as input to truck routing subproblem (step 2), in which the subproblem is solved given the demand points assigned to each truck home. Results from both step 1 (non-routing part) and step 2 (routing part) form the final solutions to the complete problem. More details of constructing the model and designing algorithms will be introduced in the following chapters.

Chapter 3

Location Inventory Subproblem

3.1 Objective

As introduced above, there are four main sections of cost to be optimized by our model. Notations in objective and constraints can be found in appendix. First section is facility cost for LSC, SHOP, MFS, and LTHOME, including facility fixed construction cost or shutdown benefit, capacity expansion cost or reduction benefit. LT is a movable facility traveling outside LTHOME and the logic is different from other four types of unmovable facilities. As a movable facility, LT has contract cost instead of fixed construction cost; it has running cost (defined as variable cost) which depends on truck trip decisions. Second section focuses on work cost when locomotives receive work from different types of facilities. It includes work cost for service, repair and maintenance and fixed regular fuel cost. And third section is about transportation cost when moving a locomotive to a facility to conduct work. Point to point transportation cost has been calculated using shortest path algorithm as coefficient input multiplied by how many locomotives are transported from one station to another station. The transportation cost in objective formulation has two parts: transportation cost for nonfuel work from demand point to LSC/SHOP/LT and transportation cost for fuel work from demand point to LSC/MFS/LT. The last section talks about fuel variable cost (fuel material cost), including regular fuel variable cost and vendor fuel variable cost. Vendor is treated as a backup if no other option exists for locomotives to purchase fuel from regular facilities, but fuel price offered by vendor is usually twice as much as regular price. In the formulation, L, S, M, and T are short for LSC, SHOP, MFS, and LT respectively to save space. Since LTHOME is actually a location to hold LTs without other special properties, it shares the same capability code with LT just for representation convenience.

$$\begin{aligned}
\text{Minimize } & \underbrace{\sum_{\alpha \in L/S/M/T} \left(\sum_{n \in N_\alpha \setminus N_\alpha^0} x_n^\alpha F_n^\alpha + \sum_{n \in N_\alpha^0} (1 - x_n^\alpha) F_n^{\alpha,0} \right)}_{\text{Fixed cost of facilities}} + \underbrace{\sum_{\alpha \in L/S/M} \sum_{n \in N_\alpha} (q_n^{\alpha,+} C_n^\alpha + q_n^{\alpha,-} B_n^\alpha)}_{\text{Capacity cost of L/S/M}} \\
& + \underbrace{\sum_{l \in T} \sum_{n \in N_T} (h_n^l \cdot F_l + t_n^l \cdot FV_l)}_{\text{Contract and variable cost of trucks}} + \underbrace{\sum_{\alpha=L/S/T} \sum_{\beta \in R/S/M} \sum_{n \in N_\alpha} \sum_{n_1} D_{n_1}^\beta \cdot y_{n_1,\beta}^{n,\alpha} \cdot W^{\alpha,\beta}}_{\text{Work cost for service, repair and maintenance}} \\
& + \underbrace{\sum_{\alpha=L/S/T} \sum_{\beta \in R/S} \sum_{n \in N_\alpha} \sum_{n_1} D_{n_1}^\beta \cdot y_{n_1,\beta}^{n,\alpha} \cdot C_{n,n_1}}_{\text{Trans. cost for nonfuel work (LSC/SHOP/LT)}} + \underbrace{\sum_{\alpha=L/M/T} \sum_r \sum_s \sum_{n \in N_\alpha} \sum_{n_1} G_{r,s}^{n_1} \cdot d_{r,s}^{n,\alpha} \cdot C_{n,n_1}}_{\text{Trans. cost for fueling (LSC/MFS/LT)}} \\
& + \sum_r \sum_s E_r \left[\underbrace{h_r^s \sum_n (G_{r,s}^n \cdot C_n)}_{\text{Variable regular fuel}} + \underbrace{p_r^s \sum_n (G_{r,s}^n \cdot C_n^0)}_{\text{Fixed regular fuel}} + \underbrace{v_r^s \cdot CV + pv_r^s \cdot CV^0}_{\text{Fixed and variable Vendor fuel}} \right]
\end{aligned}$$

3.2 Constraints

Following the problem decomposition layers in Fig. 2.2, we formulate constraints within each layer and constraints between layers to obtain a near-optimum solution to the original problem. There are about 40 constraints in total in the model of location inventory subproblem, and they are divided into four subsections: location and assignment (only related to unmovable facilities), truck (movable facility), number of truck trips estimation, and fueling strategy. In addition, there are constraints in each subsection that connect that subsection to other subsections. Details about constraints and connections will be discussed after each part of constraints.

3.2.1 Location and assignment

$$Q_n^{\alpha,\min} x_n^\alpha \leq q_n^\alpha \leq Q_n^{\alpha,\max} x_n^\alpha, \quad \forall n, \alpha \in \text{LSC/SHOP/MFS} \quad (3.1)$$

$$q_n^\alpha = \underbrace{Q_n^{\alpha,0} + q_n^{\alpha,+} - q_n^{\alpha,-}}_{\text{Current} + \text{increase} - \text{reduction}}, \quad \forall n, \alpha \in \text{LSC/SHOP/MFS} \quad (3.2)$$

$$0 \leq q_n^{\alpha,+} \leq \underbrace{Q_n^{\alpha,\max} - Q_n^{\alpha,0}}_{\text{Max increase}}, \quad \forall n, \alpha \in \text{LSC/SHOP/MFS} \quad (3.3)$$

$$0 \leq q_n^{\alpha,-} \leq \underbrace{Q_n^{\alpha,0}}_{\text{Max reduction}}, \quad \forall n, \alpha \in \text{LSC/SHOP/MFS} \quad (3.4)$$

$$\underbrace{\sum_{\beta} \sum_{n_1} D_{n_1}^{\beta} \cdot y_{n_1,\beta}^{n,\alpha} \cdot T_{\alpha,\beta}}_{\text{Nonfuel work (for LSC/SHOP)}} + \underbrace{\sum_r \sum_s d_{r,s}^{n,\alpha} \cdot T_{\alpha,\text{fueling}}}_{\text{Fueling (for LSC/MFS)}} \leq q_n^{\alpha}, \quad \forall n, \alpha \in \text{LSC/SHOP/MFS} \quad (3.5)$$

$$\sum_{\alpha \in \text{L/S/T}} \sum_{n \in N_{\alpha}} y_{n_1,\beta}^{n,\alpha} = 1, \quad \forall n_1, \beta \in \text{Repair/Service} \quad (3.6)$$

$$\left[\underbrace{\sum_{\alpha \in \text{LSC}} \sum_{n \in N_{\alpha}} G_{r,N_r}^n d_{r,N_r}^{n,\alpha}}_{\text{Fuel from LSC at destination}} \right] + \underbrace{\sum_{\alpha \in \text{MFS/LT}} \sum_{n \in N_{\alpha}} G_{r,s}^n d_{r,s}^{n,\alpha}}_{\text{Fuel from MFS/LT}} = E_r \cdot p_r^s, \quad \forall r \in (\text{Road, high priority}), s \quad (3.7)$$

$$\underbrace{\sum_{\alpha \in \text{LSC/MFS/LT}} \sum_{n \in N_{\alpha}} G_{r,s}^n d_{r,s}^{n,\alpha}}_{\text{Fuel from LSC/MFS/LT}} = E_r \cdot p_r^s, \quad \forall r \in (\text{Road, low priority}), s \quad (3.8)$$

$$\left[\underbrace{\sum_{\alpha \in \text{LSC}} \sum_{n \in N_{\alpha}} d_{r,N_r}^{n,\alpha}}_{\text{Fuel from any LSC at destination}} \right] + \underbrace{\sum_{\alpha \in \text{MFS}} \sum_{n \in N_{\alpha}} G_{r,s}^n d_{r,s}^{n,\alpha}}_{\text{Fuel from MFS}} + \underbrace{\sum_{\alpha \in \text{LT}} \sum_{n \in N_{\alpha}} d_{r,s}^{n,\alpha}}_{\text{Fuel from any LT}} = E_r \cdot p_r^s, \quad \forall r \in (\text{Nonroad}), s \quad (3.9)$$

$$\sum_{\alpha \in \text{L/M/T}} \sum_{n \in N_{\alpha}} db_{r,s}^{n,\alpha} = 1, \quad \forall r, s \quad (3.10)$$

$$y_{n_1,\beta}^{n,\alpha} \leq x_n^{\alpha}, \quad \forall n_1, \beta \in \text{Repair/Service}, n, \alpha \in \text{LSC/SHOP/LT} \quad (3.11)$$

$$d_{r,s}^{n,\alpha} \leq M \cdot x_n^{\alpha}, \quad \forall r, s, n, \alpha \in \text{LSC/MFS/LT} \quad (3.12)$$

$$0 \leq y_{n_1, \beta}^{n, \alpha} \leq 1, \forall n_1, \beta \in \text{Repair/Service}, n, \alpha \in \text{LSC/SHOP/LST/LFST} \quad (3.13)$$

$$y_{n_1, \beta}^{n, \alpha} = 0, \forall n_1, \beta \in \text{Repair/Service}, n, \alpha \in \text{MFS/LFT} \quad (3.14)$$

$$d_{r, s}^{n, \alpha} = 0, \forall r, s, n, \alpha \in \text{SHOP/LST} \quad (3.15)$$

$$x_n^\alpha \in \{0, 1\}, \forall n, \alpha \quad (3.16)$$

Constraint (3.1) limits the capacity for each facility of type α at yard n to make sure it is no less than given minimum capacity input $Q_n^{\alpha, \min}$ and no greater than given maximum capacity input $Q_n^{\alpha, \max}$. The capacity conservation for each current facility of type α at yard n is shown in constraint (3.2). Let $q_n^{\alpha, +}$ represent increased capacity and $q_n^{\alpha, -}$ represent decreased capacity, the model capacity q_n^α equals to the original current capacity added by increased capacity and then deducted by decreased capacity. To make $q_n^{\alpha, +}$ and $q_n^{\alpha, -}$ realistic, constraint (3.3) and constraint (3.4) give both upper bound and lower bound to them. After capacity expansion, the model capacity cannot exceed allowed maximum capacity bound.

Constraint (3.5) constructs connections between service, maintenance, and fueling by ensuring total work demand should be no more than the capacity of each type of facility. Since different types of facilities are capable of conducting specific types of work, they can choose from what work they can do by summing up all types of doable work, as long as the capacity is not surpassed. Constraint (3.6) makes sure that each service, maintenance or repair demand is fully satisfied by a LSC, SHOP or LT. The demand at a location can be done at only one type of facility or be split to two or more types of facilities. Fueling demand should also be fully satisfied, shown in constraint (3.7), constraint (3.8) and constraint (3.9). According to different realistic fueling requirements in railroad network for various trains, three general rules are implemented in those three constraints. For a road train with high priority, it is allowed to purchase fuel from a LSC if that LSC is located at its destination; or from a MFS or LT if the facility is located along its route. Let $G_{r, s}^n$ represent whether location n is the s^{th} station of route r . Terms about work (fueling, service, repair, maintenance) multiplied by $G_{r, s}^n$ represent that corresponding work can only be done at a facility that is on that s^{th} station of train route r . Let E_r be the frequency of train r , which represents the fueling demand amounts input. For a road train with low priority, it can receive fueling

from a LSC, MFS or LT (physical location of LTHOME is actually considered here) if the facility is on its route. While a nonroad train is allowed to receive fueling: i) from LSC at destination even if the LSC is not located at its destination; ii) from LT even if the related LTHOME is not located along its route; iii) from MFS along its route only if the MFS is located along its route. All fueling demand amounts that are satisfied by facilities should be equal to the original fueling demand amounts input.

$db_{r,s}^{n,\alpha}$ denotes whether locomotives at the s^{th} station of route r will receive fueling from facility type α at yard n , so constraint (3.10) ensures that fueling demand cannot be split into different parts served by different facilities. It is guaranteed that a locomotive will receive work from an existing facility in constraint (3.11) and constraint (3.12). Constraints ((3.13) - (3.16)) simply set default ranges of $y_{n_1,\beta}^{n,\alpha}$ and $d_{r,s}^{n,\alpha}$ for different types of facilities.

