Two-Echelon Facility Location Problems Using Approximation Methods

Mahyar Jahangiriesmaili
M.A.Sc. Candidate
Department of Civil Engineering, University of Toronto
35 St. George Street, Toronto, ON M5S 1A4
Telephone: +1 647 716 9372
Email: mahyar.jahangiriesmaili@mail.utoronto.ca

Sina Bahrami
Ph.D. Candidate
Department of Civil Engineering, University of Toronto
35 St. George Street, Toronto, ON M5S 1A4
Telephone: +1 647 994 4998
Email: sina.bahrami@mail.utoronto.ca

Matthew J. Roorda
Professor
Department of Civil Engineering, University of Toronto
35 St. George Street, Toronto, ON M5S 1A4
Telephone: +1 416 978 5976
Email: roordam@ecf.utoronto.ca

Word count: 184 words abstract + 9 tables/figures x 250 words (each) + 408 words reference +
3,814 words text = 6,656 words

Submission Date: November 13, 2016
ABSTRACT:
Two-echelon delivery structure is a strategy that can be implemented in urban areas to lower the delivery cost by reducing movement of heavy goods vehicles. In a two-echelon delivery structure, large trucks deliver shipments from a consolidation center to several terminals where packages are transferred to smaller trucks for last mile deliveries. This paper formulates a model that solves the two-echelon delivery structure using approximation techniques. We identify several potential terminal locations and demand areas, and we examine the optimal number and locations of the terminals as our model evaluates the most cost effective routings between consolidation center, potential terminals, and demand areas. To assess our model, we chose downtown Toronto as our case study and we performed cost analysis on the number and locations of the terminals. Our experiments show that number and locations of the terminals are greatly influenced by the opening cost of the terminals and transportation cost of the delivery trucks. We also discovered that likelihood of selecting terminals that are positioned near both the consolidation center and center of the service area is higher than any other locations.
INTRODUCTION:
For time sensitive deliveries, freight operators need to ensure that their trucks arrive on-time while dealing with road restrictions, parking restrictions, constrained roadway geometrics, and congestion (1). For this reason, freight operators often prefer to first ship merchandise to inventory facilities (distribution centers or warehouses) in or near cities for transloading shipments from large trucks to smaller, more agile vehicles (1). One of the challenges of this practice is establishing the sufficient number of warehouses at cost effective locations. Once the number and location of inventory facilities have been confirmed, carriers decide on the most economical routes. In operation research, these problems are often referred to as facility location problems.

Facility location problems were introduced by Pierre de Fermat in the mid-17th century (2). Alfred Weber, a German economist, further established the problem in the early 20th century, in what is often referred to as the Fermat-Weber problem (3). Weber defined three nodes on a continuous plane and attempted to find a fourth point on the same plane that was simultaneously close to the first three nodes (3). Weber later formulated this problem to minimize the Euclidean distance from a facility to ‘n’ other locations (2). Today, the Fermat-Weber problem is largely known as the “Uncapacitated Facility Location Problem”.

Uncapacitated Facility Location problems require a set of predefined potential facility locations, which are selected based on the number and position of demand points while trying to minimize the opening and service cost of each facility. This class of problems is further subdivided into Single-Echelon and Multi-Echelon facility location problems. Single-Echelon facility location problems occur where goods are directly delivered from the point of production to point of consumption (4). A typical example of this problem would be the movement of raw materials from the origin points to production plants (4). Multi-Echelon delivery structures, on the other hand, occur when commodities need to go through several intermediate facilities, which are commonly known as terminals, before they arrive at their destinations (4).

This paper focuses on the two-echelon facility location problem. Jacobsen and Madsen (5) first introduced the example of a newspaper company that delivers newspapers from a printing office to its customers (sale points). This company initially dispatches several primary (large) trucks to multiple terminals in order to transload the newspapers to secondary trucks. These trucks are smaller in size, and they carry the newspapers from the terminals to the sale points. This methodology reduces the movement of heavy goods vehicles within the urban areas, especially during rush hours, while it increases accessibility. Jacobsen and Madsen were keen to discover the optimal locations of the terminals and find the most cost effective primary and secondary truck-routes (5). However, since the scope of their problem was extensive, they proposed heuristics methods in order to solve the newspaper distribution problem (6).

