Abstract. This paper introduces the concept of transshipment networks, a collection of strategically located transshipment platforms, for efficient and flexible last-mile delivery in congested urban areas. By implementing transshipment platforms, logistics operators can select the locations, light-freight vehicle types and operating schedules that best fit specific distribution strategies, and, simultaneously, comply with access restrictions and overcome some of the logistics complexity of dense urban zones. A two-echelon location-routing model formulation is proposed, combining a mixed-integer linear programming model with a closed-form continuous approximation of routing cost. The model is applied to a consumer goods distribution case study in Latin America. Results suggest that, given proper fleet type and capacity, transshipment platforms can significantly improve delivery process efficiency.

Keywords: last-mile network design; location-routing; city logistics

1 Introduction

Every city has zones of major economic activity. Often referred as downtown, central business district or city centers, these zones concentrate large amounts of retail establishments, and consequently, drive major levels of freight flows. Reaching these dense zones is both important and difficult for consumer goods manufacturers and logistics operators. They generally account for significant portions of consumers demand. Yet, road and parking infrastructure tends to be highly congested, and local government agencies often impose truck access restrictions, as a measure geared towards reducing freight vehicle externalities, namely traffic congestion, traffic disruptions and pollution.

Nevertheless, the demand for freight trips to urban areas, and, in particular, to dense zones, continues to grow and is getting increasingly complex, due to several factors. As real estate costs increase and space becomes more constrained, retailers will privilege front-of-house space and reduce space for storage to its minimum. From a logistics standpoint, this implies that retailers will need to be replenished more frequently, increasing the flow of freight vehicles into these zones. Furthermore, as consumption patterns change and diversify, companies expand product offerings that need to be
displayed in the store’s limited shelf space. Diversity of products and packaging formats adds complexity to every delivery operation. Finally, the rapid growth of e-commerce and its overlap with traditional sales channels, demand ever increasing amounts of deliveries in different transportation modes to commercial and residential buildings.

Innovative urban logistics solutions are needed to address the unique challenges of logistics operations in congested urban areas. This work introduces a new approach for delivering freight in dense urban zones that fosters light-freight vehicles, leverages existing infrastructure, and minimizes investments needed from public and private sector. In particular, we make the following contributions: 1) we introduce an adapted urban logistics solution, transshipment platforms, inspired by existing industry practices and designed particularly for congested/restricted urban zones; 2) we propose an analytical framework to guide the design of a network of transshipment platforms and explore its performance over a range of operational scenarios; 3) we propose a set of context-based circuity factors that capture specific road network features.

2 Multi-Echelon Distribution Systems in Urban Logistics

Multi-echelon distribution systems have emerged as an alternative to serve complex urban zones. Fundamentally, multi-echelon systems imply using: 1) different freight transportation modes along the delivery route, and 2) intermediate depots or urban logistics spaces to consolidate and transship freight. Such systems allow companies to still leverage economies of scale form larger shipments originated at warehouses in the outskirts of the city, but also to comply with regulations that aim to reduce the environmental and social footprints of logistics operations in dense urban areas [1].

2.1 Traditional Urban Logistic Spaces

In broad terms, urban logistics spaces (ULS) can be defined as dedicated physical areas holding the equipment needed to enable transshipment and/or consolidation of urban freight deliveries. ULS are generally implemented on top of existing infrastructure, such as surface or underground parking lots. Common types of ULS include urban consolidation centers (UCCs) on the district-level and nearby delivery areas (NDAs) on the neighborhood-level.

Financial limitations constitute the major drawback of UCCs. Investment costs are extremely high and very few success histories without major government subsidies have been reported [2], [3]. For carriers, the additional cost of transshipment and changes in operational procedures generally surpass the financial benefits of consolidation [4]. Overall, UCCs can be feasible under very specific conditions such as carriers’ willingness to consolidate shipments and the availability of major public subsidies. Compared to UCCs, NDAs include low investment cost and are relatively easy to replicate. In general, NDAs are better suited for parcel delivery systems [5]. Limited coverage area has been signaled as one of their key limitations [4].
2.2 The Need for Flexible Last-Mile Operations

Additional considerations limit the transferability of UCCs and NDAs, especially to emerging markets. In particular, the prevalence of the traditional retail channel with a large number of so-called nanostores to be served [6] and the limited availability of public infrastructure suggest the need for alternative solutions in these parts of the world.

