Designing integrated urban delivery systems using public transport

Irecis Azcuy a, Niels Agatz b, Ricardo Giesen a,∗

a Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Chile
b Rotterdam School of Management, Erasmus University, Netherlands

A R T I C L E   I N F O
Keywords:
Goods transport by urban transit and last-mile vehicles
Freight on public transport
Urban logistics
Last-mile distribution
Strategic design model
Location-routing problem

A B S T R A C T
The growth of online retail leads to increasing last-mile delivery operations that contribute to various negative externalities, such as traffic congestion and air pollution, especially in urban areas. One way to improve urban delivery operations is to use public transport capacity to move goods to intermediate transfer locations from which they can be delivered by (small) vehicles to the final customers. We study the distance savings that can be achieved by such a two-tier urban delivery system. In particular, we focus on determining which transit stop is best located to be used as a transfer location. We present several special cases to get insights into the transfer location decisions. Moreover, we present a mixed-integer linear programming formulation and a heuristic to solve it. To evaluate the different approaches, we run several computational studies. We also perform a sensitivity analysis to assess the impact of different system parameters on the location decisions and system performance. For very conservative benchmarks, the results show that savings up to 7.1 percent are possible from using public transport capacity to support urban delivery. The savings increase with the distance to the depot, tighter deadlines and customers that are clustered around the transit line.

1. Introduction

Global e-commerce sales amounted to 3.5 trillion dollars in 2019, with an average growth rate of 21 percent over the past five years (Statista, 2020). The Covid-19 pandemic has further accelerated the growth of e-commerce in 2020 (eMarketer, 2020). The proliferation of internet commerce also means that more goods are delivered directly to consumers’ homes. As compared to the consolidated supply of retail stores, the ‘last-mile’ of online retail thus involves many small deliveries to geographically dispersed home locations. As such, home delivery contributes to higher costs (Joerss et al., 2016) and is associated with several negative externalities such as traffic congestion and air pollution (Sisson, 2019). At the same time, online retailers are increasingly offering faster delivery options. Not satisfied with next-day delivery, a growing group of consumers demands same-day delivery (PwC’s, 2018). Tighter delivery deadlines further reduce opportunities for consolidation and exacerbate the negative externalities of urban delivery operations. The negative effects are compounded by the fact that fulfillment centers tend to move away from urban areas to suburban areas to save location-related costs (Dablanc et al., 2014). This so-called ‘logistic sprawl’ is associated with additional truck-miles and vehicle trips to serve urban delivery areas.

One possible solution to improve the efficiency of the urban delivery system is to transport goods using existing public transport capacity. Combining goods and people is common in interurban transportation systems (Cochrane et al., 2017) and airline operations (Derigs et al., 2009). For example, Greyhound Courier Express (Greyhound, 2020) offer drop-off and delivery services in Canada and the United States between bus terminals. In recent years, there have been several initiatives to apply these concepts to

* Corresponding author.
E-mail addresses: ilazcuy@uc.cl (I. Azcuy), nagatz@rsm.nl (N. Agatz), giesen@ing.puc.cl (R. Giesen).

https://doi.org/10.1016/j.tre.2021.102525
Received 23 February 2021; Received in revised form 10 September 2021; Accepted 20 October 2021
Available online 22 November 2021
urban transportation using bus, tram, subway (Schafer, 2003; Kikuta et al., 2012; Trentini et al., 2010; Cochrane et al., 2017), taxi (Li et al., 2014) and water bus (Bruzzone et al., 2021). For example, Go ahead, one of the largest bus operators in London, launched the idea of transporting goods in buses and using their depots as delivery centers (Collingridge, 2017). CITIPOST, a private postal service in northern Germany, uses trams to move mail to mailboxes placed on and off the trams (Cleophas et al., 2019). The main motivation for these initiatives is to reduce traffic congestion and air pollution in urban areas (Agatz et al., 2021). Moreover, sharing the capacity between passenger and goods transportation is potentially more cost-efficient. This is especially relevant since most public transport systems worldwide are subsidized because passenger fees are insufficient to cover the high costs of the system (Sciar et al., 2016). The transportation of goods could provide additional revenues for public transport providers. These revenues could facilitate new opportunities to increase the level of service by increasing public transport frequency and coverage.

In this paper, we consider a stylized transportation system that captures the main features of an integrated urban transportation and public transport system, as seen in practice and proposed in the literature. In particular, we focus on a two-tier distribution system that uses existing public transport routes to move packages into the urban area and subsequently perform deliveries by (small) last-mile vehicles. To facilitate transfers of goods between the transit vehicle and the last-mile vehicle, we consider using automated parcel lockers at designates transit stops. This means that the transit vehicle and last-mile vehicle do not have to be at the transfer location simultaneously. This reduces the complexity and increases the robustness of the system. The transfer locations also facilitate the consolidation of shipments when the delivery vehicle’s capacity and the transit vehicle are not aligned. That is, several goods coming in different transit arrivals can be consolidated to ship last-mile vehicles less often to the delivery area. As these transfer locations would be located at existing transit stops, we refer to them as transfer stations. Therefore, similar to intercity travel flows, we assume that goods move in a dedicated space inside the transit vehicle, see, for example (Cochrane et al., 2017). Once at the transfer stations, the driver or one of the passengers could move the goods to the parcel locker, see for example Gatta et al. (2019).

We focus on the question of which transit stop to use as a transfer station. We study a stylized system for a single depot, a transit line, and a set of potential transfer stations. Several papers focused on routing problems in integrated passenger and goods systems (Fatnassi et al., 2015; Ghilas et al., 2016a; Masson et al., 2017; Pimentel and Alvelos, 2018). Others on network design problems focused on routing freight through the network (Dong et al., 2018; Zhao et al., 2018; Ji et al., 2020; Kizil and Yildiz, 2021). However, we are not aware of any research that integrates hub location and last-mile routing decisions in this context. To fill this gap, we focus on studying location decisions in an urban delivery system from one depot to many customer locations that use idle transit line capacity. With that, we aim to better understand the impact of the transfer location decisions on the operational performance of the system. To minimize the operational handling time and coordination costs associated with transferring the goods, we consider opening a single transfer station on a single line. This provides a logical starting point and building block for studying larger systems. Given the specific characteristics of this system, we believe that larger systems (e.g., with multiple depots, lines, or possible transfer stations) can often be decomposed into smaller sub-problems as urban delivery systems often operate with different service areas corresponding to different depots (Huang et al., 2018).

We can summarize the main contributions of this paper as follows: (i) We introduce a new problem focused on planning the transfer location for an urban delivery system that uses public transport, (ii) We analyze a number of special cases to obtain relevant insights, (iii) We formulate a mixed-integer linear program and provide an efficient heuristic to solve the associated planning problem, (iv) We run an extensive set of experiments for different problem parameters to evaluate the impact of the transfer location on the performance of the system.

The remainder of the paper is structured as follows. In Section 2, we present the literature review. Section 3 formally presents the problem and underlying assumptions. In Section 4, we present a few special cases on a line. In Section 5, we provide an exact and heuristic solution approach. Section 6 describes our experimental study and discuss the results. Finally, we summarize our conclusions and present some possible extensions in Section 7.

