(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Transportation Research Part B 51 (2013) 1–16

Contents lists available at SciVerse ScienceDirect

Transportation Research Part B

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

A model for estimating the optimal cycle length of demand responsive feeder transit services

Shailesh Chandra, Luca Quadrifoglio *

Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136, United States

ARTICLE INFO

Article history:
Received 7 June 2012
Received in revised form 22 January 2013
Accepted 23 January 2013

Keywords:
Optimal cycle
Demand responsive
Feeder transit
Transit performance
Continuous approximations
First/last mile

ABSTRACT

The general lack of first/last mile connectivity is one of the main challenges faced by today's public transit. One of the possible actions towards a solution to this problem is the planning, design and implementation of efficient feeder transit services. This paper develops an analytical model which allows for an easy computation of near optimal terminal-to-terminal cycle length of a demand responsive feeder service to maximize service quality provided to customers, defined as the inverse of a weighted sum of waiting and riding times. The model estimates the recommended cycle length by only plugging in geometrical parameters and demand data, without relying on extensive simulation analyses or rule of thumbs. Simulation experiments and comparisons with real services validate our model, which would allow planners, decision makers and practitioners to quickly identify the best feeder transit operating design of any given residential area.

1. Introduction and background

The US Department of Transportation recently identified the general lack of connectivity as one of the main challenges faced by public transit. Policies which encourage the desired reduction of Vehicle Miles Traveled (VMT), reduction of greenhouse gases and even an increase of "livability" depend on solutions to the issue of first/last mile access to transit and multimodal connectivity. One of the possible actions towards providing a solution to this problem is the planning, design and implementation of efficient Demand Responsive Feeder Transit services, connecting residential areas to major fixed-route transit networks, also known as Demand Responsive Connectors (DRCs).

Demand Responsive Transit (DRT) systems, also known as dial-a-ride transit (DART) or call-n-ride (CnR) systems, have been proven effective in responding to the need of low demand density areas and are welcomed by passengers as they provide an increased flexibility and higher service level compared to 'regular' fixed-route services. These services are, however, much more costly to deploy for the operators. Typical examples, other than feeder services, include ADA Paratransit services, rural services or other flexible hybrid systems, like "route deviation" services.

Feeders are one of the most often used types of flexible transit service, especially within low density residential areas. They have been operating in quite a few cities (Koffman, 2004); examples can be found in Denver (CO), Raleigh (NC), Akron (OH), Tacoma (WA), Sarasota (FL), Portland (OR) and Winnipeg (Canada). They are characterized by having a flexible routing and scheduling strategy for serving passengers. The bus/shuttle operates as a shared-ride mode in which the passengers are transported to/from the connecting terminal from/to their desired location within a predetermined service area. Typically, pre-scheduled departure times are set at the terminal and passengers have a means by internet or phone call to book their ride. Usually, they are fully or partially funded by the local transit authority in the form of socially necessary transport. At

* Corresponding author. Tel.: +1 979 458 4171; fax: +1 979 862 8498.
E-mail address: quadrifo@tamu.edu (L. Quadrifoglio).

0191-2615/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trb.2013.01.008

other times, private transit operators or managers are responsible for their operations. Over the recent years, feeders have experienced a surge in ridership among commuters. However, transit managers often discontinue operating them mainly for lack of proper strategies related to scheduling at different demands during the day or season. In fact, a recent survey conducted among 1100 transit managers, representing public transit systems of different sizes and types, indicated that flexible transit services were discontinued mainly due to ''Problems with scheduling—can't make time points when demand for flexible trips is high or have too much extra time when demand for flexible is low'' (Potts et al., 2010).

Better planning, design and operation of these services may provide a potential solution to the first/last mile problem, which ultimately contribute towards greater goals like reducing traffic, mitigating related pollution and congestion problems and ultimately increase the overall ''livability''. What stops this from happening is the lack of existing research and information for transit managers in improving the state-of-the-art practice in operating these flexible transit services, especially related to the 'problems with scheduling.' In this paper, we are focusing our research efforts to partially fill the above gap. More specifically, we propose a model to describe the relationship between the level of service provided to passengers and the terminal-to-terminal cycle length (one of the decision variables of the service), allowing for a manageable computation of its optimal value that could otherwise be estimated only using extensive simulation analyses or by trial-and-error or rule-of-thumb procedures, as often done in practice. Several sets of simulation experiments and some comparison with existing services appear to validate our model, which would allow planners, decision makers and practitioners to quickly identify a good estimate of the best feeder transit operating design within a given service area under given circumstances.

2. Literature review

Demand responsive systems resemble dial-a-ride transit in terms of its operations, which have been extensively studied (Roos, 1971). Savelsbergh and Sol (1995), Desaulniers et al. (2002), and Cordeau and Laporte (2007) provided comprehensive reviews of the proposed methodologies and solutions for dial-a-ride's pick-up/deliver problems. Dial-a-ride used as door-to-door transportation services can be frequently found in the form of ADA paratransit services (Diana and Dessouky, 2004; Melachrinoudis et al., 2007; Quadrifoglio et al., 2008). These services have often been modeled with time window settings for passenger requests (Ropke et al., 2007; Cordeau, 2006). Rural transit providers also rely on DRT systems, due to low population density not allowing traditional fixed-route service to be sufficiently efficient. These low demand density areas are especially those that often lack a reasonable transportation infrastructure. A small community called El Cenizo along the US–Mexico border in Texas serves as a good example (Quadrifoglio et al., 2009).

Mathematical formulations are often carried out for optimizing and managing fleet size for demand responsive or dial-a-ride systems (Horn, 2002; Diana et al., 2006, among others). A transit operator would prefer to minimize fleet to lower operating costs (Ahouissoussi and Wetzstein, 1998). Often, a single shuttle could suffice if the demand responsive is meant for residential areas and those in particular used among the elderly community (TriMet Paratransit Program, 2012). Single vehicle study for dial-a-ride can be found in the work of Psaraftis (1980) as an exact dynamic programming solution. The single vehicle case deals with mostly minimizing route duration, ride time, and waiting time of the passengers. Chandra et al. (2011) studied a simulation based approach using a single vehicle case to estimate optimal cycle length for the feeder services within a residential community. These kinds of studies, however, require considerable data input a priori that often are not available in a new area.

Scheduling problems of flexible transit such as DRT or DART often fall in the category of TSP that is known to be NP-hard. Previously, Jaw et al. (1986) described a heuristic for a time-constrained version of the many-to-many DART Problem. The algorithm describes the Advanced Dial-a-Ride Problem with Time Windows (ADARTW) with service quality constraints and identifies feasible insertions of passengers into vehicle work-schedules. As a test on the performance of the heuristic, Barr et al. (1995) has provided reporting guidelines for computational experiments performed using heuristic methods. Later, Campbell and Savelsbergh (2004) developed an efficient 'insertion heuristics' for vehicle routing and scheduling problems that was computationally fast and could easily handle complicated constraints.