3.2.2 Truck

$$t_n^l = \sum_{n_1} t_{n,n_1}^l, \quad \forall n, l \in \text{LT} \quad (3.17)$$

$$h_n^l T_l^{\min} \leq t_n^l \leq h_n^l T_l^{\max}, \quad \forall n, l \in \text{LT} \quad (3.18)$$

$$\underbrace{\sum_n \sum_{l \in \text{LST/LFST}} z_{n,n_1}^l}_{n_1 \text{ assigned to service truck}} \leq 1, \quad \forall n_1 \quad (3.19)$$

$$\underbrace{\sum_n \sum_{l \in \text{LFT/LFST}} z_{n,n_1}^l}_{n_1 \text{ assigned to fueling truck}} \leq 1, \quad \forall n_1 \quad (3.20)$$

$$T_{n,n_1} z_{n,n_1}^l \leq 3, \quad \forall n_1, n, l \in \text{LT} \quad (3.21)$$

$$y_{n_1,\beta}^{n,l} \leq z_{n,n_1}^l, \quad \forall n_1, \beta \in \text{Repair/Service}, n, l \in \text{LST/LFST} \quad (3.22)$$

$$\underbrace{\sum_r \sum_s G_{r,s}^{n_1} \cdot d_{r,s}^{n,l}}_{\# \text{ of locos assigned to truck}} \leq M \cdot z_{n,n_1}^l, \quad \forall n_1, n, l \in \text{LFT/LFST} \quad (3.23)$$

$$z_{n,n_1}^l \leq x_n^{\text{LT}}, \quad \forall n_1, n, l \in \text{LT} \quad (3.24)$$

$$h_n^l \leq Q_n^l \cdot x_n^{\text{LT}}, \quad \forall n, l \in \text{LT} \quad (3.25)$$

Constraint (3.17) constructs the equation that total truck working time equals to the summation of variable time assigned to various demand locations. Both t_n^l and t_{n,n_1}^l are useful in further constraint representation. Constraint (3.18) gives an upper bound and lower bound to truck total working time, which varies by LT type. For example, the total working time of a LST168 per week should be in 0 - 168 hours. Furthermore, constraint (3.19) and constraint (3.20) ensure that demand of one location (service or fueling) can only be assigned to at most one LT type from one LTHOME. Because the LT driver has work shift that is usually 12 hours, there is a hard constraint (3.21) representing that travel time should be less than 6 hours per round trip (another 6 hours for doing work). The number can be changed as a user parameter input. Constraint (3.22) simply represents that a yard will be served by a LT of type l from LTHOME n only if it is assigned to that LT. Similar with constraint (3.22), constraint (3.23) represents that a yard will be fueled by a LT of type l from home n only if it is assigned to that LT. The difference is that $d_{r,s}^{n,l}$ is the real number of locomotives assigned to that truck, but $y_{n_1,\beta}^{n,l}$ is only the assigned ratio of locomotives ranging from 0 to 1. We also need to make sure that a locomotive can only be assigned to an existing LTHome, indicated by constraint (3.24). Constraint (3.25) shows that the number of trucks at a LTHOME should not exceed the maximum number input of trucks. From all the constraints about trucks introduced above, we can observe that truck location decision variables always has one more dimension indicating the serving locations than decision variables of unmovable facilities.

3.2.3 Number of truck trips estimation

It can be observed that each truck trip only serves one location from the truck trip constraints in last section. Since VRP problem is NP-hard and not suitable for large-scale problem size (Toth and Vigo, 2001), this paper only considers serving one location per truck trip in step 1 model. And step 2 model will solve heuristic VRP for each LTHOME with assignment decisions from step 1 as input. However, assignment decisions to each truck in step 1 are significantly affected by corresponding truck cost, therefore, we need accurate estimation about the truck cost(e.g., truck fixed cost, truck working hour per week) in step 1. There is a reservice or refueling time (e.g. 1 hour) between two consecutive truck trips, which is

counted in truck total working time in step 2, so we need to estimate how many reservices and refuelings are needed without knowing the service or fueling assignments in advance. Constraint (3.26) and constraint (3.27) give an accurate estimation of reservice number and refueling number. $D_{n_1}^{\text{NonRoad}} \cdot T_{l,\text{NonRoad}} \cdot y_{n_1,\text{NonRoad}}^{n,l}$ estimates the total work time for truck to do work of amount $D_{n_1}^{\text{NonRoad}} \cdot y_{n_1,\text{NonRoad}}^{n,l}$. $12 - 2T_{n,n_1}$ is work time per truck trip, where T_{n,n_1} denotes one-way travel time from LTHOME n to serving location n_1 . Refueling estimation involves another limited resource constraint of truck tank capacity. More specifically, each truck has maximum tank capacity (e.g., 7000 gallons), a LFT can only serve one locomotive if the served locomotive needs more than 7000 gallons fuel. Therefore, t_{n,n_1}^l can be obtained more accurately using estimation in constraint (3.26) and constraint (3.27), including round trip travel time, reservice time, service time, refueling time, and fueling time. Constraints ((3.29) - (3.31)) give bounds to $d_{r,s}^{n,\alpha}$ and $w_{r,s}^{n,l}$ by involving binary variable $db_{r,s}^{n,\alpha}$ and a large number M .

$$N_{n,n_1}^{l,\text{service}} = \frac{\overbrace{D_{n_1}^{\text{NonRoad}} \cdot T_{l,\text{NonRoad}} \cdot y_{n_1,\text{NonRoad}}^{n,l}}^{\text{Total work time needed}}}{\underbrace{12 - 2T_{n,n_1}}_{\text{Work time available per trip}}} \quad (3.26)$$

$$N_{n,n_1}^{l,\text{fueling}} = \underbrace{\sum_{r \in \text{Road}} \sum_s \frac{E_r \cdot w_{r,s}^{n,l}}{7000}}_{\text{Road: 7000 gallons per trip}} + \frac{\overbrace{\sum_{r \in \text{NonRoad}} \sum_s d_{r,s}^{n,l} \cdot T_{l,\text{fueling}}}^{\text{NonRoad: Total work time needed}}}{\underbrace{12 - 2T_{n,n_1}}_{\text{Work time available per trip}}} \quad (3.27)$$

$$\begin{aligned} t_{n,n_1}^l = & \underbrace{(2T_{n,n_1} + T_{\text{re}}^l) \cdot N_{n,n_1}^{l,\text{service}}}_{\text{Travel time and reservice time}} + \underbrace{\sum_{\beta} D_{n_1}^{\beta} \cdot y_{n_1,\beta}^{n,l} \cdot T_{l,\beta}}_{\text{Service time}} \\ & + \underbrace{(2T_{n,n_1} + T_{\text{re}}^l) \cdot N_{n,n_1}^{l,\text{fueling}}}_{\text{Travel time and refueling time}} + \underbrace{\sum_r \sum_s G_{r,s}^{n_1} \cdot d_{r,s}^{n,l} \cdot T_{l,\text{fueling}}}_{\text{Fueling time}} \end{aligned} \quad (3.28)$$

$$db_{r,s}^{n,\alpha} \leq d_{r,s}^{n,\alpha} \leq M \cdot db_{r,s}^{n,\alpha}, \forall \alpha = \text{LSC/MFS/LT} \quad (3.29)$$

$$-M \cdot db_{r,s}^{n,l} \leq w_{r,s}^{n,l} \leq M \cdot db_{r,s}^{n,l} \quad (3.30)$$

$$h_r^s + M \cdot (1 - db_{r,s}^{n,l}) \leq w_{r,s}^{n,l} \leq h_r^s + M \cdot (db_{r,s}^{n,l} - 1) \quad (3.31)$$

3.2.4 Fueling strategy

Fueling strategy for each train is different, including where to stop, how much fuel to purchase, which facility to purchase fuel from. Trains are allowed to purchase fuel from LSC, MFS and LT (LFT/LFST) under some train fueling rules, demonstrated by the following constraints.

$$p_r^s \leq h_r^s \leq Q_r p_r^s, \forall r, s = 1, 2, \dots, N_r \quad (3.32)$$

$$pv_r^s \leq v_r^s \leq Q_r pv_r^s, \forall r, s = 1, 2, \dots, N_r \quad (3.33)$$

$$p_r^{N_r} = h_r^{N_r} = pv_r^{N_r} = v_r^{N_r} = 0, \forall r \quad (3.34)$$

$$v_r^s \geq 2000 \cdot pv_r^s, \forall r \quad (3.35)$$

$$\underbrace{\sum_n G_{r,0}^n h_n}_{\text{Initial fuel level}} + \underbrace{\sum_{s=1}^N (h_r^s + v_r^s - U_r^s)}_{\text{Fueling amount - Consumption}} \leq Q_r, \forall r, N = 1, 2, \dots, N_r \quad (3.36)$$

$$\underbrace{\sum_n G_{r,0}^n h_n}_{\text{Initial fuel level}} + \underbrace{\sum_{s=1}^N (h_r^s + v_r^s - U_r^s) - U_r^{N+1}}_{\text{Fueling amount - Consumption}} \geq 0, \forall r, N = 0, 1, 2, \dots, N_r - 1 \quad (3.37)$$

$$\underbrace{\sum_n G_{r,0}^n h_n}_{\text{Initial fuel level}} + \underbrace{\sum_{s=1}^{N_r} (h_r^s + v_r^s) - \sum_{s=1}^{N_r} U_r^s}_{\text{Fueling amount - Consumption}} = \underbrace{\sum_n G_{r,N_r}^n h_n}_{\text{Initial fuel for next train}}, \forall r \quad (3.38)$$

Binary variable p_r^s denotes whether a train r stops at s_{sth} station to receive normal fueling. Binary variable pv_r^s indicates the same for vendor fueling. Constraint (3.32) and constraint (3.33) clearly show that a locomotive can be fueled at s^{th} fuel station of train r only if it stops at that station. At the same time, the two constraints bound the range of h_r^s

and v_r^s , where h_r^s is normal fueling amount at s^{th} station of train r and v_r^s is corresponding vendor fueling amount. Constraint (3.34) is a boundary condition for fueling. According to realistic rules, locomotives cannot purchase fuel at the destination station, they are preferred to purchase fuel at the origin station. Meanwhile, there is a realistic rule for vendor fueling that the amount of vendor fueling to a locomotive should be at least 2000 gallons (parameter input), shown in constraint (3.35). Constraint (3.36) ensures that the fuel level of a locomotive of train r after each fueling event should never exceed its tank capacity. It is obvious but essential that no locomotive runs out of fuel before arriving at the next fuel station, indicated by constraint (3.37). And we need to guarantee that initial fuel level at each origin node plus the total fuel purchased at consecutive stations should exceed the total fuel consumption in the path, which is shown in constraint (3.38).

3.3 Overview

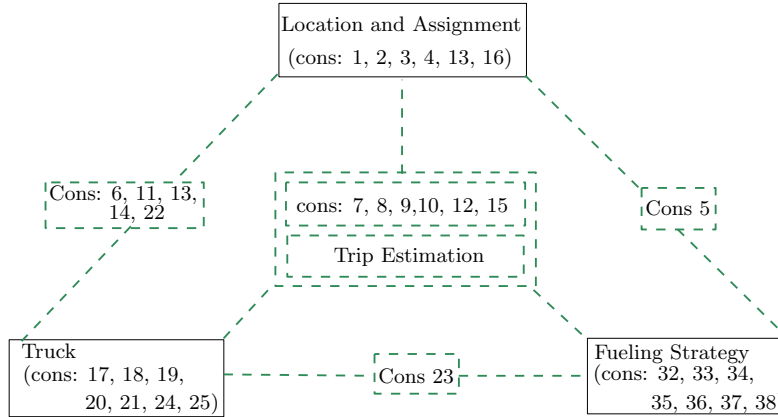


Figure 3.1: A framework of constrains relations in step 1.