From an operational perspective, UCCs and NDAs imply outsourcing the last-mile operation to a single operator. This limits companies’ flexibility to establish company-specific operational procedures. Outsourcing to a single operator may thus not always be consistent with the company’s own distribution approach. Furthermore, UCCs and NDAs imply consolidating products across carriers. However, reaching the necessary agreements to enable consolidation across companies is far from trivial, particularly when competing vendors are involved. In emerging markets, consolidation is even more difficult. In the nanostore channel, drivers will frequently collect payments and actively engage in marketing and sales efforts. Therefore, the nature of this channel requires that companies closely oversee the last-mile operation, diminishing the possibility to engage in consolidation efforts.

In terms of infrastructure, traditional ULS have utilized existing road network infrastructure, such as parking lots, to reduce the allocation of resources needed. Still, allocating some of this infrastructure solely for logistics purposes might not be feasible given the highly congested infrastructure in these zones, aggravated by the limited attention to freight from policy makers in emerging markets [7].

2.3 Transshipment Platforms

We introduce the concept of transshipment or micro-deconsolidation platforms, defined as shared urban logistic spaces to enable freight transfer between large trucks and smaller, light-freight vehicles to access congested/restricted urban areas. As any other urban logistics spaces, transshipment platforms (TPs) aim to reduce the number of and distance traveled by large freight vehicles, and consequently, reduce the social and environmental footprint of freight operations. Nonetheless, TPs ‘share’ infrastructure (off-street/on-street parking lots) generally used for passenger vehicles during specific time windows. Also, we emphasize the network perspective of TPs, as we envision a collection of feasible spaces to be chosen based on evolving business needs and operational conditions.

TPs provide greater flexibility to logistics operators. No outsourcing is imposed and companies can choose the locations, vehicle types and operating times that best fit their logistics strategies. TPs solely require simple rental agreements with parking lot proprietors. Financially, companies will need to cover the cost of renting the parking spaces, and will have to invest in light-freight vehicles. Lastly, infrastructure requirements of TPs are minimal: space for vehicles loading/unloading with basic weather protection.

A few challenges of TPs need to be addressed. Companies operating TPs need to devise proper coordination mechanisms between drivers. Furthermore, in some cases
companies also have to implement means to minimize the risks of product theft and product damage due to additional manual handling.

3 Modeling Framework

This section introduces a mathematical formulation for the problem of locating and operating transshipment platforms, hereafter referred as the transshipment network problem (TNP). The model presented in this paper was adapted from a two-echelon capacitated location-routing problem (2E-CLRP) formulation developed by Winkenbach, Kleindorfer, & Spinler [8]. In general terms, this mixed-integer linear programming model determines the cost-optimal network configuration for last mile delivery by obtaining: 1) the location of a city distribution center (CDC); 2) the number and location of intermediate depots (IDs) given a set of candidate locations; and 3) the optimal fleet configuration (size and vehicle types) to serve a specific urban area or district.

3.1 Location-Routing Models Applied to Urban Logistics

Location-routing problems (LRPs) seek to determine the network configuration, i.e. location of de-/consolidation facilities and vehicle routing sequences, that minimizes the fixed and operating cost of vehicles and facilities, and satisfies customer demand and system constraints, which are generally related to vehicle/facility capacity or service time [9]. The literature on LRPs is fairly extensive and mostly focused on single-echelon distribution systems. Several surveys on LRP developments have been made available over time including [10], [11], and more recently [12], [13] and [14].

Recently, specific applications of location-routing models to city logistics contexts have been proposed. In [15], authors explore location-routing extensions for the two-echelon city logistics distribution systems described in [1], by introducing three mixed-integer programming formulations. This work is extended with metaheuristic procedures to solve larger problem instances in [16], [17], and [18].