Literatures can be found that deal with a single cycle (or headway) optimization for performance of operating a 'fixed' feeder transit services (Chowdhury and Chien, 2011; Zhao and Zeng, 2008). The same for flexible transit systems, such as DRT, is still not properly addressed due to complex service request times and demand uncertainty. Surely, researchers have emphasized the importance of using an optimal bus dispatch policy by varying shuttle capacity and under stochastic lead times with fixed stepwise costs (Ignall and Kolesar, 1974; Alp et al., 2003). There are other models for transit bus dispatch at intermodal transfer stations using several techniques such as bus tracking technology, dynamic vehicle dispatch and using archived bus dispatch systems data (Dessouky et al., 1999; Bertini and El-Geneidy, 2004). It is emphasized that linking passenger-waiting times with feeder frequency is very critical in designing an optimal schedule. Longer and unreliable feeder bus headways or cycle lengths lead to increased passenger waiting times (Chang and Hsu, 2004; Özekici, 1987).

Analytical modeling, simulation and continuous approximations have been used to analyze flexible transit services. Cortés and Jayakrishnan (2002), Pagès et al. (2006), Aldaihani et al. (2004), Clarens and Hurdle (1975), Langevin et al. (1996) and Szplett (1984) provide good examples and reviews on these type of research approach. Research specifically on decision tool for the design of feeder transit services is somewhat limited. Quadrifoglio and Li (2009) approached the problem to estimate the critical demand density at which a switch from fixed-route to demand responsive policy would be appropriate. Other works on these services can be found in Cayford and Yim (2004) and Khattak and Yim (2004).

Serving passengers with a shuttle in demand responsive operations can be treated as a problem similar to those found in queuing theory, where the shuttles can be considered as dynamic service windows, passengers as service objects, and dispatch time from service windows as waiting times (Daganzo, 1990). However, the influence of service area characteristics cannot be easily integrated with performance in these studies. This paper fits in this category by proposing a model to estimate the most appropriate cycle length to operate the service given the demand and the shape of the service area.

3. Methodology

3.1. Service area, demand and optimal cycle

We are considering a generic residential area whose shape can be approximated with a rectangle with length L and width W, served by a single shuttle. The terminal (designated as D) with coordinates $(0,0)$ is located in the middle of an edge (at $W/2$). The shuttle is departing from D at constant time intervals (cycle) C to serve passengers, which can request to go from D to the service area ("drop-off" customers) or vice versa ("pick-up" customers). Demand within the area is assumed to be spatially uniformly distributed as well as temporally uniformly distributed (Poisson process) within a target time interval T of the day. This demand assumption is reasonable if at the terminal there are enough transit lines with high frequency and/or if the terminal is also a trip production/attraction site, such as a shopping mall or similar.

3.2. Existence of optimal cycle and performance definition

In a given time slot of the day T, the shuttle performs a number of cycles to serve the given demand. The longer the cycle length, the more ride sharing will be used, but passengers will experience longer average riding/waiting time, so shorter cycle lengths are intuitively more desirable for customers. However, the shorter the cycle length, the more the cycles, the more *extra* driving will be needed to go to/from the terminal and fewer customers can be served. At the limit, one customer at a time can be served, not taking advantage of any ride sharing. With negligible demand, this taxi-like operating practice would be plausible and best for customers; but with high enough demand, the service would be oversaturated and queues and spillovers will more likely occur, average waiting time will increase, thus lowering the level of service. These two combined effects cause an optimal cycle length to exist with the right amount of ride-sharing (minimizing disutility U, thus maximizing the average level of service), at which the system will be able to optimally serve all customers and that we aim to find in this paper.

It is generally difficult to identify a unique definition of performance for a transit system as priorities differ among stakeholders. Several authors have used measures such as passenger cost, passengers per vehicle hour, vehicle miles per operator, cost per vehicle mile, cost per vehicle hour, the ratio of cost to fare box revenue and fleet fuel efficiency for the urban public transit (Gleason and Barnum, 1982; Fielding et al., 1985; Badami and Haider, 2007). However, all seem to agree that transit performance can generally be identified as a combination of operating costs and service quality. In our model, the service quality is expressed as passengers' disutility (cost): a weighted sum of expected waiting time and the expected in-vehicle travel time of passengers (there is no walking time for door-to-door services). Thus, the disutility (cost) $U = \gamma_1 w_t + \gamma_2 r_t$ (w_t = expected waiting time and r_t = expected riding time; γ_1 and γ_2 are the weight factors for the waiting time and riding time respectively, to be properly calibrated). This is used to identify the optimal C to maximize serving quality. Within each operating cycle C, the operations are then conducted so that total distance is minimized (like a Traveling Salesman Problem) to minimize operating costs.

The proposed methodology consists of building an analytical queuing model for estimating the optimal cycle length in two steps. The first step consists of identifying equations linking a given cycle length C and the average number of passengers n that can be served at most by the shuttle within C; the second part consists of building the model of our objective function (disutility) U as a function of C to find the optimum.

3.3. $C(n)$ models

We are summarizing four models, found in the literature, which can be potentially used to describe the relationship between C and n. Note that these models provide an approximation of the optimal C needed to serve n passengers by minimizing the total distance, just as in a Traveling Salesman Problem. These routes are devoid of any consideration for the waiting time or riding time of the individual passengers. We call these relationships between average C and n as designs and address them as Design (I)–(IV).

Design (I). Nearest-neighbor approach: The average closest distance between two uniformly distributed passenger demands, as developed by Quadrifoglio et al. (2006) for a very high demand density ρ, can be used for this approach. This model provides a good approximation for the optimal C and n relationship with very high demand density. After adapting the model to our case by including a needed newly developed mathematical derivation, the relationship can be written as (see Appendix A for details),

$$C = \left(\frac{1.15}{V\sqrt{\rho}}\right) + \frac{0.63n}{V\sqrt{\rho}} + (n+1)t_s \tag{1}$$

where t_s is the time taken at each stop or terminal to pick-up/drop-off of a passenger, and V is the average speed of the feeder bus. For a very high demand density ρ, the expression in (1) can be approximated as,

$$C \approx \frac{0.63\sqrt{n}\sqrt{LW}}{V} + (n+1)t_s \tag{2}$$

Design (II). Approximate TSP solution approach: Results from work of Beardwood et al. (1959) and Jaillet (1988) can also be used to express another approximate relationship between C and very large n, with

$$C = \frac{\sqrt{nLW}}{V} + (n+1)t_s \tag{3}$$

Design (III). No-backtracking approach: Exploiting the scheduling guideline as proposed by Daganzo (1996) and Quadrifoglio and Li (2009), a strategy can be set for serving passengers for a lower passenger demand (low n) and a high length-to-width ratio of service area (see Fig. 1). In this case, the vehicle would move through the upper half of the region in a no-backtracking policy left-to-right, and move through the bottom half in a no-backtracking policy right-to-left. The relationship between C and n is expressed in this case by the following,

$$C = \frac{2L\left(\frac{n}{n+1}\right) + \frac{2W}{3} + \frac{Wn}{6}}{V} + (n+1)t_s \tag{4}$$

With the further assumption $\left(\frac{n}{n+1}\right) \approx 1$, the relationship between C and n in (4) becomes linear and can be expressed as

$$C \approx \frac{2L + \frac{2W}{3} + \frac{Wn}{6}}{V} + (n+1)t_s \tag{5}$$

Using (5) for computing C does not change its values significantly with respect to (4), especially for higher values of n (Quadrifoglio and Li, 2009). This model also closely estimates the optimal cycle length and performs better for high L/W ratios.