Although constraints are divided into sections, it is complicated to form a clear view of relations among them. Therefore, a brief sketch in Fig. 3.1 is given to show the constraint framework of step 1. Modules with solid line are main constraint sections introduced above, including location and assignment, truck, and fueling strategy. As shown in the figure, each section has relatively independent constraints which are not obviously connected to other sections. While modules with dashed line are constraints that represent connections between

sections. Constraint (3.5) performs a connection between location and assignment section and truck section. Constraint (3.23) only connects truck section and fueling strategy section. Constraints (3.6), (3.11), (3.13), (3.14), (3.22) are connections between truck section and location and assignment section. At the same time, constraints (3.7), (3.8), (3.9), (3.10), (3.12), (3.15) and all constraints in trip estimation section connect three sections together. Observe that all connection constraints are related to work demand fulfillment.

3.4 Free Move

In real-world train operation, road locomotives with service demand can still provide power to lead normal trains. Therefore, the transportation cost of moving locomotives with service demand to LSC should be counted as normal train operation cost, which indicates there is almost no extra transportation cost to move those locomotive around. Since the transportation cost of one-unit work assignment is calculated based on normal train network without free move, a number 0.8 to represent the portion of free move in the whole network is applied to estimate the real transportation cost before the model makes any decisions. The real free move part is done after model makes decision about assignments to LSCs. For each service assignment from demand destination to LSC, we try to move the demand on the free locomotive network with zero moving cost by train. Each link of train path has a capacity, representing the number of free locomotives it can carry multiplied by frequency per week of that train. Each time we move at most 2 road locomotives from a demand destination to a available (capacity is not fully utilized) and reachable (can use free move with less cost to reach it) LSC. The two locomotives may be moved to the same LSC as assigned in the model or moved to another LSC, depending on which LSC saves most transportation cost. If the cost of this recalculation is smaller than original cost, we accept this recalculation. Then the demand of destination is reduced by 2 and the path links capacity is also reduced by 2. Otherwise, keep the model decision about current service assignment and move to next uncalculated service assignment. The shortest path algorithm is run repeatedly until all the road locomotive service demand are transported to a LSC with free move or the capacities of all free move links are exhausted.

Chapter 4

Truck Routing Subproblem

4.1 Objective

For each selected home location n , we have the following objective

$$\text{Minimize} \quad \sum_{l=T} (h^l \cdot F_l + t^l \cdot FV_l)$$

where h^l is the number of type l truck and t^l is the working variable time of type l truck. Corresponding coefficients of those two variables are fixed cost of type l truck and variable work cost of type l truck. We don't need to take transportation cost into consideration because the travel time has already been counted in t^l .

4.2 Constraints

t^l denotes the total time for truck of type l to serve demands at its LTHOME and at other surrounding yards. No transportation time is considered when LT serves its own LTHOME, therefore, the model forms different constraints for those two cases. Constraint (4.1) represents summation of two parts of truck time. T_{re}^l denotes the rework time (i.e., reservice time, refueling time) per truck trip for LT type l .

$$t^l = \underbrace{\sum_m \overbrace{w^{m,l}}^{\text{Trip time}} + \sum_m \overbrace{z^{m,l} T_{re}^l}^{\text{Rework time}}}_{\text{Truck time for serving other yards}} + \underbrace{t_{n \rightarrow n}^l + \overbrace{\frac{s_{n \rightarrow n}^l}{12} T_{re}^l}^{\text{Reservice}} + \overbrace{\frac{h_{n \rightarrow n}^l}{H_{\max}} T_{re}^l}^{\text{Road refueling}} + \overbrace{\frac{d_{n \rightarrow n}^l}{12} T_{re}^l}^{\text{NonRoad refueling}}}_{\text{Truck time for serving its home location}} \quad (4.1)$$

4.2.1 Serving other yards

$$t^m = \underbrace{\sum_l \sum_{n_1} \sum_{\beta} D_{n_1}^{\beta} \cdot \hat{y}_{n_1, \beta}^{m, l} \cdot T_{l, \beta}}_{\text{Time for doing service}} + \underbrace{\sum_l \sum_r \sum_s \hat{d}_{r, s}^{m, l} \cdot T_{l, \text{fueling}}}_{\text{Time for doing fueling}} + \underbrace{2 \sum_{n_1} u_{n_1}^m \cdot T_{n, n_1}}_{\text{Time for travelling}} \quad (4.2)$$

$$t^m \leq 12 \sum_l z^{l, m} \quad (4.3)$$

$$t^m + M \cdot (z^{m, l} - 1) \leq w^{m, l} \leq t^m \quad (4.4)$$

$$0 \leq w^{m, l} \leq M \cdot z^{m, l} \quad (4.5)$$

$$h^l T_l^{\min} \leq t^l \leq h^l T_l^{\max} \quad (4.6)$$

$$\sum_l z^{m, l} \leq 1 \quad (4.7)$$

$$\sum_{n_1} u_{n_1}^m \leq 1 \quad (4.8)$$

$$u_{n_1}^m \leq \sum_l z^{m, l} \quad (4.9)$$

$$\sum_l \sum_m \hat{y}_{n_1, \beta}^{m, l} = 1 \quad (4.10)$$

$$\underbrace{\sum_l \sum_m (1 - G_{r, n}^s) \hat{d}_{r, s}^{m, l}}_{\# \text{ of loco supply to other yards}} = \underbrace{\sum_l (1 - G_{r, n}^s) D_{r, s}^l \cdot P_r^s}_{\# \text{ of loco demand from other yards}} \quad (4.11)$$

$$\underbrace{\sum_{n_1} \sum_{\beta} D_{n_1}^{\beta} \sum_l \hat{y}_{n_1, \beta}^{m, l}}_{\text{Service assigned to trip } m \text{ of truck type } l} \leq D_{\max} \cdot \sum_l z^{m, l} \quad (4.12)$$

$$\underbrace{\sum_r \sum_s H_r^s \sum_l \hat{d}_{r, s}^{m, l}}_{\text{Fuel amount provided by trip } m \text{ of truck type } l} \leq H_{\max} \cdot \sum_l z^{m, l} \quad (4.13)$$

$$\sum_{n_1} \sum_{\beta} \hat{y}_{n_1, \beta}^{m, l} \leq M \cdot z^{m, l}, \forall l \in \text{LST/LFST} \quad (4.14)$$

$$\sum_r \sum_s \hat{d}_{r, s}^{m, l} \leq M \cdot z^{m, l}, \forall l \in \text{LFT/LFST} \quad (4.15)$$

$$\sum_{\beta} \sum_l \hat{y}_{n_1, \beta}^{m, l} \leq u_{n_1}^m \quad (4.16)$$

$$\sum_r \sum_s \sum_l G_{r,s}^{n_1} \cdot \hat{q}_{r,s}^{m,l} \leq M \cdot u_{n_1}^m \quad (4.17)$$

When LT travels out of LTHOME to serve other yards, the total time of truck trip m , t^m , should be equal to the summation of servicing time, fueling time and truck travel time, shown in constraint (4.2). In addition, the total time for a trip should be within 12 hours. Constraint (4.4) and constraint (4.5) bound $w^{m,l}$ with binary variable $z^{m,l}$ and t^m . The total travel time of one truck of type l should be within a given range $[T_l^{\min}, T_l^{\max}]$. Therefore, when there are h^l trucks of type l , the total time range for type l LT is shown in constraint (4.6). Constraint (4.7) ensures that each trip from current LTHOME can only be operated by one type of truck. Also, each trip from home n can only pass through one demand location, indicated in constraint (4.8). Relations between $u_{n_1}^m$ and $z^{m,l}$ are represented in constraint (4.9), which means that a yard n_1 will be served by a truck trip m only if it is assigned to truck trip m . Constraint (4.10) makes sure that service demand at location $n1$ should be fully satisfied by any type of trucks or any trips. And constraint (4.11) ensures that fueling demand should also be fully satisfied by trucks. As discussed before, each truck trip has resource limit (e.g., 12 hour time limit, work time limit, tank capacity limit). Therefore, we need constraints to make sure that LT needs to come back LTHOME if any type of resource is exhausted. Constraint (4.12) and constraint (4.13) bounds maximum number of service and and maximum amount of fuel in one truck trip respectively. The service demand is assigned to truck trip m of truck type l only if trip m is served by truck type l , which is restricted in constraint (4.14). And constraint (4.15) ensures the same restriction for fueling demand. Furthermore, there are constraints about relations between demand location and truck trip. Constraint (4.16) and constraint (4.17) mean that demand at location $n1$ or s^{th} station of train r is served by truck trip m only if the location $n1$ or s^{th} station of train r is assigned to truck trip m , which is decided by binary variable $u_{n_1}^m$.

4.2.2 Serving home location

$$\begin{aligned} \text{Total time} \quad t_{n \rightarrow n}^l &= \underbrace{\sum_{\beta} D_n^{\beta} \cdot \hat{y}_{n,\beta}^l \cdot T_{l,\beta}}_{\text{Service time}} + \underbrace{\sum_r \sum_s G_{r,s}^n \cdot \hat{d}_{r,s}^l \cdot T_{l,\text{fueling}}}_{\text{Fueling time}} \end{aligned} \quad (4.18)$$

$$\text{Service time} \quad s_{n \rightarrow n}^l = \sum_{\beta} D_n^{\beta} \cdot \hat{y}_{n,\beta}^l \cdot T_{l,\beta} \quad (4.19)$$

$$\text{Fueling amount} \quad f_{n \rightarrow n}^l = \underbrace{\sum_r \sum_s G_{r,s}^n \cdot \hat{d}_{r,s}^l \cdot H_r^s}_{\text{Fueling amount}} \quad (4.20)$$

$$\text{Road train fueling amount} \quad h_{n \rightarrow n}^l = \underbrace{\sum_{r \in \text{Road}} \sum_s G_{r,s}^n \cdot \hat{d}_{r,s}^l \cdot H_r^s}_{\text{Fueling amount of Road train}} \quad (4.21)$$

$$\text{NonRoad train fueling time} \quad d_{n \rightarrow n}^l = \underbrace{\sum_{r \in \text{NonRoad}} \sum_s G_{r,s}^n \cdot \hat{d}_{r,s}^l \cdot T_{l,\text{fueling}}}_{\text{Fueling time of NonRoad train}} \quad (4.22)$$

$$\sum_l \hat{y}_{\beta}^l = 1 \quad (4.23)$$

$$\sum_l G_{r,s}^n \cdot \hat{d}_{r,s}^l = \sum_l G_{r,s}^n \cdot D_{r,s}^l \cdot P_r^s \quad (4.24)$$

This subsection introduces constraints about serving home location. The calculation of total time and service time is different from previous subsection of serving other yards. There are brief explanations about each notations in the equation constraints ((4.18) - (4.22)), which simply define how each variable is calculated. Constraint (4.23) and constraint (4.24) makes sure that the service demand and the fueling demand at the home location should be all satisfied by itself respectively.

4.3 Truck Trip Merge

In model step 2, the truck routing problem is not formulated as the traditional VRP. Instead, it's more like assignment problem with movable facilities. With resource constraint on each truck trip, one truck trip usually serves one demand location and has to return to LTHOME to prepare for reservice or refueling. Therefore, for each demand location, there will be at

most a truck trip that is not fully utilized. Regarding those truck trips, merging them into one trip can save time and cost. Therefore, we use saving heuristics to merge trips. The left hand side of Fig. 4.1 shows that part of assignments from demand location A , B and C to LTHOME. The truck takes three different trips to serve those three locations and the path of trip is respectively $a1 \rightarrow a2$, $b1 \rightarrow b2$ and $c1 \rightarrow c2$. Let $T_{a1 \rightarrow a2}$ be the total time of trip $a1 \rightarrow a2$, t_{a1} be one way trip travel time of on link a_1 , tw_A be work time working at A . Therefore, $T_{a1 \rightarrow a2} = t_{a1} + t_{a2} + tw_A$. The same is for $T_{b1 \rightarrow b2}$ and $T_{c1 \rightarrow c2}$. Therefore, In order to check whether those three trips can be merged into one trip, we firstly check whether $t_{a1} + t_{e1} + t_{e2} + t_{c2} + tw_A + tw_B + tw_C$ is less than total trip time limit (e.g. 12 hours). If yes, we can directly merge the trip. If no, we check $t_{a1} + t_{e1} + t_{b2} + tw_A + tw_B$ or $t_{b1} + t_{e2} + t_{c2} + tw_B + tw_C$ and choose the one that saves more travel time if they are both within trip time limit. After merging truck trips, we obtain the final results of truck routing problem.

