

# Connecting E-Hailing to Mass Transit Platform: Analysis of Relative Spatial Position

Peng (Will) Chen<sup>1</sup> and Yu (Marco) Nie <sup>\*1</sup>

<sup>1</sup>Department of Civil and Environmental Engineering  
Northwestern University  
2145 Sheridan Road, Evanston, IL 60208, USA

February 10, 2017

---

\*Corresponding author, E-mail: [y-nie@northwestern.edu](mailto:y-nie@northwestern.edu); Phone: 1-847-467-0502.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Abstract

This paper analyzes and compares two different relative spatial position (RSP) designs in an integrated e-hailing/fixed-route transit system: a zone-based design that operates e-hailing vehicles within a zone, and a line-based design that operates e-hailing vehicles along a fixed-route transit line and with a stable headway. To conduct a meaningful comparison, the optimal design problems for both systems are formulated using a same analytical framework based on the continuous approximation approach. A comprehensive numerical experiment is performed to compare various cost components corresponding to the optimal designs, and a discrete-event simulation model is developed to validate the analysis. The analytical and simulation results agree with each other well, with a discrepancy in the total system cost less than 5% in most test scenarios. These results also suggest that the line-based system consistently outperforms the zone-based system in terms of both agency and user costs, for all scenarios tested. Compared to the zone-based design, the line-based design features a sparser fixed-route network (resulting in larger stop spacing) but a higher dispatching frequency. It is concluded that the higher efficiency of the line-based design is likely derived from the strategy of operating e-hailing vehicles with a more regular route/headway structure and allowing ride-sharing.

*Keywords:* e-hailing service; fixed-route service; relative spatial position; continuous approximation; discrete-event simulation

# 1 Introduction

Technologies are driving an unprecedented wave of innovations in mobility services. Ubiquitous real-time communications and peer-to-peer interaction enabled by mobile computing promise to more effectively match transport supply and demand at low transaction costs, thereby giving rise to numerous new e-hailing services for personal mobility (e.g., Uber<sup>1</sup>, Bridj<sup>2</sup>, SpotHero<sup>3</sup>) and freight delivery (e.g., Postmates<sup>4</sup>, Roadie<sup>5</sup>). In the long run, the rapidly evolving vehicle automation technology may not only enhance profitability and competitiveness of these on-demand services, but also promote the shift from private car ownership to pay-per-use mobility models (see e.g. Bertonecello and Wee, 2015). It is widely speculated that the future personal travel market may feature *new transit services* offered by companies that operate a large number of driverless cars. Given its keen interest in driverless car technology<sup>6</sup> and recently announced partnership with Toyota Inc. on vehicle leasing<sup>7</sup>, there is little doubt that Uber is positioning itself to compete in the above futuristic scenario, so are the other eager players such as General Motor/Lyft<sup>8</sup> and Apple Inc./Didi (<http://www.xiaojukeji.com/>)<sup>9</sup>.

Despite their strong appeal, the likes of Uber and Bridj still largely rely on the niche market of door-to-door services. To scale up and succeed as a mass transport platform as envisioned above, greater ride consolidation (both temporally and spatially) and tradeoff between efficiency and level-of-service must be considered. Uber’s recent partnership with TransLoc<sup>10</sup> represents a timely move to this direction. The idea is to integrate e-hailing services into public transportation networks by using e-hailing as a demand-responsive feeder for existing transit services. *The purpose of this study is to explore and analyze design options for such an integration.* The main question asked here is how a transit operator can best allocate its resources to fixed-route and e-hailing services in order to meet demand. It is worth noting that the current line of thinking about integration appears to occur mainly in one direction, where e-hailing services are matched against the operations of fixed-route services. As the proposed TransLoc-Uber partnership puts it, the goal is to provide the users with a “personalized journey that incorporates the optimal combination of walking, transit and Uber”. In contrast, the premise of this paper is that a significant improvement in system efficiency can only be gained when the design and operation of both services are tightly coordinated. In essence, this means that the fixed-route services have to be re-designed in light of enhanced accessibility associated with e-hailing.

A crucial design decision, which motivates this study, has to do with the *relative spatial position of e-hailing services*. By relative spatial position (RSP), we mean the way by which e-hailing services are disaggregated in space and matched with relevant components of transit networks. Determining RSP is fundamental because it has to precede many other design decisions such as fixed-route headway and line spacing, as well as the number of e-hailing vehicles. While RSP may be arranged in many different ways, this paper will closely examine two simple RSP designs under idealized conditions for useful insights and

---

<sup>1</sup>The pioneer of the ride-sharing/e-hailing service (see [www.uber.com](http://www.uber.com)), which has become a very competitive market with several major players such as Lyft and the China-based Didi.

<sup>2</sup>Known as “the Uber of bus transit” (see [www.uber.com](http://www.uber.com)), it currently operates in Boston, Washington D.C. and the Kansas City. A similar service, called Via ([www.via.com](http://www.via.com)), operates in Chicago and New York.

<sup>3</sup>An on-demand parking service that operates in more than 10 US major cities (see <http://spothero.com/>).

<sup>4</sup>An on-demand express delivery service (see <https://postmates.com/>).

<sup>5</sup>A crowd-sourcing on-the-way delivery community (see [www.roadie.com](http://www.roadie.com)).

<sup>6</sup><http://www.govtech.com/fs/perspectives/Ubbers-Plan-for-Self-Driving-Cars-Bigger-Than-Its-Taxi-Disruption.html>

<sup>7</sup><http://www.wsj.com/articles/toyota-and-uber-reach-investment-lease-partnership-1464122403>

<sup>8</sup><http://techerunch.com/2016/03/14/lyft-gm-express-drive/>

<sup>9</sup><http://www.reuters.com/article/us-apple-china-idUSKCN0Y404W>

<sup>10</sup><http://transloc.com/transloc-and-uber-partner-to-pioneer-a-new-standard-in-public-transit>.

1  
2  
3  
4 design guidelines. The first design, referred to as the *zone-based RSP*, assigns a set of e-hailing vehicles to  
5 feed a given transit stop, and operates them to serve passengers within a relatively small zone surrounding  
6 the stop. The other design is called *line-based RSP* because it pairs each fixed transit line with e-hailing  
7 vehicles, which are dispatched on predetermined headway but are allowed to deviate from the fixed route  
8 to accommodate passengers who need a ride between their origin/destination and the closest transit stop.  
9 The prototypes of the zone- and line-based RSP designs have been studied respectively by Aldaihani et al.  
10 (2004) and Chen and Nie (2016). This work contributes to the literature by presenting a unifying analysis  
11 framework based on the continuous approximation approach, which enables the first comparative study of  
12 RSP designs. In addition, an agent-based, event-driven simulation platform is developed for both zone-  
13 and line-based RSP designs. The platform is then applied to validate analytical results, and to reveal and  
14 compare the performance of the two RSP designs under realistic operational conditions not fully captured  
15 by the analysis.

16  
17 The next section reviews the literature of on-demand transit services, most of which precede the era of e-  
18 hailing. Section 3 presents the setting of the analysis framework, with all simplifying assumptions. Section  
19 4 gives the formulation of the optimal design problem for both zone-based and line-based systems. The  
20 focus is given to the zone-based system because the line-based system is mostly adopted from the literature.  
21 In Section 5, numerical experiments are conducted to compare the performance of optimally configured  
22 zone-based and line-based systems. Section 6 first describes the details of the simulation platform, and then  
23 presents and discusses simulation results. Section 7 concludes the paper with major findings and possible  
24 directions for future research.

## 2 Literature review

25  
26 E-hailing, like taxi, is a special form of demand-responsive transit (DRT) that has been practiced and studied  
27 long before the era of smart phones. The reader is referred to He and Shen (2015); Wang et al. (2016); Zha  
28 et al. (2016) for recent research regarding e-hailing. The apparent lack of efficiency of taxi services had  
29 fueled enthusiasm for more advanced forms of DRT such as dial-a-ride transit (DART), which allows ride-  
30 sharing and even transferring through pre-arrangement (Wilson et al., 1976; Stein, 1978). At its core DART  
31 is a many-to-many pickup and delivery problem with time window, which is a special class of vehicle routing  
32 problems (VRP) known to be NP-hard (Cordeau and Laporte, 2003a).

33  
34 Many had attempted to tackle the computational challenges associated with DART (e.g. Psaraftis, 1980;  
35 Cordeau and Laporte, 2003b; Cordeau, 2006; Melachrinoudis et al., 2007; Ropke and Cordeau, 2009). Yet,  
36 the success of real DART systems is often limited by how fast a practically satisfactory solution to a combi-  
37 natorial problem can be obtained and properly implemented in real time. Black (1995) found many DART  
38 systems in operation suffer from high per-capita operating cost associated with low occupancy. This ob-  
39 servation has inspired proposals of “hybrid systems” that strive to integrate DRT into fixed-route services,  
40 an idea similar to what is pursued in this paper. Stein (1978) considers a system which divides the service  
41 area into zones connected by fixed-route bus services. Passengers use DART to get to designated “transfer  
42 points”, from where they travel to other zones on fixed-route buses. This system was implemented in Ann  
43 Arbor, MI (Potter, 1976).

44  
45 Malucelli et al. (1999) and Crainic et al. (2001) propose a demand adaptive system (DAS) that combines  
46 fixed-route service FRS with a restricted DART. The DART in DAS is restrictive in that vehicles have  
47 to follow a schedule at compulsory stops, which are designated transfer points. However, optional stops  
48  
49  
50  
51  
52  
53  
54  
55  
56

1  
2  
3  
4 between consecutive compulsory stops are served only by request. The DART system in DAS has been  
5 studied independently as a Mobility Allowance Shuttle Transit (MAST) (Quadrioglio et al., 2006, 2007,  
6 2008) or flexible-route transit (Pratelli et al., 2001; Alshalalfah, 2009). High-coverage point-to-point transit  
7 system (HCPPT) proposed by Cortés and Jayakrishnan (2002) may be viewed as a variant of the zone-based  
8 system of Stein (1978). Instead of using line-haul buses to connect zones, HCPPT connects them with the  
9 same vehicles that make pickup and delivery trips. The motivation is that this design eliminates the out-  
10 bound transfer. However, using the same vehicle to collect demand and provide line-haul travel may also  
11 lead to the use of a large number of small vehicles and extended in-vehicle time at low speeds, especially  
12 for high demand levels.

13  
14  
15  
16 Aldaihani et al. (2004) studied the zone-based system of Stein (1978) in an idealized grid network.  
17 They derived closed-form formulae to compute the total system cost and solved the optimal design problem  
18 analytically under simplified conditions. Recently, Chen and Nie (2016) proposed a demand adaptive paired-  
19 line hybrid transit (DAPL-HT) system, in which a demand-responsive service is used to connect passengers  
20 from their origin/destination to the fixed-route transit stop. It proposes a novel RSP design that operates the  
21 demand-responsive service with a stable headway to cover all stops along a paired fixed-route line.

22  
23  
24 It remains an open question which RSP design, zone-based or line-based, strikes a better balance be-  
25 tween cost and efficiency in general. To answer this question, we first enhance the zone-based hybrid  
26 design of Aldaihani et al. (2004) so that it is consistent with the line-based model of Chen and Nie (2016).  
27 Specifically, the following changes are made: replacing the Euclidean distance with the Manhattan distance,  
28 considering actual vehicle operating speeds instead of nominal speeds, and including passengers' waiting  
29 time for e-hailing vehicles. Then, a comparative study is conducted to determine the relative performance  
30 of the two RSP designs, using both analytical and simulation approaches.