Design (IV). Random approach: This service strategy is shown mainly for comparison purposes with above models, as it would intuitively become more inefficient and not appropriate with increasing demand. The shuttles would serve passengers in a FIFO fashion without using any optimization algorithm (see illustration in Fig. 2). Using an average rectilinear distance traveled between uniformly random distributed demand within a rectangle as $(L + W)/3$ (Gaboune et al., 1993), the cycle length for this case is,

$$C = \frac{L}{V} + \frac{W}{2V} + \frac{(n-1)(L+W)}{3V} + (n+1)t_s \tag{6}$$

3.4. Model comparison

It is important to compare the relative applications of each of the design methodologies discussed above to select the best C and n for a single cycle relationship according to the actual service area and demand. The four relationships are compared to outputs obtained from an insertion heuristic policy (Jaw et al., 1986; Quadrifoglio et al., 2007) and optimality (for lower demand). This comparison for outputs is made with respect to the total travel time, which is C, of the DRT shuttle for a single service cycle. Insertion heuristic can be used for good solutions as compared to optimality obtained using optimization software. In fact, as the demand becomes higher, it is quite impractical to compute optimal solutions, simply because of unreasonable computational times involved in the process (Li and Quadrifoglio, 2010). Besides, insertion heuristic is widely used in practice for most scheduling problems.

Cycle length values are computed for three different shapes of a one square mile service area (shown in Table 1): Case 1 is a square with $L = W = 1$ mile; Case 2 is a rectangle with $L = 2$ miles and $W = 0.5$ miles; Case 3 is a slimmer rectangle with $L = 3$ miles and $W = 0.33$ miles. An average of 1000 replications were performed in *MATLAB R2010b* for each value of C and the

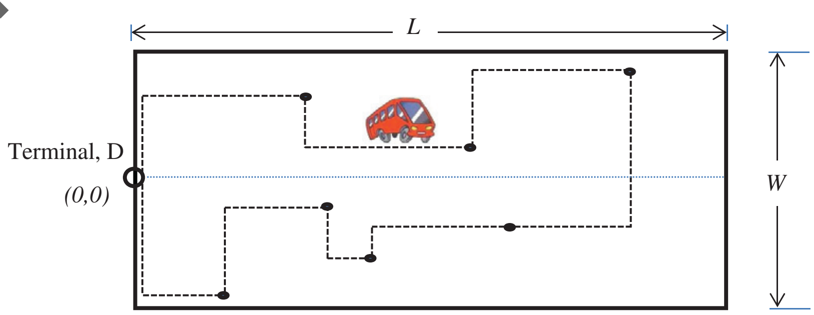

Fig. 1. Shuttle pick-up/drop-off strategy.

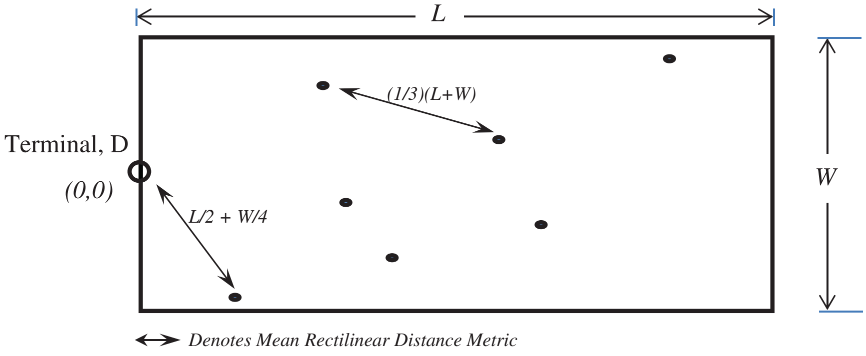

Fig. 2. Shuttle service operated using random pick-up/drop-off strategy.

Table 1
Cycle lengths (in minutes) for different demand and service area dimensions.

