Transportation Research Part B 46 (2012) 204–216

Contents lists available at SciVerse ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

A structured flexible transit system for low demand areas

Seyed Mohammad Nourbakhsh, Yanfeng Ouyang *

Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Received 5 August 2010
Received in revised form 29 July 2011
Accepted 30 July 2011

Keywords:
Flexible transit system
Hybrid network
Low demand
Optimal design

ABSTRACT

Public transit structure is traditionally designed to contain fixed bus routes and predetermined bus stations. This paper presents an alternative flexible-route transit system, in which each bus is allowed to travel across a predetermined area to serve passengers, while these bus service areas collectively form a hybrid "grand" structure that resembles hub-and-spoke and grid networks. We analyze the agency and user cost components of this proposed system in idealized square cities and seek the optimum network layout, service area of each bus, and bus headway, to minimize the total system cost. We compare the performance of the proposed transit system with those of comparable systems (e.g., fixed-route transit network and taxi service), and show how each system is advantageous under certain passenger demand levels. It is found out that under low-to-moderate demand levels, the proposed flexible-route system tends to have the lowest system cost.

1. Introduction

Traditionally, transit systems are designed to contain fixed bus routes, and passengers move to predetermined stations to gain access to the system. The bus network design problem were studied as early as in the 1960s, and the search for a cost-minimizing urban transit system structure (e.g., route positions and headways) explored possible grid networks (Holroyd, 1965), radial systems (Byrne, 1975), hub-and-spoke systems (Newell, 1979), and general bus network design (e.g., Ceder and Wilson, 1986). Wirasinghe et al. (1977) found the optimal network parameters that minimize transit operating cost and passenger travel time in coordinated rail and bus transit systems. Desaulniers and Hickman (2007) summarized a variety of optimization problems that are related to public transit, while a number of other studies focused on specific aspects of transit network design, such as express transit design (Barnett, 1970), bus priority at signalized intersections (Balke et al., 2000), and bus lane priority (Eichler and Daganzo, 2006). Very recently, Daganzo (2010) proposed an innovative transit network design framework that determines the adequate structure of the network as well as the optimal headway for a range of transit modes. It was shown that such a hybrid network structure (see Fig. 1) nicely inherits the advantages of both a hub-and-spoke structure (e.g., low infrastructure investment) and a grid structure (e.g., low travel time).

Fixed routes in a transit system provide clarity and regularity to the transit service, and it is known to work very well for densely populated cities (i.e., with high passenger demand) in general. However, in low demand regions (e.g., sprawled suburban areas), the optimal spacing between bus routes often tends to be relatively large so as to reduce the total system cost. As such, the transit system exposes passengers to the open (sometime adverse) environment for a long time while they walk to and from the bus stops. In such cases, adding flexibility to transit route and schedule seems desirable. For example, Quadrifoglio et al. (2006) demonstrated the potential advantages of allowing transit vehicles to travel within a low-demand geographical region to pick up and drop off passengers. Properties of such operations (e.g., bounds on the maximum longitudinal velocity) are obtained, often via simulations, and incorporated into the overall network design via mixed-integer

* Corresponding author. Tel.: +1 217 333 9858.
E-mail address: yfouyang@illinois.edu (Y. Ouyang).

0191-2615/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2011.07.014

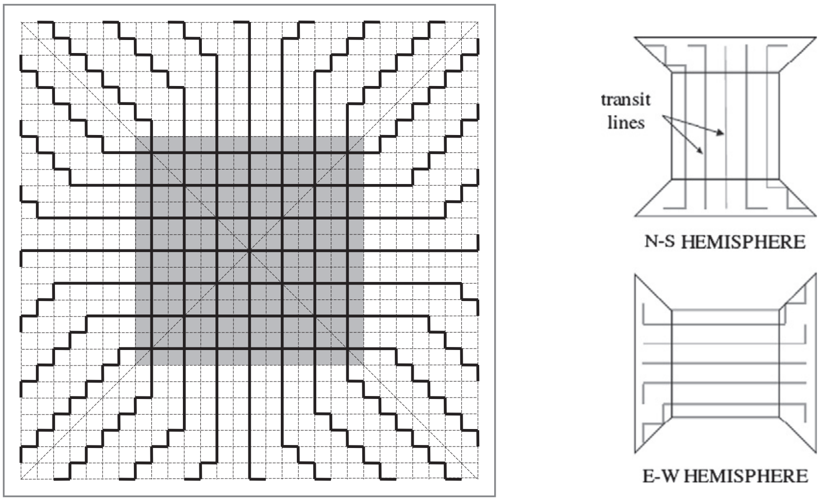

Fig. 1. A hybrid network structure (adapted from Daganzo (2010)).

programming models (e.g., Quadrifoglio et al., 2008). For feeder buses, the suitable demand density for the flexible-route operation is also studied (Quadrifoglio and Li, 2009, among others).

This paper aims to integrate these interesting ideas (e.g., hybrid network structure, flexible route) into the design of a new structured "flexible-route transit system." Individual buses operate without fixed routes or predetermined stops, but rather they can travel around within their own service regions to pick up or drop off passengers. At the macroscopic level, however, the buses (or their service regions) collectively form a suitable network structure to provide reliable spatial and temporal service coverage to the entire demand area. In this paper, instead of relying on complex mathematical programs to determine the optimal network structure numerically, we express the system's operating performance into analytical functions of a few key design variables, and solve for the optimal design as a simple constrained nonlinear optimization problem. These analytical functions also cast important insights into the impacts of these design variables on the overall system performance. Numerical examples and comparison with comparable alternative systems (e.g., the fixed-route transit system and taxi service) show that the proposed flexible-route transit system is advantageous under a range of low-to-moderate demand levels. This is encouraging because the proposed transit system can be used in a number of real-world applications. One possibility, for example, is to let the transit system switch among different operating modes according to the demand level at specific hours (e.g., at night or during weekend). In addition, the flexible network can be used for design of "safe ride" or "dial-a-ride" systems (Daganzo, 1984) in which passengers call and wait for pick-ups.

The exposition of this paper is as follows. Section 2 introduces the notation, concept and formulation of the flexible-route transit system. Section 3 examines other transit system structures (including the fixed-route transit system and a taxi system) that are comparable to the proposed one. In Section 4, the proposed system is numerically compared with the other systems at different demand levels. Finally, Section 5 provides conclusions.

2. Methodology

2.1. Notation and definition

Similar to Daganzo (2010), we consider a square service region of side D (km) that generates λ passenger trips per hour per unit area. The trip origins and destinations are uniform and independently distributed in the region according to a homogeneous spatial Poisson process. The local streets in this service region align along a grid network with constant spacing s.

We consider designing a new flexible transit system to provide service to the passengers when λ is relatively small. Unlike the traditional transit system where buses travel along a fixed route and make stops at predetermined stations, in the flexible transit system buses pick up passengers at their origins or drop them off at their destinations. Each bus now serves the passengers in a narrow elongated area, which we call it a "bus tube," as shown in Fig. 2a. The bus makes lateral movements while sweeping longitudinally back and forth through the tube. The exact bus trajectory obviously depends on the realization of passenger locations, but fixed transfer points are planned along each bus tube, and we assume buses always stop at those points.