Various heuristic approaches have been suggested to solve multi-echelon delivery problems, but most of them are not computationally efficient. Current literature recommends a two level approach (decomposition method) for solving two-echelon delivery problems where the vehicle routing problem (VRP) is applied at two separate levels: one between the consolidation center and the terminals and one between the terminals and the customers. The concern with this technique is that the computational running time increases remarkably as the facility location problem becomes extensive.
This paper uses an approximation method to solve the two-echelon delivery problem for the optimal deployment of terminal locations and the vehicle routing. Approximation methods are less computationally intensive as they solve the VRP between consolidation centers, terminals, and customers in one level. This method is applied to a case study in downtown Toronto. For this case study we perform cost analyses on the number and locations of the terminals, given variation in transportation costs, terminal operating costs and the selection of demand area segmentation.

**LITERATURE REVIEW**

Methods to solve urban freight distribution network problems can be classified into two categories: exact and approximate. The exact approach provides explicit solutions, whereas the approximation method generates results that are not fully determined. An exact solution to a facility location problem determines locations of the terminals and order of deliveries. If the order of the deliveries was left undetermined, results of the problem would be considered to be approximate. Articles that have used exact or approximate methods to solve facility location problems are summarized in Table 1.

Crainic et al. (7) solve a two-echelon delivery structure for a synchronized, scheduled, multi-depot, multiple-tour, heterogeneous vehicle routing problem with time windows (SS-MDMT-VRPTW) by decomposing it into two phases (7). The first phase performs vehicle routing techniques between terminals and customers; the second phase deals with the routings between consolidation centers and terminals (7). The first stage of the decomposition method provides the inputs to the second phase where the number of the required vehicles is determined (7). This model is suitable for noncomplex networks with known travel times between terminals and demand points. A drawback of this method is that it is only applicable to problems with short-term strategic plans that are effective for less than a day.

Crainic et al. (8) address the delivery structure of Crainic et al. (7) and conduct three experiments that provide various measures to obtain the optimal locations of the terminals. The first experiment provides a measure to find the optimal locations of the terminals by adjusting the average transportation cost and consumer distribution (8). The second experiment verifies the locations of the terminals by adjusting the customer and the terminal distributions, and the final experiment provides an estimate of the best terminal locations by relocating the positions of the consolidation center (8). All three experiments apply the same vehicle routing technique that was introduced by Crainic et al. (7). The primary limitation of this approach is that these experiments are mostly applicable to small-sized delivery networks.

Hemmelmayr et al. (9) and Estrada-Romeu and Robusté (10) solve several network problems using the model of Crainic et al. (7). They test various vehicle routing solution methods at each phase of the two-phase problem to obtain the most optimal solution with the shortest computational running time. Their solution methods, however, can only manage problems with limited numbers of customers and terminals.

Campbell (11) tackles the two-echelon facility location problem from a different angle. He partitions the total demand area into compact and non-elongated sub-regions. He then specifies one terminal inside each sub-region to transfer goods from line haul vehicles to local vehicles. Line haul vehicles operate between origin points and terminals, whereas local vehicles circulate between terminals and demand points. Campbell also generates a formula to estimate
the average distance between customers and terminals. His formula, which is based on
Daganzo’s (12) approximation cost method, attempts to optimize the transportation, inventory,
and terminal cost (11). Campbell assumes that line haul trucks are allowed to visit only one
terminal. The shortcoming of this assumption is that the routing between terminals is neglected.