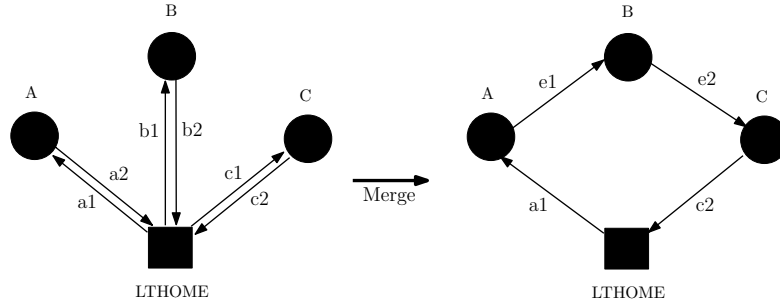


Figure 4.1: Merge truck trips into one trip

Chapter 5

Data Processing Methods

5.1 Network Consolidation

Define that important node is the station that train is allowed to stop by and allowed to do fueling, service and maintenance in its route, and unimportant node is the station that train is not allowed to stop by. To make the constructed network smaller and improve computation speed, we only consider links between important nodes, which involves the following network consolidation methods.

Train network consolidation: For each train route r , if s^{th} station is unimportant, store cost information of link (n_{s-1}, n_s) from $(s-1)^{th}$ station to s^{th} station and delete node n_s and link (n_{s-1}, n_s) . Keep doing so until reach an important node at t^{th} station and create new link (n_{s-1}, n_t) with cost equal to summation of cost of all deleted links from s^{th} station. Simple algebra shows that this algorithm runs $O(|V| + |E|)$ where $|V|$ denotes the cardinality of node set V and $|E|$ denotes the cardinality of link set E .

Truck network consolidation: For an undirected graph $G(V, E)$, let $v \in V$ be an unimportant node and $Ne(v)$ denotes its neighbors (i.e., $i \in Ne(v)$ and (i, v) is an edge). For each $i, t \in Ne(v)$, if they are not neighbors then add a new edge (i, t) to the network; otherwise, update the length of (i, t) . After this operation, remove node v and go to another unimportant node until there is no unimportant node in the network. Simple algebra shows that this algorithm runs $O(|E|^2)$.

5.2 Network Generation

Before introducing the algorithm of calculating the transportation cost, we need to preprocess the consolidated nodes and train routes so that the shortest path algorithm can run on the network. There are four types of network in the model, including train network, free move network, light move network and truck network.

- (1) Train network: Train route information provided by the railroad network is used to construct train network. Observe that a link (B, D) can be passed by multiple trains and may involve different costs, we need a mechanism to distinguish the same link for different trains. Xie et al. (2014) proposed a method to construct the network. Similar to that, we separate the multi-graph by adding new nodes and new edges shown in Fig. 5.1. For a directed link (B, D) , there are two trains a and b passing it with transportation costs t_a, t_b from B to D . To distinguish those two trains on the same link, we add 2 new nodes a_B, a_D for train a and 2 new nodes b_B, b_D for train b . Besides the new virtual nodes, we also add 5 more links for each train, which are $(a_B, B), (a_B, a_D), (B, a_B), (a_D, D), (D, a_D)$ for train a and $(b_B, B), (b_B, b_D), (B, b_B), (b_D, D), (D, b_D)$ for train b . The length of (a_B, a_D) is t_a and the length of (b_B, b_D) is t_b , and the lengths are all zeros for other links. Moreover, this method can also deal with the dropping cost; e.g., there is a penalty if train a drops locomotives at node D which is equivalent to that link (a_D, D) has an extra penalty cost. Read links from consolidated train route table, the train network can be constructed according to that mechanism.

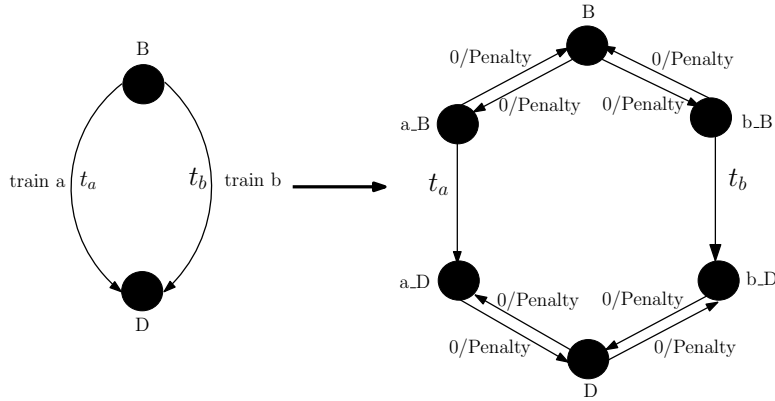


Figure 5.1: Create the network (Xie et al., 2014)

- (2) Free move network: Locomotives will have no extra transportation cost if they are used to provide energy for trains. Each train needs two working locomotives, which means first two locomotives that a train carries are for free. But the first two locomotives are restricted to work from the origin station to destination station without being dropped off in intermediate stations. Based on the definition and rules of free move, we add only origin-destination link (B_{ori}, D_{dest}) for each train a based on the same mechanism introduced in constructing train network. Besides, each link (a_B_{ori}, b_D_{dest}) has a capacity of $(2 * \text{train frequency per week})$, representing that the link can carry $(2 * \text{train frequency per week})$ locomotives for free per week.
- (3) Light move network: We treat the light move as a special type of train which can run everywhere in the network. It's used for connecting locomotives from one station to another station if no suitable trains can carry them. Therefore, the light move network needs to be combined with both train network and free move network. By using the same mechanism introduced in constructing train network, links like $(light_move_B, light_move_D)$ will be added to light move network without any capacity limit. Note that the light move is usually much more expensive but faster than normal train to carry locomotives.
- (4) Truck network: Since the truck is different from train that it does not need to travel on the rail, truck network is constructed with highway nodes and links. Different from train network, there are no distinctions between trucks in truck network. Therefore, original highway links and nodes will be added to truck network without any capacity limit on links.

5.3 Shortest Path Algorithm

After creating the network, we can use Dijkstra's algorithm to compute transportation cost between any two nodes. Suppose that the new network is $G(V, E)$. For the single node shortest path algorithm with priority queue, the running time is $O((|V| + |E|) \log |V|)$.

The pseudo-code for shortest path algorithm is as follows, s is the start node and V is the set of nodes.

```

1.  $Q \leftarrow \text{makePQ}()$ 
2.  $\text{Insert}(Q, (s, 0))$ 
3.  $\text{prev}(s) = \text{null}$ 
4. for each node  $u \neq s$  do
5.    $\text{Insert}(Q, (u, \infty))$ 
6.    $\text{prev}(u) = \text{null};$ 
7. end for
8. while  $Q$  is not empty do
9.    $(v, \text{dist}(s, v)) = \text{extractMin}(Q)$ 
10.  for each  $u$  in  $\text{Adj}(v)$  do
11.    if  $\text{dist}(s, v) + (v, u) < \text{dist}(s, u)$  then
12.       $\text{decreaseKey}(Q, (u, \text{dist}(s, v) + (v, u)))$ 
13.       $\text{prev}(u) = v$ 
14.    end if
15.  end for
16. end while

```

In each type of network, the $\text{dist}(s, u)$ in shortest path algorithm is calculated by different equations. In general, the calculation formula is

$$\text{cost} = \text{connection dwell time cost} + \text{travel cost} + \text{connection pick drop cost}$$

Connection dwell time cost happens: i) when a locomotive is waiting to be attached to a train for the first time; ii) when a locomotive is detached from previous train and transferred to another train; iii) when a locomotive arrives at destination and needs to wait in queue for being served. Travel cost includes: i) fuel consumption cost; ii) crew cost; iii) fixed running cost. There is connection pick drop cost when train needs to stop to drop locomotive or pick locomotive from other trains. The formulas of calculating each item differ in different networks.

(1) **Train + light move network**

When calculating connection dwell time cost, it has different rules to deal with various cases. When it is the first train connection for locomotive, if the next train is light move, Connection dwell time cost = Travel time cost factor $\times 12$; if the next train is normal train, Connection dwell time cost = Travel time cost factor $\times (24 - 12 \times \text{Train frequency per day}) / 2$. Travel time cost factor denotes cost per hour per locomotive. When it is not the first connection and the locomotive is currently attached to a light move, if the next train is light move, no dwell time cost is added; if the next

train is normal train, Connection dwell time cost = Travel time cost factor $\times$ 12. When it's not first connection and the locomotive is currently attached to a normal train, if the next train is light move, Connection dwell time cost = Travel time cost factor $\times$ 12; if the next train has the same train ID with current train, Connection dwell time cost = Travel time cost factor $\times$ Dwell time; if the next train has different train ID with current train: i) if train frequency of any two trains equals to 7, number = Frequency of first train + Frequency of second train - 7 ; ii) otherwise, number = Number of common operation day. Then Dwell time = $24 + 8 \times \text{number} / 7 - 12 \times (\text{First train frequency} + \text{Second train Frequency}) / 7$, and Connection dwell time cost = Dwell time $\times$ Travel time cost factor.

Travel cost is less complicated than connection dwell time cost. Add locomotive transportation cost to travel cost firstly, Travel cost+ = Travel time cost factor $\times$ Runtime. Then if next train is light move train, extra cost should be added to Travel cost: i) if start station of the link and end location of the link have same group milepost, Travel cost = (Light move crew cost + Cost fuel) $\times$ Distance $\times$ 0.2 + Travel time cost factor $\times$ Runtime; ii) otherwise, Travel cost = $3 \times ((\text{Light move crew cost} + \text{Cost fuel}) \times \text{Distance} \times 0.2 + \text{Travel time cost factor} \times \text{Runtime})$.

When a current segment has no previous segment: i) if current segment has penalty to drop a locomotive, Connection pick drop cost += Cost Penalty Pick Drop, ii) if current segment has penalty to pick a locomotive, Connection pick drop cost += Cost Penalty Pick Drop. When a current segment has a previous segment, it has two subcases to consider: when current segment and previous segment have same train ID and when they have different train ID. But in realistic calculation, those two cases have the same rules as following: i) if current segment has penalty to drop a locomotive, Connection pick drop cost -= Cost Penalty Pick Drop, ii) if current segment has penalty to pick a locomotive, Connection pick drop cost -= Cost Penalty Pick Drop.

(2) Free move + light move network

It's the same with train + light move network to calculate the connection dwell time cost. The differences between two networks lie in calculation of travel cost. When next train is light train, Travel cost = (Light move crew cost + Cost fuel) $\times$ Distance

$\times 0.2 + 2 \times \text{Travel time cost factor} \times \text{Runtime}$. When next train is non-free normal train, $\text{Travel cost} = \text{Travel time cost factor} \times \text{Runtime}$. When next train is free normal train, $\text{Travel cost} = \text{Dwell cost for first two road loco} \times \text{Runtime} + \text{Penalty Dist} \times \text{Distance}$. Connection calculation logic is the same with train + light move network, but the cost parameter involved is Dwell cost for first two road loco instead of Cost Penalty Pick Drop.

- (3) **Truck network** Truck network is quite different from other two networks introduced above. Since it is running on the highway instead of railroad, it has no obvious connection dwell time cost. And by using the coordinate information provided with each station in railroad network, this paper uses ArcGis to calculate travel cost between any two stations. The ArcGis uses its own highway network information to calculate the distance. It involves finding nearest highway nodes to given station coordinates, and then calculates shortest path algorithm between all corresponding highway nodes. And there is no connection pick drop time or cost to trucks.