Fig. 2b shows the grand overview of the structured flexible-route transit system, including the layout of tubes and the location of transfer points. Based on the intended service level, the whole demand area can be divided into the central square and the peripheral quadrants. The transit system includes N–S and E–W hemispheres each containing N equal transit tubes

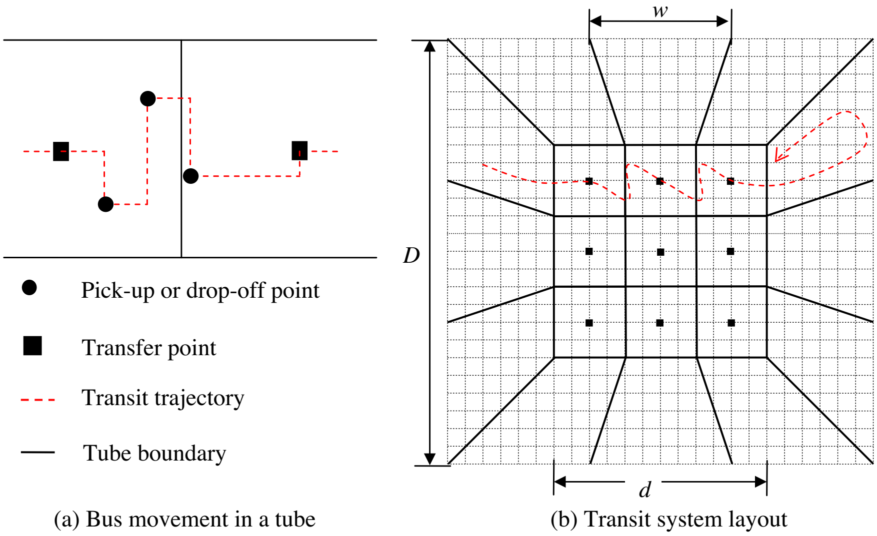

Fig. 2. General scheme of the structured flexible transit system.

(with variable width), providing double coverage in the central square and single coverage in the peripheral part.[1] Obviously, there are N transfer points in each tube (in the inner square) where passengers may transfer to other buses. The central square has a side length of d, and we use the ratio $\alpha = \frac{d}{D}$ to indicate the relative size of the inner square. The maximum width of the tubes (at the boundary of the entire area) is $w = \frac{D}{N}$, and the width of the tube inside the central square equals to αw. We further assume that each bus travels in its tube with headway H.

Our design problem is to find the optimal combination of decision variables, $\alpha \in [0, 1]$, $H \geqslant 0$, and $N \in \{1, 2, \ldots\}$, that minimize the total system costs for the proposed transit system. We further assume that the bus cruising speed is v, the time needed to make one stop (i.e., the delay for acceleration and deceleration) is τ_1, the time needed to pick up or drop off a passenger is τ_2. Since we assume one stop per passenger, the time needed per passenger stop is τ, where $\tau = \tau_1 + \tau_2$. As a result, the average bus travel speed (along its flexible trajectory) reduces to $v_c \leqslant v$. Let C be the capacity of the buses, and δ be the factor that captures the discomfort associated with transfers. Passenger can walk at speed v_w.

2.2. Formulation

The total system cost mainly consists of two parts: user costs and agency costs. In the following subsections we present formulas for these cost components. Detailed derivations of these formulas can be found in Appendix A.

2.2.1. 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. Unlike the fixed-route transit system, the proposed flexible-route system has no specific bus routes or specific locations for picking up or dropping off passengers, and thus the need for capital infrastructure investments (e.g., building bus stations) is minimal.[2] Appendix A shows that the following formula holds for Q:

$$Q = \frac{2N}{H}\left[D\sum_{i=2}^{\infty}(i-1)[\alpha P_c\{i\} + (1-\alpha)P_p\{i\}] + 2D + \frac{2\lambda H D^3\alpha^3}{3N^2} + \frac{2\lambda H D^2(1-\alpha^2)l_p}{N}\right], \quad (1)$$

where

$$P_c\{i\} = \left(\frac{\alpha D s\lambda H}{N}\right)^i e^{-\frac{\alpha D s\lambda H}{N}} \Big/ i!, \quad (2)$$

[1] Even though the passenger origins and destinations are assumed to be homogeneous over space and time, it is reasonable to design a higher density of transit network infrastructures near the central part of the city (where most of the traffic is expected to traverse). In the context of many-to-many freight logistics systems, Campbell (1990) showed that the optimal transshipment terminals shall be evenly spaced and yet clustered at the center of the service region.

[2] There may be need for user interfaces (e.g., Internet or phone service systems) and driver communication devices (e.g., radio). But their costs are normally negligible compared with traditional roadway infrastructure investments.

$$P_p\{i\} = \left(\frac{(1+\alpha)Ds\lambda H}{N}\right)^i e^{-\frac{(1+\alpha)Ds\lambda H}{N}} \Big/ i!, \quad \text{and} \tag{3}$$

$$l_p = \begin{cases} \frac{(1+\alpha)D}{6N} + \frac{2N^3}{(1+\alpha)^3 D^3 \lambda^2 H^2} - \frac{4N^5}{3D^5(1+\alpha)^5 \lambda^3 H^3}, & \text{if } D^2(1+\alpha)^2 \lambda H \geqslant 2N^2 \\ \frac{N}{D(1+\alpha)\lambda H}, & \text{otherwise} \end{cases}. \tag{4}$$

The bus fleet size, M, is simply given by

$$M = Q/v_c, \tag{5}$$

where the bus average travel speed v_c is given by

$$\frac{1}{v_c} = \frac{1}{v} + \frac{2\tau\lambda D^2 H(1+\alpha^2)/N}{QH/2N}. \tag{6}$$

Interested readers are referred to Appendix B to see qualitatively how decision variables α, N, and H influence agency cost components Q and M.

2.2.2. User costs

The user costs are associated with (i) the total time for passengers to travel from their origins to their destinations and (ii) other comfort-related factors such as the number of transfers. In the flexible transit system, the total travel time includes the passenger's waiting time at the origin and possibly at the transfer point(s), and the in-vehicle riding time. Because the passengers are picked up from their exact origins and dropped off at their exact destinations, there is no walking time.

Similar to Daganzo (2010), we assume that the passengers prefer fast and efficient travel. They always choose the travel plan with the least number of transfers, and if transfer is necessary, they transfer at the first opportunity. If there are choices with equal user costs (e.g., regarding the initial direction of travel), they break ties arbitrarily. We further assume that headway control strategies are implemented to stabilize bus schedules such that headway variations at the check points are negligible.[3] Under these assumptions, it is shown in Appendix A that the expected waiting time, W, the expected number of transfers, e_T, the expected travel distance, E, the expected in-vehicle travel time, T, for a generic passenger, and the maximum expected vehicle occupancy, O, are given by the following formulas:

$$W = \frac{H}{2}\left(\frac{N-1}{N}\left(1-\frac{\alpha^4}{N}\right) + \frac{(1-\alpha^2)^2}{2} + 1\right), \tag{7}$$

$$e_T = \frac{N-1}{N}\left(1-\frac{\alpha^4}{N}\right) + \frac{(1-\alpha^2)^2}{2}, \tag{8}$$

$$E \approx (4\phi(\alpha) + 5\varphi(\alpha))D + 2D\sum_{i=2}^{\infty}(i-1)[\phi(\alpha)P_c\{i\} + \varphi(\alpha)P_p\{i\}] + \frac{2\lambda HD^2}{N}\left(\frac{D\alpha^2}{3N}\phi(\alpha) + (1+\alpha)l_p\varphi(\alpha)\right), \tag{9}$$

$$T = \frac{E}{v_c} \tag{10}$$

$$O = \frac{\lambda HD^2}{N}\max\left\{\frac{1-\alpha^2}{2\alpha}; \frac{3+2\alpha^2-3\alpha^4}{8\alpha} + \frac{D}{w}\frac{(1-\alpha^2)^2}{32}\right\}, \tag{11}$$

where $\phi(\alpha) = \frac{1}{12}(11\alpha - \alpha^3 - \alpha^5)$ and $\varphi(\alpha) = \frac{1}{18}(2 - 3\alpha + \alpha^3)$.