2. Related literature

The work in this paper is related to two streams of literature: routing and scheduling using existing transportation flows and network design and facility location. We briefly review both streams below.

2.1. Goods distribution using existing transportation flows

Several papers have focused on route planning decisions related to combining flows of goods and passengers. Masson et al. (2017) study a two-tier distribution system with a depot and a single bus line similar to ours. They focus on finding urban delivery routes that minimize transportation costs given a certain set of transshipment points taking into account synchronization constraints, unloading capacities, and time windows. In a similar vein, Pimentel and Alvelos (2018) consider route planning using multiples transit lines to minimize the total completion time of serving all customers. Chebhi and Chauachi (2016) solve a node routing problem in a Personal Rapid Transit (PRT) network subject to PRT vehicles’ battery capacity. They study a two-tier distribution system, where good and passenger move by PRT and then passenger use a Bicycle sharing system (BSS) and freight move by a city freighter to the final destination. On the other hand, Ghilas et al. (2016a) study route planning for a pickup and delivery problem with time windows using a scheduled transit line. Goods from multiple origins and destinations can move by a scheduled transit vehicle and pickup and delivery routes to and from a known set of transit terminals. They also deal with this problem for uncertain demand through scenario-based planning using Sample Average Approximation in Ghilas et al. (2016b). Other papers have focused on train scheduling and cargo allocation using existing passenger trains (Behiri et al., 2018) or dedicated freight trains (Li et al., 2021).
The use of commuter trips and taxis for moving goods in an urban area has also been studied in several papers. Archetti et al. (2016) study a delivery routing problem in which occasional drivers can make deliveries along their route. Arslan et al. (2019) consider the use of commuter trips for delivery operations given a certain detour flexibility. Chen et al. (2018) also consider transfers between drivers. Similar to our paper, Qi et al. (2018) consider a two-tier distribution system; however, goods are transported to certain transfer locations by trucks and then by private vehicles to the final destinations. Li et al. (2014, 2016a,b) consider using taxis in urban last-mile delivery. The dynamic case is validated through experimental cases using operational data from a taxi company in Tokyo in Nguyen et al. (2015). In the grocery retail context, Paul et al. (2019b,a) consider the dynamic sharing of capacity between scheduled and flexible transportation flows.

2.2. Network design in urban logistics

The facility location problem focuses on the optimal placement of facilities in order to minimize a distance-based objective (Klose and Drexl, 2005). One problem related to our work is the $p$-median problem (PMP), which considers the decision to open $p$ facilities to minimize the weighted average distance between demand and the closest designated facilities. In the context of facility location, the demand that needs to be served from the facility is often not known with certainty but is stochastic (Snyder, 2006). Mirchandani and Odoni (1979) study PMP and deal with uncertainty by minimizing total expected travel cost between facilities and customers across multiple demand realizations. Facility location problems have also been studied in the context of more complex multi-tier systems in which goods are moved from a depot to an intermediate point. These involve decisions on where to locate facilities at one or more tiers, potentially also including routing decisions. For an overview of the multi-tier distribution system, see Ortiz-Astorquiza et al. (2018) and Cattaruzza et al. (2017). Closely related to our setting, there are some recent papers that focus on hub location problems (Dong et al., 2018; Zhao et al., 2018; Ji et al., 2020; Kizil and Yildiz, 2021) in the context of routing freight flows through public transport systems.

Related to our simplified special cases, several have focused on the $p$-facility problem on a line (Jackson et al., 2007; Gastner, 2011). Hsu et al. (1997) study the uncapacitated problem and Brimberg et al. (2001) consider facilities with homogeneous capacity. Larson and Odoni (1981), and Anderson and Fontenot (1992) study a setting which they locate service units on a line in which demand occurs along that line. The aim is to find locations that minimize the expected distance to the demand points. In contrast to our paper, the extant work simplifies vehicle routing decisions in the last mile, e.g., by assuming unit-capacity vehicles or full-truckload operations.

As we consider delivery vehicles that can make multiple stops on a route, our work is also related to the location-routing problem (LRP). The LRP (Schneider and Drexl, 2017) aims to find the set of facility locations and route plans to visit all customer in a way that minimizes cost. It is well known that simultaneously solving location and routing problems leads to better solutions than considering both decisions separately (Salhi and Rand, 1989; Shen and Qi, 2007). Several papers have looked at capacitated facilities (Baldacci et al., 2011; Contardo et al., 2014; Lopes et al., 2016) and uncapacitated facilities (Chien, 1993).

3. Problem statement

We consider a two-tier delivery system where we can use a public transit vehicle, e.g., bus, tram, or subway, to move goods from a depot to a transfer station from which final customers are served with a last-mile delivery vehicle. We consider one depot located at the start of the transit line and customers that are spread uniformly in a delivery area around the transit line. Let $0$ and $J$ represent the depot and the set of customer locations, respectively. All customers must be served before a deadline $T$.

We assume that transit vehicles have a dedicated space reserved for goods that is large enough to ship all goods. This means that the transit-vehicle capacity is not a constraining factor and does not interact with the transit capacity for passengers. We consider a scheduled transit service that stops at $i \in O_T$ transit stops at given times. Transit vehicles have an average operating speed $s$, which includes the dwell time at transit stops. We assume that the transit schedule is not affected by the unloading of goods at the transfer station. Each transit stop can potentially be used as a transfer station.

Last-mile vehicles can depart from the transfer station and also directly from the depot with an average speed of $s$. The set $P$ contains all vehicle departures from the depot, and the set $R$ those from the open transfer station. Each vehicle has to return to the same location that it started from, i.e., the depot or transfer station. We express the vehicle capacity $Q_v$ in terms of the number of customers that can be served in one vehicle trip. This is reasonable for settings in which weight is not a limiting factor, and there is not much variation in package sizes. We consider a single vehicle at the transfer station that can make multiple trips. Let $t_{ij}$ be the travel time between location $i$ and location $j$. Note that the effective deadline for trips from the transfer station must take into account the travel time from the depot to the transfer station ($t_{0i}$) and the transfer time at the transfer station ($r$). By using two sets of vehicles, we can easily incorporate different vehicle sizes and deadlines for each starting location.

We focus on where to open a single transfer station. To evaluate the costs to serve all customers given a certain transfer station location, we create a routing schedule for the last-mile vehicles. The objective is to find a transfer location that minimizes the expected travel distance across different possible demand realizations. In particular, we focus on minimizing the distance-related costs associated with last-mile delivery vehicles. We only consider the distance traveled by the last-mile vehicles as the distance traveled by the public transit system is already fixed and not dependent on the transfer decisions. We ignore the costs of moving goods in transit vehicles. Taking into account both the associated routing decisions and different demand scenarios gives rise to a complex optimization problem, even with one transfer location. The single location problem serves as a starting point for systems with multiple transfer locations.
4. Special cases: Customers on a line

This section focuses on several special cases that provide relevant insights into the factors determining the optimal transfer location. In particular, we study two simplified transit network configurations as depicted in Fig. 1: (a) a line network, and (b) a circular network. We believe that these two cases capture the general structure of many transit networks in practice (Kuo, 2014). To simplify further, we assume all customers are uniformly distributed on the transit line. A depot is located at the beginning of the transit line. Moreover, for these special cases, we ignore capacity restrictions on the last-mile and transit vehicles. Conceptually, this means that the delivery routes from the depot and the transfer station can be represented by a Traveling Salesman Problem (TSP) on a line. Note that the solution approaches as presented in Section 5 and the numerical results in Section 6 consider the more general case in which we incorporate vehicle capacities.