5.4 Reduce Number of Candidate Locations

5.4.1 Gravity algorithm

A heuristic gravity algorithm is implemented to reduce the number of candidate locations for LSC and SHOP. The heuristic gravity algorithm includes two methods, local search method and interchange method. It is stable and convergent since given a set of candidate locations for facilities, the algorithm can always converge to a state that there is no more cost reduction, and outputs same candidate locations for facilities each time it runs.

Local Search Method: We define the set of demand destinations served by each facility as the neighborhood within this facility; then for each node which is a candidate location for LSC or SHOP location within the neighborhood of this facility, we replace the location of the facility with this node and reassign the demand points to the node if the total cost (including connection dwell cost, travel cost and connection pick drop cost) can be reduced; repeat the same steps for every facility.

Interchange Method: This method scans each of the remaining facilities other than

selected facilities and attempts to replace a selected facility with an unselected facility. If there is a total cost reduction, then do the replacement; otherwise, go to another unselected facility. Keep on doing this until there is no cost reduction for all remaining facilities.

5.4.2 P-Median scaling algorithm

We use p-median based heuristic method to reduce the number of LTHOME locations. The procedures are described as follows:

1. Select one center point of each state randomly and add it to sample LTHOMES. And then solve the following p-median model to get minimal P candidate locations for LTHOMES from sample LTHOMES that can cover all demand locations:

$$\begin{aligned}
& \min && D \\
& \text{s.t.} && \sum_{j \in N_h} x_j = P \\
& && y_{i,j} \leq x_j, \forall 1 \leq i \leq N_h, \forall 1 \leq j \leq N_d \\
& && \sum_{i \in N_h} y_{i,j} = 1, \forall 1 \leq j \leq N_d \\
& && \sum_{i \in N_h} d_{i,j} y_{i,j} \leq D, \forall 1 \leq j \leq N_d \\
& && x_i, y_{i,j} \in \{0, 1\}
\end{aligned}$$

where P is the size of candidate locations for LTHOME of P-median problem, N_h is the size of total candidate locations for LTHOME and N_d is the size of total demand locations. x_i is 1 if we locate a LTHOME at i and 0 otherwise. $y_{i,j}$ is 1 if demand j is covered by a LTHOME at i and 0 otherwise. $d_{i,j}$ is the travel cost from LTHOME i to demand location j .

2. Solve step 1 model with sample truck homes to obtain unused LTHOMES.
3. Delete all truck homes in the state where there is an unused LTHOME.

4. Iteratively solve P-median model and step 1 model until there are no unused LTHOMES in step 1 model.

5.4.3 Maximum throughput algorithm

Since mainline fueling station (MFS) has properties that it is cheap, fast to fuel locomotives and it can only provide fuel for trains whose routes includes it, MFS can benefit more trains if more trains pass by it. We define throughput of station i as T_i , T_i is the total number of trains that has station i in route. Therefore, to reduce the number of candidate locations for MFS, we compute the train throughput at each station and choose top-K stations that have largest throughput as the MFS candidate locations.

Chapter 6

Case Study

The proposed model and algorithms are programmed in *C#*, and the decomposed subproblems are solved by Gurobi. All numerical cases are performed on a desktop with 3.07 GHz CPU and 8 GB RAM. There are about 3,000 train trips per week running in a U.S. Class-1 railroad's network (including about 11,000 nodes and links), and detailed information on the shipment schedules and paths is available. These trips generate about 500 work demand points (i.e., rail yards) and 10,000 permitted travel links. In the current railroad network, the company runs about 30 LSCs, 10 SHOPS, 30 LTHOMEs, and 20 MFSs¹. In the case study, a group of input parameters are generated based on realistic data².

The first part of parameters are facility construction cost and shutdown benefit, capacity expansion cost and reduction benefit. The corresponding costs/benefits are shown in Table 6.1. Labor expansion means that no physical machines are needed for expansion; only the labor cost is considered. However, labor and physical expansion requires both labor and physical machines to expand the facility capacity.

Table 6.1: Facility cost/benefit

	Construction	Shut down	Labor expansion per unit	Labor and Physical expansion per unit	Reduction per unit
LSC	600,000	-600,000	70,000	170,000	-170,000
SHOP	800,000	0	70,000	170,000	-170,000
MFS	500,000	0	0	170,000	-170,000
LTHOME	300,000	0	0	0	0

A road locomotive will cost \$40 per hour when it is running, while moving nonroad locomotive incurs twice the hourly cost. Additionally, a cost of \$10,000 every time incurs

¹The full scale implementation has about 10 times as many candidate locations, but details of the data are omitted to protect confidentiality

²Realistic data of CSX railroad is not allowed to protect confidentiality

when there is a pick up or drop off along the route. An extra locomotive cost of \$40 per hour, a crew cost of \$5 per mile, and a fuel cost of \$10 per mile are applied to light train movement expenses. There is a \$1,000 additional penalty for each train stop. Fuel material cost is \$2 per gallon on average but varies based on location.

In addition to various costs related to locomotives, facilities consume capacity to perform work for locomotives and involves work cost. In the model, one unit of capacity for any facility is equal to 24 hours or work per day. A LSC takes 2 hours to perform service work and 26 hours to finish repair or maintenance work for one locomotive, which costs \$500 and \$1000 respectively. A SHOP takes 24 hours for repair/maintenance work for one locomotive and consumes \$1000. The work costs at different facilities are shown in Table 6.2. The work time for each type of work at different facilities is shown in Table 6.3. A "No" label in a cell means that corresponding type of work cannot be done at the facility.

Table 6.2: Work cost at different facilities

	Service	Maintenance	Repair	Fueling
LSC	500	1000	No	1000
Shop	No	1000	1000	No
MFS	No	No	No	1000
Truck	1500	No	No	1000

Table 6.3: Work time (hr) at different facilities

	Service	Maintenance	Repair	Fueling
LSC	2	26	No	1
Shop	No	24	24	No
MFS	No	No	No	0.5
Truck	2	No	No	1

The following case studies will use the parameters specified above as model input and analyze the results. First, a base case is run to provide a baseline for other cases. The first group of cases are to shut down LSCs and analyze the changes in cost components. They contain three subcases including shutting down 1 LSC, shutting down 2 LSCs, and shutting down 3 LSCs. The second group of case allows the model to build new LSCs providing with large number of candidate locations for LSC . By performing analysis between cases, this paper evaluates the correctness and efficiency of the model.

6.1 Base Case

In the base case, all locations, types and capacities of facilities are read from the current realistic network and not allowed to be changed. There are 28 LSCs, 10 SHOPS, 30 LTHOMEs, and 2 MFSs as the input. Our model will allocate demands based on fixed facilities and generate an optimized result. The base case is used to compare with other cases and the base results will be used as input for other cases. It takes 1800s to reach a 1.24% gap and the gap is reduced to 0.18% within 3600s. The location results are shown visually in Figure 6.1. The black circle represents that several types of facility share the same location. Light blue lines in the map are the CSX train network. From the map, we can observe that all facilities are spread across the network and are located along train routes.

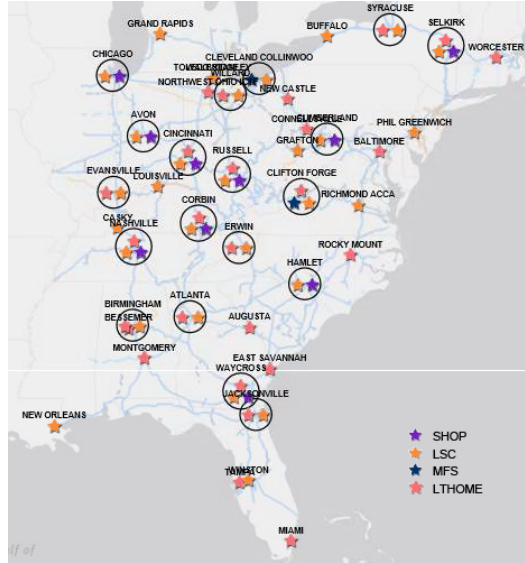


Figure 6.1: Location base result illustration

Results related to LSC and the corresponding capacity at each location are shown in Table 6.4. The capacities will be the baseline for other cases. And Table 6.5 shows similar results related to Shop. There is only one MFS at CLEVELAND COLLINWOO with 1 capacity in the result. Since the model does not allow facilities to shut down in the base case, another MFS at CLIFTON FORGE has 0 capacity. The capacity results are the minimal amount to ensure that all work demand is satisfied and the total cost involved is minimized.

Table 6.4: Base result: LSC location and capacity

City	Capacity	City	Capacity
ATLANTA	3	HAMLET	6
AVON	5	JACKSONVILLE	2
BIRMINGHAM	3	LOUISVILLE	3
BUFFLO	3	NASHVILLE	4
CHICAGO	6	NEW ORLEANS	2
CINCINNATI	4	PHIL GREENWICH	2
CLEVELAND	1	RICHMOND	4
CLFITON	1	RUSSEL	5
CORBIN	2	SELKIRK	6
BUMBERLAND	6	SYRACUSE	3
ERWIN	2	TOLEDO	2
EVANSVILLE	2	WAYCROSS	6
GRAFTON	1	WILLARD	4
GRAND RAPIDS	1	WINSTON	3

Table 6.5: Base result: Shop location and capacity

City	Capacity
AVON	8
CHICAGO	6
CINCINNATI	7
CORBIN	8
CUMBERLAND	20
HAMLET	4
NASHVILLE	9
RUSSELL	12
SELIRK	19
WAYCROSS	27

LTHOME and LT results are shown in Table 6.6. Each LTHOME location can hold different types of LFT or LST depending on which combination of LTs minimize related costs. From the Table 6.6, we can observe that the serving cities can be different to the same type of LT if there are two LTs of the same type at one LTHOME. For example, ATLANTA has two LTs of type LFT168, but the first LFT168 serves two cities, CARTERSVILLE and ATLANTA, but the second LFT 168 only serves ATLANTA. The difference is caused by the demand amount variation between different cities. The first LFT168 can satisfy all demands from CARTERSVILLE but can only satisfy part of demands from ATLANTA. Details about each trip of LTs (e.g., ATLANTA $\rightarrow$ CARTERSVILLE $\rightarrow$ ATLANTA) are not shown in this paper due to the large amount of information about trips. Several examples will be given to illustrate how the truck trips work.