### 34 35 **3 Integrated e-hailing/fixed-route system**

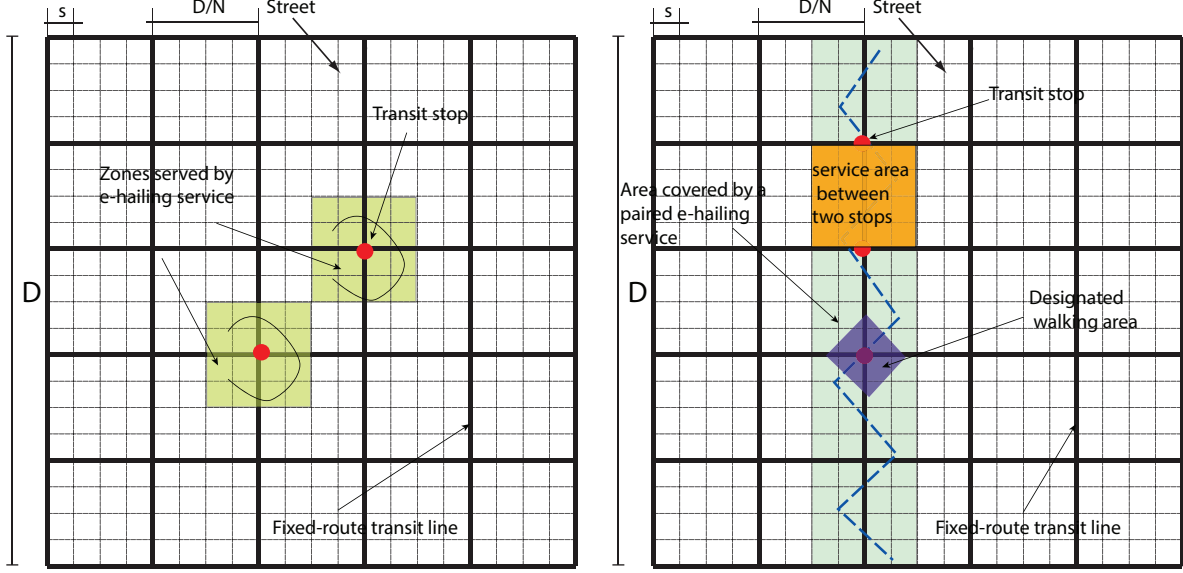
36  
37 For simplicity we shall consider a square service area with a side length of  $D$ , and a grid street network with  
38 a constant spacing of  $s$ . Following Daganzo (2010b), we assume that the service area generates  $\lambda$  passenger  
39 trips per hour per unit area, and their origins and destinations are uniformly and independently distributed in  
40 the area according to a homogeneous Poisson process. The transit system that serves the passengers consists  
41 of both a fixed-route service and an e-hailing service. Figure 1 presents two different RSP designs to couple  
42 these two services. As shown in the figure, the fixed-route service is always operated in both North-South  
43 (N-S) direction and East-West (E-S) direction. The following assumptions are needed regardless of the RSP  
44 design in order to facilitate the analysis.

45  
46  
47  
48 **Assumption 1.** *When they need to use transit to complete a trip, passengers always use the stops closest to*  
49 *their origin or destination.*

50  
51  
52 **Assumption 2.** *Passengers send their request for the e-hailing service prior to the desired departure time*  
53 *from the origin. Their request will be processed in a first-come-first-serve basis.*

54  
55  
56 **Assumption 3.** *The driving/walking distance between two points is measured by the Manhattan distance.*

57  
58 We note that Aldaihani et al. (2004) uses the Euclidean distance. Yet, the Manhattan distance is more  
59 appropriate in a grid network, which is the focus of the present study.



(a) Zone-based design

(b) Line-based design

Figure 1: Two relative spatial position (RSP) designs for an integrated e-hailing/fixed-route system

**Assumption 4.** *Passengers travel between fixed-route transit stops with the least possible number of transfers and as directly as possible, i.e., at most one transfer if the origin station and the destination station are not in the same line. When transfer is needed, passengers randomly choose the initial direction of travel.*

Let  $N$  be the number of fixed-route lines running in either of the N-S or E-W direction. Accordingly, the line spacing can be determined as  $S = D/N$ . The headway for the fixed-route lines is denoted as  $H$ . The service is provided by vehicles with following characteristics: the cruising speed including stops due to traffic and pedestrian interference,  $v$  (km/h); the time lost per stop due to the deceleration and the acceleration,  $\tau_1$  (seconds/stop); the time added per boarding passenger,  $\tau'_1$  (seconds/pax); and the additional pick-up and drop-off time required per passenger for the demand-adaptive service,  $\tau_2$  (seconds/pax). Passengers can walk at speed  $v_w$  (km/h).

### 3.1 Line-based design

In this design, the fixed-route lines are operated in both N-S and E-W directions, whereas the e-hailing services are paired with the fixed-route lines in the N-S direction<sup>11</sup> (see Figure 1b). We define the *access distance* as the distance between a passenger's origin/destination to the nearest transit stop. A unique feature of the line-based design, introduced in Chen and Nie (2016), is that the e-hailing service, by design, only serves passengers whose access distance exceeds certain percentage of the line spacing  $S$ , denoted as  $\beta$ , which itself is a design variable. This is formally stated as the following assumption:

**Assumption 5.** *Passengers always use the stops closest to their origin and destination. If the access distance is less than  $\beta D/N$  (where  $\beta \in (0, 1]$  is a design variable), passengers will walk to/from the stop; otherwise, passengers will request the e-hailing service.*

<sup>11</sup>Note that operating the e-hailing service in one direction is enough to cover all stops. We choose the N-S direction in this study, following Chen and Nie (2016).

1  
2  
3  
4 Both the fixed-route and e-hailing services are operated with a pre-selected headway, denoted as  $H_1$  and  
5  $H_2$ , respectively. The fixed-route vehicles travel along the designated lines and make pre-scheduled stops.  
6 The e-hailing vehicles, on the other hand, make lateral movements to pick up and drop off passengers while  
7 sweeping longitudinally back and forth along the fixed-route line that they are paired with, as shown by the  
8 zigzagging trajectory in Figure 1b. Operational characteristics of the e-hailing vehicles in the line-based  
9 design are summarized below.

- 10 • E-hailing vehicles always follow the fixed route line when they are not diverted by pickup and/or  
11 drop-off requests.
- 12 • Passenger requests are assembled by a central operator and dispatched to each vehicle in real-time.  
13 Prior to arriving at a stop  $i$ , each vehicle receives the requests for service for the area between stop  $i$   
14 and the next stop  $i + 1$  (see Figure 1b).
- 15 • The vehicle takes as many requests as possible while ensuring the arrival time at the next stop ( $i + 1$ )  
16 meets the requirement to maintain a stable headway. Extra requests will be passed on to the next  
17 vehicle in the line.
- 18 • The vehicle makes and executes a routing plan to accommodate all agreed requests, and it always  
19 returns to stop  $i + 1$  to start a new service cycle afterwards. The exact trajectory of the vehicle depends  
20 on the requests and route planning.

### 31 **3.2 Zone-based RSP design**

32  
33 In a zone-based system, as depicted in Figures 1a and 2, the operation of e-hailing vehicles is confined to  
34 each zone centering around a transit stop. These vehicles, upon request, can transfer passengers from their  
35 origins to the closest transit stations within the zone or to their destinations if the destinations are within the  
36 same zone. The on-demand vehicles can also pick up passengers at the transit stops and deliver them to their  
37 final destinations within the zone. To simplify the analysis, a few additional assumptions are introduced to  
38 model the zone-based system. We note that these assumptions are similar to those made in Aldaihani et al.  
39 (2004), with modest modifications.

40  
41 **Assumption 6.** *Ride sharing on e-hailing vehicles is not allowed.*

42  
43  
44  
45 Allowing ride-sharing would complicate the routing operations significantly, and likely lead to in-  
46 tractability in analysis. Since most of the current e-hailing services have just begun to explore the idea  
47 (e.g. Uber-pool), we plan to exclude ride-sharing from the present study.

48  
49 **Assumption 7.** *Passengers will be assigned to the closest idle e-hailing vehicle.*

50  
51 **Assumption 8.** *The e-hailing vehicles are assumed to wait at the last delivery spot until the next demand  
52 arrives.*

53  
54  
55 For simplicity, walking is not considered in the zone-based system. Note that if a walking area is  
56 designated within a zone, pick-up and drop-off points must lay outside that area, which gives rise to a donut-  
57 alike service area. Estimating the average distance between two random points in such an irregular shape  
58 using a close-form function turns out to be quite challenging, especially because its shape also depends  
59 on a decision variable. In contrast, the analysis in the line-based system is more straightforward, because  
60  
61  
62

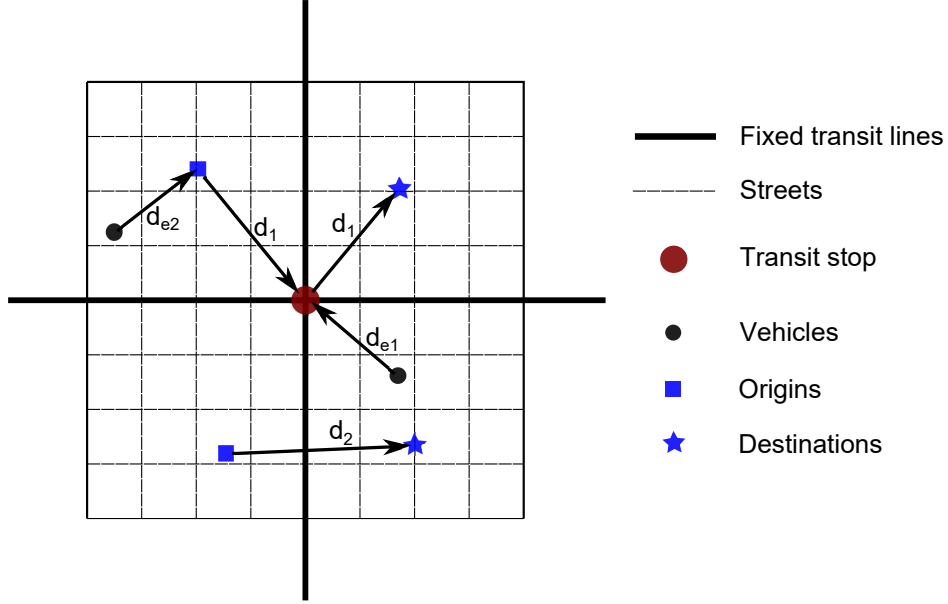


Figure 2: Illustration of the zone area around the transit station

the distance traveled by an e-hailing vehicle between two stops can be decomposed into longitudinal and lateral components, which can then be easily estimated regardless of the walking area (Chen and Nie, 2016). Without a closed form objective function, it would be difficult to determine the optimal design for the zone-based system. Thus, we opt for ignoring walking in the zone-based system in favor of tractability. While this creates certain modeling inconsistency, we shall show through numerical studies that the total walking cost in such hybrid systems is quite small, and would most likely have negligible impact on the relative performance of the RSP designs.