Still, due to size and complexity, real-world applications of LRPs are computationally infeasible for such metaheuristic procedures. Different solution approaches are needed, that reduce the computational cost of the problem, particularly the routing component, but still provide near-optimal results [19] [14]. Winkenbach et al. [8] introduce an augmented routing cost estimation (ARCE) formula, which expands previous analytical approximations introduced in [19] and [20], to account for heterogeneous vehicle capacity and service time constraints. Their approximation proves to significantly reduce the computational times while maintaining the quality of the solution within a 10% of those provided by tabu-search metaheuristics [8].

Furthermore, to the best of our knowledge, only in [8] a real industry problem is addressed. The authors propose a mixed-integer linear programming formulation for the capacitated 2E-LRP (2E-CLRP), along with a 2-stage solution heuristic and a closed form approximation for optimal routing cost. The model is applied to a postal service operation in two citywide case studies. Their results suggest that network design considerations are heavily dependent on specific demand characteristics of the city, and
on targeted service levels; motivating therefore the need for analytical frameworks adapted to real-world city logistics decisions. The authors also suggest comparing direct versus multi-echelon networks in specific city segments, to better accommodate differences in demand patterns and urban conditions. Our paper contributes to the latter research gap.

3.2 Problem Statement

In general terms, we seek to determine the optimal network configuration (i.e. number, location and capacity of transshipment platforms), fleet composition and routing guidelines to serve critical urban zones. These zones can be served using two competing, but non-mutually exclusive network configurations: direct and multi-echelon delivery. In the direct delivery configuration, customers are served directly from a distribution center (DC) located in the city outskirts using large freight vehicles (trucks) with homogeneous capacity. This is the existing network configuration.

In the transshipment, multi-echelon network configuration, trucks are used to serve TPs located in the surroundings of the zone from the DC (first echelon). At the TPs, freight is deconsolidated and transferred to light-freight (and possibly less-polluting) vehicles, suitable for dense urban environments, to reach customers or other transshipment spaces (additional echelons). These types of vehicles include vans, motorized/non-motorized 2-3 wheelers, and handcarts for pedestrian deliveries. Using the notation introduced by [10] and extended by [15], the transshipment network configuration can be represented as a $3/R/\bar{T}$ system, whereas the direct delivery network is represented as a $2/T$ system.

3.3 Mathematical Formulation

The TNP considers dense urban districts with intense logistics flows. The area size ranges from approximately one to two square kilometers, partitioned in equally sized segments. The TNP assumes static and deterministic demand within the area. We also assume that locations for TPs are available within the district. Similarly, at each available location, space is constrained but sufficient for vehicle parking and transshipment operations. Furthermore, the model assumes different speed ranges per vehicle type and a maximum daily service time.

The 2E-CLRP introduced in this section is inspired by the formulation proposed in [8]. Differences arise mostly from problem settings and the consideration of road-network features through accessibility and road circuity factors. Tables 1 – 5 in the Appendix summarize the notation used for model parameters and variables.

Accessibility factor $k^a$ and road circuity factor $k_v'$ are scalar multipliers that adjust city–level ($L_1$ or $L_2$) and district-level ($L_4$) distance estimations respectively, to account for road network features that impact travel directness. These features include road density, urban form and regulations, and need to be captured for each case, at city and district levels [21]. In addition, the parameter $k$ is another dimensionless, norm-specific multiplier used to estimate tour distances [20].

Two sets of binary decision variables are defined:
\[ X_i \begin{cases} 
1 & \text{if location } i \text{ accommodates a TP} \\
0 & \text{otherwise} 
\end{cases} \\
Y_{i,s,v} \begin{cases} 
1 & \text{if TP located in } i \text{ serves city segment } s \text{ using vehicle type } v \\
0 & \text{otherwise} 
\end{cases} \\
\]

The objective function accounts for the total daily distribution cost and includes the following components: the fixed cost of enabling a TP \( (1) \), the transportation cost from the DC to TPs, or first-echelon distribution cost \( (2) \); the transportation cost from TPs to customers, or second-echelon distribution cost \( (4) \), and the cost of physical capacity utilization at each TP location \( (11) \).