n	L = 1 mile, W = 1 mile						L = 2 miles, W = 0.5 miles						L = 3 miles, W = 0.33 miles					
	I	II	III	IV	V	C_s	I	II	III	IV	V	C_s	I	II	III	IV	V	C_s
1	2.9	4.0	6.5	5.5	5.1	5.0	2.9	4.0	8.2	7.7	7.5	7.4	2.9	4.0	10.8	10.5	10.6	10.5
2	4.2	5.7	8.5	8.0	7.2	6.5	4.2	5.7	11.0	10.7	10.4	10.1	4.2	5.7	14.5	14.3	14.2	14.1
3	5.3	7.2	10.0	10.5	8.6	8.1	5.3	7.2	12.7	13.7	12.1	11.9	5.3	7.2	16.7	18.2	16.1	15.8
4	6.3	8.5	11.3	13.0	10.0	9.6	6.3	8.5	14.1	16.7	13.5	13.4	6.3	8.5	18.2	22.0	17.6	17.2
5	7.2	9.7	12.5	15.5	11.0	10.4	7.2	9.7	15.2	19.7	14.5	14.3	7.2	9.7	19.5	25.8	19.1	18.5
6	8.1	10.8	13.6	18.0	12.1	11.8	8.1	10.8	16.3	22.7	15.6	15.4	8.1	10.8	20.6	29.7	19.9	19.3
7	9.0	11.9	14.7	20.5	13.0	12.8	9.0	11.9	17.2	25.7	16.4	16.3	9.0	11.9	21.6	33.5	21.0	20.4
8	9.8	13.0	15.8	23.0	14.0	13.3	9.8	13.0	18.1	28.7	17.6	17.5	9.8	13.0	22.5	37.3	21.9	20.7
9	10.7	14.0	16.9	25.5	14.9	14.6	10.7	14.0	19.0	31.7	18.2	18.1	10.6	14.0	23.3	41.1	22.8	22.1
10	11.5	15.0	17.9	28.0	15.8	15.3	11.5	15.0	19.9	34.7	18.9	18.9	11.5	15.0	24.2	45.0	23.5	22.8
11	12.2	15.9	19.0	30.5	16.8	–	12.2	15.9	20.7	37.7	19.7	–	12.2	15.9	25.0	48.8	24.2	–
12	13.0	16.9	20.0	33.0	17.7	–	13.0	16.9	21.6	40.7	20.5	–	13.0	16.9	25.8	52.6	25.1	–
13	13.8	17.8	21.0	35.5	18.5	–	13.8	17.8	22.4	43.7	21.3	–	13.8	17.8	26.5	56.5	25.8	–
14	14.5	18.7	22.1	38.0	19.4	–	14.5	18.7	23.2	46.7	22.4	–	14.5	18.7	27.3	60.3	26.5	–
15	15.3	19.6	23.1	40.5	20.2	–	15.3	19.6	24.0	49.7	22.4	–	15.3	19.6	28.0	64.1	27.2	–
16	16.0	20.5	24.1	43.0	21.0	–	16.0	20.5	24.8	52.7	23.0	–	16.0	20.5	28.7	68.0	28.0	–
17	16.8	21.3	25.1	45.5	21.8	–	16.8	21.3	25.5	55.7	24.5	–	16.8	21.3	29.5	71.8	28.7	–
18	17.5	22.2	26.1	48.0	22.7	–	17.5	22.2	26.3	58.7	25.7	–	17.5	22.2	30.2	75.6	29.3	–
19	18.2	23.0	27.2	50.5	23.5	–	18.2	23.0	27.1	61.7	26.0	–	18.2	23.0	30.9	79.5	30.0	–
20	18.9	23.9	28.2	53.0	24.4	–	18.9	23.9	27.9	64.7	26.8	–	18.9	23.9	31.6	83.3	30.6	–
21	19.6	24.7	29.2	55.5	25.1	–	19.6	24.7	28.7	67.7	27.1	–	19.6	24.7	32.3	87.1	31.3	–
22	20.3	25.5	30.2	58.0	26.0	–	20.3	25.5	29.4	70.7	27.8	–	20.3	25.5	33.0	90.9	32.0	–
23	21.0	26.3	31.2	60.5	26.7	–	21.0	26.3	30.2	73.7	28.5	–	21.0	26.3	33.7	94.8	32.7	–
24	21.7	27.1	32.2	63.0	27.5	–	21.7	27.1	31.0	76.7	29.3	–	21.7	27.1	34.4	98.6	33.3	–
25	22.4	27.9	33.2	65.4	28.3	–	22.4	27.9	31.7	79.7	30.0	–	22.4	27.9	35.1	102.4	34.0	–
26	23.1	28.7	34.2	67.9	29.1	–	23.1	28.7	32.5	82.7	30.7	–	23.1	28.7	35.8	106.3	34.7	–
27	23.8	29.5	35.2	70.4	30.0	–	23.8	29.5	33.3	85.7	31.4	–	23.8	29.5	36.5	110.1	35.3	–
28	24.4	30.3	36.2	72.9	30.7	–	24.4	30.3	34.0	88.7	32.1	–	24.4	30.3	37.1	113.9	36.0	–
29	25.1	31.1	37.2	75.4	31.4	–	25.1	31.1	34.8	91.7	32.8	–	25.1	31.1	37.8	117.8	36.6	–
30	25.8	31.9	38.2	77.9	32.3	–	25.8	31.9	35.6	94.7	33.5	–	25.8	31.9	38.5	121.6	37.2	–
31	26.5	32.6	39.2	80.4	33.0	–	26.5	32.6	36.3	97.7	34.2	–	26.5	32.6	39.2	125.4	37.9	–
32	27.1	33.4	40.3	82.9	33.8	–	27.1	33.4	37.1	100.7	34.9	–	27.1	33.4	39.9	129.3	38.5	–
33	27.8	34.2	41.3	85.4	34.5	–	27.8	34.2	37.8	103.7	35.6	–	27.8	34.2	40.6	133.1	39.2	–
34	28.5	34.9	42.3	87.9	35.2	–	28.5	34.9	38.6	106.7	36.3	–	28.4	34.9	41.2	136.9	39.9	–
35	29.1	35.7	43.3	90.4	36.0	–	29.1	35.7	39.3	109.7	36.9	–	29.1	35.7	41.9	140.7	40.5	–
36	29.8	36.4	44.3	92.9	36.8	–	29.8	36.4	40.1	112.7	37.6	–	29.8	36.4	42.6	144.6	41.1	–
37	30.4	37.2	45.3	95.4	37.5	–	30.4	37.2	40.9	115.7	38.3	–	30.4	37.2	43.3	148.4	41.7	–
38	31.1	37.9	46.3	97.9	38.2	–	31.1	37.9	41.6	118.7	38.9	–	31.1	37.9	44.0	152.2	42.4	–
39	31.7	38.7	47.3	100.4	38.9	–	31.7	38.7	42.4	121.7	39.6	–	31.7	38.6	44.6	156.1	43.0	–
40	32.4	39.4	48.3	102.9	39.7	–	32.4	39.4	43.1	124.7	40.3	–	32.4	39.4	45.3	159.9	43.7	–

I – *Nearest-neighbor* [Eq. (2)], II – *Approx. TSP* [Eq. (3)], III – *No-backtrack* [Eq. (5)], IV – *Random* [Eq. (6)], V – *Insertion heuristic*, C_s – *Optimal* for a single cycle.

values of n vary between 1 and 40 to cover any variability in demand within each cycle. The average feeder speed is assumed to be 20 mile per hour, posted speed limit found in most residential areas where shuttles operate, and the value for average time t_s spent at a passenger stop is assumed to be 30 s. Optimal cycle lengths are obtained using a TSP code in MATLAB and computing optimal using CPLEX 12.1. The optimal cycle lengths are denoted as C_s in Table 1. Following the values in table, a close match between the insertion heuristic and C_s is found. We did not compute C_s for $n > 11$ as optimal TSP were too long to obtain and insertion heuristic (closely matching C_s for $n \leqslant 11$) is used as benchmark for all the four design methodologies.

The results in Table 1 clearly show that the relationship from design method (III), *No-Backtrack*, matches well with insertion heuristic (V) cycle lengths for higher L/W and lower n, which is the more commonly found shape for real residential areas. The cycle length values from design method (II) are closer to insertion heuristic outputs for the square shaped area for almost all values of n. The values for C using design methods (I) and (II) match more closely with insertion heuristic values as n increases as compared to other design methods. This match is further improved as the demand density becomes higher (as expected). Also observe that (I) and (II) show identical cycle length values for all three cases for the same n, as they depend on the total area (LW, which is the same for all) and not on the shape (identified by the L/W ratio). Cycle lengths from the *Random* design method (IV) match the insertion heuristic (V) only for low demands ($n < 5$), as expected, and do not show a significant edge over the other methods. Thus, we discard *Random* design (IV) from here on and use design methods (I), (II) and (III) only to be included in the second step model building.

These results set the stage for building our analytical model and choosing the most appropriate $C(n)$ expression depending on the conditions, as they are useful in estimating cycle lengths for any given service area geometry and demand.

3.5. Feeder operations

The feeder shuttle is assumed to operate daily for a fixed duration of time. We focus on a portion of the day (T hours), during which $\left[\frac{T}{C}\right] = [\tau]$ cycles ($[\cdot]$ stands for the nearest integer value of $\cdot$) of constant duration C will be performed to serve a total of N passengers. Passenger requests for a given cycle are collected during the previous cycle and considered in a FIFO fashion. The above situation is explained through sets of slots (each of duration C) corresponding to each of the dispatch times of the feeder bus in the sketch of Fig. 3.

The variables t_1 and T_1 are the start time of the passenger request times and the feeder bus service times, respectively. Similarly, t_{k+1} and T_{k+1} are the end times of the passenger request times and the feeder bus service times, respectively, for a total of k slots in a day formed equal to $[\tau]$. Note that $(T_1 - t_1) = C$ (and $(T_{k+1} - t_{k+1}) = C$) with $(t_{k+1} - t_1) = T$ (and $(T_{k+1} - T_1) = T$). However, in case of passenger spillovers carried from the previous cycles, the shuttle operates an extra cycle to serve all at end time = T_{k+2}.