Fig. 1. Special case structures.

The length of the TSP tour starting from the depot is determined by the round trip distance of the furthest customer. For \(|J|\) uniformly distributed customers on a unit-length line segment, the expected tour length is \(E[L]/|J|\). Similarly, we can calculate the tour lengths from the depot and the transfer station within their respective service intervals.

Let \(L_{1/2} = \frac{1}{2}(l + 1)\) be the expected value of the TSP tour for a given interval of length \(l \leq 1\). The expected length of the tour as a function of the probability of having \(m\) customers within the interval of length \(l\) is \(E[L \mid l] = \sum_{m=0}^{\lfloor l/2 \rfloor} \frac{m!}{m!(l-m)!} p^m(1-p)^{l-m}\), where \(p = l\) in the unit-line case. This is the sum of the expected tour length for \(m\) customers multiplied by the probability of having \(m = \{0, \ldots, |J|\}\) customers in interval of length \(l\). Taking the sum over all the possible realizations, we can express the expected tour length for a service interval of length \(l\) as:

\[
E[L \mid l] = \frac{2(1-l)^{|J|} \left(\frac{1}{2} |J| + l - 1\right)}{|J| + 1}
\]

(1)

We can use Eq. (1) to compute the expected travel distance of TSP in an interval on a line and circular network. We consider a transit line segment of unit length. Let \(x\) be the location of the transfer station at \(x\) units from the depot.

4.1. Customers on a line network without deadlines

We can analytically express the expected travel distance on the line network without deadline as follows. The total expected travel distance is the sum of the expected distances of the vehicle starting from the depot and the vehicle starting from the transfer station. All customers that are located \(\leq \frac{x}{2}\) are closer to the depot and hence are served by the depot. All customers that are located at \(\geq \frac{x}{2}\) are closer to the transfer station and thus served by this location. Recall that moving the goods to the transfer station on the public transit vehicle does not result in additional distance.

\[
E[L_{j}] = 2E[L \mid l = \frac{x}{2} + E[L \mid l = 1-x]
\]

(2)

Let \(0, \frac{x}{2}\) be the service interval of the depot and \(\frac{x}{2}, 1\) be the service interval of the transfer station. Therefore, the expected length of a tour that originates from the depot is \(E[L \mid l = \frac{x}{2}]\). As there is a symmetric split at \(\frac{x}{2}\) between the transfer and the depot, we can use the same expression to calculate the expected length of the tour originating from the transfer station to serve customers located between the depot and the transfer station. On the other hand, customers located on the side of the transfer station opposite the depot are served only from the transfer station. In this case, the expected distance for the transfer station is \(E[L \mid l = 1-x]\). Eq. (2) shows the expected travel distance for the TSP on a line \((E[L_{j}])\) as a function of transfer station location \(x\).

**Proposition 1.** The transfer location that minimizes the expected system-wide distances on the line network is located at \(\frac{x}{2}\) of the line segment from the depot.
The derivative of Eq. (2) is defined as
\[
\frac{dE[L_c]}{dx} = 2^{n+1} \left( \left( \frac{1}{2} - x \right)^n - 1 \right) \left( 2 - x \right)^n - 2 \left( \frac{1}{x} \right)^n - 1 \right) x^n
\]
(3)

By taking Eq. (3), and setting it equal to 0 with \( n \geq 1 \), we can find the value of \( x \) that minimizes the total expected distance is at \( \frac{2}{3} \) of the line segment from the depot.

The maximum potential savings of using a transfer location compared to shipping everything directly from the depot depend on the length of the line segment and the number of customers. See Appendix A.1 for the expected savings \( E[S(M)] \) on a line segment for the optimal transfer station location.

**Property 1.** The savings in vehicle distance of deliveries from the transfer station using the transit line as compared to direct deliveries from the depot decrease with an increasing number of customers on the line segment.

**Proof.** Taking the limit of the expected saving function when the number of customers, denoted by \( M \), goes to infinity, we see that the expected saving tends to 0, i.e., \( \lim_{M \to \infty} E[S(M)] = 0 \). As the number of customers per trip increases, the vehicle needs to travel the whole line segment, both in the case with and without using the transit line.

Note that Property 1 only focuses on distance savings. Taking into account the weight of the goods, there could still be savings in terms of emissions from serving goods via the transfer station.

### 4.2. Customers on a circular network without deadlines

For a circular transit network, the depot and transfer station can serve customers on both sides. We denote the transfer station’s location \( x \) as a function of the angle \( \alpha \) between the depot and transfer station as depicted in Fig. 1. The distance between the depot and transfer station traveling on the circle is equal to \( x = ar \), where \( r \) is the circular transit radius. Similarly, we proceed to obtain the analytical expression based on the expected distance in each interval. The expected travel distance for TSP on a line for a circular network \( E(L_c) \) as a function \( a \) is as follows.

\[
E(L_c) = 2E[L \mid l = \frac{ar}{2}] + 2E[L \mid l = \frac{1-ar}{2}]
\]
(4)

Since the last-mile vehicle can move bidirectionally, customers are always split between depot and transfer. The expected distance for customers located in the interval \([0, \frac{\pi}{2}) \) or \([\frac{\pi}{2}, ar] \) is \( E[L \mid l = \frac{ar}{2}] \). On the other hand, the expected distance for the customers located in the interval \([ar, \frac{\pi+ar}{2}) \) or \([\frac{\pi+ar}{2}, 1] \) is \( E[L \mid l = \frac{1-ar}{2}] \).

**Proposition 2.** The transfer location that minimizes the expected system-wide distances in a circular transfer system is at \( \frac{1}{2} \) of the length of the transit line.

**Proof.** The derivative of Eq. (4) is defined as
\[
\frac{dE[L_c]}{dx} = 2^{n+1} \left( \left( \frac{1}{2} - x \right)^n - 1 \right) \left( 2 - x \right)^n + 2^{n+1}(x + 1)^n \left( 1 - 2^n \left( \frac{1}{x + 1} \right)^n \right)
\]
(5)

By taking Eq. (5), and setting it equal to 0, with \( x = ar \), \( r = \frac{1}{2} \) and \( n \geq 1 \), we find that minimizes the expected distance value by \( a^* = \frac{1}{2} \).

If we look at the savings for the circular line, we also confirm Property 1. See Appendix A.1 the expected savings on a circular line at the optimal transfer station location.