The maximum expected occupancy is not directly related to user costs. However, it could be useful to the transit agency such that buses with proper capacity (e.g. a van or a mini-bus) can be used.

2.2.3. Design

Appendix B shows that the user costs generally decrease with α, N and increase with H, but the agency costs generally increase with α, N and decrease with H. The optimization problem for the proposed system is to find the best decision

[3] Examples of such bus headway control strategies include adaptive holding (Daganzo, 2009) and adaptive cruise speed control (Daganzo and Pilachowski, 2011). Without headway control, the expected passenger waiting time at the transfer points may be larger than half of the headway due to length time bias (Daganzo, 1997).

variables α (central square size), N (the number of tubes in each direction), and H (bus headway) that balance the trade-offs between the agency and user costs.

We convert the agency costs into travel time equivalents. Suppose $\$_Q$ is the agency operation cost per vehicle-distance, $\$_M$ is the agency cost per vehicle hour, and μ is the average monetary value of one passenger-hour, 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 travel time equivalent per passenger (Daganzo, 2010). We let δ measure transfer discomfort, and then $\frac{\delta}{v_w}$ converts the expected transfer number e_T into passenger riding time. The optimization problem becomes the following:

$$\begin{aligned} &\text{Min} \quad z = \pi_Q Q + \pi_M M + W + T + \frac{\delta}{v_w} e_T, \\ &\text{s.t.}, \alpha \in \left[\frac{1}{N}, 1\right], \qquad H \geqslant 0, \quad N \in \{1, 2, \ldots \lfloor D/s \rfloor\}. \end{aligned} \tag{12}$$

Note that we enforce $\alpha \geqslant \frac{1}{N}$ and $N \leqslant \lfloor D/s \rfloor$ (where $\lfloor \cdot \rfloor$ is the floor operation) because the central square should contain at least one transfer point, and each tube should contain at least one local street in the longitudinal direction.

To find the optimal value of the objective function, z, we may allow N to take a continuous value and apply numerical nonlinear optimization method (such as the steepest decent method). Numerical approximation is used to estimate the gradient of the objective function.

3. Other transportation system structures

In this subsection, we compare the performance of other service providing systems that are comparable to the proposed flexible-route transit system.

3.1. Fixed route transit network

In the fixed-route transit system, the total system costs include those related to infrastructure investment, total vehicle distance, bus fleet size, passenger walking time to and from the bus stations, passenger waiting time, and transfer discomfort. To be consistent with our flexible system, and to be conservative (i.e., favoring the fixed-route system), we disregard any possible infrastructure investment in the fixed-route system. The optimization objective can be expressed in similar notations (while replacing each bus tube by a fixed bus route) as follows (Daganzo, 2010):

$$\text{Min } z = \pi_Q Q + \pi_M M + A + W + T + \frac{\delta}{v_w} e_T, \tag{13}$$

where $Q = \frac{2D^2}{wH}(3\alpha - \alpha^2)$, $M = \frac{Q}{v_c}$, $W = H\left(\frac{2+\alpha^3}{3\alpha} + \frac{(1-\alpha^2)^2}{4}\right)$, $E = \frac{D}{12}(12 - 7\alpha + 5\alpha^3 - 3\alpha^5 + \alpha^7)$, $T = \frac{E}{v_c} \approx \frac{D}{12}\left(\frac{1}{v} + \frac{\tau_1}{w}\right)(12 - 7\alpha + 5\alpha^3 - 3\alpha^5 + \alpha^7)$, and the expected walking time is $A = \frac{w}{v_w}$. Note that in the fixed-route system, $\frac{D}{N}$ represents the spacing between adjacent bus routes at the boundary of the region.

3.2. Taxi

For comparison, we consider an idealized situation where "chartered" vehicles (e.g., taxi) deliver each passenger directly from its origin to its destination, and we ignore passenger waiting time at the origin. Hence, the user costs only include the travel time from origin to destination. In the $D \times D$ square area, the expected travel distance for each passenger is obviously $E = 2D/3$, and hence the expected travel time per trip is:

$$T = \frac{E}{v} = \frac{2D}{3v}. \tag{14}$$

Here we have assumed that travel speed v equals the vehicle cruising speed (i.e., ignoring the stopping time to pick up the only passenger). The agency costs depend on the total vehicle-distance per hour of operation, Q, and the fleet size, M. We further ignore any vehicle distance traveled between delivery trips (e.g., the distance traveled to pick up the next passenger),[4] and hence

$$Q = \lambda D^2 E = \frac{2}{3}\lambda D^3, \quad \text{and} \quad M = \frac{Q}{v} = \frac{2\lambda D^3}{3v}. \tag{15}$$

Then the objective function for the taxi system is:

$$\text{Min } z = \pi_Q Q + \pi_M M + T. \tag{16}$$

[4] This assumption and the other simplified assumptions are in favor of the taxi service and hence conservative for our comparison.

Table 1
Parameters setting for the numerical examples.

Parameter	s (km)	μ ($/h)	τ (s)	τ_1 (s)	τ_2 (s)	v (km/h)	v_w[a] (km/h)	δ	$\$_Q$ ($/veh-km)	$\$_M$ ($/veh-h)
Fixed transit	0.15	20	–	12	1	25	2	0.03	2	40
Flexible transit	0.15	20	13	–	–	25	2	0.03	2	40
Taxi	–	20	–	–	–	25	–	–	2	40

[a] The average walking distance is normally 3–5 km/h but we consider a lower speed to address the discomfort and delay associated with walking.

It shall be noted that when $\pi_Q = \frac{\$_Q}{\lambda D^2 \mu}$ and $\pi_M = \frac{\$_M}{\lambda D^2 \mu}$, the objective function (16) becomes $\frac{2D}{3v\mu}(\$_Q v + \$_M + \mu)$, which is independent of the decision variables and demand density. Also note that the minimum user cost, T, by itself provides an absolute lower bound of the system cost for any transportation mode with speed v.

3.3. Discussion

Qualitatively, the proposed flexible transit system could be beneficial because it eliminates the need for passengers to walk to the bus stations. This is desirable because it implies a higher level of service to the passengers. In some circumstances (e.g., transit service during night), the benefit also comes from enhanced passenger safety. In addition, the flexible transit system reduces possible need for infrastructure investment, because there are no specific bus lines or stops. However, these advantages may come at a cost: the total vehicle travel distance (and consequently the passenger riding time) could be higher than that in a fixed-route system, because buses also move laterally in their tubes.

As the lateral distance is highly dependent of the number of passengers in the tube, we would expect the flexible-route transit system to be relatively desirable under low demand, while the traditional fixed-route transit system to be desirable under high demand. This intuition is quantitatively verified with the numerical examples in Section 4.