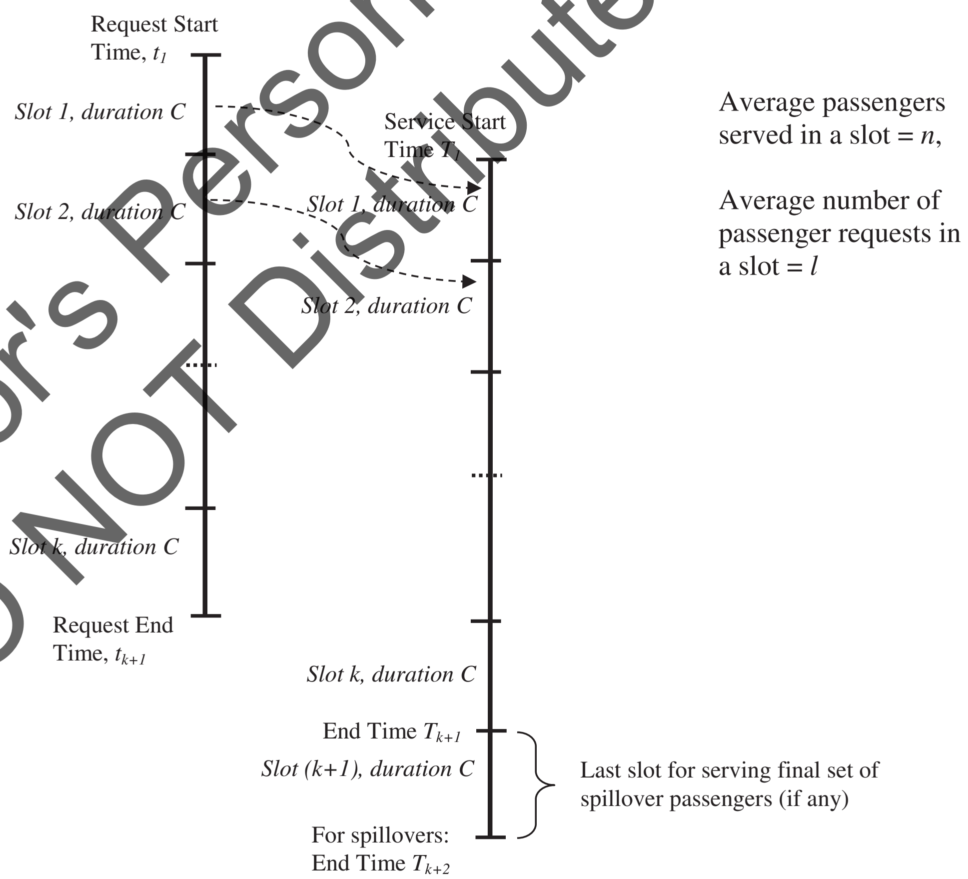

Fig. 3. Illustration of passenger requests and service times using slots of duration C.

3.6. Service disutility

In this section we develop a continuous approximation model assuming a deterministic process with constant (average) time intervals between service requests and compare them to simulated realistic Poisson process. There are $\frac{NC}{T} = l$ requests/cycle on average. We know that n is the numbers of passengers which can be served at most on average in a cycle C (as estimated by the models illustrated in the previous Section 3.3). If $l \leqslant n$, the system is under saturation and can comfortably serve all demand in every cycle. We assume that this collected demand (l) is scheduled for service in a given cycle following a near optimal sequence, generated by an insertion heuristic algorithm (this scheduling procedure is the same adopted by SuperShuttle (2012). If instead $l > n$, the system is oversaturated and the (average) residual ($l - n$) passengers would need to be served in the next cycle, causing a stable queue to build up in the system and additional waiting time to exist.

The disutility (cost) experienced by passengers $U = \gamma_1 w_t + \gamma_2 r_t$ can be modeled as

$$U = \begin{cases} Q_1 & \text{if } l > n \\ Q_2 & \text{if } l \leqslant n \end{cases} \tag{7}$$

We aim to model the above cases as a function of any demand level and L/W geometric dimension of the service area and the decision variable C, to derive its optimal. Derivation of w_t and r_t in Q_1 and Q_2 involves careful accounting for passenger requests served in a cycle. We define α as the fraction of pick-up passengers (going to service area to terminal).

For Q_1, the average riding time will be easily computed for all passengers as $r_t = C/2$. The waiting time is more complicated. $1 - \alpha$ drop-off passengers (going from terminal to service area) need to wait an average of $C/2$ time (from their show-up time to the beginning of the shuttle ride from the terminal); α pick-up passengers need to wait an average of time C ($C/2$ from their show-up time to the beginning of the shuttle ride from the terminal and an additional $C/2$ till their pick-up time). Thus, average waiting time for all customers served would be $(1 + \alpha)C/2$ (see also Quadrifoglio and Li (2009)). However, only n passengers per cycle can be served at their requested cycle; the remaining $l - n$ passengers (both pick-up and drop-off ones) in each cycle will have to be served in the following cycle and there will also be an additive spillover effect. Thus, $l - n$ passengers in the first demand cycle will be served in the second service cycle and would need to wait for an additional average time of C. In the second demand cycle, $2(l - n)$ customers will be served in the third service cycle (since additional $l - n$ customers will be pushed back by the first service cycle's spillover). In general, at demand cycle k, $k(l - n)$ customers will be served in service cycle $k + 1$. Thus, the additional waiting time to be considered will be $C\{[(l - n) + 2(l - n) + \ldots + \tau(l - n)]/N\} = C[\tau(1 + \tau)/2](l - n)/N$, considering only $\tau = \lfloor \tau \rfloor$. Eventually, the overall average waiting time will be $w_t = (1 + \alpha)C/2 + C[\tau(1 + \tau)/2](l - n)/N$. Note that the queuing effect we are trying to capture and model is significant, but small enough that we can safely assume that there will not be situations in which customers are pushed back more than one extra cycle within T, since this would not be realistic, as the system would be excessively oversaturated and would need more shuttles, not just an operational fix. In particular, this assumption would be acceptable when $\tau(l - n) < n$, which ensures that for the number of spilled-over passengers from the last slot τ are less than n and thus all served in the very next cycle. By substituting $\tau = T/C$ and $l = NC/T$, we need to have

$$n > \frac{N}{\left(1 + \frac{T}{C}\right)} \tag{8}$$

For Q_2, when demand $l < n$, the shuttle would take an average time $t < C$ to complete every cycle (and wait a slack time $C - t$ at the terminal before the next service cycle). An estimate of t that can be computed using the design methodologies (I), (II), and (III) discussed earlier, plugging in l instead of n and taking the resulting C value as t. Average riding time for all passengers is $r_t = t/2$. The waiting time for drop-off passengers is the same as for Q_1; for pick-up passengers it is $C/2 + t/2$, where the second part takes in account the actual travel time $t < C$ at each cycle. There will not be any stable queue formation in this Q_2 case, so no extra waiting time.