### 4.3. Customers on a line network with deadlines

Now, we focus on a system in which all customers must be served before deadline \( T \). Recall we let \( c_l \) and \( c_s \) be the average speed of last-mile and transit vehicles, respectively. We consider a setting in which the effective travel speed on the transit line is lower than by last-mile vehicle from the depot, i.e., \( c_l > c_s \). The effective speed of the transit lines may be lower as it needs to make a detour and typically has multiple intermediate stops that take time. We assume that we can always serve all customers directly from the depot within the deadline. Given these assumptions, the deadline only affects the customer that are served through the transfer station.

If we serve a customer via the transfer station, the effective deadline is \( T \) minus the travel time from the depot to the transfer station via public transport. There are no waiting times to consolidate goods before shipping the last-mile vehicle, as it is an uncapped system. Therefore, there is a maximum distance that can be reached from each potential transfer station within the deadline. For transfer stations further down the line, there is less time available. We express the maximum coverage intervals for a transfer station located at \( x \) as \( x \pm \Delta \), where \( \Delta \) represents the maximum distance from the transfer station within the deadline, expressed as \( (T - \frac{1}{c_s})s_l \).
As we stated before, the coverage range is reduced by moving the transfer station away from the depot because time is restricted. Based on this, we can identify relevant points that show the cases where the total expected system-wide travel distance function changes depending on the transfer station’s location. Fig. 2 illustrates some of these special points in a line network. These relevant points are as follows. $x'$ is the furthest location from the transfer station where we can still split customers located between the depot and the transfer station equally, but customers located between the transfer station and the end of the line are now served also by the depot in the line case. See Fig. 2(a). Note that special point $x$ at that point is equal to $x = \frac{T_s + 1}{2}$. The value of $x'$ at that point is equal to $x' = \frac{T_s}{2}$. $x''$ is the furthest location that can still serve customers on time from the transfer station but affects both sides’ coverage, causing that customers between the depot and the transfer are not served at the same cost. This refers to the line case in Fig. 2(b). Finally, $x^{\text{max}}$ is the transfer station location beyond which customers cannot be served on time from the transfer station. Observe that $x' < x'' < x^{\text{max}}$ and any point beyond $x^{\text{max}}$ can never be a feasible location for the transfer station because all the time is spent moving orders in transit vehicles.

Similarly to the case without deadline, we can derive some structural properties on the optimal transfer station location. Under the assumptions that $T = \frac{1}{s}$, the expected distance function in a line network $\mathbb{E}[L^T]$ is defined by expression (6).

$$
\mathbb{E}[L^T] = \mathbb{E}\left[L \mid l = \max\left\{\frac{x}{2}, x - \Delta\right\}\right] + \mathbb{E}\left[L \mid l = \min\left\{\frac{x}{2}, \Delta\right\}\right] + 
\mathbb{E}\left[L \mid l = \min\left\{\Delta, 1 - x\right\}\right] + \mathbb{E}\left[L \mid l = 1 - \min\left\{x + \Delta, 1\right\}\right] + P(m > 0)(x + \Delta)
$$

(6)

The two first terms in Eq. (6) represent the expected distance to serve customers located between the depot and the transfer station. Here, unlike in the case without time, the expected distance depends on the effective deadline. The last three terms are the expected distance to serve customers located between the transfer station and the depot’s opposite side. The third term, $\mathbb{E}[L \mid l = \min\{\Delta, 1 - x\}]$, is the expected distance for customers served from the transfer station. Finally, the fourth and the fifth terms, represent the expected distance to visit customers from the depot who cannot be served from the transfer station within the deadline. $\mathbb{E}[L \mid l = 1 - \min\{x + \Delta, 1\}]$ is the expected distance to visit customers in the service interval $[x + \Delta, 1]$. Note that a vehicle from the depot has to travel a distance of $x + \Delta$ to reach this interval. This ‘stem’ distance is traveled if at least one customer needs to be visited in this interval. Then, the expected stem distance is $(x + \Delta)$ multiplied by $P(m > 0)$, which is the sum of the binomial probabilities of having at least one customer of the total in this interval.

By solving graphically, we see that the optimal location of the transfer station varies according to the deadline, and speeds of transit and last-mile vehicles. Overall, we observe that including deadlines moves the optimal transfer station closer to the depot as compared to the situation without deadlines.

Property 1 shows that, without deadlines, the expected distance savings decreases with the number of customers in the line segment. We expect this to hold for deadlines as well, but we could not find a closed-form expression for the expected distance savings in this case.

4.4. Customers on a circular network with deadlines

In the circular transit network, we can identify the same relevant points found in Fig. 2. The customers between the depot and the transfer station can still be served by the depot and the transfer, but the potential range from the transfer station is constrained by the deadline. The expected distance in a circular network with deadline $\mathbb{E}[L^T]$ is given as:

$$
\mathbb{E}[L^T] = \mathbb{E}\left[L \mid l = \max\left\{\frac{ar}{2}, ar - \Delta\right\}\right] + \mathbb{E}\left[L \mid l = \min\left\{\frac{ar}{2}, \Delta\right\}\right] + 
\mathbb{E}\left[L \mid l = \min\left\{\Delta, 1 - \frac{ar}{2}\right\}\right] + \mathbb{E}\left[L \mid l = \min\left\{1 - ar - \Delta, \frac{1-ar}{2}\right\}\right]
$$

(7)

The two first terms of Eq. (7) are the expected distance from the depot and the transfer respectively in the arc $ar$. The two last terms are the expected distance from the transfer station and the depot respectively in the arc side $2\pi - ar$, or in this case $1 - ar$. Note that even though the transit line runs in one direction on a circular line, the last-mile vehicle can move bidirectionally, minimizing the expected travel distance.

The insights of the circular network are similar to the line network case. If the transit vehicle is faster than the last-mile vehicle, the optimal location matches the case with no deadline. However, with a restricted deadline such as $T = \pi r$, and transit vehicle slower than last-mile vehicle ($s \leq s_t$), then the optimal transfer location depends on the speed of transit and last-mile vehicles, the deadline, and the radius of the transit line.
5. Solution approach

We first provide a mixed-integer linear programming formulation to model our location-routing problem. Next, we provide a heuristic to solve the problem.

5.1. Mixed-integer linear programming formulation

This section presents a mixed-integer linear programming formulation for our problem. We model the problem on a graph \( G = (V, A) \), where \( V \) is the set of vertices that represent the customer locations \( J \), the depot \( (0) \), and the potential transfer locations \( O_T \) and \( A \) is the set of arcs between all locations. The set of possible vehicle departures is \( K = P \cup R \). Each vehicle departure from the depot represents a vehicle, but it is a single vehicle that can make multiple vehicle departures from the transfer station. We let \( O_T \) define the set of possible transfer stations where a vehicle departure \( k \in R \) can begin. Let \( N \) be the set of scenarios, where each scenario represents a number of \( J \) randomly distributed customer locations. We denote the travel distance from location \( i \) to location \( j \) in scenario \( n \) by \( c_{ij}^n \). Let binary variable \( w_i \) be 1 if transfer station \( i \in O_T \) is open and 0 otherwise. Let binary variable \( x_{ijk}^n \) be 1 if arc \((i, j) \in A \) is used by vehicle departure \( k \in K \) in scenario \( n \in N \), and \( y_{ik}^n \) be a binary variable equal to 1 if the vehicle departure \( k \in K \) starts or visits the node \( i \in V \) in scenario \( n \in N \).