4. Numerical analysis

4.1. Cost comparison

We compute the optimal system costs of three transit service systems: fixed-route transit, flexible-route transit, and taxi systems, to serve a square area with D = 10 km and λ varying from 1 to 500 passengers per hour per km^2. For comparison, we use parameter values that are either realistically assumed or consistent with those in the literature,[5] as shown in Table 1.

Fig. 3 plots the optimal cost (per passenger) versus demand density λ for the three systems. The cost of the taxi system is independent of λ. The proposed flexible-route transit system has a lower cost than the fixed-route system when demand is low, while the cost of the fixed-route transit system decreases faster as demand increases. Among the three systems, taxi is desirable under extremely low demand (e.g., $\lambda < 4$ passengers per hour per km^2), while the fixed-route transit system is favorable under high demand. In the middle range when demand is low-to-moderate (e.g., λ = 4–40 passengers per hour per km^2), the flexible-route transit system is advantageous. If we realize that the cost functions for the taxi system and for the fixed-route transit system are derived under relatively favorable conditions (e.g., ignoring taxi trips between deliveries and infrastructure investment), then the flexible transit system is probably the best choice for a larger range of demand densities.

Table 2 shows the detail of the optimal design for the proposed flexible-route transit system. When travel demand increases, the optimal size of the central square slightly increase, but the optimal number of tubes increases and the optimal headway decreases considerably as expected. The maximum expected bus occupancy increases at a relatively slow rate, and the average bus velocity decreases only moderately. Almost all agency and user cost components (per passenger) decrease monotonically, except that the number of transfers (and hence the associated cost) remains almost the same.

For comparison, Table 3 shows the optimal design for the fixed-route system. When travel demand increases, the optimal size of the central square and the optimal number of bus tubes both increase. From Tables 2 and 3 we see that the flexible transit system seems to require relatively fewer tubes under lower demand and relatively more tubes under high demand. The maximum expected bus occupancy and average bus speed follow similar trends as those of the flexible transit system. All the agency and user cost components (per passenger) decrease with demand except that the in-vehicle travel time increases slightly (due to a lower bus speed v_c and a higher number of stops N).

As discussed in Section 3.3, for all λ, the agency costs in the flexible transit system (Q and M) are larger than their counterparts in the fixed transit system, and each passenger spends more time (T) in the vehicle, because the buses in the flexible system travel extra distances to pick up and drop off passengers. The advantage of the flexible system is that the passenger walking time (A) is completely eliminated, while the waiting time (W) is also reduced.

[5] Most parameters in Table 1 are taken from Daganzo (2010), except that we use a smaller value for bus deceleration and acceleration time per stop, τ_1 = 12 s––per Levinson (1983), τ_1 = 11.13 s on average. We also assume that the spacing between adjacent local streets is s = 0.15 km (http://www.Wikipedia.com).

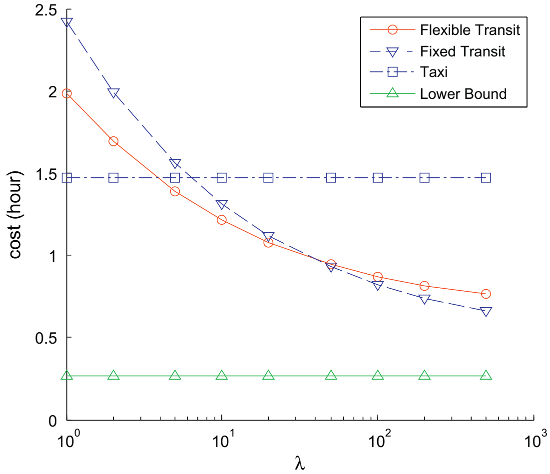

Fig. 3. System cost versus demand density (log-scale).

Table 2
Network design and cost components for the flexible-route transit system.

λ	α	N	H (h)	O	v_c	$\$_Q Q$	$\$_M M$	W	T	$\delta \frac{e_T}{v_w}$	z
1	0.27	3	0.45	25.75	23.45	0.59	0.50	0.47	0.41	0.02	1.99
2	0.31	4	0.35	25.51	23.09	0.48	0.41	0.38	0.41	0.02	1.69
5	0.35	6	0.26	27.89	22.49	0.36	0.32	0.29	0.40	0.02	1.39
10	0.36	8	0.21	34.22	21.95	0.29	0.27	0.24	0.40	0.02	1.22
20	0.35	10	0.15	41.50	21.54	0.25	0.23	0.17	0.40	0.02	1.08
50	0.34	14	0.11	59.53	20.68	0.19	0.19	0.13	0.42	0.02	0.94
100	0.37	19	0.08	64.39	20.18	0.17	0.17	0.09	0.41	0.02	0.87
200	0.36	24	0.06	84.12	19.54	0.15	0.15	0.07	0.43	0.02	0.81
500	0.38	35	0.04	106.40	18.81	0.13	0.13	0.05	0.44	0.02	0.77

Table 3
Network design and cost components for fixed-route transit.

λ	α	N	H (h)	O	v_c	$\$_Q Q$	$\$_M M$	A	W	T	$\delta \frac{e_T}{v_w}$	z
1	0.82	7	0.65	4.54	23.57	0.38	0.32	0.37	0.69	0.29	0.02	2.42
2	0.83	9	0.52	5.73	23.23	0.30	0.25	0.29	0.55	0.29	0.02	1.99
5	0.85	11	0.38	7.62	22.72	0.22	0.19	0.22	0.40	0.30	0.02	1.56
10	0.86	14	0.30	9.67	22.27	0.17	0.15	0.18	0.31	0.30	0.02	1.31
20	0.88	17	0.24	12.01	21.74	0.13	0.12	0.15	0.25	0.31	0.02	1.12
50	0.90	21	0.17	15.99	21.03	0.09	0.09	0.12	0.17	0.32	0.02	0.93
100	0.92	24	0.13	19.38	20.45	0.07	0.07	0.10	0.13	0.33	0.02	0.82
200	0.93	29	0.10	24.53	19.79	0.06	0.06	0.09	0.10	0.34	0.02	0.74
500	0.96	33	0.07	32.05	19.01	0.04	0.04	0.08	0.07	0.35	0.02	0.66

For the taxi system, we can easily compute from (16)-(18) that the optimal system cost components are $Q = 0.67$, $M = 0.53$, $T = 0.27$, and hence the system cost per passenger is $z = 1.47$. The agency costs in the flexible transit system (Q and M) are much smaller than that of the taxi system (even under our favorable assumptions), while the taxi system incurs the smallest in-vehicle travel time.