## 4 Optimal design problem

The design problem aims to determine the optimal design parameters (namely  $N$  and  $H$  for the zone-based design, and,  $N$ ,  $H_1$ ,  $H_2$  and  $\beta$  for the line-based design) such that the total system cost, inclusive of both user and agency costs, is minimized. We first examine the two components of the system cost: agency and user costs. The formulation of the optimal design problem will follow. For the convenience of the reader, the notations used throughout the paper are listed in Table 1.

### 4.1 Zone-based design

#### 4.1.1 Agency costs

The agency costs include the expected total vehicle distance traveled per hour of operation,  $Q$ , and the expected total fleet size in operation,  $M$ . Below, we analyze these components separately.

#### Vehicle distance travelled $Q$

In the zone-based system,  $Q$  consists of two parts: the expected total vehicle distance traveled per hour of operation by fixed-route buses,  $Q_1$  and that by on-demand vehicles,  $Q_2$ .

For the fixed-route buses, the total vehicle distance is related to the number of lines, the expected travel



Table 1: Description of notations used in the paper

Notation	Description
$D(km)$	Side length of the grid network
$s(km)$	Distance between two adjacent streets (street spacing)
$N$	Number of lines in N-S(E-W) direction
$\lambda(pax/h/km^2)$	Passenger trips generated per hour per unit area
$H(h)$	Headway of the fixed-route lines in zone-based system
$H_1(h)$	Headway of the fixed-route lines in line-based system
$H_2(h)$	Headway of the demand adaptive lines in line-based system
$\beta$	Parameter that determines maximum walking distance in line-based system
$\mu(\$/h)$	Value of time
$\tau_1(s)$	Time lost per stop due to deceleration and acceleration
$\tau'_1(s)$	Time added per boarding passenger for fixed-route vehicles
$\tau_2(s)$	Additional pick-up and drop-off time required per passenger
$v(km/h)$	Vehicles' cruising speed
$v_w(km/h)$	Walking speed
$\delta(km)$	Transfer penalty expressed in terms of the equivalent distance walked
$\$Q(\$/veh \cdot km)$	Operation cost per vehicle distance
$\$M(\$/veh \cdot h)$	Operation cost per vehicle hour
$Q(km/h)$	Expected total distance traveled per hour of operation
$M$	Number of vehicles required
$A(h)$	Average total walking time per passenger
$W(h)$	Average total waiting time per passenger
$T(h)$	Average total in-vehicle travel time per passenger
$e_T$	Expected number of transfers required per passenger

distance of one bus round trip, and the dispatch frequency. Since the round-trip distance per line is  $2D$ , the expected distance traveled per hour by all fixed-route buses is

$$Q_1 = \frac{4ND}{H}. \quad (1)$$

For e-hailing vehicles, their total distance traveled is the sum of the distance traveled by each passenger, because ride-sharing is not allowed as per Assumption 6. To estimate the distance traveled for each passenger, we first distinguish three types of passengers (Aldaihani et al., 2004):

1. Passengers with their origin and destination located in the same zone. They are served only by the e-hailing service.
2. Passengers with their origin and destination located in two separate zones that share a common fixed bus line. They are served by exactly one fixed-route line and two different e-hailing vehicles, one in each zone.
3. Passengers with their origin and destination located in two zones that do not share a common bus line. They are served by exactly two fixed-route lines and two e-hailing vehicles.

The proportion of each type of passengers is given as (Aldaihani et al., 2004):

$$P_1 = \frac{1}{N^2}, P_2 = \frac{2(N-1)}{N^2}, \text{ and } P_3 = \frac{(N-1)^2}{N^2}. \quad (2)$$

Let  $d_1$  denote the average distance traveled by a passenger on an e-hailing vehicle between any pickup/delivery point in the zone and the bus stop, and  $d_2$  denote the average distance traveled by a passenger between any pickup and delivery point within a zone (for Type 1 passenger). Also, let  $d_{e1}$  denote the average distance traveled by an empty e-hailing vehicle from a delivery point to the bus stop in the zone, and  $d_{e2}$  denote the average distance traveled by an empty e-hailing vehicle from a delivery point to the next pickup point within the zone. The above four types of travel distance are illustrated in Figure 2. Since we consider the Manhattan distance, we have

$$d_1 = \frac{1}{4} \frac{D}{N} + \frac{1}{4} \frac{D}{N} = \frac{1}{2} \frac{D}{N}, d_2 = \frac{1}{3} \frac{D}{N} + \frac{1}{3} \frac{D}{N} = \frac{2}{3} \frac{D}{N}. \quad (3)$$

Note that  $d_1$  is the average Manhattan distance between a random point in a square area and the center of the square and  $d_2$  is the average Manhattan distance between two random points in a square area. These distances are longer than those in Aldaihani et al. (2004) because they consider the Euclidean distance.

When a passenger's request is assigned to the closest idle e-hailing vehicle, the e-hailing vehicle must travel from its last delivery point to the new pickup point. This travel distance is related to the number of idle e-hailing vehicles in the zone, denoted by  $n_i$ . When  $n_i$  is small, the distances  $d_{e1}$  and  $d_{e2}$  are well approximated by  $d_1$  and  $d_2$ . For large  $n_i$ ,  $d_{e1}$  and  $d_{e2}$  can be approximated by  $0.5 \frac{D}{N \sqrt{n_i}}$  and  $0.63 \frac{D}{N \sqrt{n_i}}$  (Daganzo, 1978). The later analysis and numerical results will show that, in our model, the number of e-hailing vehicles assigned to each zone is relatively small under most demand scenarios, because our zones only feed a single stop. Clearly, the number of e-hailing vehicles in the idle state is even smaller. Therefore, for simplicity, we approximate  $d_{e1}$  and  $d_{e2}$  with  $d_1$  and  $d_2$ , respectively.

The average on-demand vehicle distance traveled can be determined for each type of passenger as follows:

- For a Type 1 passenger, the average distance is  $d_{e2} + d_2 \approx 2d_2$ , where  $d_{e2}$  is the distance travelled to pick up the passenger, and  $d_2$  is the distance travelled to drop off the passenger.
- For Type 2 and Type 3 passengers, the distance is  $d_{e2} + 2d_1 + d_{e1} \approx 3d_1 + d_2$ . In the pickup zone, the distance traveled to pick up the passenger ( $d_{e2}$ ) and the distance travelled to drop off the passenger at the bus stop ( $d_1$ ); in the drop-off zone, the distance travelled to pick up the passenger at the bus stop ( $d_{e1}$ ) and the distance traveled to drop off the passenger at the destination ( $d_1$ ).

The total distance traveled per hour of operation by e-hailing vehicles can be computed as

$$Q_2 = \lambda D^2 [P_1 \times 2d_2 + (P_2 + P_3) \times (3d_1 + d_2)] = \frac{13N^2 - 5}{6} \frac{\lambda D^3}{N^3}. \quad (4)$$

Hence, the overall expected vehicle distance traveled per hour is

$$Q = Q_1 + Q_2 = \frac{4ND}{H} + \frac{13N^2 - 5}{6} \frac{\lambda D^3}{N^3}. \quad (5)$$

### Fleet size

The fleet required to run the integrated the system consists of two types of vehicles: the fixed-route buses (with the size of  $M_1$ ), and the e-hailing vehicles (with the size of  $M_2$ ). Because  $Q_i$  by definition is the distance traveled by  $M_i$  vehicles in unit time, the size for each type can be estimated as

$$M_i = \frac{Q_i}{v_{ci}}, i = 1, 2, \quad (6)$$

where  $v_{ci}$  is the average operating speed for fixed-route vehicles ( $i = 1$ ) and e-hailing vehicles ( $i = 2$ ). For the fixed-route buses, the time required to travel distance  $Q_1$  in one hour ( $\frac{Q_1}{v_{c1}}$ ) should include the time spent on (i) overcoming distance ( $\frac{Q_1}{v}$ ); (ii) stopping ( $\frac{4\tau_1 N^2}{H}$ ); and (iii) collecting passengers ( $\tau'_1(P_2 + P_3 + e_T)\lambda D^2$ ). Hence, we have

$$M_1 = \frac{Q_1}{v_{c1}} = \frac{Q_1}{v} + \frac{4\tau_1 N^2}{H} + \tau'_1 \left( \frac{N^2 - 1}{N^2} + e_T \right) \lambda D^2. \quad (7)$$

For relatively large  $N$  (e.g.,  $N \geq 10$ ), we shall approximate the last term in the above as  $\tau'_1(1 + e_T)\lambda D^2$ .

For the e-hailing vehicles, the time required to travel distance  $Q_2$  in one hour ( $\frac{Q_2}{v_{c2}}$ ) should include the time spent on (i) overcoming distance ( $\frac{Q_2}{v}$ ); and (ii) picking up and dropping off passengers ( $\lambda D^2[2\tau_2 P_1 + 4\tau_2(P_2 + P_3)]$ ). Accordingly,

$$M_2 = \frac{Q_2}{v_{c2}} = \frac{Q_2}{v} + \lambda D^2[2\tau_2 P_1 + 4\tau_2(P_2 + P_3)] = \frac{Q_2}{v} + \frac{2(2N^2 - 1)\tau_2 \lambda D^2}{N^2}. \quad (8)$$

The above estimation assumes that the system will just have enough number of e-hailing vehicles such that neither the vehicles nor the passengers need to wait to find a match. In reality, because there are always demand fluctuations, passenger waiting and vehicle idling are unavoidable, and ignoring these effects all together may underestimate  $M_2$ . Such an effect is difficult to estimate and its impact is expected to be relatively small since the demand is homogenous in both space and time in our analysis. In addition, the underestimation may be offset by the fact that the distance travelled by the e-hailing vehicles is overestimated (by approximating  $d_{ei}$  with  $d_i$ ).

#### 4.1.2 User costs

In the zone-based design, important metrics of user costs include three main components: waiting time ( $W$ ), in-vehicle travel time ( $T$ ), and the penalty cost for each transfer taking place between two fixed-route lines. Note that the transfer between e-hailing service and fixed-route service is not penalized because passengers can arrange e-hailing service ahead of time as per Assumption 2.