\[
K^F = \sum_i c_i^F X_i \\
K^U = \sum_i X_i \left[ 2 \frac{k^F t^w}{\sigma^F} (w + c^{O,v}) \right] + (N)w + N t^a w + (N)c^{F,U}.
\]

\[
N = \sum_i X_i.
\]

\[
K^V = \sum_i \sum_s \sum_v Y_{i,s,v} f_{i,s,v}.
\]

\[
f_{i,s,v} = q_{i,s,v} m_{i,s,v} t^s_w + q_{i,s,v} n_{i,s,v} t^s_w + \]

\[
q_{i,s,v} m_{i,s,v} n_{i,s,v} t^s_w (w + c^{O,v}) + q_{i,s,v} m_{i,s,v} \left[ 2 \frac{k^V r_{i,v}}{\sigma_v} (w + c^{O,v}) \right] + q_{i,s,v} c^{p,v}.
\]

\[
t^s_p = \xi^s_v \left( \frac{k^s_p}{\sigma_v^p Y_s} \right) + t^w_v + t^p_v.
\]

\[
\delta_{i,s,v} = \frac{t^s_p + t^w_v + 2k^V r_{i,v}}{\sigma_v^p}.
\]

\[
n_{i,s,v} = \xi^s_v \min \left[ 1, \delta_{i,s,v} \right].
\]

\[
m_{i,s,v} = \max \left[ 1, \delta_{i,s,v} \right].
\]

\[
q_{i,s,v} = \frac{\gamma s}{\xi^s_v \delta_{i,s,v}}.
\]

\[
K^G = \sum_i c_i^F \left[ \sum_s \sum_v Y_{i,s,v} (f_{i,s,v} Y_s) \right] + \varphi^{U}.
\]

Equations 5-10 represent the ARCE method from [8]. By obtaining \( t^s_p \), the average tour time needed to serve a segment \( s \) using a specific vehicle type \( v \) (6), the method calculates \( \delta_{i,s,v} \), the number of full-load tour a vehicle departing from a TP at \( i \) can complete within the maximum service time (MST) \( (7) \). Then, it is possible to obtain the average number of customers served per tour \( n_{i,s,v} \) \( (8) \), the average number of tours needed \( m_{i,s,v} \) \( (9) \), and the average fleet size needed \( q_{i,s,v} \) \( (10) \) given vehicle capacity constraints, for vehicle type \( v \) serving segment \( s \) from a TP at \( i \). Finally, we obtain the average total distribution cost to serve \( s \) from a TP at \( i \) using vehicle type \( v \), \( f_{i,s,v} \) \( (5) \).

Then, the mixed-integer linear programming model can be formulated as follows:

\[
\min K = K^F + K^U + K^V + K^G
\]

Subject to

\[
\sum_i \sum_v Y_{i,s,v} = 1 \quad \forall \ s.
\]
\[ \sum_v Y_{i,s,v} \leq X_i \quad \forall i, s . \quad (14) \]
\[ Y_{0,s,v} = 0 \quad \forall s, \forall v \neq 1 . \quad (15) \]
\[ Y_{i,s,1} = 0 \quad \forall i \neq 0, \forall s . \quad (16) \]
\[ \sum_v \sum_{i,s,v} Y_{i,s,v} q_{i,s,v} \varphi_{v,s} \leq (\varepsilon_i - \varphi_i^U)X_i \quad \forall i . \quad (17) \]
\[ \sum_i X_i \geq X^L . \quad (18) \]
\[ Y_{i,s,v}, X_i \in \{0,1\} . \quad (19) \]

Constraints (13) ensure that all segments are served and constraints (14) restrict the allocation of segments and vehicles to active TPs only. A fictional TP located at the DC, location \( i = 0 \), is used to model the direct delivery network. Therefore, only trucks are available at this location (15) and are not available at all other TP candidate locations (16). Constraints (17) restrict the space usage in every TP and constraint (18) imposes a lower bound on the number of active TPs, if needed.

4 Case Study

We explored the impact of transshipment platforms to serve the highly congested historic center of Bogota, Colombia, for a large consumer packaged goods (CPG) company. The selected district covered a one-square-kilometer area and hosted 160 out of its 42,000 customers across the city. The company operates in this multifaceted zone on a daily basis. The district was divided in 4 segments, each with the same store density, i.e. \( \gamma_i = 40 \).