In summary,

$$Q_1 = \gamma_1 C\left(\frac{1+\alpha}{2} + \frac{\tau(1+\tau)}{2}\frac{(l-n)}{N}\right) + \gamma_2\left(\frac{C}{2}\right) \tag{9}$$

$$Q_2 = \gamma_1\left(\frac{C}{2} + \frac{\alpha t}{2}\right) + \gamma_2\left(\frac{t}{2}\right) = \gamma_1\left(\frac{C}{2}\right) + (\alpha\gamma_1 + \gamma_2)\left(\frac{t}{2}\right) \tag{10}$$

As mentioned earlier (Section 3.2), the existence of an optimal C to minimize U is intuitive; we are attempting to find it analytically and verify our results.

Keep in mind that n is a function of C and can be obtained by using the reverse expressions from methods (I)–(III) by using approximate expressions wherever necessary. Skipping the easy but cumbersome mathematical passages, we can write $n \approx (hC + f\sqrt{C} + g)$ with the values of h, f and g listed in Table 2.

It is difficult to work with non-linear expressions (due to square root of C involved) if design methods used are (I) and (II) as $f \neq 0$. This also makes it hard to derive closed form expressions as roots of cubic polynomials are not trivial to estimate. Thus, we proceed to give some closed form results only for design method (III) when $f = 0$ and only report optimal cycle lengths obtained graphically for the other two methods along with simulation results later. From now on, unless specified, assume we are only working with design method (III). Thus, writing Q_1 in (9) for $f = 0$, $l = NC/T$ and $\tau = T/C$ we get

Table 2
Expressions for h and g.

Design method	h	f	g
(I)	$\frac{1}{t_s}$	$-\sqrt{\left(\frac{0.63LW}{t_s^3V^2}\right)}$	$-1+\frac{0.31LW}{t_s^2V^2}$
(II)	$\frac{1}{t_s}$	$-\sqrt{\left(\frac{LW}{t_s^3V^2}\right)}$	$-1+\frac{0.5LW}{t_s^2V^2}$
(III)	$\frac{1}{\left(\frac{W}{6V}+t_s\right)}$	0	$-\left(\frac{12L+4W+6Vt_s}{W+6Vt_s}\right)$

Table 3
Expression for right hand side coefficients of Eq. (13).

Design method	m	b
(III)	$\frac{\gamma_1}{2}+\frac{(\alpha\gamma_1+\gamma_2)}{2}\left(\frac{NW}{6VT}+\frac{Nt_s}{T}\right)$	$\frac{(\alpha\gamma_1+\gamma_2)}{2}\left(\frac{6L+2W}{3V}+\boldsymbol{t}_s\right)$

$$Q_1=\frac{\gamma_1}{2}\left((2+\alpha)C-\frac{(hC+g)T}{N}+T-\frac{(hC+g)T^2}{NC}\right)+\gamma_2\left(\frac{C}{2}\right) \tag{11}$$

which is a convex function for $C > 0$ with minimum attained at

$$C\approx\sqrt{\frac{-gT^2}{N\left((2+\alpha)-\left(\frac{hT}{N}\right)+\left(\frac{\gamma_2}{\gamma_1}\right)\right)}} \tag{12}$$

which would be the optimal C, should Q_1 still be valid at this value. This would be possible theoretically, but not in practice, as at this value either the spillover condition does not hold anymore (and thus Q_2, not Q_1, would describe the system) or it would require demand higher than the system could sustain (over saturation) or it would violate Eq. (8) and therefore our model assumptions.

Q_1 is monotonically decreasing with increasing C and representing the system till the first intersection with Q_2 (10), which is a monotonically increasing linear function of C. This intersection represents the point at which the spillover condition stops and is also the estimate for the optimal C, which we aim to find. The equality $Q_1 = Q_2$ would be a quadratic equation of the following form:

$$\left(\frac{u}{C}+vC+r\right)=(mC+b) \tag{13}$$

where $u=\frac{\gamma_1}{2}\left(-\frac{gT^2}{N}\right)$, $v=\frac{\gamma_1}{2}\left((2+\alpha)-\frac{hT}{N}\right)+\left(\frac{\gamma_2}{2}\right)$, $r=\frac{\gamma_1}{2}\left(T-\frac{hT^2}{N}-\frac{gT}{N}\right)$, with h and g as defined in Table 2 for method (III).

Depending on the relationship between C and n from design methods (III), the expressions for m and b are as shown in Table 3.

Solving (13) gives the smallest value of cycle length which is C optimal (C_{opt}),

$$C_{opt}=\frac{(b-r)-\sqrt{(b-r)^2-4u(v-m)}}{2(v-m)} \tag{14}$$

with all variables defined earlier. The other root is to be discarded as it would have a higher disutility, could also very well occur beyond realistic capacity of the shuttle, but primarily Q_1 would simply not hold anymore (no more spillover).

For (14) to hold true, it is essential that $(b-r)^2-4u(v-m)\geqslant 0$ (say, Constraint 1) and $(v-m)>0$ (say, Constraint 2). If Constraint 1 is violated it would mean that there is no real intersection point for the spillover and the no-spillover curves. Therefore, C_{opt} cannot be obtained by the intersection of spillover and no-spillover case. Moreover, this situation would arise when the demands are much higher than sustainable causing a large number of spillovers at every cycle, violating our model assumptions. If Constraint 2 is violated then it would result in a negative value of C_{opt} which is again impractical. However, this situation would never arise for any DRT system since this would occur at very high variable demand (for a very high N in m to make $(v-m)$ non-positive).

4. Simulation experiments

The purpose of this section is to validate the derived analytical expressions for optimal cycle length estimation. Disutility values are obtained for different cycle lengths with three different assumptions of dimensions for rectangular service area, denoted by Case 1 (L = 1 mile, W = 1 mile), Case 2 (L = 2 mile, W = 0.5 mile) and Case 3 (L = 3 mile, W = 0.33 mile).

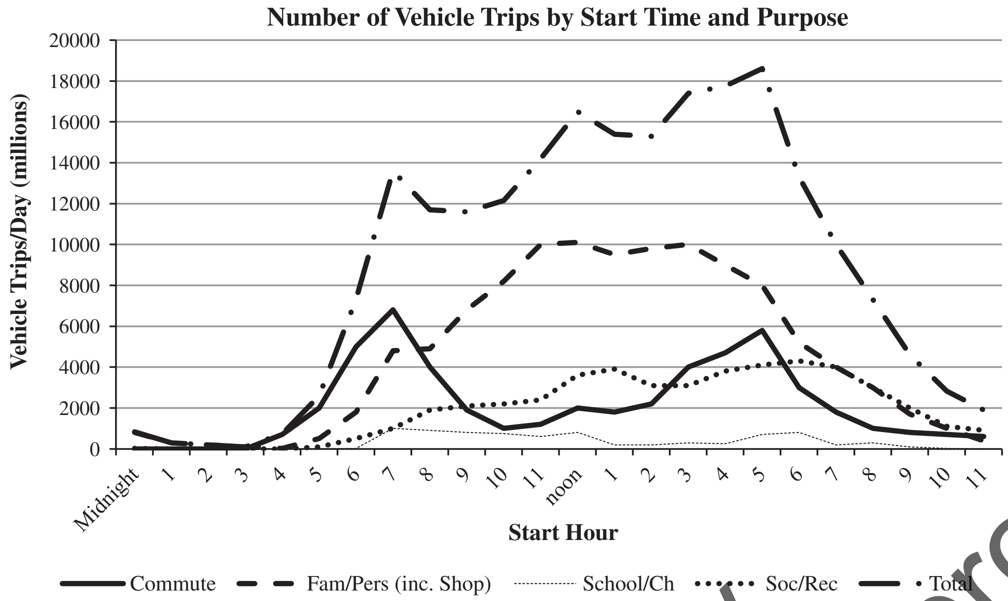

Fig. 4. Trip data for entire US. *Source*: Santos et al., 2011.