Ouyang and Daganzo (13) believe that an area served by a terminal should be determined
by the location of the adjacent terminals. They assume that the number of terminals is fixed and
that customers are served by the terminals. Then, they move the terminals and divide the area
using the weighted-Voronoi tessellation method until the total cost of the distribution network is
minimized. These methods are fast and provide near optimal solutions. However, application in
real world problems is restricted because of the assumption that cities are continuous flat planes
where terminals can be positioned anywhere.

<table>
<thead>
<tr>
<th>Author(s):</th>
<th>Method</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Crainic et al. (7)                | Exact       | 1- Computationally intensive  
2- Applicable to noncomplex problems with short operating time (e.g. mid-day deliveries)  
3- Requires input of travel time |
| Crainic et al. (8)                | Exact       | 1- Applicable to noncomplex problems  
2- Computationally intensive  
3- Requires input of travel time |
| Hemmelmayr et al. (9) Estrada-Romeu & Robusté (10) | Exact       | 1- Computationally intensive  
2- Functions under limited number of terminals and consumers |
| Campbell (11)                     | Approximation | 1-Provides approximate solution  
2-Ignores vehicle routing between terminals |
| Ouyang and Daganzo (13)           | Approximation | 1-Provides approximate solution  
2-Any location can be a terminal (assumes a continuous plane) |

**METHODOLOGY**

Demand for urban freight fluctuates day by day or hour by hour. Thus, the method to solve the
two-echelon delivery facility location problem needs to be adaptable to the encountered demand
variations. However, since exact methods are computationally challenging and limited by the
size and complexity of the problem, we decided to use an approximate method, recognizing its
limitations in providing a fully determined result.

We develop our model based upon the idea of Campbell (11). We segregate the total
demand area into smaller regions assuming that the density of the consumers is uniform across
each demand area while the shape and the size of each partitioned area are assumed to be
distinct. Nevertheless, we, divide the total demand area into homogenous rectangular sub-regions
which makes the calculation of average interregional distance easier. Potential terminal locations
are then identified. Any sites that allow for the transfer of merchandise from the first level
( primary trucks) to the second level (secondary carriers) are nominated as terminals. Unlike
Ouyang and Daganzo (13), our approach identifies only the locations that can be used as terminals. Due to limited empty space in urban areas, public parking lots are assumed to be used as potential terminals.

Our model evaluates the optimal number and location of the terminals and allocates demand points to the selected terminals. This process minimizes the total logistic cost, which is composed of transportation cost between consolidation centers and terminals, delivery cost between terminals and customers, and terminal opening cost. Each terminal can serve more than one partitioned demand area whereas each demand area can only be served by one particular terminal. If an area is served by a terminal that is not located within that area, that terminal plays a similar role as a consolidation center while the center of the area resembles a terminal; as a result, the delivery cost is calculated in the same way as the delivery cost between the terminals and the consolidation center. Once the total logistic cost has been minimized using integer linear programming, our model generates minimum logistic cost, primary and secondary vehicle (truck) routings, and optimal terminal locations.

**FORMULATION**

The formulation for the two-echelon facility location problem is presented below:

\[ R = \text{set of potential terminal locations and consolidation centers}; \]
\[ K = \text{set of potential terminal locations} \]
\[ L = \text{set of demand areas} \]
\[ c_{ij} = \text{travel cost from facility } i \in R \text{ to facility } j \in R; \]
\[ c'_i = \text{average daily operating cost at terminal } i \in K; \]
\[ L_{ij} = \text{distance between terminal } i \in K \text{ and center of demand area } j \in L; \]
\[ a_s = \text{operating cost of secondary (small) trucks (dollars per mile)}; \]
\[ c_s = \text{capacity of secondary (small) trucks (number of shipments)}; \]
\[ d_j = \text{demand density of demand area } j \in L \text{ (demand points per square mile)}; \]
\[ m_j = \text{number of demand points in demand area } j \in L; \]
\[ E(D_j) = \text{expected distance between demand points and center of demand area } j \in L; \]
\[ M = \text{a large positive constant}; \]
\[ S = \text{a sub tour}; \]