To solve the problem, we compute the expected travel distance using a Sample Average Approximation (SAA) approach (Verweij et al., 2003). To do this, we generate a set of \( N \) demand location scenarios from a uniform distribution. Then, we minimize the expected travel distance across this set of scenarios as shown in Eq. (8).

\[
\min \left \{ \frac{1}{|N|} \sum_{n \in N} \sum_{k \in K} \sum_{(i,j) \in A} \sum_{t \in T} c_{ij}^n x_{ijtk}^n \right \}
\]

\[\forall i \in J, \forall n \in N (9)\]

\[\sum_{k \in K} y_{ik}^n = 1, \forall i \in J, \forall n \in N (10)\]

\[\sum_{i \in O_T} w_i = 1, \forall n \in N (11)\]

\[y_{in}^n \leq w_i, \forall \in O_T, k \in R, \forall n \in N(12)\]

\[\sum_{j \in S^+(i)} x_{ijk}^n - \sum_{j \in S^-(i)} x_{ijk}^n = 0, \forall i \in V, k \in K, \forall n \in N (13)\]

\[x_{ijk}^n = \sum_{j \in S^+(i)} x_{ijk}^n, \forall i \in V, k \in K, \forall n \in N(14)\]

\[\sum_{i \in S, j \notin S, j \neq j} x_{ijk}^n \leq (|S| - 1), \forall S \subseteq V, 2 \leq |S| \leq |V| - 2, k \in K, \forall n \in N (15)\]

\[\sum_{i \in S, j \notin S, j \neq j} x_{ijk}^n \leq Q_k, \forall k \in K, \forall n \in N (16)\]

\[\sum_{(i,j) \in A, i \neq d} r_{ij} x_{ijk}^n \leq T, \forall k \in P, \forall n \in N (17)\]

\[\sum_{k \in K} \sum_{(i,j) \in A, i \neq d} r_{ij} x_{ijk}^n \leq T - \sum_{l \in O_T} (t_{0l} + \tau) y_{lk}^n, \forall n \in N (18)\]

\[y_{ik}^n \leq y_{0k}^n, \forall i \in J, k \in P, \forall n \in N (19)\]

\[x_{ijk}^n \in \{0, 1\}, \forall i \in J, k \in R, \forall n \in N (20)\]

\[y_{ik}^n \in \{0, 1\}, \forall i \in V, k \in K, \forall n \in N (21)\]

\[w_i \in \{0, 1\}, \forall i \in O_T (22)\]

Constraints (9) guarantee that each customer \( i \in J \) is served in each scenario \( n \in N \). Constraint (10) states that exactly one transfer station is opened. Note that this constraint could be extended to consider \( p \) locations. Constraints (11) ensure that we cannot start a vehicle departure from a transfer station that is not open. Constraints (12) are flow conservation constraints that state that for each node visited, there is an outgoing and an incoming arc. Constraints (13) state that there is an arc going out from the node for every node that visits or starts a vehicle. \( \delta^+(i) \) and \( \delta^-(i) \) are the sets of incoming and outgoing arcs of node \( i \). Constraints (14) are subtour elimination constraints. Constraints (15) ensure that the total number of customer visits per vehicle \( k \) does not exceed the vehicle capacity \( Q_k \). Constraint (16)–(17) guarantees that all customers are served before deadline \( T \). Recall \( r_{ij} \) is the travel time between location \( i \) and location \( j \) in scenario \( n \in N \), \( t_{0l} \) is then the travel time from the depot \( (0) \) to the transfer station \( (l \in O_T) \) and travel time at the transfer station is \( \tau \). Deadline constraints (16) are associated with the vehicles departing from the depot and constraints (17) with the vehicle departing from the transfer station. Since the delivery deadline only applies to the customer visits, we can exclude the final leg back to the depot \( (0) \) or the transfer station. As multiple vehicle trips are allowed at
the transfer station, we identify the end node of the last trip to the transfer station as “d”. Constraints (18)–(19) guarantee that customer can only be served from trips that take place from the depot or transfer station. These constraints are not strictly necessary but help to decrease run times. Note that each vehicle departure \(k\) can either start from the depot, i.e., \(k \in P\) or from one of the potential transfer locations, with \(k \in R\). Finally, constraints (20)–(22) represent integrality conditions.

5.2. A metaheuristic

As our problem generalizes the multi-depot, multi-trip vehicle routing problem, it is unlikely that we will be able to find exact approaches that can solve larger instances in a reasonable amount of time. Therefore, we present a heuristic in this section. Building on Ropke and Pisinger (2006), we propose an Adaptive Large Neighborhood Search (ALNS) heuristic using a Greedy Randomized Adaptive Search Procedure (GRASP). The general idea is to find the best solution for each transfer location and every scenario and pick the transfer location that is associated with the lowest costs across all location-scenarios. Therefore, for every possible transfer location and demand scenario, we create an initial solution and then improving this solution by using local search. Algorithm 1 provides an overview of the main steps.

Given a set of customers and a transfer location, we get an initial feasible solution by applying a “Cluster First and Route Second” approach. To determine the customer clusters, we assign each customer to the nearest location, i.e., to the depot or to the transfer location. To build routes per cluster, we use a randomized savings algorithm based on Clark and Wright (C&W) (Prins et al., 2006; Radojičić et al., 2018). Similar to the classic C&W (Clarke and Wright, 1964), we start with location. To build routes per cluster, we use a randomized savings algorithm based on Clark and Wright (C&W) (Prins et al., 2006; Radojičić et al., 2018). Similar to the classic C&W (Clarke and Wright, 1964), we start with location.

Algorithm 1: Adaptive Large Neighborhood Search (ALNS)

```
1 \( \bar{c}(S^*) \leftarrow \infty \)
2 for \( i \in O_T \) do
3     for \( n \in N \) do
4         \( S(n) \leftarrow S_0(n) \)
5         /* Equal weights for operators \( W^- \) and \( W^+ \) */
6     while stop criterion is not met do
7         \( S_T = \text{Repair}(\text{Destroy}(S(n))) \)
8         if \( c(S_T) < (1 + \delta)c(S(n)) \) then
9             \( S_T = \text{Local Search}(S_T) \)
10         end if
11         if \( c(S_T) < c(S(n)) \) then
12             \( S(n) \leftarrow S_T \)
13     end if
14     end while
15     /* Update weights for operators \( W^- \) and \( W^+ \) */
16     if \( \frac{1}{N} \sum_{n=1}^{N} c(S(n)) < \bar{c}(S^*) \) then
17         \( S^* \leftarrow (S(1) \ldots S(N)) \)
18         \( \bar{c}(S^*) = \frac{1}{N} \sum_{n=1}^{N} c(S(n)) \)
19 end if
20 return \( S^* \)
```

Given an initial set of solutions, we improve the solution for each scenario \( n \in N \) based on local search, see algorithm 1. In particular, we use different destroy and repair methods that are randomly selected. The probability of choosing a specific method depends on the success of the method given by the weights \((W^- \text{ or } W^+)\). We update the weights based on their performance. We apply the following destroy and repair operators.