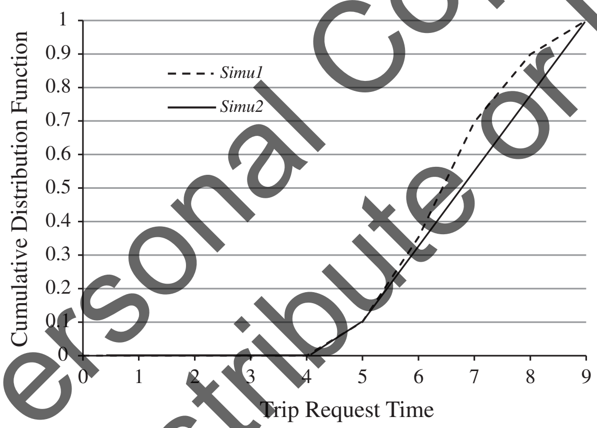

Fig. 5. CDF representing trip travel time for commuting.

Many residential areas can be approximated by square and rectangular shapes with dimensions shown in the three cases. The simulation is performed by coding the Feeder operations in MATLAB R2010b. The shortest street based path between any two demand points is computed using the Dijkstra's algorithm using the rectilinear distance between the points. The rectilinear distance is also a good approximation for street based distances between two demand points (Quadrifoglio and Li, 2009). This is particularly true for residential areas, most of which have a grid-like rectilinear street pattern, especially the ones using feeders. There were 20 replications used for each simulation.

The passenger demands used are N = 50, 80, 100 and 240 for T = 4 h of operations (assumed to be within peak-hours). The first three demands fall within those that are found in practice from several call-n-ride systems (Potts et al., 2010).

We perform two sets of simulations: one (*Simu1*) with realistic demand distribution taken from the report by Santos et al. (2011). The curves in Fig. 4 show different types of trips by start and purpose. We focus on the morning peak hour distribution periods of 5 am to 9 am for the *Total* distribution.

The other one (*Simu2*) is performed with the above N considered constant over T to closely match the assumption of our analytical model.

For simplification, the *Simu1* trip data for the commuters are converted to cumulative distribution functions (CDFs) as shown in Fig. 5. The actual CDF (a polynomial of higher degree and hence, difficult to invert) is slightly modified into an easily invertible piecewise linear function for random generation of travel times for passengers. In simulation, this is achieved by generating random real numbers between 0.1 and 1, and simply computing the corresponding travel request times using the linear function. A linear CDF is obtained that expresses the uniform random generation of passengers between 5 am to 9 am (*Simu2*).

Table 4
Input settings for the simulation model.

Parameter	Input numerical values		
(L,W)	Case 1: (1,1)	Case 2: (2,0.5)	Case 3: (3,0.33)
V	20 miles per hour (i.e. the average speed in a residential area)		
(γ_1/γ_2)	1.8 (Wardman, 2004)		
Fleet size	1 (A single feeder shuttle)		
Shuttle capacity	Infinite (or large enough) to accommodate all passengers within a given cycle		
Dwell time (t_s)	30 s (at the depot as well as at the passenger pick-up/drop-off location)		
Cycle length C	Minimum: $C_m = (2L + W)/V + 2t_s$ (time needed to reach the farthest point of the service area from the terminal and back in a rectilinear fashion) Maximum: Long enough to capture lowest disutility for each simulation case		

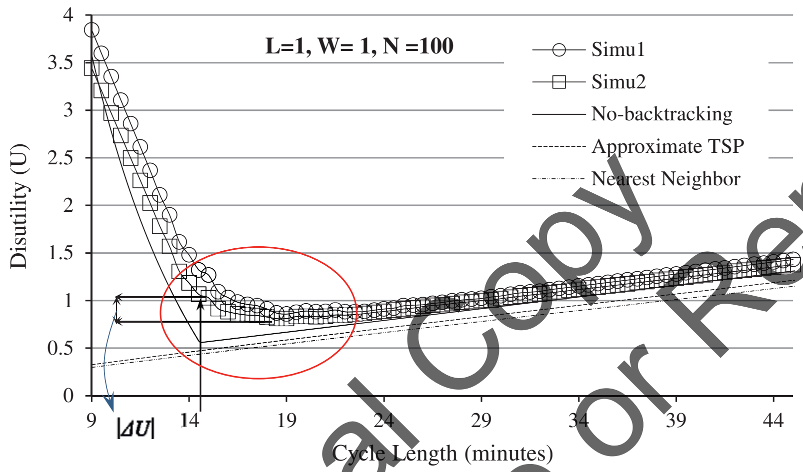

Fig. 6. Case 1 disutility versus C for passenger demand N = 100.

Note that the simulated morning peak-hour period will reasonably consist of mostly pick-up passengers ($\alpha \cong 1$). A similar analysis can be performed for the afternoon peak hours with $\alpha \cong 0$. However, our model allows for any kind of combination, just by choosing appropriately the value of α included in the analytical modeling earlier.

The collected demand at each cycle is scheduled with insertion heuristic using up all the cycle time C. Customers are considered in a FIFO fashion and simulated as pick-up or drop-off depending on a random number generated depending on α. Any customer not been able to be served in a cycle will be considered first in the next cycle.

The shuttle is assumed to travel at constant speed V, leave the terminal at the beginning of each cycle, stop a time t_s at each requested location following the generated schedule, and travel back to the terminal.

Other simulation settings used are outlined in Table 4.

4.1. Simulation results and discussion

Results are shown in Figs. 6–8. We are pleased to observe a very close match of our queuing model Eqs. (9), (10), and (14) to simulation results, validating our approach and allowing us to adopt the analytical model for design purposes. We are also pleased to observe a close overlap between *Simu1* and *Simu2*. This is particularly important to show that assuming constant demand over T as opposed to more realistic distribution is not a very strong assumption.

Three of the curves are used in the charts of Figs. 6–8 for disutility from three different design methodologies adopted for relationship between C and n: *Nearest Neighbor* uses method (I), *Approximate TSP* uses method (II) and *No-backtracking* uses method (III). As previously mentioned, we did not include the Random method (IV) as its results are far off compared to the others, as expected, for all realistic cases. Figs. 6–8 only show results for passenger demand = 100 with other optimal cycles and minimum disutility values for other demand levels being tabulated in Tables 5a–5c. The term C_m stands for the minimum cycle length needed (see Table 4).