\[ \begin{align*}
\min z &= \sum_{i \in R} \sum_{j \in R} c_{ij} x_{ij} + \sum_{i \in K} c'_i t_i + \sum_{i \in K} \sum_{j \in L} y_{ij} \left[ \frac{a_s}{c_s} \left( 2L_{ij} + \frac{0.57m_j}{\sqrt{d_j}} + 2E(D_j) \right) \right] \\
\sum_{i \in R} x_{ij} &\leq Mt_j \quad \forall j \in K \\
\sum_{i \in R} x_{ij} - \sum_{i \in R} x_{ji} &= 0 \quad \forall j \in R \\
\sum_{i \in S} \sum_{j \in S} x_{ij} &\leq |S| - 1 \quad \forall S \subseteq \{1, ..., n\}
\end{align*} \]
\[ \sum_{i \in K} y_{ij} = 1 \quad \forall j \in L \]  \hspace{1cm} (5)

\[ \sum_{j \in L} y_{ij} \leq M \sum_{j \in R} x_{ji} \quad \forall i \in K \]  \hspace{1cm} (6)

\[ x_{ij} = \begin{cases} 1 & \text{if primary vehicles travel from facility } i \in R \text{ to facility } j \in R \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} (7)

\[ t_i = \begin{cases} 1 & \text{if the terminal is selected at location } i \in K \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} (8)

\[ y_{ij} = \begin{cases} 1 & \text{if area } j \in L \text{ is served by the terminal at location } i \in K \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} (9)

Eq. 1 is the objective function which minimizes the total logistic cost. The first term of the objective function corresponds to the cost of freight movements between terminals and the consolidation center while the second term denotes the terminal daily operating cost. The final term of the objective function approximates the delivery cost of the secondary trucks. This cost is evaluated using the formula that was developed by Campbell (11). The first expression of this formula \((2a_s L_{ij} / c_s)\) calculates the transportation cost between terminal and demand areas, whereas the second expression \((0.57 m_j a_s / \sqrt{d_j c_s} + 2E(D_j) a_s / c_s)\) approximates the cost of door-to-door delivery within each demand area. Constraint 2 ensures that primary vehicles can travel through the selected terminals, and Constraint 3 represents the enclosed loop between the consolidation center and the selected terminals. Constraint 4 ensures that sub tours are eliminated from the feasible solution of primary vehicle tour. Constraint 5 makes sure that each demand area is served by one terminal. Constraint 6 ensures that all demand areas are assigned to selected terminals. Constraints 7, 8, and 9 confirm that the decision variables are binary. In the above formulation, transportation costs are assumed to be linear and terminals are considered to be homogeneous. Figure 1 demonstrates the model graphically.

**FIGURE 1** Graphical illustration of the mathematical formulation
The algorithm employed to solve the optimization problem starts by reading the operating cost of secondary trucks, terminals, and primary trucks. These costs are nested in order to examine different combinations of operating costs. Other inputs such as capacity of secondary trucks, demand size, network data, sub tours, and size of demand areas remain unchanged for every combination of operating cost. The network data consist of number and location of terminals, consolidation center, and demand areas. Since decision variables are restricted to be integers while the objective function and constrains are linear, a branch-and-bound algorithm is employed to solve the integer linear programming (ILP) problem. Output of the algorithm specifies the selected terminals and the truck routings. Solution algorithm of the ILP problem is shown in Figure 2.

**FIGURE 2** Solution algorithm of the ILP problem

**CASE STUDY**

Downtown Toronto is the primary central business district in the City of Toronto, Ontario, the most populated city in Canada (14). Daytime congestion in downtown Toronto is inevitable as more than 500,000 people travel through this region every day (15). The presence of large volumes of vehicles on streets in downtown Toronto increases the delivery time, fuel consumption, and unauthorized truck parking (16). Two-echelon delivery structure is able to moderate the delivery problems in downtown Toronto by minimizing the movement of heavy goods vehicles within congested areas. Given the ongoing delivery challenges and significance of delivery in downtown Toronto, we have selected this area as our case study to solve the two-echelon facility location problem and assess the validity of our model.