Note that Type 1 and Type 2 passengers require no transfer, and Type 3 passengers require one transfer. The expected number of transfers is then

$$e_T = 0 \times (P_1 + P_2) + 1 \times P_3 = \frac{(N - 1)^2}{N^2}. \quad (9)$$

The total waiting time includes waiting for e-hailing service and waiting for fixed-route service. Note that Aldaihani et al. (2004) did not consider waiting time associated with the e-hailing vehicles in their model. Waiting time for e-hailing vehicles indeed consists of two parts: (1) the time until an e-hailing vehicle becomes available and is assigned to serve a passenger, and (2) the time needed for the e-hailing vehicle to travel to the pickup point. As discussed above, the fleet size of the e-hailing service is such selected that, without demand fluctuations, the first component in the passenger waiting time is negligible. Thus, for simplicity, these waiting times are ignored in the analytical model. We will, however, record these costs in simulation. The second component can be approximated by  $d_{e2}/v$  when passengers wait at the origin and by  $d_{e1}/v$  when passengers wait at the transit station to their final destination. Again, since  $d_{e1}$  and  $d_{e2}$  are overestimated, this may offset the effect of excluding the first component. Waiting time at each stop for

the fixed-route service is approximately half of the headway. To summarize, the expected waiting time per passenger is approximately calculated as

$$\begin{aligned} W &= \frac{de_2}{v} + (P_2 + P_3)\frac{de_1}{v} + \frac{H}{2} + \frac{H}{2}P_3 \\ &= \frac{2D}{3Nv} + \frac{D}{2Nv}\left(1 - \frac{1}{N^2}\right) + \frac{H}{2}\left[1 + \frac{(N-1)^2}{N^2}\right]. \end{aligned} \quad (10)$$

Passengers' in-vehicle travel distance includes the distance traveled in the fixed-route bus ( $E_1$ ) and the distance traveled in the e-hailing vehicle ( $E_2$ ). Let  $d$  denote the average distance traveled by a passenger on only one fixed-route bus, the average distance traveled by passengers with and without transfer in fixed-route buses are  $2d$  and  $d$ , respectively. Aldaihani et al. (2004) showed that the average distance  $d$  converges to  $0.34D$  for large  $N$  values. As in Aldaihani et al. (2004), we set the average distance  $d$  to its limit of  $0.34D$ . Consequently, we have

$$E_1 = dP_2 + 2dP_3 = \frac{0.68DN(N-1)}{N^2}.$$

Depending on the passenger type, the average distance travelled on e-hailing vehicles is different. The average value can be estimated as

$$E_2 = P_1 \times d_2 + (P_2 + P_3) \times 2d_1 = \frac{(3N^2 - 1)D}{3N^3}. \quad (11)$$

Therefore, the expected in-vehicle travel distance per passenger is

$$E = E_1 + E_2 = \frac{0.68DN(N-1)}{N^2} + \frac{(3N^2 - 1)D}{3N^3}. \quad (12)$$

Accordingly, the expected in-vehicle travel time per passenger is

$$T = \frac{E_1}{v_{c1}} + \frac{E_2}{v_{c2}}. \quad (13)$$

### 4.1.3 Formulation

We are now ready to present the optimal design model. To define the objective function consistently, all cost components are first converted into travel time equivalents. Let  $\$Q$ ,  $\$M$  and  $\mu$  be the agency operation cost per vehicle-distance, the agency cost per vehicle hour, and the average monetary value of one passenger-hour, respectively. Then,  $\pi_Q = \frac{\$Q}{\lambda D^2 \mu}$  and  $\pi_M = \frac{\$M}{\lambda D^2 \mu}$  convert the corresponding agency costs into the equivalent travel time per passenger (see Daganzo, 2010b). Let  $\delta$  measure the transfer penalty expressed in terms of the equivalent distance walked. Then  $\frac{\delta}{v_w}$  gives the equivalent travel time per transfer per passenger.

The total generalized cost of the zone-based hybrid transit system is defined as

$$z(N, H) = \pi_Q Q + \pi_M M + W + T + \frac{\delta}{v_w} e_T. \quad (14)$$

The corresponding optimization problem reads:

$$\min z(N, H) \quad (15)$$

$$\text{s.t. } H > 0, N \in \{1, 2, \dots, \left\lfloor \frac{D}{s} \right\rfloor\}. \quad (16)$$

Problem (14) is a mixed-integer optimization problem that is NP-hard. However, since it only has two decision variables, many numerical schemes can solve the problem efficiently. Our previous study (Chen and Nie, 2016) compared several methods, and concluded that Matlab's built-in genetic algorithm (GA) is capable of quickly producing high quality solutions for this problem. In the following, this meta-heuristic will again be used to obtain optimal design solutions.

## 4.2 Line-based design

We now briefly discuss the line-based design. Similarly, the agency costs depend on the expected total vehicle distance traveled per hour of operation,  $Q$ , and the expected total fleet size in operation,  $M$ . For the user costs, in addition to those considered for the zone-based design, the line-based system also considers walking time ( $A$ ), because per Assumption 5, passengers may be required to walk to their desired stop if they are within the walking zone. Let  $p_y$  denote the percentage of passengers who use the e-hailing service in the line-based system and  $p_y = 1 - p_n$ , where  $p_n$  denotes the percentage of passengers who have to walking, which is given

$$p_n = \begin{cases} 2\beta^2, & 0 < \beta \leq 0.5, \\ 1 - 2(1 - \beta)^2, & 0.5 < \beta \leq 1. \end{cases} \quad (17)$$

It is worth noting that in the line-based system Type 1 passenger are not singled out as in the zone-based system. Instead, they are assumed to use the hybrid transit system the same way as other passengers do. Alternatively, these passengers may be assumed to only ride e-hailing or simply walk to the destination. Since they account for a small percentage of all passengers when  $N$  is not too small, we expect that these different assumptions would only lead to negligible discrepancies in the optimal solutions.

To be concise, the details of the derivation are ignored here. The reader is referred to Chen and Nie (2016). Table 2 compares the main analytical results for the zone-based and line-based designs side by side.

Table 2: Summary of cost components in zone-based and line-based systems

Metrics	Zone-based system	Line-based system
$Q_1$	$\frac{4ND}{H}$	$\frac{4ND}{H_1}$
$Q_2$	$\frac{13N^2-5}{6} \frac{\lambda D^3}{N^3}$	$\frac{5ND}{2H_2} + \frac{2p_y \lambda D^3}{3N}$
$\frac{1}{v_{c1}}$	$\frac{1}{v} + \frac{4\tau_1 N^2}{Q_1 H} + \frac{\tau'_1(1+e_T)\lambda D^2}{Q_1}$	$\frac{1}{v} + \frac{4\tau_1 N^2}{Q_1 H_1} + \frac{\tau'_1(1+e_T)\lambda D^2}{Q_1}$
$\frac{1}{v_{c2}}$	$\frac{1}{v} + \frac{2(2N^2-1)\tau_2 \lambda D^2}{N^2 Q_2}$	$\frac{1}{v} + \frac{2\tau_2 p_y \lambda D^2}{Q_2}$
$M$	$\frac{Q_1}{v_{c1}} + \frac{Q_2}{v_{c2}}$	$\frac{Q_1}{v_{c1}} + \frac{Q_2}{v_{c2}}$
$A$	Not considered	$\begin{cases} \frac{4p_n \beta D}{3N v_w}, & 0 < \beta \leq 0.5, \\ \frac{3-4(1-\beta)^2(1+2\beta)}{6-12(1-\beta)^2} \frac{2p_n D}{N v_w}, & 0.5 < \beta \leq 1. \end{cases}$
$W$	$\frac{7D}{6Nv} + \frac{H}{2} \left[ 1 + \frac{(N-1)^2}{N^2} \right]$	$p_y H_2 + \frac{H_1}{2} \left[ 1 + \frac{(N-1)^2}{N^2} \right]$
$T_1$	$\frac{0.68DN(N-1)}{N^2 v_{c1}}$	$\frac{0.34D(2N^2-2N+1)}{N^2 v_{c1}}$
$T_2$	$\frac{(3N^2-1)D}{3N^3 v_{c2}}$	$\begin{cases} \frac{3-8\beta^3}{12-24\beta^2} \frac{p_y Q_2 H_2}{N^2 v_{c2}}, & 0 < \beta \leq 0.5, \\ \left( \frac{1}{6} + \frac{1}{3}\beta \right) \frac{p_y Q_2 H_2}{N^2 v_{c2}}, & 0.5 < \beta \leq 1. \end{cases}$
$e_T$	$\frac{(N-1)^2}{N^2}$	$\frac{(N-1)^2}{N^2}$

The optimal design problem of the line-based system can be formulated similarly as follows.

$$\min z(N, H_1, H_2, \beta) = \pi_Q Q + \pi_M M + W + A + T + \frac{\delta}{v_w} e_T \quad (18)$$

$$\text{s.t. } H_1 > 0, H_2 > 0, N \in \{1, 2, \dots, \lfloor \frac{D}{s} \rfloor\}, 0 < \beta \leq 1. \quad (19)$$

## 5 Numerical analysis

In this section, we conduct numerical experiments to compare the two RSP designs. The default values of the input parameters are listed in Table 3. The operation cost per distance  $\$_Q$  for the e-hailing vehicles in both the zone-based system and the line-based system is set to be 30% of the value in Table 3 because this service can be operated with smaller vehicles, as explained in Chen and Nie (2016). Also worth reiterating is that no walking cost is involved in the zone-based system. And we choose a small walking speed to penalize the walking time. As mentioned before, both optimal design problems are solved using Matlab’s built-in genetic algorithm.

Table 3: Values of input parameters in the numerical experiment

Notation	Value	Description
$s(km)$	0.15	the distance between two adjacent streets (street spacing)
$\mu(\$ / h)$	20	value of time
$\tau_1(s)$	12	time lost per stop due to deceleration and acceleration
$\tau'_1(s)$	1	time added per boarding passenger for fixed-route vehicles
$\tau_2(s)$	13	additional pick-up and drop-off time required per passenger
$v(km/h)$	25	vehicles’ cruising speed
$v_w(km/h)$	2	walking speed
$\delta(km)$	0.03	transfer penalty expressed in terms of the equivalent distance walked
$\$_Q(\$ / veh \cdot km)$	2	operation cost per vehicle distance
$\$_M(\$ / veh \cdot h)$	40	operation cost per vehicle hour

### 5.1 Base scenario

Figure 3 plots the optimal cost (per passenger) of the two systems against various demand levels  $\lambda$  for two different network sizes,  $D = 10km$  and  $D = 20km$ , chosen to represent a small- and mid-size US city respectively (Daganzo, 2010b). Both plots show that the line-based system has a lower system cost across all demand levels. Moreover, the absolute cost difference of the two design remains flat as the level of demand changes. Because the size of the study area does not seem to matter to the relative performance, we shall focus on the mid-size city in what follows.

Figure 4 compares the optimal design parameters for two systems when  $D = 20km$ . Figure 4a shows that the optimal zone-based design requires more fixed-route lines, hence smaller stop spacing. This is expected because the e-hailing vehicles in the zone-based system must travel to the pickup location without passengers on board, and the smaller stop spacing (which determines the zone size) help decrease such undesirable travel distance. Figure 4b shows that the fixed-route service operates on a smaller headway in the line-based system than in the zone-based counterpart, which implies that passenger would incur shorter

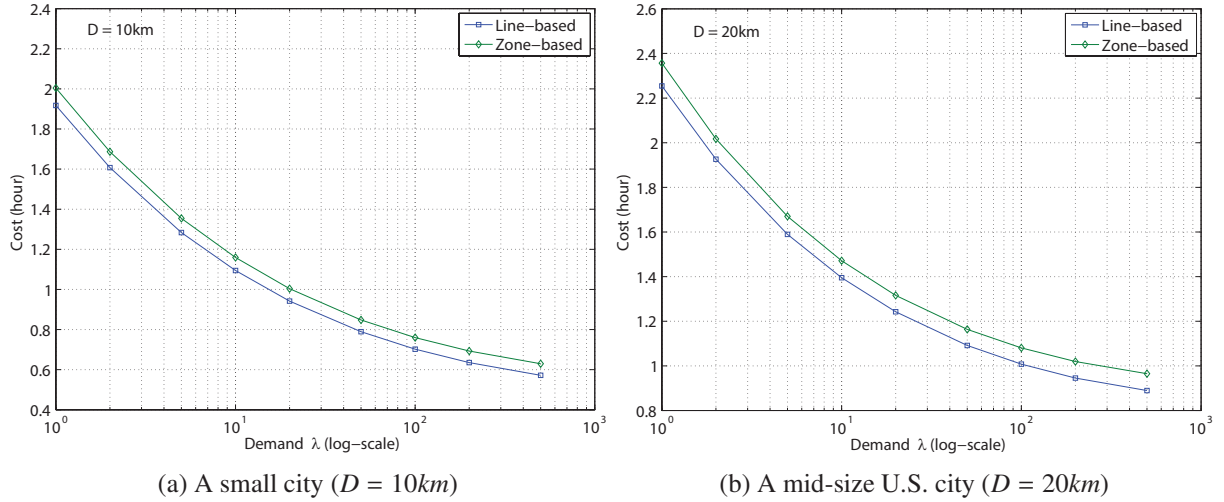


Figure 3: Optimal system cost vs demand level with default input parameters

waiting time for the fixed-route service in the line-based system. It also suggests that the e-hailing service in the line-based system must be operated more frequently than the fixed-route service at optimum.