**Destroy operators:** Remove \( \beta \) percent of the customers following the following rules: (i) Random destroy: randomly remove \( \beta \) percent of the customers. When the solution is destroyed to create a capacity-feasible initial solution, the \( \beta \) is given. (ii) Origin destroy: Randomly select a transfer station or depot and remove all routes starting at this origin. (iii) Worst destroy: Remove \( \beta \) percent of the customers that worsen the solution, i.e., with a greater increase in distance. For all destroy methods, \( \beta \% \) change in a randomly predefined range.

**Repair operators:** Re-insert the unrouted customers back into the destroyed solution following the next repair operator rules. (i) Cheapest insertion: Re-insert pool nodes with the cheapest cost at its best position. (ii) Random insertion: Choose unrouted customers in random order and insert them at the best position. (iii) Forbidden insertion: It behaves similarly to random insertion but does not allow reinsertion at the starting point of the route, e.g., depot or transfer station, from which it was removed. This operator rule is valid when the solution feasibility is met. Otherwise, it operates as a random insertion. The destruction of a route can be gradual or total. Therefore, for all repair operators, we try to insert the customer in one of the existing route or a new one to ensure feasibility in terms of the capacity or deadline.
To carry out a more exhaustive search, we proceed to a local search at intra-route and inter-route to improve the current solution after destroying and repairing the current solution. This local search is performed if the cost of the temporary solution is less than the cost of the best solution by a certain percentage. We use 2-opt, 3-opt, and or-opt as intra-route, and 2-opt* as inter-route. We implement every feasible improvement. The repetitive local neighborhood search stops when \( M \) number of iterations has been performed, or no improvement has been found in \( m \) iterations. After determining the route plans for each demand scenario and transfer location, we select the solution with the lowest cost.

6. Computational results

This section presents numerical studies to evaluate the benefits of integrating public transit in urban home delivery. First, we detail the experimental setup in Section 6.1. Then, we evaluate the proposed heuristics using small instances solved to optimality in Section 6.2. Finally, in Section 6.3, we present the results and discuss the impact of different parameter values. We implemented the experiments in Python 3.8 using Gurobi 9.0.2 as a mixed-integer programming solver (Gurobi-Optimization, 2020) and ran all instances on an Intel Core i7-10510U computer with 16 GB RAM.

6.1. Experimental setup

We present our base case experiments for the two stylized settings as introduced in Section 4: (a) a rectangular delivery area with a line transit network, and (b) a circular delivery area with a circular transit network. To facilitate comparison, we consider an area of 100 km\(^2\) for both settings, i.e., a rectangular area of 10 km by 10 km, and a circular area with a radius of \( R = 5.6 \) km. Fig. 3 shows both networks in which the depot is located at the beginning of the transit line. We consider three transit stops that can be used as a potential transfer station. For the line setting, these transit stops are located at one-third, two-third, and the end of the line from the depot. In the circular case, these transit stops are located at \( \frac{\pi}{2} \), \( \pi \), and \( \frac{3\pi}{2} \) angle from the depot on a circular network with a \( r = 3 \) km radius.

![Fig. 3. Baseline case.](image)

(a) Line network in a rectangular service area
(b) Circular network in a circular service area
\( \blacktriangle \) Depot and \( \square \) Potential transfer

We consider 50 customers that are uniformly spread across the delivery region. We assume that we can use a fleet of homogeneous last-mile vehicles with a capacity to serve 80% of total customers per vehicle for both locations. That is, \( Q_0 \) and \( Q_T \) equal 40 customers per vehicle in all of our experiments, unless otherwise stated. In the baseline experiments, we consider a setting without deadlines, so we ignore travel times. For every experiment, we generate \( N = 10 \) demand location realizations, where the locations are uniformly distributed across the delivery region. We compare all baseline results to the benchmark in which all customers are served from the depot. We use the Manhattan metric to compute the distance between two customer locations. We also ran experiments using the Euclidean metric and the results are similar (see Appendix A.2).

6.2. Heuristic validation

To evaluate the performance of the heuristic, we use smaller instances as we could not find proven optimal solutions in reasonable time for larger instances using the MIP formulation in Section 5.1. In particular, we compare the heuristic solutions to the optimal solutions for ten instances with ten customer and the same characteristics as the base instances for both the line and the circular network. That is, we again use a capacity of 80% of the total number of customers (8–8) for the base case. Moreover, we also experiment with a number of instances in which the last-mile vehicles at the transfer station have less capacity (4 and 2 customers).

Table 1 shows the minimum, average (\( \mu \)), and maximum optimality gap (\( % \text{Gap} \)). Let \( Z^* \) be the objective value (distance) of the optimal solution and \( Z_h \) the objective value (distance) of the heuristic solution. The optimality gap is computed by

\[
% \text{Gap} = 100 \times \frac{Z_h - Z^*}{Z^*}
\]

Also, the table provides the number of instances for which the heuristic finds the optimal location for the transfer station (Opt. location) and an optimal solution for both the location and the route schedule (Opt. solution). We use a stop criterion of 100 consecutive iterations without improvement or a maximum of 5000 iterations for the heuristic.

The results show that the heuristic performs well. The optimal transfer location is found in almost all cases. Moreover, we see that the average optimality gap is small, i.e., 0.7 percent for the line network and 0.5 percent for the circular network.
Table 1
Heuristic solution compared to the exact solution (averages across ten random instances of ten customers).

<table>
<thead>
<tr>
<th>System</th>
<th>$Q_0 - Q_T$</th>
<th>% gap</th>
<th>Opt. location</th>
<th>Opt. solution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min $\mu$ max</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line network</td>
<td>8-8</td>
<td>0</td>
<td>1.1</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>8-4</td>
<td>0</td>
<td>0.8</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>8-2</td>
<td>0</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Circular network</td>
<td>8-8</td>
<td>0</td>
<td>0.4</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>8-4</td>
<td>0</td>
<td>0.4</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>8-2</td>
<td>0</td>
<td>0.7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

6.3. Numerical tests

In this section, we discuss the results for the baseline case and the sensitivity analysis.

6.3.1. Baseline results

Table 2 reports the minimum (min), average ($\mu$), maximum (max) and standard deviation ($\sigma$) over ten random demand realizations for the following performance indicators. Total km: the system-wide distances in kilometers. % km Transfer: percentage of distance traveled by the last-mile vehicles from the transfer station. % km Depot: percentage of distance traveled by the last-mile vehicles from the depot. % Orders Transfer: the proportion of customers served via the transfer station. % Savings: percentage of kilometers saved compared to the setting without transfer. Transfer Location: the best transfer location.

Table 2
Baseline results for different transit line structures (values across ten customer instances).