Most model parameters were estimated from company data or previous research. Additionally, for this case study we set \( k^a = 1 \) since the real distances between the DC and the districts were used. Also, the literature suggests \( k \approx 1.15 \) if the number of stores is significantly larger than the number of tours needed to serve the area, which is assumed to apply to most TNPs [20]. Finally, the road circuity factors per vehicle type, \( k'_{v,i} \), were obtained as described in [10]. For trucks, vans, bicycles, and handcarts we obtained values of 1.83, 1.83, 1.45, and 1.12, respectively.

The 2E-CLRP formulation stated in section 3.3 was implemented in PYTHON 2.7 and solved using GUROBI 5.6 x64 on an Intel ® Core ™ i7-3540 M CPU @ 3.0 GHz and 8.0 GB of physical memory running a 64-bit operating system.

4.2 Results and Discussion

Results for a baseline scenario suggest daily distribution cost reductions between 20% and 45%, depending on the vehicle type used, if the transshipment network configuration is chosen. Given mainly its low fixed costs, bicycles represent the most cost-effective option. In this case, a fleet size of approximately three bicycles should suffice to serve the area using only one TP on a daily basis.
However, these results are highly dependent on the density of the segment. For lower customer density ranges (0 – 50/km²), the direct delivery network using trucks is the most cost-effective option (Figure 1). Considering that most areas in the city do not exhibit such large customer density levels as observed in downtown (160/km²), our results validate the hypothesis that TPs are primarily suited for highly dense and congested zones.

Fig. 1. Optimal fleet type selection for each segment over a range of customer density values. For lower density ranges, trucks represent the optimal solution (red line), whereas for denser segments, bicycles constitute the most cost-effective solution (black line).

We conducted sensitivity analyses on several key parameters to further inform the design of the network. To remain the most cost-effective option, bicycle capacity should at least accommodate the average demand of eight customers (or approximately 70 kg.), which can be easily obtained given current cargo-bike design standards. Also, the cost optimality of bicycles proves to be robust to variability in the vehicle cost parameters. For example, only after a threefold increase in the fixed operational cost we assume for bicycles, vans become the optimal vehicle choice. Finally, considering the TP capacity assumption of four loading/unloading spaces at each TP, the optimal decision to open only one TP to serve the area did not change after testing several store concentration scenarios, nor after adding a second one-square-kilometer district of similar characteristics.

Finally, if road circuity factors would have been excluded, the store density range in which trucks are selected as the optimal fleet type is twice as large. This significant difference in the results due to the inclusion of road circuity factors highlights the importance of incorporating context-based parameters and opens an interesting avenue for further research.

5 Conclusions

This paper introduces a solution for last-mile delivery in dense urban areas. For companies, transshipment platforms offer operational efficiency and flexibility, overcoming several of the limitations observed in other urban logistics solutions. For
cities, this solution reduces the amount of kilometers traveled by large freight vehicles, stimulates the use of light-freight vehicles, leverages existing infrastructure and does not rely on intense public co-funding.

The location-routing formulation for the TNP presented in this paper is inspired by previous contributions and combines a mixed-integer linear programming formulation with continuous approximation methods. Specific contributions of this paper include a set of factors to capture urban context features such as road density, access, and directionality constraints.

In the case study we explore, results suggest that for urban zones with high retail density, bicycle deliveries are the most cost effective delivery option, mostly because of fixed cost considerations. Also, proper capacity levels for vehicles and transshipment platforms are important to keep the solution robust and need to be determined on a case-by-case basis. The solution is less sensitive to factors such as the concentration of stores and the location of the transshipment centers.

The research problem presented in this paper can be extended in multiple ways. First, this model could be scaled to an entire city operation to further explore the impact of different density levels and urban conditions in the network configuration. Second, several of the parameters in the model could be refined such as parking, service and loading times per vehicle type, using data from vehicle telematics or GPS. Finally, the modeling framework could also incorporate demand variability and uncertainty using robust optimization or simulation techniques.

Acknowledgments. We are grateful to Dr. Christopher Mejía, from LOGYCA/Research for engaging the company that facilitated the case study.