The study area of our case study is shown in Figure 3. Boundaries of the total demand area are represented by dash-dotted lines. This area is 6 km by 3 km, and it is bounded by Bloor Street to north, Gardiner Expressway to south, Don Valley Parkway to east, and Dufferin Street to west. This area covers the entire downtown Toronto. We randomly picked out one depot that is outside of the demand area in a commercial/warehousing district of west Toronto and identified that as the consolidation center of our case study. We then nominated 10 public parking lots that are suitable for transloading goods from primary to secondary trucks and...
specified them as our potential terminals. Locations of the consolation center and potential terminals have been geocoded onto the map below. The hollow square represents the consolidation center, and filled circles portray the potential terminals. Each parking lot is defined with an identification number that is shown in Figure 3.

FIGURE 3 Location of consolidation center, potential terminals, and total demand area in the downtown Toronto case study

Cost analysis was conducted by varying the parameters of the model, including the segmentation of total demand area into smaller demand areas, and cost elements. The total demand area is divided into: (a) 2 areas of 3 km by 3 km; (b) 4 areas of 3 km by 1.5 km; and (c) 8 areas of 1.5 km by 1.5 km. Cost analysis is conducted on each area segmentation to determine the optimal number and location of the terminals as the value of the terminal operating cost ($c'_i$), travel cost between terminals and consolidation center ($c_{ij}$), and secondary operating cost ($a_s$) are allowed to vary. All terminals are assigned a single operating cost that varies between $0$ and $5000$, for different scenarios. Travel cost between terminals, consolidation center, and demand areas vary between $0$ per mile and $2.5$ per mile. Customer population of each demand area is randomly generated between 1,000 and 5,000 customers, and customer locations are assumed to be uniform. One primary truck provides service to the selected terminals. Secondary trucks carry no more than 50 packages, and their expected distance from the center of the demand areas to customer locations ($E(D_j)$) is calculated using the formulation of Daganzo (12). The model was coded on MATLAB using a computer with an Intel Core i7-2.40 GHz CPU and 16 GB RAM.

RESULTS

3146 terminal and travel cost combinations were tested for each area segmentation to determine the optimal number of terminals. Figures 4, 5 and 6 represent the optimal number of terminals for different combinations of terminal and travel cost if the number of demand areas were 2, 4, and 8 respectively. Horizontal axes of these figures denote the terminal operating cost, and their vertical axes depict the travel cost/mile between facilities, i.e., the cost/mile between two terminals or the cost between the consolidation center and a terminal. Terminal costs are labeled in multiples of $100$, and the travel costs between facilities are labeled as dollars per mile. Figures 4a, 5a, and 6a display the number of selected terminals if the operating cost of the
secondary trucks is 0.5 ($/mile). Figures 4b, 5b, and 6b represent the number of chosen terminals if the operating cost of the secondary trucks is 1.5 ($/mile). Figures 4c, 5c, and 6c demonstrate the same information once the operating cost of the secondary trucks is 2.5 ($/mile). Darker sections of the figures portray higher number of selected terminals.

**FIGURE 4** Number of terminals with respect to terminal operating cost and travel cost between two facilities when the total demand area is segregated into 2 demand areas and the secondary trucks’ operating costs are (a) 0.5 ($/mile), (b) 1.5 ($/mile), and (c) 2.5 ($/mile)