Table 6.6: Base Result: LTHOME locations, LT types and serving cities

City	Truck Type	Serving City
ATLANTA	LFT168	CARTERSVILLE, ATLANTA
	LFT168	ATLANTA
AUGUSTA	LFT168	AUGUSTA, NEWBERRY
	LST120	FLORENCE, ORANGEBURG, AUGUSTA, NEWBERRY, GREENWOOD
BALTIMORE	LFT168	JESSUP, FTGEORGE, HAGERS, BENNING
	LFT168	BRUNSWI, BALTIMORE, JESSUP, FTGEORGE, HANOVER, HAGERS, BENNING
	LST168	BRUNSWI, BALTIMORE, JESSUP, FTGEORGE, HANOVER, HAGERS, BENNING
BESSEMER	LFST168	BESSEMER
BIRMINGHAM	LFT168	DECATUR, CULLMAN, BIRMINGHAM, TALLADEGA, COOSA PINES
	LFT168	BIRMINGHAM
CINCINNATI	LFT168	CINCINNATI TOFC, CINCINNATI
	LST60	WORTHVILLE, CONNERSVILLE, LIMA, MIDDLETOWN
CLIFTON FORGE	LFT168	CLIFTON FORGE, LYNCHBURG
CONNELLSVILLE	LFT60	CONNELLSVILLE
CORBIN	LFT168	CORBIN, MORLEY, KNOXVILLE, BLACKKEY
	LFT60	CORBIN, MORLEY, BLACKKEY
EAST SAVANNAH	LFT120	EAST SAVANNAH
ERWIN	LFT168	KINGSPORT, ERWIN
EVANSVILLE	LFT168(2)	EVANSVILLE
JACKSONVILLE	LFT168	THOMASVILLE, JACKSONVILLE TOFC, BALDWIN, STARKE, TALLAHASSEE
	LFT168	JACKSONVILLE, THOMASVILLE, STARKE, TALLAHASSEE
MIAMI	LFT120	FT LAUDERDALE, MIAMI
	LST120	FT LAUDERDALE, MIAMI
MONTGOMERY	LFT120	MONTGOMERY
NASHVILLE	LFT168	GALLATIN, NASHVILLE, MURFREESBORO, TULLAHOMA, BRUCETON, GUTHRIE
	LFT168	NASHVILLE
NEW CASTLE	LFT168	NEW CASTLE
	LST120	DEMMLER, NEW CASTLE, LORDSTOWN, ALIQUIPPA
NORTHWEST OHIO	LFST168(2)	NORTHWEST OHIO
ROCKY MOUNT	LFT168	ROCKY MOUNT
	LFT168	ROCKY MOUNT, WILSON, GREENVILLE, RALEIGH
RUSSELL	LFT120	HUNTINGTON, DANVILLE, PEACH GREEK, PAINTSVILLE
	LFT168	HUNTINGTON, DANVILLE, ELKRUN, PEACH GREEK, PAINTSVILLE
	LFT168	HUNTINGTON, RUSSELL, DANVILLE, ELKRUN, PAINTSVILLE
	LST120	HUNTINGTON, DANVILLE, ELKRUN, PEACH GREEK, PAINTSVILLE, SHELBY
SELKIRK	LFT168	SELKIRK
SYRACUSE	LFT120	SYRACUSE, WATERTOWN
TAMPA	LFT120	MULBERRY, TAMPA
	LFT168	TAMPA
	LST60	TAFT, ROCKPORT

6.2 Shut Down LSCs

6.2.1 Shut down 1 LSC

By fixing the location of other facilities (SHOP, MFS, LTHOME), we force the model to shut down 1 LSC by adding a user constraint to make the total number of LSCs are 1 less than the current number. The model reaches a gap of 0.44% within 1800s and the gap reduces to 0.32% within 3600s. Fig. 6.2 illustrates the location result on the CSX train network. Since the locations of LTHOMEs are not allowed to change in this case, LTHOMEs are not shown on the map to make it easier to see changes to LSCs. The "U" shapes in the figure represent a capacity change at that location. Observe that only capacities of LSCs at CLEVELAND COLLINWOOD and BUFFLO have a change. Actually, the capacity of CLEVELAND COLLINWOOD reduces from 1 to 0 and capacity of BUFFLO increases from 3 to 4. Therefore, CLEVELAND COLLINWOOD is considered to be shut down and demands assigned to CLEVELAND COLLINWOOD shift to BUFFLO. In addition, WINHAVEN becomes a new serving city to LTHOME at TAMPA. Both of them are not shown on the map, but they are close to WINSTON in Florida. This small change in the

most southern part of the network has no obvious impacts on the total cost.

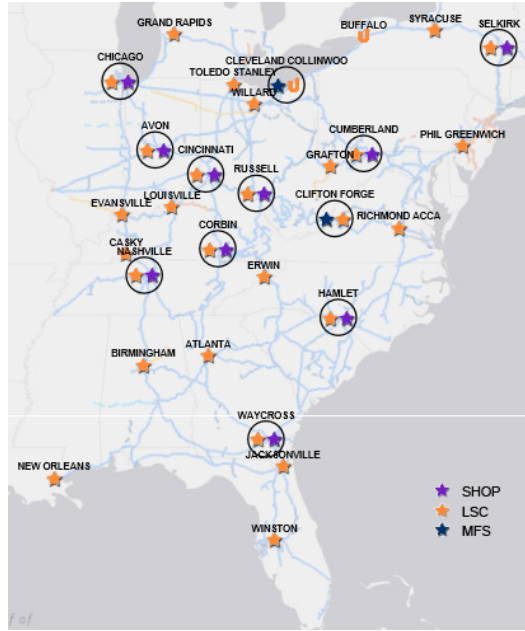


Figure 6.2: Location result illustration when 1 LSC is shut down

Table 6.7 shows the comparison between current case results and base case results. After shutting down 1 LSC, the transportation cost increases by \$1,049,315. And there is one unit LSC capacity expansion cost of \$170,000 and one unit LSC capacity reduction benefit. Therefore, with the current network data as input, the capacity reduction costs of LSCs in this case compared with base case are the same. When a LSC is shut down, it has a benefit of \$60,000. At the same time, fuel total cost (fixed cost and variable cost) decreases by about \$250,000. The rest of costs remain the same with base case since there are no changes in those items. To sum the costs up, current case costs \$630,502 more than base case, which is caused by increased transportation cost. From the comparison, we conclude that transportation cost is most significantly affected in this case.

6.2.2 Shut down 2 LSCs

In order to explore the trend of cost changes, we force the model to shut down 2 LSCs in this case. A 0.36% gap is fulfilled in 1800s, and it decreases to 0.25% after another 1800s. Fig. 6.3 shows the location results for LSC, SHOP and MFS. Model chooses to shut down LSCs

Table 6.7: Cost comparison between base case and current case (Shut down 1 LSC)

Cost Type	Base	Model	Diff
Transportation Cost	\$257,642,419	\$258,691,734	\$1,049,315
SHOP - Capacity Change Cost	-\$1,190,000	-\$1,190,000	\$0
LSC - Capacity Change Cost	-\$5,440,000	-\$5,440,000	\$0
MFS - Capacity Change Cost	-\$850,000	-\$850,000	\$0
SHOP-Construction Cost	\$0	\$0	\$0
LSC-Construction Cost	\$0	-\$60,000	-\$60,000
MFS-Construction Cost	\$0	\$0	\$0
Truck Contracting Cost	\$24,550,000	\$24,550,000	\$0
Truck Operating Cost	\$6,236,514	\$6,231,604	-\$4,910
Fueling Fixed Cost	\$500,890,000	\$500,084,000	-\$806,000
Fueling Variable Cost	\$1,906,483,150	\$1,906,935,247	\$452,097
Total	\$2,688,322,083	\$2,688,952,584	\$630,502

at CLEVELAND COLLINWOO and CLIFTON FORGE and increases capacity of LSCs at BUFFALO and RUSSELL. CLEVELAND COLLINWOO and CLIFTON FORGE both reduce capacity from 1 to 0. BUFFALO increases its capacity from 3 to 4 and RUSSELL raises its capacity from 5 to 6. Meanwhile, MFS at CLIFTON FORGE increases capacity from 0 to 1 to satisfy fueling demand generated at CLIFTON FORGE and surrounding locations. In addition, LTHOME at JACKSONVILLE, located in Florida, adds one more LFT168 and has 3 LFT168s in total.

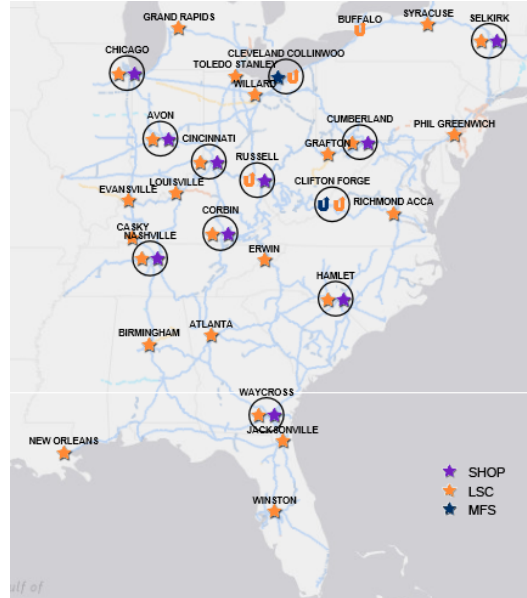


Figure 6.3: Location result illustration when 2 LSCs are shut down

It is obvious that capacities increase at RUSSELL and BUFFALO, because RUSSELL is close to CLIFTON FORGE and BUFFALO is close to CLEVELAND COLLINWOO. But JACKSONVILLE is quite far from those four LSCs. We infer that there is an indirect impact on southern locations from shutting down LSCs from the northern part of the network. More interesting facts will be explored in next case.

The comparison of cost components between base case and current case is shown in Table 6.8. The transportation cost increases by \$2,974,298 after shutting down 2 LSCs, . There is no change in the total capacity of LSCs but MFS expands one unit of capacity which involves a cost of \$170,000. In addition, the model adds one truck at cost of \$550,000 and the truck operating cost increases by \$4,770. When 2 LSCs are shut down, it has a benefit of \$120,000. Meanwhile, fuel total cost has an increase of about \$700,000. In total, current case costs \$4,279,508 more than base case. Transportation cost is still affected significantly.

Table 6.8: Cost comparison between base case and current case (Shut down 2 LSCs)

Cost Type	Base	Model	Diff
Transportation Cost	\$257,642,419	\$260,616,717	\$2,974,298
SHOP - Capacity Change Cost	-\$1,190,000	-\$1,190,000	\$0
LSC - Capacity Change Cost	-\$5,440,000	-\$5,440,000	\$0
MFS - Capacity Change Cost	-\$850,000	-\$680,000	\$170,000
SHOP-Construction Cost	\$0	\$0	\$0
LSC-Construction Cost	\$0	-\$120,000	-\$120,000
MFS-Construction Cost	\$0	\$0	\$0
Truck Contracting Cost	\$24,550,000	\$25,100,000	\$550,000
Truck Operating Cost	\$6,236,514	\$6,241,284	\$4,770
Fueling Fixed Cost	\$500,890,000	\$501,124,000	\$234,000
Fueling Variable Cost	\$1,906,483,150	\$1,906,949,590	\$466,440
Total	\$2,688,322,083	\$2,693,281,591	\$4,279,508

6.2.3 Shut down 3 LSCs

To explore more hidden facts in the model, we force model to shut down 3 LSCs. Within 1800s, the model reaches 1.23% gap and the gap continues to decrease to 0.29% in the following 1800s. The model decides to shut down LSCs at CLEVELAND COLLINWOO, CLIFTON FORGE and JACKSONVILLE, which involves a reduction of 3 capacities in total. In order to compensate for the service and fueling demand assignments to those three LSCs, LSCs at BUFFALO and CORBIN both increase 1 capacity and MFS at CLIFTON

FROGE expands 1 capacity. As predicted from the trend, LTHOMEs in the southern parts should add more serving yards or increase capacities. However, the LTHOME at JACKSONVILLE removes TALLAHASS and THOMASVIL from its serving cities. The reason is that CORBIN attracts the demand flow from the southern part of network to its LSC. We can infer from three cases that one unit change in capacity will have affects in its surrounding demand assignment but the change is usually visualized at the edge region of the network because the demand flow will be relayed continuously from the edge region to changed LSCs.

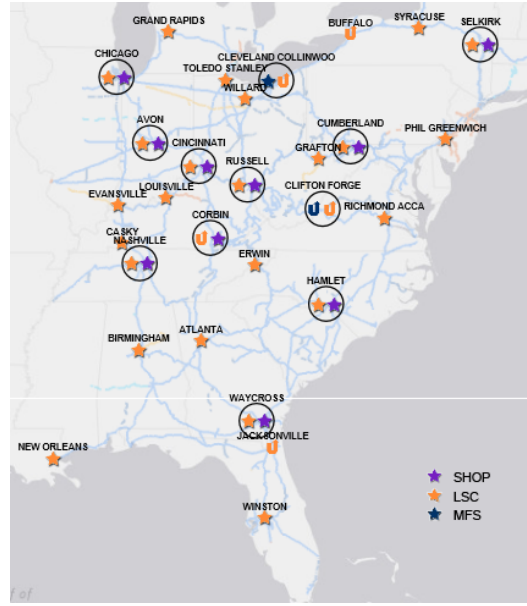


Figure 6.4: Location result illustration when 3 LSCs are shut down

After shutting down 3 LSCs, the transportation cost increases by \$8,431,476 and fueling total cost decreases by \$2,073,428, shown in Table 6.9. The capacity reduction benefit at LSC is \$340,000 ,and the capacity expansion cost at MFS is \$170,000. Trucks operating cost increases \$10,675, which indicates more work is assigned to trucks even though the LTHOME at JACKSONVILLE removes some serving cities. Overall, the total cost increases by \$6,018,723 compared with the base case.