Appendix

Table 1. Objective function components

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>Total delivery cost per day</td>
</tr>
<tr>
<td>$K^F$</td>
<td>Facilities fixed cost per day</td>
</tr>
<tr>
<td>$K^U$</td>
<td>First echelon (1E) transportation cost per day</td>
</tr>
<tr>
<td>$K^V$</td>
<td>Second echelon (2E) transportation cost per day</td>
</tr>
<tr>
<td>$K^G$</td>
<td>Capacity utilization cost per day</td>
</tr>
</tbody>
</table>

Table 2. General model parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_s$</td>
<td>Customer density in city segment $s$</td>
</tr>
<tr>
<td>$\delta_s$</td>
<td>Average customer demand in segment $s$</td>
</tr>
<tr>
<td>$w$</td>
<td>Wage per hour</td>
</tr>
<tr>
<td>$c^F_i$</td>
<td>Fixed cost of enabling a TC at location $i$ per day</td>
</tr>
</tbody>
</table>

Table 3. First echelon parameters and endogenous variables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^D_i$</td>
<td>$L_1$ or $L_2$ norm distance from $DE_0$ to TC at location $i$</td>
</tr>
</tbody>
</table>
Table 4. Second echelon parameters and endogenous variables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{i,s}$</td>
<td>$L_1$ norm distance from TC at location $i$ to the centroid of segment $s$</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>2E speed (within district) for vehicle type $v$</td>
</tr>
<tr>
<td>$k_v$</td>
<td>2E road circuity factor for vehicle type $v$</td>
</tr>
<tr>
<td>$t_a$</td>
<td>Tour factor for 2E routing</td>
</tr>
<tr>
<td>$c^o_{v,u}$</td>
<td>2E vehicle type $v$ operating cost per hour</td>
</tr>
<tr>
<td>$c^f_{v,u}$</td>
<td>2E vehicle type $v$ fixed cost per day</td>
</tr>
<tr>
<td>$f_{i,s,v}$</td>
<td>Average total distribution cost to serve segment $s$ from TC at $i$ using vehicle type $v$</td>
</tr>
<tr>
<td>$m_{i,s,v}$</td>
<td>Average number of 2E tours per vehicle type $v$ starting from TC at location $i$</td>
</tr>
<tr>
<td>$n_{i,s,v}$</td>
<td>Average number of customers served per vehicle type $v$ per tour from TC $i$</td>
</tr>
<tr>
<td>$\delta_{i,s,v}$</td>
<td>Average number of full-load tours a vehicle type $v$ can complete within the MST starting from TC located in $i$ to city segment $s$</td>
</tr>
<tr>
<td>$\xi_v$</td>
<td>Capacity of 2E vehicle type $v$ in terms of average number of stores that can be served</td>
</tr>
<tr>
<td>$q_{i,s,v}$</td>
<td>Number of type $v$ vehicles needed to serve segment $s$ from TC at location $i$</td>
</tr>
<tr>
<td>$t^b_{v,s}$</td>
<td>Average time per tour of vehicle type $v$ in segment $s$</td>
</tr>
<tr>
<td>$t^p_v$</td>
<td>Average operational setup time per tour using vehicle type $v$</td>
</tr>
<tr>
<td>$t^s_v$</td>
<td>Average service time at each customer using vehicle type $v$</td>
</tr>
<tr>
<td>$t^p_{v,s}$</td>
<td>Average time to park vehicle type $v$ at each customer location</td>
</tr>
</tbody>
</table>

Table 5. Physical capacity parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_v$</td>
<td>Space requirement for 2E vehicle type $v$ at a TC</td>
</tr>
<tr>
<td>$\phi^o$</td>
<td>Space requirement for 1E vehicle at a TC</td>
</tr>
<tr>
<td>$e_i$</td>
<td>Cost of unit of space at TC located in $i$</td>
</tr>
<tr>
<td>$e_l$</td>
<td>Physical space capacity at TC located in $i$</td>
</tr>
</tbody>
</table>

References

4. S. Verlinde, C. Macharis and F. Witlox, "How to consolidate urban flow of goods without setting up an urban consolidation centre?," in International Conference on City Logistics, 2012.