**FIGURE 5** Number of terminals with respect to terminal operating cost and travel cost between two facilities when the total demand area is segregated into 4 demand areas and the secondary trucks’ operating costs are (a) 0.5 ($/mile), (b) 1.5 ($/mile), and (c) 2.5 ($/mile)
As is shown in Figures 4 to 6, the optimal number of terminals increases as the number of demand areas increases. Since Constraint 5 of the formulation ensures that each demand area is only serviced by one terminal, the maximum number of selected terminals is equal to the number of partitioned areas. Therefore, 2, 4, and 8 are the maximum number of selected terminals for the demand area with 2, 4, and 8 partitions. It is also evident that as the operating cost of the secondary trucks increases, the number of chosen terminals increases, given fixed values of terminal operating cost and travel cost between two facilities. It is efficient to have more terminals near the demand areas if the operating cost of the secondary trucks is high, relative to the transportation cost between facilities. Even though in Figure 6c, operating cost of the secondary trucks is $2.5/mile, one terminal is always selected if the cost of transportation between facilities is greater than $2/mile. Moreover, the optimal number of terminals depends on the terminal operating cost. As the terminal operating cost increases, it is economical to operate fewer terminals. This is apparent in Figures 4, 5, and 6 as the color of the figures becomes pale on the right hand side of the graphs.

Terminal selections were recorded for every cost analysis combination to find optimal locations for terminals. Figure 7 illustrates the selection frequency of each nominated terminal if the total demand area were subdivided into 2, 4, and 8 areas. This figure suggests that Terminals 7 and 10 are the ideal locations for terminal operations as they have the highest selection frequencies. Since terminal operating cost is assumed to be invariable across all terminals, selection of terminal locations relies on the primary and secondary trucks’ transportation cost. As a result, it is cost effective to select terminals that are a short distance away from both the consolidation center and the demand areas, such as Terminals 7 and 10. With the increase in the number of demand areas, the likelihood of choosing more terminals in addition to Terminals 7 and 10 is higher due to Constraint 5 of the developed formulation. However, Figure 7 shows that
the terminal selection frequency of Terminal 10 reduces as the number of demand areas increases. This is because the expected distance between Terminal 10 and center of demand areas grows larger with the increase in number of demand areas, and therefore, other terminals, like Terminals 4 and 7, that are closer to center of the total demand area are selected. The results also show that as the travel cost between facilities decreases, locations of the selected terminals are more likely going to be apart. Figures 8a and 8b represent the selected terminals for the same secondary truck operating cost if the travel cost between two facilities were $0.25/mile and $0.75/mile respectively.

FIGURE 7 Selection frequency of the nominated terminals for different area segmentations

Figure 8a and 8b represent the selected terminals for the same secondary truck operating cost if the travel cost between two facilities were $0.25/mile and $0.75/mile respectively.
**CONCLUSION**

The two-echelon delivery structure enables carriers to operate at a lower cost within the congested urban areas. Carriers ship their deliveries to terminals where packages are unloaded from heavy trucks. Unloaded packages are later carried to customer locations via small trucks, bicycles, or on foot. This reduces road congestion, fuel consumption, number of parking violations, delivery time and emissions and increases accessibility. Despite the limitations of the existing literature, we propose an approximate model that is able to solve large and complex two-echelon facility location problems within a short computational time. The proposed model simultaneously solves the vehicle routing problem between the consolidation center, terminals, and demand areas, and it generates the optimal number and locations of the terminals at the lowest cost. This model is adjustable to demand variation, and it is applicable to mobile distribution center problems where the location of the terminals vary every day.

We test our model on a case study in downtown Toronto. Our analysis indicates that the trade-off between operating cost of the terminals and trucks significantly influences the number and location of terminals. Terminals are positioned near the demand areas when the operating cost of secondary trucks is greater than the operating cost of primary trucks. This is the case for downtown Toronto where congestion is prominent and parking is expensive and limited. Operating cost of terminals also impacts the number and location of the terminals. The cost of opening and operating a terminal in downtown Toronto, where vacant space for transloading shipments is scarce, is costly. Thus, it is economical to locate terminals on the outskirts of congested regions where terminal operating cost is low, while considering the transportation cost from the consolidation centre to the terminals. Our model is able to consider these costs and provide a definitive result.
REFERENCES