<table>
<thead>
<tr>
<th>System</th>
<th>Total km</th>
<th>% km transfer</th>
<th>% km depot</th>
<th>% orders transfer</th>
<th>% Savings</th>
<th>Transfer location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line network</td>
<td>min</td>
<td>70.2</td>
<td>27.4</td>
<td>21.9</td>
<td>22</td>
<td>$\frac{3}{4}L$</td>
</tr>
<tr>
<td></td>
<td>$\mu$</td>
<td>73.4</td>
<td>55.5</td>
<td>44.5</td>
<td>54.2</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>77</td>
<td>78.1</td>
<td>72.7</td>
<td>78.0</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>2.4</td>
<td>17.3</td>
<td>17.3</td>
<td>18.4</td>
<td>1.8</td>
</tr>
<tr>
<td>Circular network</td>
<td>min</td>
<td>66.5</td>
<td>17.9</td>
<td>25.0</td>
<td>20</td>
<td>$\pi$</td>
</tr>
<tr>
<td></td>
<td>$\mu$</td>
<td>71.2</td>
<td>54.2</td>
<td>45.8</td>
<td>56.8</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>74.6</td>
<td>75</td>
<td>82.1</td>
<td>80</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>2.5</td>
<td>17.6</td>
<td>17.6</td>
<td>19.4</td>
<td>1.2</td>
</tr>
</tbody>
</table>

We observe that the transfer location that minimizes the expected distance is at $\frac{3}{4}L$ of the line network and at an angle of $\pi$ for the circular network. These results coincide with the analytical results in Section 4 in which we simplified routing by assuming all customers to be located on the transit line.

Looking at the benefits of using public transit, we observe average distance savings of 4.9 percent for the line network and 3.4 percent for the circular network with a low deviation of 1.8 and 1.2 respectively. We see savings of up to 7.1 (5.4) percent for the line (circular) network. Although the customer density is equal in both cases, the transit line structure and the shape of the service area lead to lower overall distances and savings in the circular network. In line with the analytical results, roughly half of the customers are served via the transfer station. On average, that is 54.2 percent for the line network and 56.8 percent for the circular network.

6.3.2. Effect of capacity of last-mile vehicles

To further reduce the environmental impact, it may be beneficial to use smaller, low or zero emission last-mile vehicles at the transfer station. In this section, we consider the impact on the system performance of using such smaller vehicles. In particular, we vary the capacity of the last-mile vehicles at the transfer station $Q_T$ to 40 (base case), 20 and 10 customers per vehicle. We keep the baseline settings for all other parameters.

Fig. 4 shows box plots of the total distance for different vehicle capacities at different transfer locations for the line network 4(a) and the circular network 4(c). We see that the median results are relatively stable across different locations within a certain capacity scenario. This suggests that the performance of the system is relatively insensitive to the specific transfer location. One of the reasons for this is that the system is flexible in deciding how many customers are served via the transfer station.

Overall, we observe that the system-wide distances increase with smaller vehicles. This effect is stronger for the line network than for the circular network. The increase is associated with the fact that less customers are transported via the transfer station, e.g., from 55% (base case) to 8% for both transit network. With smaller vehicles, the differences between different transfer locations are more pronounced for the line network.

Fig. 4(b) and (d) show the average distance savings for each capacity scenario and line structure for the best solution shown in 4(a) and (c). Both bar graphs show similar behavior; savings increase with capacity. This is intuitive as smaller vehicles at the transfer station means that more orders are handled by direct shipment. Also, there are more mileage savings of using the transfer station on the line network due to the characteristic of the delivery area around the transit line.

Note that in some settings it may not be beneficial to use the transit station. Fig. 5 provides more insight into this by plotting how often the transfer station is used across the ten instances for different capacity settings. The results show that for the base case...
the transfer location is always used. However, with smaller vehicles at the transfer station it is not always beneficial to use the transfer station.

6.3.3. Effect of depot location

In the base case, we consider a depot at the start of the transit line within the delivery region. Here, we consider the depot is at the start of the transit line but outside the delivery region, i.e., at 5 and 10 kilometers from the start of the delivery region (base case). Fig. 6 shows the impact of the depot location on the performance indicators for both line and circular structures.
As expected, we see that the savings increase when the depot is outside the delivery area. We see average savings up to 36.2 percent with a depot at 10 km from the delivery area for the line network. Now, most orders are transported via the transfer station to save on the stem distance from the depot to the first customer. For the circular network, the average savings increase by up to 32.3 percent. The best location remains at $\frac{2}{3}$ of the line network and $\frac{1}{2}$ of the circular network.

6.3.4. Effect of customer density

In this section, we study the impact of customer density on the performance of the system. For our base delivery area, we set the number of customers to 25, 50 (base case) and 100. Table 3 reports the average, minimum, maximum and standard deviation distance savings of using the transit station as compared to serving all customers from the depot.

<table>
<thead>
<tr>
<th>System</th>
<th># customers</th>
<th>% savings</th>
<th>min</th>
<th>$\mu$</th>
<th>max</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line network</td>
<td>25</td>
<td>4.2</td>
<td>8.2</td>
<td>15.1</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1.2</td>
<td>4.9</td>
<td>7.1</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>2.3</td>
<td>5.8</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Circular network</td>
<td>25</td>
<td>0</td>
<td>5.7</td>
<td>11</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1.2</td>
<td>3.4</td>
<td>5.4</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>1.8</td>
<td>4.1</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

The results on kilometers savings are in line with the analytical results from Section 4. We see that higher customer densities correspond to lower savings. That is, the average saving decreases from 8.2% to 2.3% when increasing the density for the line network. Similarly, we observe a drop from 5.7% to 1.8% for the circular network case.

6.3.5. Effect of clustered customers

Up to now, we considered uniformly spread customers across the delivery region. Next, we check out different clusters. In particular, we divide the rectangular service area into two parts: the area around the transit line and the outside area, away from the line. We consider the two extreme cases: 0% of customers in the delivery region center, or 100%. For the line network, the delivery area center is around the transit line, while the circular network is within the transit line’s inner radius.

The bar chart in Fig. 7 shows the average savings for these two cases. As a benchmark, we also present the baseline uniform case. We observe that the behavior is similar for the line network and the circular network. As expected, we see higher savings of using the transit line when the customers are clustered around the transit line. On average, we observe 8.2 percent savings for the line network and 7.4 percent savings for the circular network. Overall, greater savings are found for the transit line than for the circular network.

To get some more insights into the actual routes, Fig. 8 illustrates which customer are served from the depot (‘‘Depot’’) or the transfer station (‘‘Transfer’’) for the ten instances. Fig. 8(a)–(c) show the cases in which all customers are centered around the transit line, while Fig. 8(b)–(d) show the cases in which none of customers are centered around the transit line.

Fig. 8(a)–(c) indicate that most customer locations are consistently served by either the transfer station or the depot. In the ‘‘centered case’’ (Fig. 8(a)–(c)), we see that the locations on the left of the line, close to the depot, are served from the depot. The locations further to the right are served by the transfer station. This pattern is less clear for the case in which the locations are not centered around the line (Fig. 8(b)–(d)). One potential reason for this is that in this case the routes from the depot and transfer station are clustered horizontally (top/bottom) instead of vertically (left/right).
6.3.6. Effect of delivery deadline

Thus far, we have ignored the impact of time constraints in our experiments. In this section, we take into account the delivery deadline and evaluate its impact on the different performance measures as introduce earlier. Moreover, we also report the average number of required vehicles across the different scenarios and instances (# vehicle).