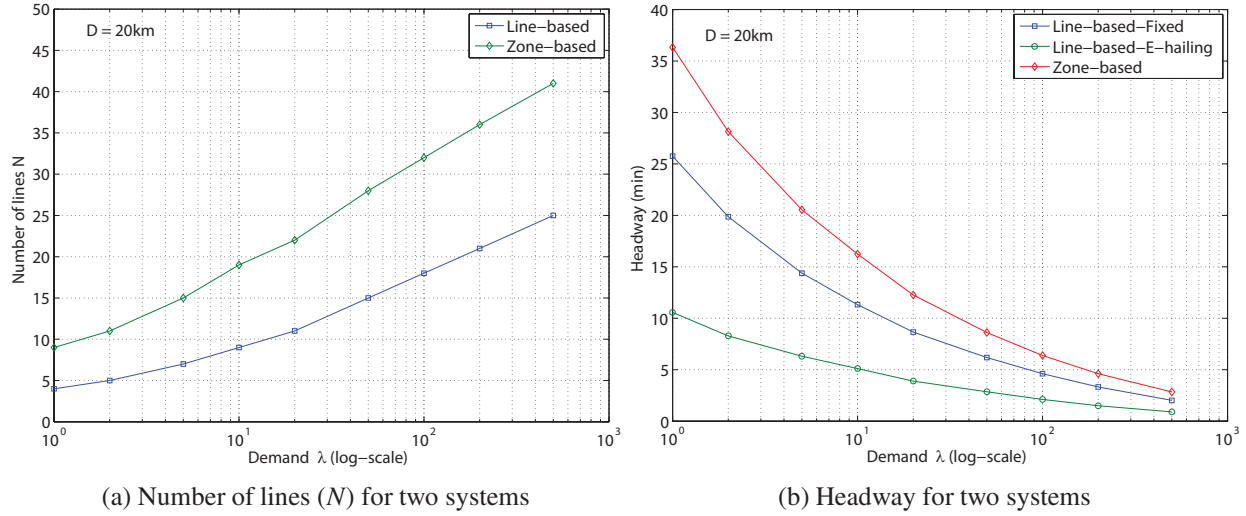


Figure 4: Optimal design parameters for two RSP designs

Figure 5a shows that the walking range parameter  $\beta$  in the line-based system remains relatively flat as the level of demand rises. Yet, since the number of lines increases with demand, the maximum walking distance decreases. Figure 5b highlights the fact that the walking cost constitutes a rather small portion of the total cost: it is always lower than 1.5% of the total cost, and for higher levels of demand, lower than 1%.

We proceed with more detailed analysis of various cost components. Figure 6 compares the agency and user costs in the two RSP designs. These plots first show that both agency and user costs per passenger decrease as the level of demand increases, thanks to the increasing returns to scale. Also, the average agency cost is significantly more sensitive to the demand level than is the user cost. When  $\lambda$  increases from 1 to 100, for example, the average agency cost drops almost 70%, whereas the average user cost only drops about 40%. Notably, the line-based system has lower agency and user costs across all demand levels. For

higher demand levels, the advantage of the line-based system seems greater: the absolute gap in the average user cost is enlarged by the increase in the demand, and because the total cost is decreasing, even the same absolute difference would imply a greater benefit in the relative sense.

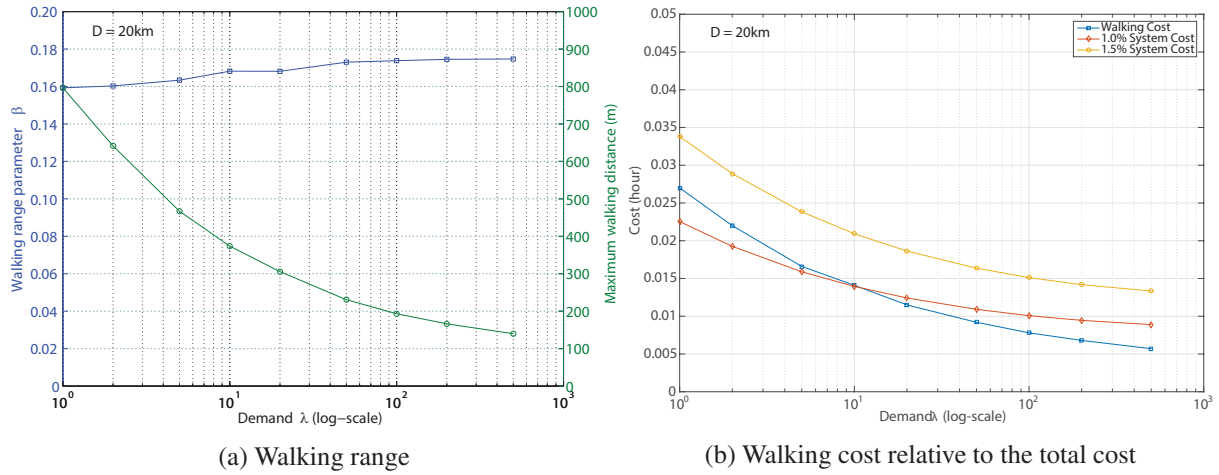


Figure 5: Walking range and relative cost for the line-based system

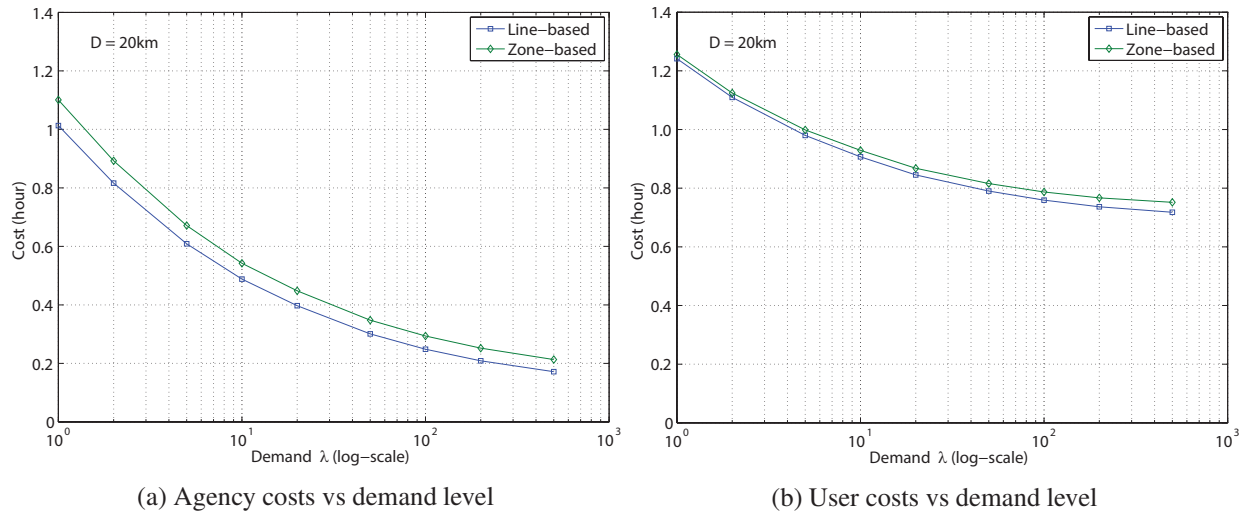


Figure 6: Agency and user costs for two RSP designs

Figure 7 shows that the two agency cost components, the total fleet size and distance travelled, have a rather consistent linear relation with the demand density on a log scale for both systems. A closer look suggests that the slope of this linear relation, denoted as  $\alpha$ , is a value between 0.5 and 0.6. This implies that the total agency cost is a function of  $\lambda^\alpha$ , or that the average agency cost is a function of  $\lambda^{\alpha-1}$ . For comparison, Daganzo (2010a) (Chapter 4) shows that for an idealized two-dimensional fixed-route transit system  $\alpha$  is  $2/3$ .

Figure 7a further shows that in both systems e-hailing vehicles travel a longer distance than fixed-route vehicles. Also, the fixed-route vehicles in the zone-based system travel more than those in line-based system across all demand levels, mainly because the zone-based system requires operating more fixed-route lines. Another observation is that the zone-based system requires a larger fleet of e-hailing vehicles for all demand



levels (see Figure 7b), but on average each e-hailing vehicle seems to travel less distance in the zone-based system. That the line-based design leads to longer e-hailing distance is not surprising, given the design requires e-hailing vehicles to travel along a fixed route regardless of the level of demand. In contrast, e-hailing vehicles in zone-based systems can circulate within a relatively small area, which helps reduce the average travel distance, especially when the demand is light.

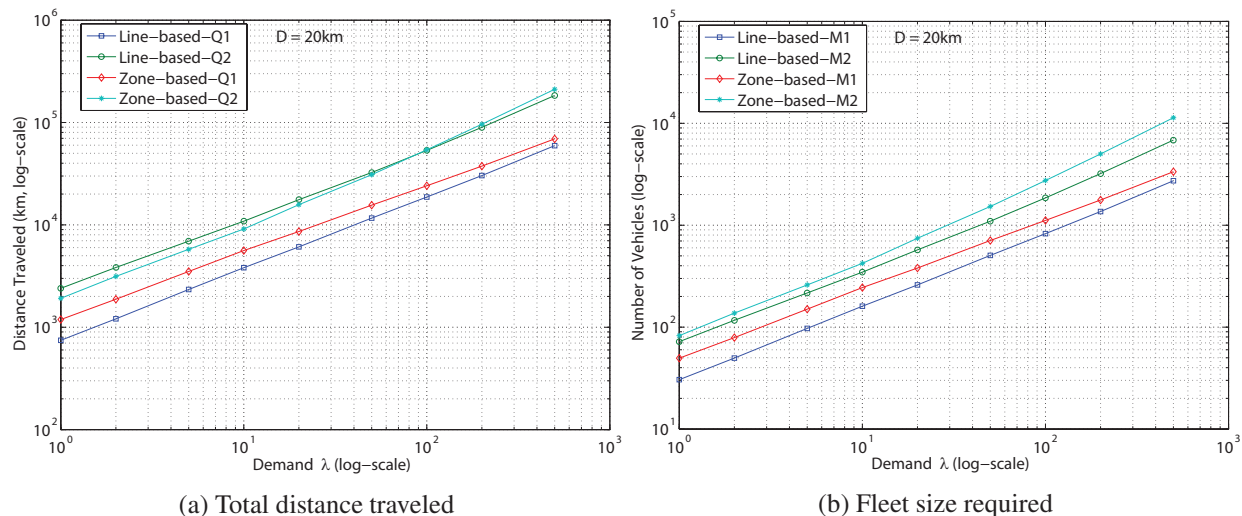


Figure 7: Components related to agency costs

Figure 8 plots the components related to the user costs for both systems. Since the walking time and transfer penalty are relatively small, we do not compare them here. Figure 8a shows that the zone-based system has longer waiting time across all demand levels, although the difference is smaller for higher demand levels. The longer waiting time in the zone-based system is mainly attributed to the fact that its fixed-route service operates on a larger headway. Figure 8b indicates that passengers actually spend slightly more time sitting in fixed-route vehicles when the demand increases in both systems. On the other hand, the time spent on the e-hailing vehicles keep decreasing with the demand level. The reason for this trend, again, is because a higher demand level requires a fixed-route service with smaller spacing. On one hand, smaller spacing means each e-hailing vehicle has shorter distance to traverse on average, and the impact is much greater for the zone-based design than for the line-based design, as clearly shown in the figure. On the other hand, smaller spacing also means passengers are more likely to transfer, which contributes to the longer riding time in the fixed-route service. Finally, compared to the line-based design, the zone-based system offers shorter in-vehicle time for the e-hailing service, but longer in-vehicle time for the fixed-route service.