4.2. Sensitivity analysis

In Section 4.1, we assumed that $\tau_1 = 12$ s (Levinson, 1983) for both fixed and flexible transit systems. In Daganzo (2010) a more conservative value of $\tau_1 = 30$ s is assumed. Intuitively, longer bus delay per stop will have a direct impact on the total number of buses in operation and the expected travel time per passenger. Such impact will be especially significant for the flexible transit system, because it makes a stop for each passenger visit. To this end, we conduct a sensitivity analysis to see how the three systems perform under $\tau_1 = 30$ s, and the results are shown in Fig. 4a. We can see

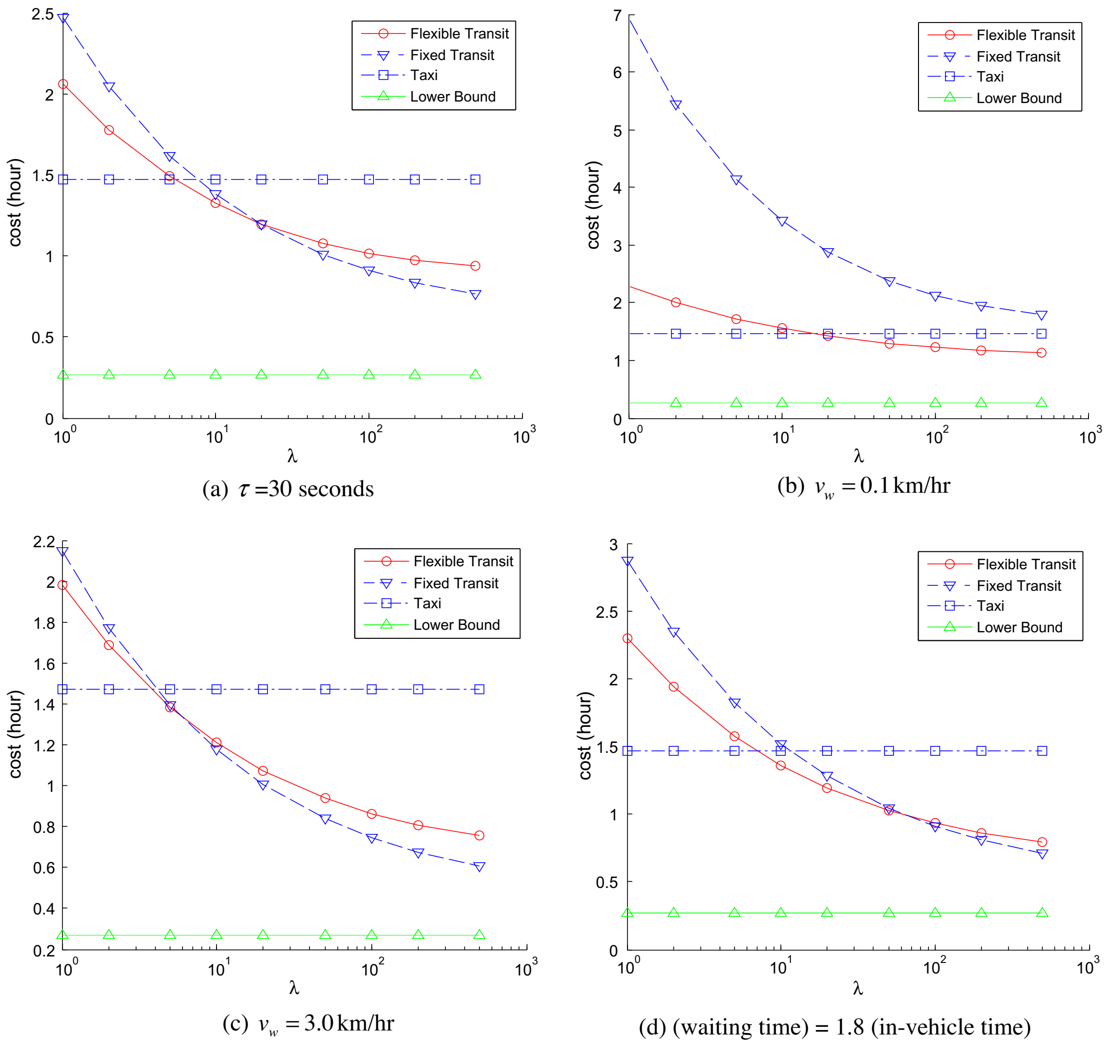

Fig. 4. System cost versus demand density: (a) long bus dwell time; (b) inconvenient walking; (c) higher walking speed; (d) higher weight of waiting time.

that all curves follow the same trend as those in Fig. 3, but the optimal costs for the flexible- and the fixed-route transit systems have increased (more so for the flexible system). Although the range of suitable demand for the flexible system has shrunk, we nevertheless notice that the flexible transit system remains the best choice when $\lambda = 5–20$ passengers per hour per km^2.

Under certain circumstances (e.g., adverse weather, night time), passenger walking is associated with significant discomfort or passenger safety becomes a major concern. This situation can be addressed by a dramatic reduction in walking speed v_w. We consider the extreme case when v_w reduces to 0.1 km/h, and the cost curves are shown in Fig. 4b. The total cost for the fixed transit system increases dramatically, mainly due to the increase in walking time (A). The flexible transit system, in contrast, only bears a slight increase in the total cost. The cost for the taxi system remains unchanged. In this case, the flexible transit system becomes favorable for a much larger range of customer demand densities.

Fig. 4c shows the system cost curves under a higher walking speed, 3.0 km/h. Even with this walking speed the proposed system shows a lower cost for some demand densities. We would like to highlight that the flexible transit system is particularly targeting situations in which passenger walking is inconvenient (e.g. severe weather condition, night time, or unsafe area). Also note that the actual taxi cost probably much higher than the conservative one shown in this figure; the actual demand range for the flexible system to be optimal is probably much larger.

In the base case the expected waiting time and in-vehicle travel time are assumed to be the same to the passengers. In reality, waiting time is sometimes considered to be less convenient. To study this effect, we consider a case where one unit of waiting time is 1.8 times as costly as in-vehicle travel time. The resulting cost curves are plotted in Fig. 4d. Since the relative significance of waiting time in the fixed and flexible route systems are approximately the same, the relative comparison across different systems is similar to that of the base case.

5. Conclusion

This paper proposed a structured flexible-route transit system in which buses pick up or drop off passengers within their predetermined service areas (i.e., bus tubes) with no specific routes. Collectively, the bus tubes form a ‘‘grand’’ structure, which includes a grid tube network that provides double coverage to passengers in the central part of the city, and a hub-and-spoke tube network that provides a single coverage in the peripheral part. We analyzed all relevant agency and user costs associated with the system design, and optimized the system by determining the layout of the grand network, the bus tube size, and the bus headway.

The proposed transit system is shown to have the potential to provide a higher level of service to the passengers, mainly by eliminating walking time to and from bus stations. This is particularly desirable under situations where concerns over pedestrian safety are high (e.g., at night or under adverse weather). Quantitative analysis of the proposed system shows that under low-to-moderate passenger demand the system incurs lower cost than other conventional counterparts such as the fixed-route transit system and the chartered taxi system. In practice, the proposed flexible transit system is also easy to implement. The grand structure can be used to guide the tube layout while details of the local streets and neighborhoods are considered. Existing spatial partitioning models (e.g. the disk models in Ouyang and Daganzo (2006) and Ouyang (2007)) can be potentially adapted to help align bus tubes and adjust the size.

Future research could be conducted to explore other network structures (e.g., a ring-radial network) or any hybrid combination of them (e.g. multi-hub-and-spoke). Other shapes of the service region (e.g., rectangle, circle) or heterogeneous demand distributions (e.g., monocentric or multicentric) could be interesting to analyze, although these considerations may add extra complexity to the model.

Acknowledgments

This research was supported in part by the US National Science Foundation through Grant CMMI #0748067. We thank Professor Carlos Daganzo (UC Berkeley) and the two anonymous reviewers for their helpful comments.

Appendix A

Result 0. We list the following useful geometry quantities without proof.