By comparing Table 6.7, 6.8 and 6.9, we can observe that the transportation cost increases as the total number of existing LSCs decreases. There is no obvious trend in fuel total cost and truck operating cost. Meanwhile, the transportation cost change is much

larger than the changes in other cost components.

Table 6.9: Cost comparison between base case and current case (Shut down 3 LSCs)

Cost Type	Base	Model	Diff
Transportation Cost	\$257,642,419	\$266,073,895	\$8,431,476
SHOP - Capacity Change Cost	-\$1,190,000	-\$1,190,000	\$0
LSC - Capacity Change Cost	-\$5,440,000	-\$5,780,000	-\$340,000
MFS - Capacity Change Cost	-\$850,000	-\$680,000	\$170,000
SHOP-Construction Cost	\$0	\$0	\$0
LSC-Construction Cost	\$0	-\$180,000	-\$180,000
MFS-Construction Cost	\$0	\$0	\$0
Truck Contracting Cost	\$24,550,000	\$24,550,000	\$0
Truck Operating Cost	\$6,236,514	\$6,247,189	\$10,675
Fueling Fixed Cost	\$500,890,000	\$499,356,000	-\$1,534,000
Fueling Variable Cost	\$1,906,483,150	\$1,905,943,722	-\$539,428
Total	\$2,688,322,083	\$2,695,870,806	\$6,018,723

6.3 Allow to Add Candidate Locations for LSCs

In this case, about 400 candidate locations of LSCs are provided and 88 locations (including 28 current locations and 60 candidate locations) are the model input after heuristically reducing candidate numbers. The model allows all current LSCs to shut down and change capacities, with no limit to the total number of LSCs, so the transportation cost will be minimized, which is discussed in previous cases that it is the main part of changing cost.

Due to a significant size increase of LSCs, the model reaches a gap of 5% within 1800s and the gap reduces to 2.43% within 3600s. The location results are shown in Fig. 6.5, some adjacent city names are overlapped due to resolution limit of the map. The check marks in the figure represent that a new facility is constructed at that location. There are 54 new LSCs in addition to the 28 original LSCs, detailed information of capacities and locations can be found in the appendix Table A.3. There are no changes in decisions about SHOPS and MFSs.

It can be observed from Table 6.10 that only LSC construction cost increases and all other cost components decrease dramatically. Compared with the saved fueling total cost, transportation cost saving is much smaller. Therefore, fueling total cost is the significant factor that controls the objective cost. Also, truck operating cost decreases, indicating that part of fueling work or service work is transferred from trucks to LSCs. From this case

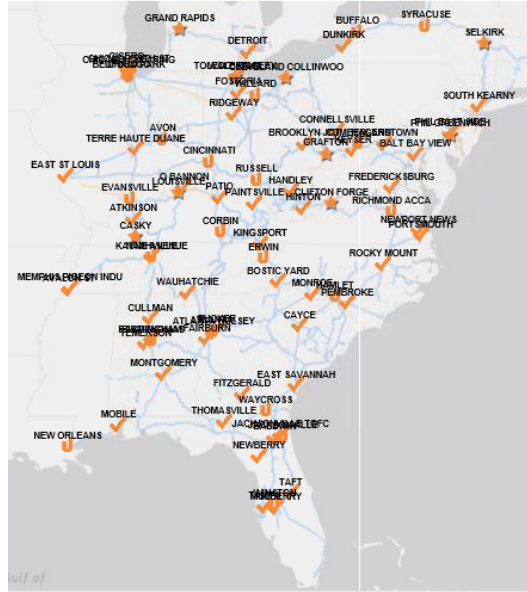


Figure 6.5: Location result when adding candidate locations for LSCs

result, we conclude that it is strongly suggested to sacrifice LSC construction cost to reduce fuel cost and transportation cost.

Table 6.10: Cost comparison between base case and current case

Cost Type	Base	Model	Diff
Transportation Cost	\$257,642,419	\$189,744,111	-\$67,898,308
SHOP - Capacity Change Cost	-\$1,190,000	-\$1,190,000	\$0
LSC - Capacity Change Cost	-\$5,440,000	-\$7,120,000	-\$1,680,000
MFS - Capacity Change Cost	-\$850,000	-\$850,000	\$0
SHOP-Construction Cost	\$0	\$0	\$0
LSC-Construction Cost	\$0	\$27,600,000	\$27,600,000
MFS-Construction Cost	\$0	\$0	\$0
Truck Contracting Cost	\$24,550,000	\$23,900,000	-\$650,000
Truck Operating Cost	\$6,236,514	\$6,056,782	-\$179,732
Fueling Fixed Cost	\$500,890,000	\$313,196,000	-\$187,694,000
Fueling Variable Cost	\$1,906,483,150	\$1,689,113,384	-\$217,369,766
Total	\$2,688,322,083	\$2,240,450,277	-\$447,871,806

Chapter 7

Conclusions

To make planning for locations, work assignments, truck routing trips and fuel strategy for thousands of locomotives in CSX railroad network, this paper has introduced multi-step models and multiple problem solving methods. Firstly, we give an introduction to current problem statement. A background of the current researches related to our problem is then discussed to emphasize the complexity of current problem. That helps readers form a better understanding of the problem and the reason why this paper is doing meaningful and challenging work. Then, to make the model developed in easier but still practical ways, we list several assumptions and some business rules. After that, we develop the formulation of a large-scale linear mixed-integer mathematical model, including step 1 and step 2. We also design a framework consists of several problem solving algorithms, including network consolidation, network generation, reduce candidate location number, free move, truck trip merge, to solve large-scale problem efficiently. Empirical case studies show that the proposed model and algorithms are capable of providing near-optimum solutions effectively. Some of the results have been applied in practice.

Although the results in case study show the efficiency and correctness of our approach, we can improve the models and data processing methods in the future. In the model part, an attempt to merge location inventory subproblem and truck routing problem into one model will be made. Lagrangian Relaxation method will be developed for the complete problem to compare efficiency with the current model. It is possible to integrate our current model with other models to solve more complicated problem. In the part of data processing methods, modifications will be made on current methods to make them faster, and new efficient algorithms will be possibly developed. We will also use some post processing methods to make results closer to optimal solution.

Appendix A

Tables

Table A.1: SHOP capability and work type mapping

SHOP	Work															
	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	L16
S1F1M1W0						Y				Y						
S1F2M1W0					Y	Y			Y	Y						
S1F2M3W0	Y	Y			Y	Y			Y	Y	Y	Y				
S2F2M3W0	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y				
S1F1M1W1						Y				Y					Y	
S1F2M1W1					Y	Y			Y	Y			Y	Y		
S1F2M3W1	Y	Y			Y	Y			Y	Y	Y	Y	Y	Y	Y	Y
S2F2M3W1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table A.2: Truck types, changed serving cities in the case of allowing to add candidates

City	Truck Types	Add Serving Cities	Remove Serving Cities
AUGUSTA	LFT168/LST120	LUGOFF	NO
BALTIMORE	LFT168/LST168	NO	HAGERSTOW
BIRMINGHAM	LFT168	NO	CULLMAN
CORBIN	LFT168/LFT60	NO	BLACKKEY
ERWIN	LFT168	NO	KINGSPORT
JACKSONVILLE	LFT168	NO	BALDWIN, THOMASVIL
MONTGOMERY	LFT120	CALERA, GEORGIANA	NO
NORTHWEST OHIO	LFST168	NO	NO
ROCKY MOUNT	LFT168	APEX, FAYETTEVI	GREENVIL
RUSSELL	LFT120/LFT168/LST120	NO	PAINTSVIL
SYRACUSE	LFT120	LYONS, ROCHESTER	NO
TAMPA	LFT120/LFT168/LST60	WINHAVEN	NO

Table A.3: Locations and capacities of LSCs in the case of allowing to add candidates

City	Capacity	City	Capacity
ATKINSON	2	JACKSONVILLE	1
ATLANTA	1	JACKSONVILLE TOFC	4
ALTANTA HULSEY	2	KAYNE AVENUE	2
AVALON ST	2	KEYSER	2
AVON	3	KINGSPORT	2
BALDWIN	2	LOUISVILLE	3
BALT BAY	2	MEMPHIS	2
BEDFORD PARK	2	MOBILE	2
BIRMINGHAM	1	MONROE	2
BOSTIC	2	MONTGOMERY	2
BROOKLYN	2	MULBERRY	2
BUFFALO	1	NASHVILLE	2
CAYCE	2	NEW ORLEANS	1
CHICAGO	2	NEWBERRY	2
CHICAGO 59TH	2	NEWPORT NEWS	2
CHICAGO CLEAR	2	O BANNON	2
CICERO	2	PAINTSVILLE	2
CINCINNATI	2	PATIO	2
CLEVELAND COL	1	PEMBROKE	2
CLFITON FORGE	1	PHIL EAST	2
CONNELLSVILLE	2	PHIL GREENWICH	2
CORBIN	1	PORTSMOUTH	2
CULLMAN	2	RICHMOND ACCA	1
CUMBERLAND	2	RIDGEWAY	2
DETROIT	2	ROCKY MOUNT	2
DUNKIRK	2	RUSSELL	2
EAST SAVANNAH	2	SELKIRK	6
EAST ST LOUIS	2	SOUTH KEARNY	4
EAST THOMAS	2	SYRACUSE	1
ERWIN	1	TAFT	2
EVANSVILLE	1	TAMPA	2
FAIRBURN	2	TEMERSON	2
FITZGERALD	2	ERRE HAUTE	2
FOSTORIA	6	THOMASVILLE	2
FREDERICKSBURG	2	TOLEDO STANLEY	2
GRAFTON	1	TUCKER	2
GRAND RAPIDS	1	WALBRIDGE	2
HAGERSTOWN	2	WAUHATCHIE	2
HAMLET	3	WAYCROSS	3
HANDLEY	2	WILLARD	1
HINTON	2	WINSTON	1

Appendix B

Model Step 1: Notations and Variables

B.1 Set notation

$\mathcal{N}_\alpha$	set of all locations (candidate & current) for facility type $\alpha \in \mathcal{A}$
$\mathcal{N}_\alpha^0$	set of current locations for facility type $\alpha \in \mathcal{A}$
$\mathcal{A}$	set of facility types, consisting of four sub-type sets $\mathcal{A} = \bigcup_{i \in \{1,2,3,4\}} \mathcal{A}_i$
$\mathcal{B}$	set of work types, consisting of sixteen sub-type sets $\mathcal{B} = \bigcup_{i,j \in \{1,2,3,4\}} \mathcal{B}_{ij}$
$\mathcal{R}$	set of Train ID routes

B.2 Subscript notation

n	node in yard set $\mathcal{N}_\alpha^{(0)}$
α	facility type in set $\mathcal{A}$
β	work type in set $\mathcal{B}$
r	Train ID route in set $\mathcal{R}$

B.3 Parameter notation

F_n^α	Annual fixed cost of operating a facility of type α at candidate location n
F_n^α	Annual benefit of shutting a facility of type α at current location n
B_n^α	Benefit of unit capacity reduction at an facility of type α at yard n
C_n^α	Cost of unit capacity expansion at an facility of type α at yard n
F_l	Fixed cost of type l LT (including contract fee and driver labor fee)
FV_l	Variable cost of type l LT
$Q_n^{\alpha,0}$	Current capacity of facility type α located at n