## 5.2 Sensitivity analysis

In this section a sensitivity analysis is performed to test how the relative performance of the two RSP designs varies with input parameters. Because the walking zone is only designated in the line-based system, we set out to first test to what extent allowing walking contributes to the discrepancy revealed in the previous section. To this end, we revise the line-based design model to force  $\beta \leq 0.001$ , practically eliminating walking. As shown in Figure 9, such a no-walking line-based system delivers a similar total system cost as the system with an optimally configured walking area. More important to notice is the fact that the line-based system still outperforms the zone-based system with a wide margin, even when walking is prohibited

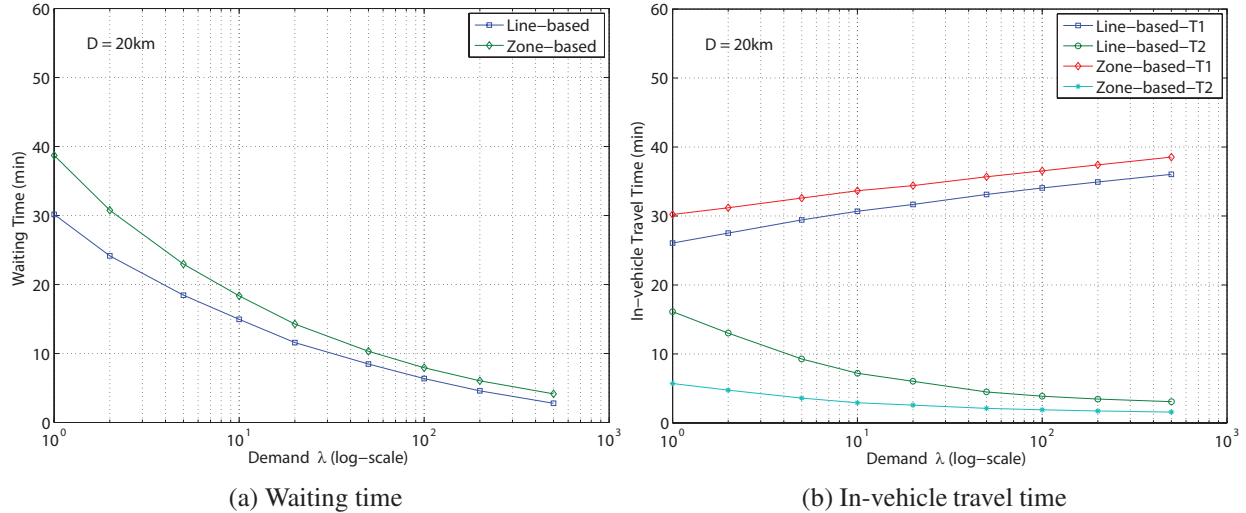


Figure 8: Components related to user costs

as in the zone-based system. This confirms that the main performance discrepancy between the line- and zone-based RSP designs cannot be attributed to walking.

Figure 10 compares the optimal system costs of the two designs with input parameters different from the default values. Figure 10a shows the case of a much longer dwell time (30 seconds vs. 18 in the base scenario); Figures 10b and 10c show the cases when the walking is penalized much more and less, respectively; and in Figure 10d waiting is given a much greater weight. In all four cases, the line-based system consistently outperforms the zone-based counterpart with a margin similar to or greater than that seen in the base scenario. Moreover, it clearly enjoys a greater lead for longer dwell times and higher waiting costs. Thus, the relative advantage of the line-based system seems robust against inputs.

## 6 Validation by agent-based simulation

To validate the analytical results, a discrete-event agent-based simulation model is developed and tested for each RSP design in this section. The simulation model is implemented in NetLogo<sup>12</sup>, a multi-agent programmable modeling environment. The simulation is conducted in a square service area with side length  $D = 20\text{km}$  and constant street spacing  $s = 150\text{m}$ . A brief description of the simulation model will be presented first, followed by a comparison between simulation and analytical results.

### 6.1 Simulation model

#### 6.1.1 Zone-based system

There are three types of agents in the zone-based system, namely, passengers, fixed-route vehicles, and e-hailing vehicles. The behavior of each agent is described below. Note that the transit stops closest to a passenger's actual origin (destination) are called the origin (destination) stops.

- **Passengers** do not walk in this system. Instead, they will always request the e-hailing service to travel from their home to the origin stop. After arriving at the origin stop, they will take the fixed-route

<sup>12</sup><https://ccl.northwestern.edu/netlogo/>

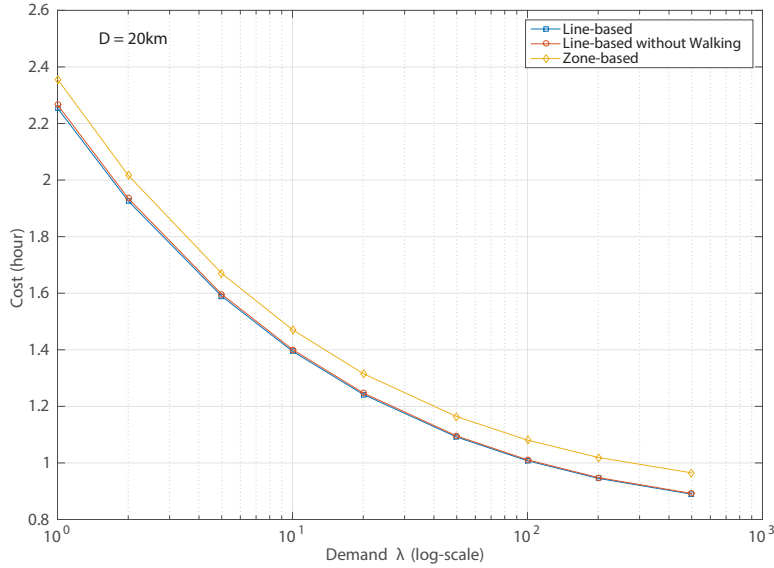


Figure 9: Optimal system cost vs demand level for three systems

transit to their destination stop (transfer will be made if necessary). After arriving at their destination stop, passengers will again use the e-hailing service to travel to the final destination. If a passenger's origin and destination are in the same zone, they will take e-hailing service from the origin directly to the destination.

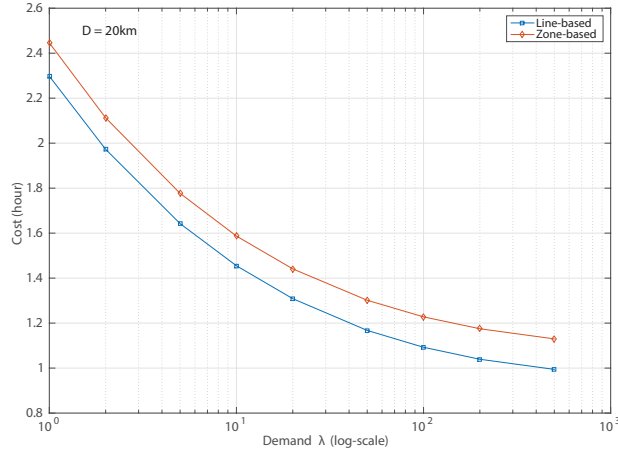
- **Fixed-route vehicles** are dispatched from the terminal stop according to the optimal headway. They will stop at each transit stop to pick up and drop off passengers.
- **E-hailing vehicles** will pick up passengers who request the service from the origin and take them to their origin stop or to their destination if the destination is in the same zone. These vehicles will also pick up passengers who need the service at each transit stop and drop them off at their destination. E-hailing vehicles only operate in their own zones.

When a passenger requests the e-hailing service, the request will be assigned to the closest idle e-hailing vehicle. If all e-hailing vehicles in the same zone are busy, the request will be put into a first-in-first-out (FIFO) queue. Once an e-hailing vehicle completes its current task, it will be assigned to the first passenger in the FIFO queue. If the queue is empty, the e-hailing vehicle will simply stop at its current location and wait for the next request.

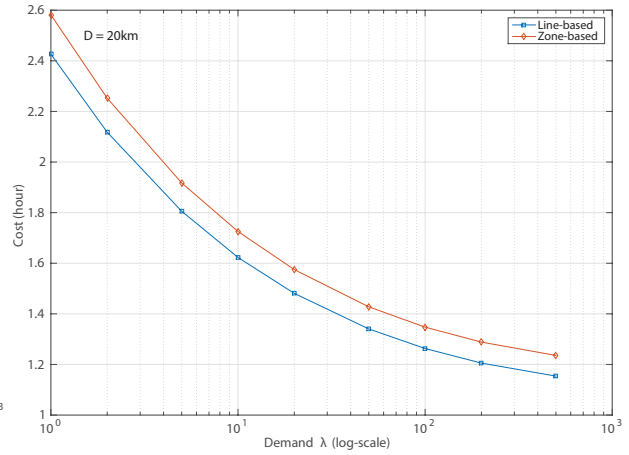
### 6.1.2 Line-based system

The behavior of the three agents (passengers, fixed-route and e-hailing vehicles) is described as follows:

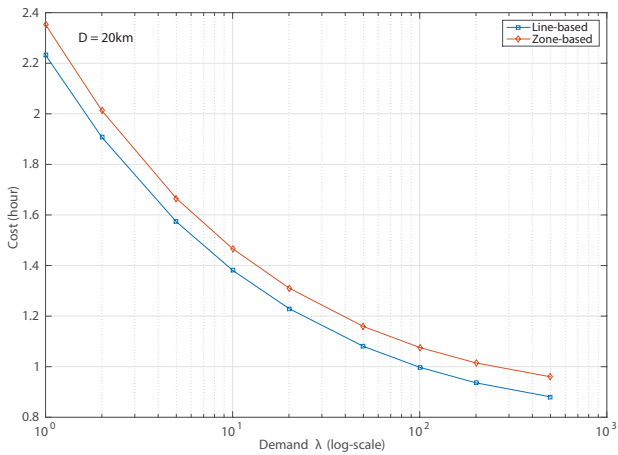
- **Passengers** will walk to their origin stop if the stop is within the walking range; otherwise, they will request and use the e-hailing service to reach their origin stop. After arriving at the origin stop, they will take the fixed-route transit to their destination stop (transfer will be made if necessary). After arriving at their destination stop, they will directly walk to the destination if it is within the walking range; otherwise, they will wait for the e-hailing service.