Area of each tube that belongs to the inner square	$\frac{D^2\alpha^2}{N}$
Area of each tube that belongs to the outer peripheral part	$\frac{D^2(1-\alpha^2)}{2N}$
Expected lateral distance per passenger in the center	$\frac{D\alpha}{3N}$
Longitudinal distance per bus in the central square during one round trip	$2\alpha D$
Longitudinal distance per bus in the peripheral part during one round trip	$2(1-\alpha)D$
Expected number of passengers to pick up and drop off in the central square during a round trip	$\frac{2\lambda HD^2\alpha^2}{N}$
Expected number of passengers to pick up and drop off in the peripheral part during a round trip	$\frac{2\lambda HD^2(1-\alpha^2)}{N}$

Result 1. The expected number of transfers is given by this formula: $e_T = \frac{N-1}{N}\left(1-\frac{\alpha^4}{N}\right)+\frac{(1-\alpha^2)^2}{2}$.

Proof. One trip may include 0, 1 or 2 transfers, depending on the relative locations of the origin and destination. The conditional probability of having zero transfer is $\frac{2}{N}-\frac{1}{N^2}$ if both the origin and destination locations are in the central square (case 1), $\frac{1}{N}$ if one is in the peripheral quadrants and the other is in the central square (case 2), or $\frac{1}{2N}$ if both are in the peripheral quadrants (case 3). Therefore the unconditional probability for a trip to have zero transfer is $\Pr[e_T=0]=\alpha^4\left(\frac{2}{N}-\frac{1}{N^2}\right)+2\alpha^2(1-\alpha^2)\frac{1}{N}+(1-\alpha^2)^2\frac{1}{2N}$. $\square$

Similarly, it can be shown that the conditional probability of one transfer is $\frac{N-2}{N}+\frac{1}{N^2}$ in case 1, $\frac{N-1}{N}$ in case 2, or $\frac{1}{2}$ in case 3. Therefore,

$$\Pr[e_T=1]=\alpha^4\left(\frac{N-2}{N}+\frac{1}{N^2}\right)+2\alpha^2(1-\alpha^2)\frac{N-1}{N}+(1-\alpha^2)^2\frac{1}{2}.$$

Two transfers may occur with conditional probability $\frac{N-1}{2N}$ only in case 3; i.e.,

$$\Pr[e_T=2]=\alpha^4\times 0+2\alpha^2(1-\alpha^2)\times 0+(1-\alpha^2)^2\frac{N-1}{2N}.$$

Therefore, the expected number of transfers is $e_T=\Pr[e_T=1]+2\Pr[e_T=2]=\frac{N-1}{N}\left(1-\frac{\alpha^4}{N}\right)+\frac{(1-\alpha^2)^2}{2}$.

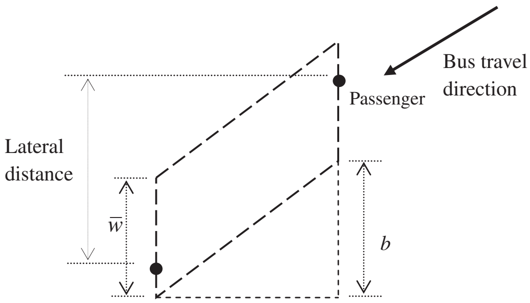

Fig. A1. The expected lateral distance to visit one passenger in a slanted tube.

Result 2. The hourly total vehicle distance is given by $Q = Q_c + Q_p$, where

$$Q_c = \frac{2N}{H}\left[\sum_{i=2}^{\infty}\alpha(i-1)DP_c\{i\} + 2\alpha D + \frac{2\lambda HD^3\alpha^3}{3N^2}\right] \quad \text{and}$$
$$Q_p = \frac{2N}{H}\left[\sum_{i=2}^{\infty}(1-\alpha)(i-1)DP_p\{i\} + 2(1-\alpha)D + \frac{2\lambda HD^2(1-\alpha^2)l_p}{N}\right]$$

are the expected distances in the central square and the peripheral part, respectively,

$$P_c\{i\} = \left(\frac{\alpha Ds\lambda H}{N}\right)^i e^{-\frac{\alpha Ds\lambda H}{N}} \Big/ i!,$$
$$P_p\{i\} = \left(\frac{(1+\alpha)Ds\lambda H}{N}\right)^i e^{-\frac{(1+\alpha)Ds\lambda H}{N}} \Big/ i!, \quad \text{and}$$

$$l_p = \begin{cases} \frac{(1+\alpha)D}{6N} + \frac{2N^3}{(1+\alpha)^3D^3\lambda^2H^2} - \frac{4N^5}{3D^5(1+\alpha)^5\lambda^3H^3}, & D^2(1+\alpha)^2\lambda H \geqslant 2N^2 \\ \frac{N}{D(1+\alpha)\lambda H}, & \text{otherwise} \end{cases}.$$

Proof. The total vehicle distance per hour is given by the product of the number of bus tubes, the expected travel distance of one bus round trip, and the inverse of the bus headway. The expected travel distance per bus round trip includes three parts: longitudinal distance, lateral distance, and extra detour distance. □

In the inner square, the round-trip longitudinal distance is $2\alpha D$ and the lateral distance is $\frac{2\lambda HD^3\alpha^3}{3N^2}$; in the peripheral quadrants, the round-trip longitudinal distance is $2(1-\alpha)D$ and the lateral distance is $\frac{2\lambda HD^2(1-\alpha^2)l_p}{N}$, where the expected lateral distance per passenger in the peripheral quadrants, l_p, is computed as follows.

Note that the tube is slanted in the peripheral quadrants, as illustrated in Fig. A1. The expected lateral distance, l_p, between the two dots (i.e., consecutive passengers) can be shown to be as follows:

$$l_p = \begin{cases} \frac{\bar{w}}{3} + \frac{b^2}{\bar{w}} - \frac{b^3}{3\bar{w}^2} & \text{if } b < \bar{w} \\ b & \text{if } b \geqslant \bar{w} \end{cases}.$$

where $\bar{w}$ is the local tube width and b is the lateral off-set of the tube between the two consecutive passengers. Considering Poisson process for passenger locations and 0–45° slant angles of the tubes, the expected value of b across all local areas of the quadrants is approximately $\frac{1}{w(1+\alpha)\lambda H} = \frac{N}{D(1+\alpha)\lambda H}$, and hence the approximate[6] formula for l_p follows.

Since the street spacing $s > 0$, the bus must travel extra distance if it needs to visit more than one passenger before advancing through one street block in the longitudinal direction; see Fig. A2.[7] The expected extra longitudinal distance to visit each additional passenger is $\frac{s}{2}$. The number of passengers in the "tube block" area (i.e., the shaded area in Fig. A2) follows a Poisson distribution with mean $\bar{w}s2\lambda H$, where $\bar{w} = \alpha D/N$ in the central square (considering double coverage in the center) and $(1+\alpha)D/2N$ in the peripheral quadrants. Therefore the probabilities of having i passengers in the center and peripheral parts are $P_c\{i\} = \left(\frac{\alpha Ds\lambda H}{N}\right)^i e^{-\frac{\alpha Ds\lambda H}{N}} \Big/ i!$ and $P_p\{i\} = \left(\frac{(1+\alpha)Ds\lambda H}{N}\right)^i e^{-\frac{(1+\alpha)Ds\lambda H}{N}} \Big/ i!$ respectively.

[6] The formula for l_p is approximate due to Jensen's Inequality.
[7] The above sweeping strategy is used in the cost derivation because of (i) simplicity for formulation (in closed-form) and implementation in the real world and (ii) conservativeness for comparison with other systems. In practice, other heuristic routing strategies (e.g., insertion heuristics) can be used for pick up and drop off passengers between two consecutive check points so as to further improve the system performance.