$Q_n^{\alpha,\max}$	Maximum allowed capacity of facility type α at yard n
$Q_n^{\alpha,\min}$	Minimum allowed capacity of facility type α at yard n
Q_n^l	Maximum number of LT of type l at yard n
$H_n^{l,0}$	Current number of LT of type l at yard n
C_{n_1,n_2}	Transportation cost for moving locomotive from yard n_1 to yard n_2
T_{n_1,n_2}	Travel time of LT from yard n_1 to yard n_2
$W^{\alpha,\beta}$	Working cost of doing work type β in facility type α .
$T_{\alpha,\beta}$	Time for facility type α to do unit work of work type β
D_n^β	Weekly demand of work type β for locomotives from yard n
$D_{\max}$	Maximum number of locomotives can be serviced by one truck trip
$H_{\max}$	Maximum amount of fuel can be purchased by one truck trip
$T_l^{\max}$	Maximum working hours of weekly trips of LT type l
$T_l^{\min}$	Minimum working hours of weekly trips of LT type l
T_{re}^l	The reservicing/refueling time per truck per trip for LT type l .
E_r	Travel frequency of locomotives in path r
N_r	Number of candidate mainline fuel stations in locomotive path r
Q_r	Capacity of tank of locomotives in path r
U_r^s	Fuel consumption amount from $(s-1)$ th station to s th station in path r
$G_{r,n}^s$	Parameters for whether a station n is the s th station passed by locomotive path r , $s = 1, 2, \dots, N_r$, $s = 0$ denotes the departing station
C_n	Variable cost of regular fueling at node n (including fuel price, pumping time cost, etc.)
C_n^0	Fixed cost of regular fueling at node n
CV	Variable cost of vendor fueling
CV^0	Fixed cost of vendor fueling
$K_{\alpha,\beta}$	Parameters for whether facility type α can do work type β , $K_{\alpha,\beta} = 1$ if α can do work type β , 0 otherwise

B.4 Decision variable notation

x_n^α	Decision variable for whether a facility of type $\alpha \in \mathcal{A}$ is open at $n \in \mathcal{N}_\alpha$. For an
--------------	--

	current facility of type α at $n \in \mathcal{N}_\alpha^0$, $x_n^\alpha = 0$ if it is closed, 0 otherwise
q_n^α	Selected capacity of the facility of type $\alpha \in \mathcal{A}$ at yard $n \in \mathcal{N}_\alpha$
$q_n^{\alpha,+}$	Expanded capacity of facility of type $\alpha \in \mathcal{A}$ at yard $n \in \mathcal{N}_\alpha$, $q_n^{\alpha,+} = 0$ if capacity is not expanded
$q_n^{\alpha,-}$	Reduced capacity of facility of type $\alpha \in \mathcal{A}$ at yard $n \in \mathcal{N}_\alpha$, $q_n^{\alpha,-} = 0$ if capacity is not reduced
$y_{n_1,\beta}^{n,\alpha}$	The portion of work demand of type β from yard n_1 serviced by facility of type α at yard n .
$d_{r,s}^{n,\alpha}$	Number of locomotives at the s th station of route r that receive fueling from facility type α at yard n .
$db_{r,s}^{n,\alpha}$	Binary: Whether locomotives at the s th station of route r will receive fueling from facility type α at yard n .
h_n^l	Number of LT of type l at yard n
Δh_n^l	Incremental number of LT of type l at yard n
t_n^l	Total time of LT of type l at yard n
t_{n,n_1}^l	Time of LT of type l at yard n serving yard n_1
z_{n,n_1}^l	Whether demand yard n_1 will be serviced by LT of type l from LT Home n
p_r^s	Decision variables for whether a locomotive stops at s th station in path r for regular fueling
h_r^s	Amount of regular fueling purchased at s th station in path r
pv_r^s	Decision variables for whether a locomotive stops at s th station in path r for vendor fueling
v_r^s	Amount of vendor fueling purchased at s th station in path r
$w_{r,s}^{n,\alpha}$	Amount of regular fueling purchase by locomotive at the s th station of route r which receives fueling from facility type α at yard n . Equal to h_r^s if $db_{r,s}^{n,\alpha} = 1$.
h_n	Fuel level at yard n if n is the origin of some train IDs
h_r	Destination arriving fuel level of train ID r

Appendix C

Model Step 2: Notation and Variables

C.1 Parameters

P_r^s	Whether a locomotive stops at sth station in path r for regular fueling
H_r^s	Amount of regular fueling purchased at sth station in path r
$D_{r,s}^l$	Number of locomotives at the sth station of route r fueled Truck type l at the home location

C.2 Decision Variables

h^l	Number of trucks of type l
t^l	Total time of trucks of type l
t^m	Total time of truck trip m
$z^{m,l}$	Decision variables for whether m th trip is operated by truck type l
$w^{m,l}$	Total time of trip m operated by truck type l . Equal to t^m when $z^{m,l} = 1$
$u_{n_1}^m$	Decision variable for whether the m th trip serves yard n_1 .
$\hat{y}_{n_1,\beta}^{m,l}$	The portion of locomotives from yard n_1 which receive work β from m th trip by truck type l
$\hat{d}_{r,s}^{m,l}$	Number of locomotives at the sth station of route r that receive fueling from m th trip by truck type l
$t_{n \rightarrow n}^l$	Total time of truck type l serving demand of the home location
$s_{n \rightarrow n}^l$	Total time of truck type l serving demand of the home location
$h_{n \rightarrow n}^l$	Total time of truck type l serving demand of the home location
$d_{n \rightarrow n}^l$	Total time of truck type l serving demand of the home location

$\hat{y}_\beta^l$	Portition of work demand of type β at home location assigned to truck type l
$\hat{d}_{r,s}^l$	Number of locomotives at home assigned to truck type l at the s th station of route r

C.3 Extra User Constraints in step 1

[1] Extra 1: Each location has most one LSC/SHOP/MFS.

$$\sum_{\alpha \in \text{LSC/SHOP/MFS}} x_n^\alpha \leq 1, \forall n \quad (\text{C.1})$$

[2] Extra 2: Shop should be with LSC, only effective with `Use_Cons_SHOP_With_LSC` is `True`.

$$\sum_{\alpha \in \text{LSC}} x_n^\alpha = \sum_{\alpha \in \text{SHOP}} x_n^\alpha, \forall n \quad (\text{C.2})$$

[3] Extra 3: Shop should not be with LSC, only effective with `Use_Cons_SHOP_Not_With_LSC` is `True`.

$$\sum_{\alpha \in \text{LSC}} x_n^\alpha + \sum_{\alpha \in \text{SHOP}} x_n^\alpha \leq 1, \forall n \quad (\text{C.3})$$

[4] Extra 4: Reduced spot number should be equal to added spot number (effective for SHOP only), effective with `Use_Cons_Spot_Reduced_Equal_Added` being `True`.

$$\sum_{\alpha \in \text{SHOP}} \sum_n (q_n^{\alpha,+} + q_n^{\alpha,-}) = 0 \quad (\text{C.4})$$

[5] Extra 5: Allow decrease capacity or not (LSC/SHOP).

$$\sum_n q_n^{\alpha,-} = 0, \forall \alpha \in \text{LSC/SHOP} \quad (\text{C.5})$$

[6] Extra 6: Total number of trucks no larger than User input.

$$\sum_{l \in \mathcal{A}} \sum_n h_n^l \leq \sum_{l \in \mathcal{A}} \text{Input}_l, \forall \mathcal{A} = \text{LST/LFT/LFST} \quad (\text{C.6})$$

[7] Extra 7: Move one truck away scenario

$$h_n^l + \Delta h_n^l \leq H_n^{l,0} \quad (\text{C.7})$$

$$\sum_{l \in \text{LT}} \sum_n \Delta h_n^l = 1 \quad (\text{C.8})$$

[8] Extra 8: Only shutting down LSC can reduce capacity and increased capacity for all LSC equals decreased capacity, effective with `Use_Cons_Only_ShutDown_LSC_Reduce_Cap`

is **True**..

$$q_n^{\alpha,-} \leq Q_n^{\alpha,0}(1 - x_n^\alpha) \quad (\text{C.9})$$

[9] Extra 9: Build new truck home at where LSC is shut down, effective with **Use_Cons_ShutDown_1_LSC** is **True**.

$$x_n^{\text{LT}} \leq 1 - x_n^\alpha \quad (\text{C.10})$$

[10] Extra 10: Number of new facilities and shutting down facilities.

[11] Extra 11: Open LT Home should have at least one truck of its current types.

$$M \cdot \sum_{l \in \mathbf{T}} h_n^l \geq \sum_{l \in \mathbf{T}} H_n^{l,0}, \quad \forall \mathcal{A} = \text{LST/LFT/LFST} \quad (\text{C.11})$$

Bibliography

- Amir Ahmadi-Javid and Amir Hossein Seddighi. A location-routing problem with disruption risk. *Transportation Research Part E: Logistics and Transportation Review*, 53:63–82, 2013.
- R. K. Ahuja, J. B. Orlin, S. Pallottino, M. P. Scaparra, and M. G. Scutella. A multiexchange heuristic for the single-source capacitated facility location problem. *Management Science*, 50(6):749–760, 2004.
- José-Manuel Belenguer, Enrique Benavent, Christian Prins, Caroline Prodhon, and Roberto Wolfler Calvo. A branch-and-cut method for the capacitated location-routing problem. *Computers & Operations Research*, 38(6):931–941, 2011.
- Ann Campbell, Lloyd Clarke, Anton Kleywegt, and Martin Savelsbergh. The inventory routing problem. In *Fleet management and logistics*, pages 95–113. Springer, 1998.
- Ann Melissa Campbell and Martin WP Savelsbergh. A decomposition approach for the inventory-routing problem. *Transportation Science*, 38(4):488–502, 2004.
- Mark S Daskin. *Network and discrete location: models, algorithms, and applications*. Wiley-Interscience, New York, 1995.
- Martin Desrochers, Jacques Desrosiers, and Marius Solomon. A new optimization algorithm for the vehicle routing problem with time windows. *Operations research*, 40(2):342–354, 1992.
- Zvi Drezner. *Facility location: a survey of applications and methods*. Springer Verlag, 1995.
- A Geoffrion and R Me Bride. Lagrangean relaxation applied to capacitated facility location problems. *AIIE transactions*, 10(1):40–47, 1978.
- Anton J. Kleywegt, Vijay S. Nori, and Martin WP Savelsbergh. Dynamic programming approximations for a stochastic inventory routing problem. *Transportation Science*, 38(1):42–70, 2004.
- Gilbert Laporte and Yves Nobert. An exact algorithm for minimizing routing and operating costs in depot location. *European Journal of Operational Research*, 6(2):224–226, 1981.
- Gilbert Laporte, Francois Louveaux, and Hélène Mercure. Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, 39(1):71–78, 1989.
- R. Levi and C. Shmoys, D. and Swamy. Lp-based approximation algorithms for capacitated facility locations. *Mathematical Programming*, 131(1):365–379, 2012.

- Seyed Mohammad Nourbakhsh and Yanfeng Ouyang. Optimal fueling strategies for locomotive fleets in railroad networks. *Transportation Research Part B: Methodological*, 44(8):1104–1114, 2010.
- Jossef Perl and Mark S Daskin. A warehouse location-routing problem. *Transportation Research Part B: Methodological*, 19(5):381–396, 1985.
- Samuel Raff. Routing and scheduling of vehicles and crews: The state of the art. *Computers & Operations Research*, 10(2):63–211, 1983.
- Wu Tai-Hsi, Chinyao Low, and Jiunn-Wei Bai. Heuristic solutions to multi-depot location-routing problems. *Computers & Operations Research*, 29(10):1393–1415, 2002.
- Paolo Toth and Daniele Vigo. *The vehicle routing problem*. SIAM, 2001.
- T. J. Van Roy and D. Erlenkotter. A cross decomposition algorithm for capacitated facility location. *Operations Research*, 34:145–163, 1986.
- Harvey M Wagner and Thomson M Whitin. Dynamic version of the economic lot size model. *Management science*, 5(1):89–96, 1958.
- Weijun Xie, Yanfeng Ouyang, and Kamalesh Somani. Planning location and capacity for multiple types of locomotive maintenance shops. *Computer-Aided Civil and Infrastructure Engineering*, *In press*, 2014.