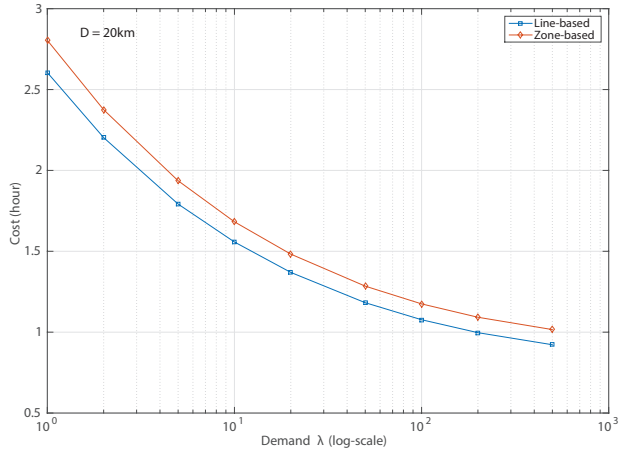
(a) Long dwell time:  $\tau_1 = 30s, \tau_2 = 31s$  ( $\tau_2 = \tau_1 + 1$  in our experiments)



(b) Slow walking:  $v_w = 0.1km/h$



(c) Fast walking:  $v_w = 3km/h$



(d) More expensive waiting: 1 unit of waiting time = 1.8 unit of in-vehicle time

Figure 10: Optimal system cost vs demand level with different input parameters

- **Fixed-route vehicles** in the line-based system behave the same way as in the zone-based system.
- **E-hailing vehicles** are also dispatched from the terminal stop according to the optimal headway and travel along the fixed-route line. E-hailing vehicles can deviate from the route to pick up passengers who request the service from the origin and take them to their origin stop. These vehicles can also pick up passengers who need the service at each transit stop and drop them off at their destination.

A key issue in the simulation of the line-based system is representing the route planning of the e-hailing vehicles. Our implementation follows operational characteristics discussed in Section 3.1. Additional details are provided below. When an e-hailing vehicle arrives at a stop, it will first drop off passengers who need to depart from the stop and collect the passengers who just arrive at the stop and need the e-hailing service. It will then determine the list of passengers who need to be picked up on its way to next transit stop. When all pick-up and drop-off locations are determined, the visit order will be simply determined based on the longitudinal distance from the current stop. Once a vehicle leaves the stop, no further requests between that stop and the next stop will be processed, which rules out the possibility of “backward” movement.

## 6.2 Simulation results

We report the simulation results for four different demand levels, i.e.,  $\lambda = 10, 20, 50, 100$ . It is worth emphasizing that, in the simulation, the number of lines may not be exactly the same as the optimal number of lines obtained in the previous section because the line spacing ( $D/N$ ) must be multiples of street spacing  $s$  (see Figure 11a). In addition, the number of e-hailing vehicles in each zone (for the zone-based RSP) is approximated by  $M_2/N^2$ , which may also not be an integer. In the simulation, we choose the closest integer to the optimal numerical value for each zone, as shown in Figure 11b. Since the number of e-hailing vehicles per zone is typically small, this approximation may cause substantial approximation errors in the total number of e-hailing vehicles. Similar approximation errors exist for the number of e-hailing vehicles in the line-based system and for the number of fixed-route vehicles in both systems. Yet, those errors are smaller because the fleet size is larger. Finally, we note that the rather small total number of e-hailing vehicles in each zone implies that the number of idle e-hailing vehicles is even smaller, which justifies approximating  $d_{e1}$  and  $d_{e2}$  by  $d_1$  and  $d_2$ .

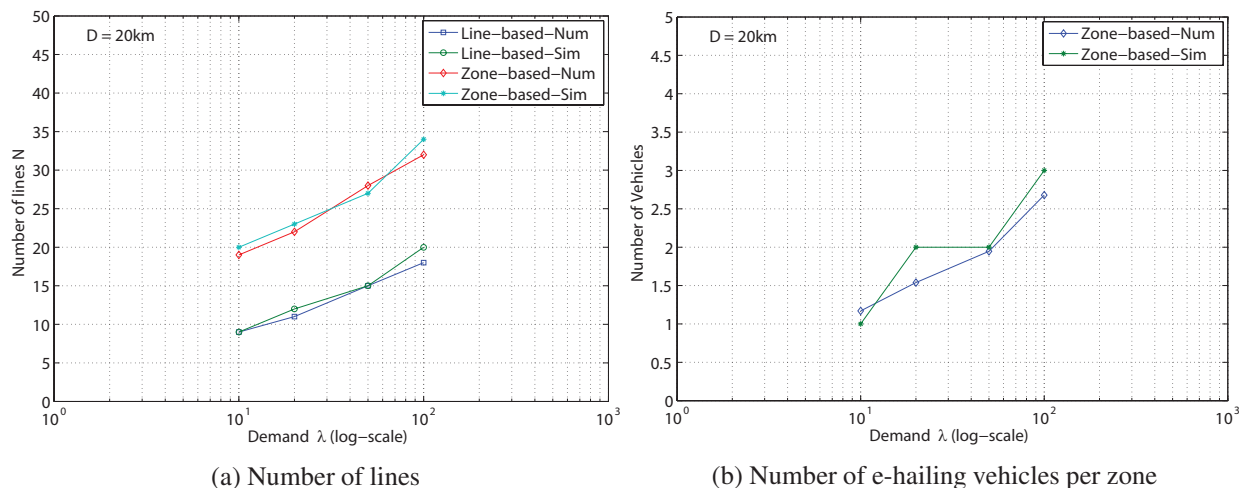


Figure 11: Comparison of design parameters in simulation and analytical results

Figure 12 compares the simulation and analytical results in terms of the total system cost for both RSP systems. The plots show that the simulation results generally match the numerical results well, despite that the input parameters are subject to nontrivial approximation errors in some cases. The discrepancy between the simulated and analytical costs is within 5% in all cases except for the line-based system when  $\lambda = 10$ . The reason for the relatively large discrepancy in this case is likely the small number of fixed-route lines. The simulation may not represent the transit network well in this case due to street spacing constraint. However, even in this case, the difference is still just around 7%. This finding indicates that the analysis based on continuous approximation provides a robust estimation of the overall system cost in our problem setting.

Figure 13 compares the simulated system costs of the two systems. It confirms that the line-based system consistently outperforms the zone-based system for all tested demand levels, with a similar margin as predicted by the analytical results.

We now examine to what extent the waiting time in the simulated results deviate from the estimation adopted in the analysis. In Table 4,  $W_o$ ,  $W_{os}$ ,  $W_{ts}$ , and  $W_{ds}$  denote the waiting time at origin, origin stop, transfer stop, and destination stop, respectively; and  $z$  and  $z_u$  denotes the total system cost and user costs,

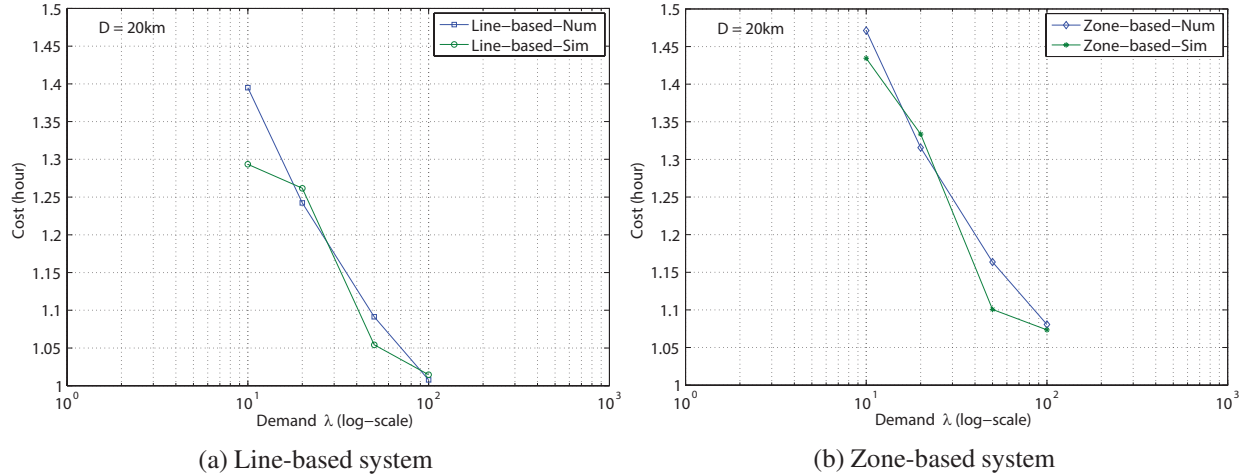


Figure 12: Simulation results vs numerical results in total system cost for two systems

respectively. As shown from the table, the waiting time at the origin is much shorter compared to the analytical results. This is because in simulation passengers are required to make the service request 15 minutes before their desired departure time, and they are considered to “begin to wait” only after that desired departure time passes. At the origin stop, the waiting times in the simulation results are significantly lower in all cases. We believe that this phenomenon is also a result of the 15 minute planning horizon implemented in the simulation. Because the planning horizon allows demand consolidation in time, passengers no longer “randomly” arrive at the stop, and accordingly their average waiting time is lower than half of the headway<sup>13</sup>. As a comparison, the simulated waiting times at transfer stops match half of the headway much better, likely because passenger arrival patterns at these stops meet the random arrival assumption well. At the destination stops, however, simulated waiting times are significantly higher. For the line-based system, the simulated waiting time is about 1.5 times of the analytical counterpart, whereas in the zone-based system, the discrepancy is much more uneven. At least some of the differences may be attributed to demand fluctuations, and the fact that the design parameters in the simulation are not exactly the same as in the analytical results. Yet, offering a convincing theory to explain these differences require a more in-depth investigation and we leave that to future research. Finally, Table 4 also reports the percentage of the waiting cost in the total system cost and the user cost. When the level of demand is low, the waiting cost accounts for up to about 30% of the user cost and up to about 20% of the total system cost. As the demand increases, the importance of the waiting cost decreases.

## 7 Conclusions

The rise of e-hailing is transforming personal mobility. However, to evolve into a true mass transport platform, this innovation must achieve greater ride consolidation in both time and space. One possibility is to use e-hailing as a feeder to conventional fixed-route transit systems. This paper analyzes and compares two

<sup>13</sup>The 15 minute planning horizon allows the demand consolidation precisely because it expands a passenger’s desired departure time from a point to a 15-minute window. Thus, unless the demand is very high relative to the supply, e-hailing vehicles can consolidate departures that would disperse more widely in time than the case without such a planning horizon - imagine two travelers whose desired departure times are 10 minutes apart, but can now be served almost simultaneously because of this consolidation effect.