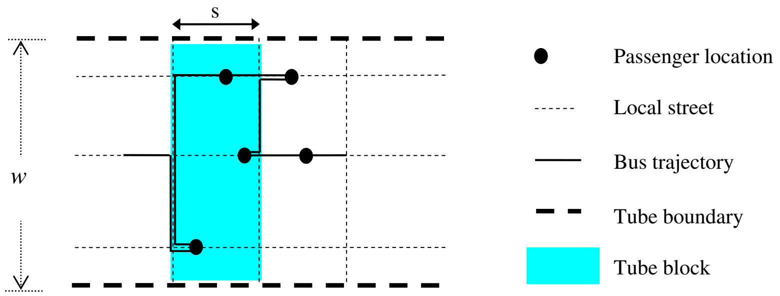

Fig. A2. The extra longitudinal distance for additional passengers.

If there are $i \geqslant 2$ passengers in one "tube block", the bus incurs an extra distance of $(i-1)\frac{s}{2}$ per passenger, and the expected extra distance per round trip is therefore $\alpha D(i-1)P_c\{i\}$ in the center part and $(1-\alpha)D(i-1)P_p\{i\}$ in the peripheral part. The expected extra distance shall be the summation of these two terms across all $i \geqslant 2$.

These longitudinal, lateral and extra distances per bus round trip add up to the formulas for Q_c and Q_p, given that there are $2N/H$ buses in operation per hour.

Result 3. The expected waiting time per passenger is given by $W = \frac{H}{2}\left[\frac{N-1}{N}\left(1-\frac{\alpha^4}{N}\right)+\frac{(1-\alpha^2)^2}{2}+1\right]$.

Proof. Note that we have assumed that (i) the passengers randomly choose the transfer location in case they have more than one option and (ii) headway control strategies are implemented to eliminate irregularities in bus arrival headways at the check points. The waiting time is obtained by multiplying the waiting time per transfer by the number of transfers, and adding it to the waiting time at the origin. Each waiting time (at the transfer or the origin) is approximately $\frac{H}{2}$, and the number of transfers is given by Result 1. □

Result 4. The average bus speed satisfies $\frac{1}{v_c} = \frac{1}{v} + \frac{2\tau\lambda D^2 H/N}{QH/2N}$.

Proof. Without losing generality, we assume that during each stop we pick up or drop off exactly one passenger. From Result 0, during each trip, the number of passengers in the central square is $\frac{\lambda}{2} a_{inner} H = \frac{\lambda}{2}\frac{D^2\alpha^2}{N}H$ and the number of passengers in the peripheral quadrants is $\lambda a_{outer} H = \lambda\frac{D^2(1-\alpha^2)}{2N}H$. Therefore the total number of stops is $2\left(\lambda\frac{D^2\alpha^2}{2N}H + \lambda\frac{D^2(1-\alpha^2)}{2N}H\right) = \frac{\lambda D^2 H}{N}$. The bus needs to overcome time over distance $\frac{1}{v}$, and stop for τ time per passenger during a round trip with distance $QH/2N$. Hence, $\frac{(QH/2N)}{v_c} = \frac{(QH/2N)}{v} + 2(\tau)\lambda D^2 H/N$, which leads to the result. □

Result 5. The expected in-vehicle travel distance and travel time per passenger are $E = \rho_p E(R_p) + \rho_c E(R_c)$ and $T = \frac{E}{v_c}$ respectively, where $E(R_p) \approx \frac{2-3\alpha+\alpha^3}{3}D$ and $E(R_c) = \left(\frac{11}{12}\alpha - \frac{1}{12}\alpha^3 - \frac{1}{12}\alpha^5\right)D$ are the expected in-vehicle longitudinal distances per passenger trip in the peripheral part and the central square, respectively, and ρ_c and ρ_p are the ratios of the expected total distance over the expected longitudinal distance for the central and peripheral parts, respectively.

Proof. Fig. A3 shows a quarter of the entire square region. The longitudinal distance between a random location (x, y) in the northern quadrant and the northern boundary of the central square is $(y - \frac{d}{2})$. Based on symmetry we conduct integral on $\frac{1}{8}$ of the peripheral region to calculate the expected longitudinal distance per passenger in the peripheral quadrants. While doing so we consider various combinations of trip origin/destination locations (i.e., in the central square or peripheral quadrants).[8] It can be verified that

$$E(R_p) \approx 2\frac{8}{D^2}\int_{\frac{d}{2}}^{\frac{D}{2}}\int_0^y\left(y-\frac{d}{2}\right)dx\,dy = 2\frac{2D^3-3D^2d+d^3}{6D^2} = \frac{2-3\alpha+\alpha^3}{3}D. \quad \square$$

[8] While calculating the expected longitudinal distance in the peripheral area, we have assumed that all trips, if only starting and ending both in the peripheral part, shall travel through the central square. In fact, there is a very small probability, $\frac{1}{4N}(1-\alpha^2)^2$, that the origin and destination are both in the same peripheral part of the same tube, such that the trip will not go through the central square. The exact formula for $E(R_p)$, when considering this possibility, would have been $\left[1-\frac{1-\alpha^2}{4N}\right]\frac{2-3\alpha+\alpha^3}{3}D + \left[\frac{(1-\alpha^2)^2}{4N}\right]\frac{2(1+2\alpha-2\alpha^2-\alpha^3)}{15(1+\alpha)^2}D$. It can be shown, however, that for most realistic examples, ignoring such a possibility would make no practical difference to the estimated value of E. In addition, it is conservative for us to ignore such a possibility, because in doing so we are (slightly) overestimating E, and hence in favor of other transit systems.

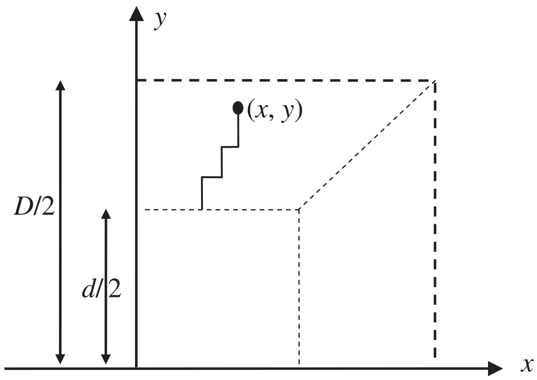

Fig. A3. The lateral distance for pick-up or drop-off passengers.

To compute $E(R_c)$, we again consider cases 1–3 regarding relative locations of the trip origin and destination (see the proof for Result 1). The expected longitudinal distances in these cases are $\frac{2d}{3}$, $\frac{5d}{6}$, and $\frac{11d}{12}$, respectively, and the corresponding probabilities are α^4, $2\alpha^2(1-\alpha^2)$, and $(1-\alpha^2)2$. Then,

$$E(R_c) = \alpha^4 \frac{2}{3} d + 2\alpha^2(1-\alpha^2)\frac{5}{6} d + (1-\alpha^2)^2 \frac{11}{12} d = \left(\frac{11}{12}\alpha - \frac{1}{12}\alpha^3 - \frac{1}{12}\alpha^5\right) D.$$

To compute the lateral movement of each bus, we define ρ_c and ρ_p as the ratio of the total distance over the longitudinal distance. In the central square the longitudinal distance during one bus round trip is $2\alpha D$ and the total distance is $Q_cH/2N$ (as shown in Result 2), and then

$$\rho_c = \frac{Q_c H/2N}{2\alpha D}.$$

In the periphery the longitudinal distance during one bus round trip is $2D(1 - \alpha)$ and the total distance is $Q_pH/2N$, then

$$\rho_p = \frac{Q_p H/2N}{3D(1-\alpha)}.$$

The analysis above yields the formula for E after simple algebraic manipulations. Then, the total expected travel time for each passenger trip is obviously

$$T = \frac{E}{v_c} = \frac{\rho_p E(R_p) + \rho_c E(R_c)}{v_c}.$$

Result 6. The vehicle occupancy is given by $O = \frac{\lambda H D^2}{N} \max\left\{\frac{1-\alpha^2}{2\alpha}; \frac{3+2\alpha^2-3\alpha^4}{8\alpha} + \frac{D}{w}\frac{(1-\alpha^2)^2}{32}\right\}$.