Table 4 presents the results for the two transit structures for different operating speed and deadline scenarios. Recall that we let $s_l$ be the last-mile vehicle speed and $s_t$ the speed of the public transport vehicle. We consider minimum speeds of 20 km/h and
maximum speeds of 30 km/h in the following scenarios: $s_j > s_l$, $s_j < s_l$, and $s_j = s_l$. We set a delivery deadline of respectively two and one hour.

Our results show that tight deadlines decrease the performance of the system, i.e., increase travel distances, and the number of last-mile vehicles needed. Moreover, we see that fewer customers are served via transfer station. In a line network, the deadlines affect the optimal location of the transfer station for a deadline of 1 h in the uneven speed scenarios. In these cases, it is better to move the transfer station to the end of the line, away from the depot. Looking at the savings, we see that when transit vehicles run at the same or a higher speed than the last-mile vehicles, the savings of using the transit line increase with tighter deadlines. The reason is that the transit vehicle is fast enough to reach the end of the line and serve customers far away from the depot. In this manner, it is possible to minimize the average distance and thus increase savings.

In the circular network, we observe that the optimal transfer location is less sensitive to the transit vehicle’s operating speed and deadline. Although the savings of the circular network are lower than in the line network, we see high savings in the case with tight deadlines. The reason for this is that with tight deadlines fewer customers can be served per route. This corresponds to more routes and more ‘stem distances’ in the case without using public transport. This provides more savings potential from using public transport.

### 7. Concluding remarks

Our study has shown that using the capacity of a public transit line can help to reduce the total system-wide distance in urban delivery. As such, it can help to reduce the negative externalities of urban delivery operations. Our analytic analysis of several special cases and our computational experiments for the general problem provide the following insights. First, the relative savings of using a transit line to support urban delivery operations decrease with customer density. Second, the savings increase with the distance of the depot to the delivery region. Third, the savings increase when customers are clustered close to the public transit line. Fourth, the savings increase with tight delivery deadline. We derive analytical expressions to determine the optimal transfer location for cases with customers on a line. Our simulation studies suggest that these locations are also optimal for the general case.

We have focused on a stylized setting with one transfer station on one public transit line. Future research can consider multiple transfer stations on multiple transit lines. Moreover, we have considered fixed transfer locations. It may be valuable to compare such a fixed system with a more flexible system in which the transfer points can change with the specific customer orders. While this would provide more flexibility, there are also additional costs as it would require the synchronization of the last-mile vehicles with the transit schedule. Also, we have captured the most constrained capacities, such as last-mile vehicle capacity and tight deadlines. Other future research can include transit vehicle capacity, loading, unloading, and storage capacities at transfer stations. A limited capacity of transit vehicles opens the possibility of exploring the consolidation of multiple transit shipments at transfer stations. Besides, we have focused on low-volume home delivery operations, so we simplify the analysis considering that size/volume is not relevant. Future works can define capacity in terms of volume and size.

### CRediT authorship contribution statement

**Irecis Azcuy**: Conceptualization, Software, Data curation, Visualization, Validation Formal analysis, Writing – original draft. **Niels Agatz**: Conceptualization, Methodology, Validation, Writing – reviewing and editing, Supervision. **Ricardo Giesen**: Conceptualization, Methodology, Validation, Writing – reviewing and editing, Supervision, Project administration, Funding acquisition.
Acknowledgments

We would like to thank the support by Fondecyt 1171049, CEDEUS, ANID/FONDAP 15110020, and the BRT+ Centre of Excellence funded by the Volvo Research and Educational Foundations (VREF). The first author acknowledge the support by the National Agency for Research and Development (ANID)/Scholarship Program/DOCTORADO BECAS CHILE/2016 - 21160726.

Appendix

A.1. Saving analysis for the special cases without deadline

If we look at the maximum potential savings from adding a transit line, these are equal to $2 \frac{M}{(M+1)} E(L) - E[L_i]$. Where the first term is the expected length of the TSP tour with just direct shipment to serve $M = |J|$ customers, and the second term is the expected length of the TSP tour considering the mixed system. The expected savings for line network case for the optimal location at $x^* = \frac{2}{3} L$ is as follows.

$$E[S(M)] = 2 \frac{M}{(M+1)} L - \left( \frac{4L}{2M} \left( \frac{2L}{3}(\frac{1}{2})^M + M(\frac{3}{2})^M - 1 \right) \right) \left( \frac{M}{(M+1)^2} \right)$$

In a circular line, the expected saving for the optimal location $\alpha = \pi$ is

$$E[S(M)] = M \left( \frac{2}{3} \pi r \left( \frac{(3\pi)^M}{(4\pi)^M} \pi M + \left( \frac{1}{3} \right)^M + \pi \left( \frac{1}{3} \right)^M + (4\pi)^M \left( \frac{1}{3} \right)^M - 1 \right) \right)$$

When it comes to a unit capacity setting, the transfer station’s optimal location is kept at $2/3$ of the line, and we obtain a savings ratio equal to $2/3$. On the contrary, in a circular line, we can satisfy half of the delivery orders via transfers station, representing a $\frac{1}{2}$ savings compared to direct shipments. That means a savings ratio of $1/2$. This result holds for any length of the transit line since the savings at the optimal location are proportional to the transfer station’s location. However, for a more significant number of orders, the savings decrease to zero, which is $\lim_{M \to \infty} E[S(M)] = 0$. This is explained because as the number of orders increases, the expected distance will increase up to the line length is thoroughly traveled. Therefore, one important insight is that the savings do not depend on the length of the line but depend on the transit system’s structure and the number of orders.

A.2. Base case results for Manhattan versus Euclidean distance metric

We provide a comparison between the Manhattan metric and the Euclidean for the base case. The following table shows the minimum, average, and maximum values of the savings compared to direct shipments and the percentage of orders fulfilled from the transfer station. Besides, it provides the location of the transfer station for each case (see Table 5).

<table>
<thead>
<tr>
<th>System</th>
<th>Distance metric</th>
<th>% savings</th>
<th>% orders from transfer</th>
<th>Transfer location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>min</td>
<td>mean</td>
<td>max</td>
</tr>
<tr>
<td>Line network</td>
<td>Euclidean</td>
<td>3.1</td>
<td>5.3</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>Manhattan</td>
<td>1.2</td>
<td>4.9</td>
<td>7.1</td>
</tr>
<tr>
<td>Circular network</td>
<td>Euclidean</td>
<td>0</td>
<td>3.3</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Manhattan</td>
<td>1.2</td>
<td>3.4</td>
<td>5.4</td>
</tr>
</tbody>
</table>

The results suggest that there are no large differences in average between metrics, especially in the circular network, and we see that the location is not affected by the distance metric. Therefore, the distance metric does not have a great impact when comparing results between the direct delivery system and the mixed system.