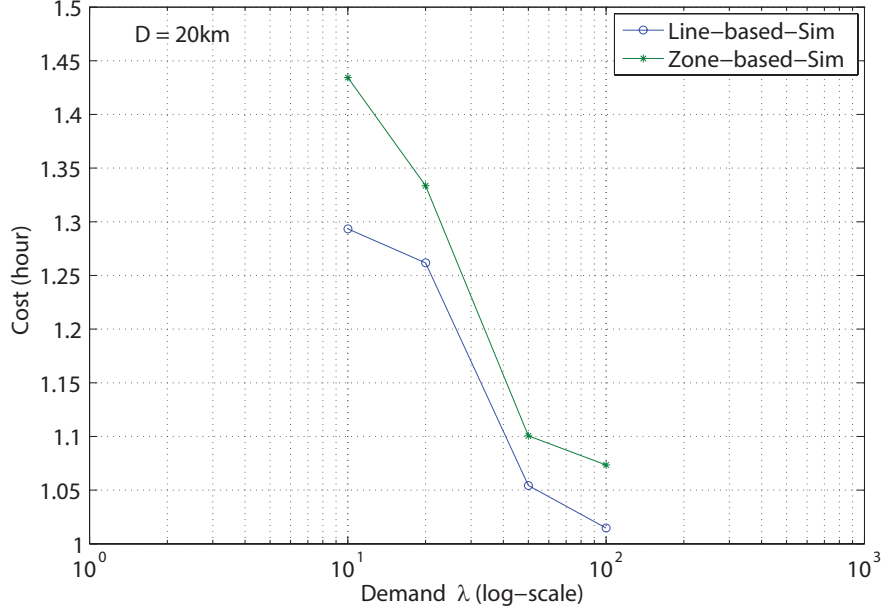


Figure 13: Comparison of system cost in simulation

different relative spatial position (RSP) designs in an integrated e-hailing/fixed-route transit system: a zone-based design that operates e-hailing vehicles within a zone, and a line-based design that operates e-hailing vehicles along a fixed-route transit line and with a stable headway. While prototypes of these integrated systems have been studied in the literature, little is known about their relative performance, to the best of our knowledge. To conduct a meaningful comparison, we formulated the optimal design problem for both systems using the same analytical framework based on the continuous approximation approach. We then performed a comprehensive numerical experiment to compare various cost components corresponding to the optimal designs. To validate our analysis, a discrete-event simulation model is developed and the results from simulation are compared against the analytical results. The main findings from these numerical and simulation experiments are summarized below.

- The analytical and simulation results match each other well, with a discrepancy in the total system cost less than 5% in most test scenarios. This finding supports that the continuous approximation approach is a valid tool for addressing high-level design decisions such as the choice of RSP.
- All results, analytical and simulation, show that the line-based system outperforms the zone-based system in terms of both agency and user costs, for all scenarios tested.
- Compared to the zone-based design, the line-based design features a sparser fixed-route network (resulting in larger stop spacing) but a higher dispatching frequency.

One naturally wonders why, intuitively, the line-based system is the winner. A plausible explanation is that the e-hailing vehicles in the zone-based system might experience relatively longer detour per passenger because they have to make dedicated pickup trips without carrying any riders. In contrast, the line-based design operates the e-hailing vehicles with a more regular route/headway structure, which promises greater efficiency. It seems that the lower efficiency in operating the e-hailing service is the reason why the optimal zone-based design dictates smaller line spacing (hence smaller zones) and more e-hailing vehicles.

Table 4: Waiting time comparison between numerical and simulation results

(a) Waiting time of line-based system

$\lambda$		$W_o(min)$	$W_{os}(min)$	$W_{ts}(min)$	$W_{ds}(min)$	$W(min)$	$W/z(\%)$	$W/z_u(\%)$
10	Analytical	2.41	5.66	4.47	2.41	14.95	17.87%	27.48%
	Simulation	0.35	4.39	3.17	3.49	11.41	14.70%	20.92%
20	Analytical	1.84	4.33	3.58	1.84	11.57	15.53%	22.83%
	Simulation	0.23	3.13	4.21	2.63	10.20	13.47%	19.32%
50	Analytical	1.34	3.09	2.69	1.34	8.47	12.93%	17.85%
	Simulation	0.28	2.17	2.77	1.98	7.20	11.38%	15.16%
100	Analytical	1.00	2.31	2.06	1.00	6.36	10.51%	13.95%
	Simulation	0.38	1.61	2.14	1.52	5.65	9.28%	12.30%

(b) Waiting time of zone-based system

$\lambda$		$W_o(min)$	$W_{os}(min)$	$W_{ts}(min)$	$W_{ds}(min)$	$W(min)$	$W/z(\%)$	$W/z_u(\%)$
10	Analytical	1.69	8.12	7.29	1.26	18.35	20.79%	32.92%
	Simulation	0.01	5.63	7.93	4.92	18.49	21.49%	33.56%
20	Analytical	1.45	6.13	5.59	1.09	14.26	18.07%	27.39%
	Simulation	0.00	4.19	5.39	1.40	10.98	13.72%	22.47%
50	Analytical	1.14	4.31	4.01	0.86	10.31	14.77%	21.07%
	Simulation	0.00	2.94	4.02	2.51	9.47	14.34%	20.16%
100	Analytical	1.00	3.19	2.99	0.75	7.93	12.23%	16.80%
	Simulation	0.00	2.18	3.04	0.82	6.04	9.37%	13.44%

We caution that other factors may contribute to the superiority of the line-based system that call for further investigation. For one thing, the line-based system does not provide the e-hailing service to passengers whose origin and destination fall into the designed walking area. While the walking cost is taken into consideration, it may not be properly valued relative to other costs. Another issue has to do with the fact that ride-sharing is allowed in the line-based system by design, but not in the current zone-based system. It would be relatively easy to incorporate ride-sharing in a simulated zone-based system, in which a real-time vehicle routing algorithm can be implemented to consolidate rides. It is much more challenging, however, to analytically account for the impact of coordinated ride-sharing in the optimal design problem.

In general, the operational aspects of e-hailing vehicles are largely simplified in the current analysis to maintain tractability. There are certainly opportunities in both systems to improve the performance through more sophisticated routing and control methods that properly consolidate rides in time and space. The question is to what extent such tactics matter to strategic decisions, and if so how they can be incorporated into the making of those decisions. Finally, the optimization problem is currently solved using a generic algorithm built in MATLAB. Developing specialized and more efficient algorithms warrants further investigation, especially if the model is to be extended to include more decision variables (such as routing and control).



## Acknowledgement

The work is partially funded by National Science Foundation under the award number CMMI-1402911. Constructive comments of two anonymous reviewers are greatly appreciated.

## References

- Aldaihani, M. M., Quadrioglio, L., Dessouky, M. M., Hall, R., 2004. Network design for a grid hybrid transit service. *Transportation Research Part A: Policy and Practice* 38 (7), 511–530.
- Alshalalfah, B. W., 2009. Planning, design and scheduling of flex-route transit service. Ph.D. thesis, University of Toronto.
- Bertoncello, M., Wee, D., 2015. Ten ways autonomous driving could redefine the automotive world. Tech. rep., McKinsey & Company.  
URL <http://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world>
- Black, A., 1995. *Urban mass transportation planning*. McGraw-Hill, New York.
- Chen, P., Nie, Y., 2016. A demand adaptive paired-line hybrid transit system. Working paper.
- Cordeau, J.-F., 2006. A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research* 54 (3), 573–586.
- Cordeau, J.-F., Laporte, G., 2003a. The dial-a-ride problem (darp): Variants, modeling issues and algorithms. *Quarterly Journal of the Belgian, French and Italian Operations Research Societies* 1 (2), 89–101.
- Cordeau, J.-F., Laporte, G., 2003b. A tabu search heuristic for the static multi-vehicle dial-a-ride problem. *Transportation Research Part B: Methodological* 37 (6), 579–594.
- Cortés, C. E., Jayakrishnan, R., 2002. Design and operational concepts of high-coverage point-to-point transit system. *Transportation Research Record: Journal of the Transportation Research Board* 1783 (1), 178–187.
- Crainic, T., Malucelli, F., Nonato, M., 2001. Flexible many-to-few+ few-to-many= an almost personalized transit system. TRISTAN IV, Miguel Azores Islands.
- Daganzo, C. F., 1978. An approximate analytic model of many-to-many demand responsive transportation systems. *Transportation Research* 12 (5), 325–333.
- Daganzo, C. F., 2010a. *Public transportation systems: Basic principles of system design, operations planning and real-time control*. Tech. rep., Institute of Transportation Studies, University of California at Berkeley.
- Daganzo, C. F., 2010b. Structure of competitive transit networks. *Transportation Research Part B: Methodological* 44 (4), 434–446.
- He, F., Shen, Z.-J. M., 2015. Modeling taxi services with smartphone-based e-hailing applications. *Transportation Research Part C: Emerging Technologies* 58, 93–106.

- 1  
2  
3  
4 Malucelli, F., Nonato, M., Pallottino, S., 1999. Demand adaptive systems: some proposals on flexible transit.  
5 Operations research in industry, 157–182.  
6  
7  
8 Melachrinoudis, E., Ilhan, A. B., Min, H., 2007. A dial-a-ride problem for client transportation in a health-  
9 care organization. Computers & Operations Research 34 (3), 742–759.  
10  
11 Potter, B., 1976. Ann arbor’s dispatching system. Transportation Research Record 608, 89–92.  
12  
13 Pratelli, A., Schoen, F., et al., 2001. A mathematical programming model for the bus deviation route prob-  
14 lem. Journal of the Operational Research Society 52 (5), 494–502.  
15  
16 Psaraftis, H. N., 1980. A dynamic programming solution to the single vehicle many-to-many immediate  
17 request dial-a-ride problem. Transportation Science 14 (2), 130–154.  
18  
19  
20 Quadrifoglio, L., Dessouky, M. M., Ordóñez, F., 2008. Mobility allowance shuttle transit (mast) services:  
21 Mip formulation and strengthening with logic constraints. European Journal of Operational Research  
22 185 (2), 481–494.  
23  
24  
25 Quadrifoglio, L., Dessouky, M. M., Palmer, K., 2007. An insertion heuristic for scheduling mobility al-  
26 lowance shuttle transit (mast) services. Journal of Scheduling 10 (1), 25–40.  
27  
28 Quadrifoglio, L., Hall, R. W., Dessouky, M. M., 2006. Performance and design of mobility allowance shuttle  
29 transit services: bounds on the maximum longitudinal velocity. Transportation science 40 (3), 351–363.  
30  
31  
32 Ropke, S., Cordeau, J.-F., 2009. Branch and cut and price for the pickup and delivery problem with time  
33 windows. Transportation Science 43 (3), 267–286.  
34  
35  
36 Stein, D. M., 1978. Scheduling dial-a-ride transportation systems. Transportation Science 12 (3), 232–249.  
37  
38 Wang, X., He, F., Yang, H., Gao, H. O., 2016. Pricing strategies for a taxi-hailing platform. Transportation  
39 Research Part E: Logistics and Transportation Review 93, 212–231.  
40  
41 Wilson, N. H., Weissberg, R. W., Hauser, J., 1976. Advanced dial-a-ride algorithms research project. Tech.  
42 rep., MIT.  
43  
44 Zha, L., Yin, Y., Yang, H., 2016. Economic analysis of ride-sourcing markets. Transportation Research Part  
45 C: Emerging Technologies 71, 249–266.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65