Proof. This formula is essentially the same as Result 8 in Daganzo (2010). □

The above analytical results for all different cost components have been validated by discrete-event simulations under the same assumptions.

Appendix B

In this appendix we illustrate how decision variables α, N, and H influence the cost components Q, M, W, and T according to (1)–(12). We use D = 10 km^2 and λ = 10 passengers per km^2 per hour.

As shown in Fig. B1a, the agency costs (Q and M) and the passenger waiting time W all decrease only slightly with α; however the in-vehicle travel time T is high for both small α (i.e., a grid network) and large α (i.e., a hub-and-spoke network). As such, the optimal value of α is likely to be somewhere in the middle between 0 and 1, implying that a hybrid combination of grid and hub-and-spoke network will be desirable.

Fig. B1b shows that the agency costs are high for either small or large N. When N is small, the tubes are wide, and the high agency cost is probably due to a large amount of lateral movements per passenger. On the other hand, when N is large, the tubes are narrow, and the agency cost could still be high due to the large number of buses needed in the system. The passengers' waiting time almost remains constant across all N, but the in-vehicle travel time decreases dramatically until

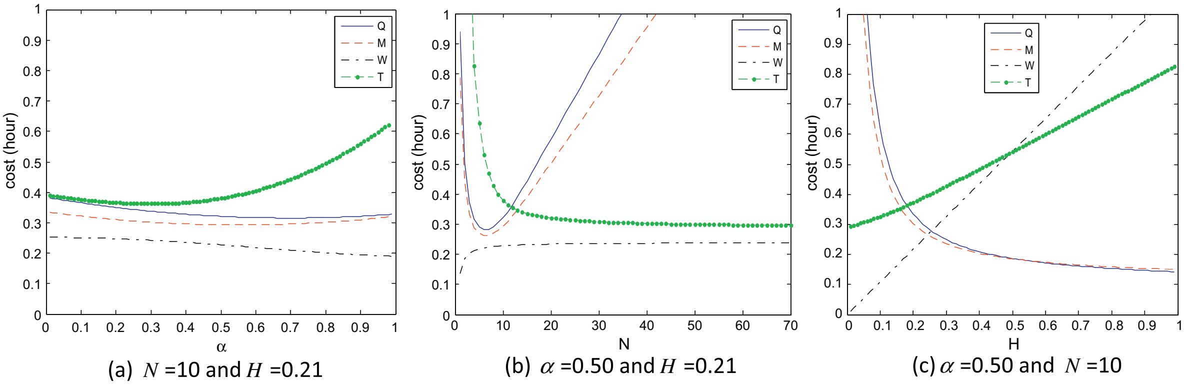

Fig. B1. Sensitivity of Q, M,W and T to α, N, and H.

N increases to a moderate number. This sharp decrease is probably because of the significant reduction in cumulative lateral travel distances that a passenger experiences. This curve flattens out when N continues to increase, implying that as long as the tube width is not too wide, the lateral travel would not significantly increase the system cost.

When headway H increases, the agency costs Q and M both decrease to a nonzero value; however the waiting time and the in-vehicle travel time increase almost linearly. These trends are shown in Fig. B1c.

References

Balke, K.N., Urbanik, C.L., Urbanik, T., 2000. Development and evaluation of intelligent bus priority concept. Transportation Research Record 1727, 12–19.
Barnett, J.F., 1970. Express bus mass transit. Journal of the Transportation Engineering Division 96 (2), 133–148.
Byrne, B.F., 1975. Public transportation line positions and headways for minimum user and system cost in a radial case. Transportation Research 9 (2), 97–102.
Campbell, J.F., 1990. Freight consolidation and routing with transportation economies of scale. Transportation Research Part B 24 (5), 345–361.
Ceder, A., Wilson, N.H.M., 1986. Bus network design. Transportation Research Part B 20 (4), 331–344.
Daganzo, C.F., 1984. Checkpoint dial-a-ride systems. Transportation Research Part B 18 (4-5), 315–327.
Daganzo, C.F., 1997. Fundamentals of Transportation and Traffic Operations. Pergamon-Elsevier, New York, NY.
Daganzo, C.F., 2009. A headway-based approach to eliminate bus bunching: systematic analysis and comparisons. Transportation Research Part B 43 (10), 913–921.
Daganzo, C.F., 2010. Structure of competitive transit networks. Transportation Research Part B 44 (4), 434–446.
Daganzo, C.F., Pilachowski, J., 2011. Reducing bunching with bus-to-bus cooperation. Transportation Research Part B 45 (1), 267–277.
Desaulniers, G., Hickman, M.D., 2007. Public transit. In: Laporte, G., Barnhart, C. (Eds.), Handbooks in Operations Research and Management Science: Transportation Systems. North-Holland.
Eichler, M., Daganzo, C.F., 2006. Bus lanes with intermittent priority: strategy formulae and an evaluation. Transportation Research Part B 40 (9), 731–744.
Holroyd, E.M., 1965. The optimum bus service: a theoretical model for a large uniform urban area. In: Edie, L.C., Herman, R., Rothery, R. (Eds.), Vehicular Traffic Science, Proceedings of the 3rd International Symposium on the Theory of Traffic Flow. Elsevier, New York, NY.
Levinson, H.S., 1983. Analyzing transit travel time performance. Transportation Research Record 915, 1–6.
Newell, G.F., 1979. Some issues relating to the optimal design of bus routes. Transportation Science 13 (1), 20–35.
Ouyang, Y., 2007. Design of vehicle routing zones for large-scale distribution systems. Transportation Research Part B 41 (10), 1079–1093.
Ouyang, Y., Daganzo, C.F., 2006. Discretization and validation of the continuum approximation scheme for terminal system design. Transportation Science 40 (1), 89–98.
Quadrifoglio, L., Li, X., 2009. A methodology to derive the critical demand density for designing and operating feeder transit services. Transportation Research Part B 43 (10), 922–935.
Quadrifoglio, L., Hall, R., Dessouky, M., 2006. Performance and design of mobility allowance shuttle transit (MAST) services: bounds on the maximum longitudinal velocity. Transportation Science 40 (3), 351–363.
Quadrifoglio, L., Dessouky, M., Ordóñez, F., 2008. Mobility allowance shuttle transit (MAST) services: MIP formulation and strengthening with logic constraints. European Journal of Operational Research 185 (2), 481–494.
Wirasinghe, S.C., Hurdle, V.F., Newell, G.F., 1977. Optimal parameters for a coordinated rail and bus transit system. Transportation Science 11 (4), 359–374.