All three models provide a good visual match for the no-spillover part (right) of the simulation curves. The *No-backtracking* method visually gives a very close match to the spillover part (left) too, overall providing a good approximation of the behavior of the system in all situations. As expected, curves obtained by the model underestimates the simulation one around the minimum, as delays and temporary queues are formed due to randomness to worsen the level of service, but not captured by our model based on expected values and continuous approximations. More sophisticated queuing models

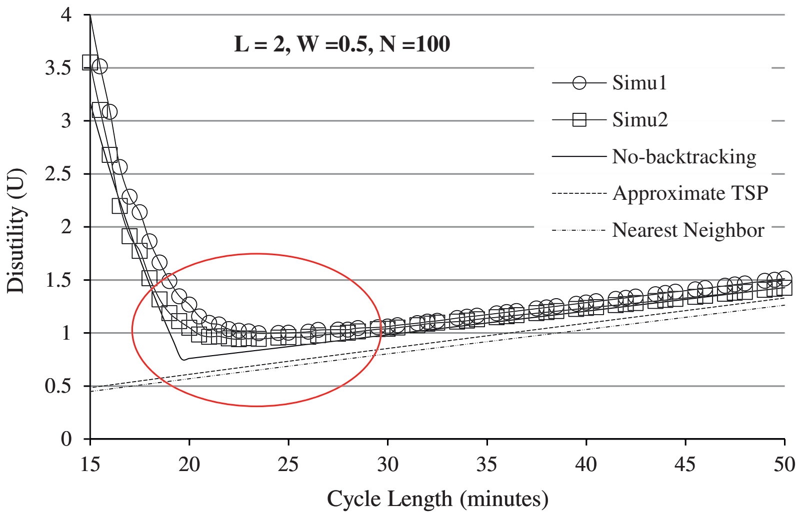

Fig. 7. Case 2 disutility versus C for passenger demand $N = 100$.

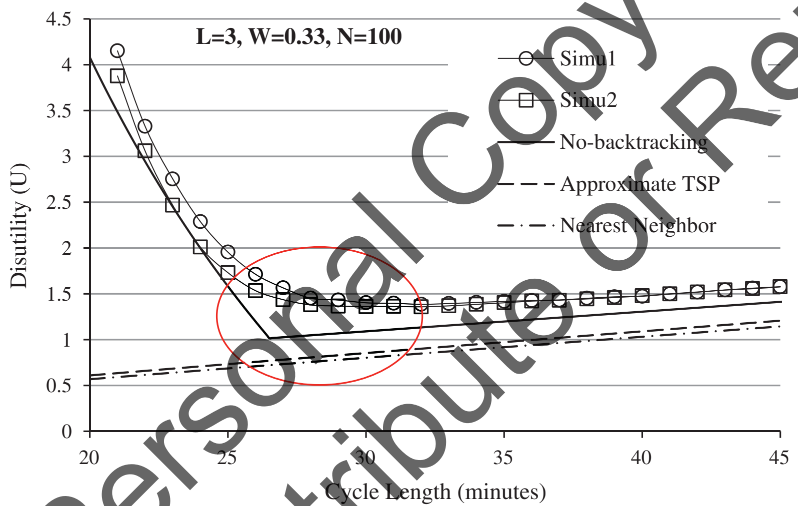

Fig. 8. Case 3 disutility versus C for passenger demand $N = 100$.

Table 5a
Optimal cycle lengths (C_{opt} in minutes) for Case 1 with different demands.

N	Case 1: L = 1 mile, W = 1 mile												
	Simu1		*Simu2*		Design method used								
					I			II			III		
	U	C_{opt}	U	C_{opt}	$^*C_{opt}$	$\|\Delta U\|\%$	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|(\%)$	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|\%$	$Err\ C_{opt}$ (%)
50	0.71	14–15	0.70	13–17	C_m	>100	31	C_m	>100	31	**10.7**	4	17
80	0.81	16–18	0.80	15–19	C_m	>100	40	C_m	>100	40	**12.7**	21	15
100	0.88	19–21	0.82	17–20	C_m	>100	47	C_m	>100	47	**14.5**	67	15
240	1.45	36–37	1.40	35–37	24.4	>100	30	**37.6**	8	16	28.3	>100	19

* All C_m values are 9 min.

might be more appropriate for close matches of the results, but they would certainly lose in practicality and simplicity, needed to more easily make planning decisions.

The column $|\Delta U|\%$ in Tables 5a–5c stands for the absolute percentage error calculated on the *Simu2* curve caused by using C_{opt} recommended by the analytical model versus the actual optimal of the curve. The expression for $|\Delta U|\%$ is calculated using ($|\Delta U| \times 100/U_{Simu2}$; see Fig. 6 as an example demonstration for *No-backtracking* design). The best C (C_{opt}) obtained from the theoretical curves are selected (**bold**) based on their $|\Delta U|\%$ values. The lower the $|\Delta U|\%$, the more preferred a particular design method is. Too high values of $|\Delta U|\%$ are reported as '>100' indicating values higher than 100 in Table 5a – it is an

Table 5b
Optimal cycle lengths (C_{opt} in minutes) for Case 2 with different demands.

N	Case 2: L = 2 mile, W = 0.5 mile													
	Simu1		*Simu2*		Design method used									
					I			II			III			
	U	C_{opt}	U	C_{opt}	$^*C_{opt}$	$\|\Delta U\|$%	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|$%	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|$%	$Err\ C_{opt}$ (%)
50	0.83	17–18	0.82	16–19	C_m	>100	13	C_m	>100	13	**16.0**	0	0
80	0.90	20–22	0.88	19–22	C_m	>100	26	C_m	>100	26	**18.0**	16	5
100	1.00	23–27	0.90	22–26	C_m	>100	36	C_m	>100	36	**19.4**	22	12
240	1.65	39–42	1.55	38–39	24.4	>100	36	**37.6**	11	1	44.2	55	16

* All C_m values are 14 min.

Table 5c
Optimal cycle lengths (C_{opt} in minutes) for Case 3 with different demands.

N	Case 3: L = 3 mile, W = 0.33 mile												
	Simu1		*Simu2*		Design method used								
					I			II			III		
	U	C_{opt}	U	C_{opt}	$^*C_{opt}$	$\|\Delta U\|$%	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|$(%)	$Err\ C_{opt}$ (%)	C_{opt}	$\|\Delta U\|$%	$Err\ C_{opt}$ (%)
50	1.10	24–25	1.00	23–25	C_m	>100	13	C_m	>100	13	**22.2**	2	3
80	1.20	27–30	1.20	25–29	C_m	>100	23	C_m	>100	23	**24.6**	6	2
100	1.40	30–33	1.40	27–32	C_m	>100	26	C_m	>100	26	**26.5**	7	2
240	1.80	44–47	1.75	42–46	24.4	>100	42	**37.6**	45	10	58.1	54	26

* All C_m values are 20 min.

indication that the corresponding design model is not accurate and others are preferred. The error terms (expressed as $Err\ C_{opt}$%) in the tables are similarly calculated with respect to the C_{opt} on *Simu2* but along the C axis. Since we are aiming to identify the best estimate of C_{opt} in order to minimize the disutility (U), it is appropriate to select model performance based on $|\Delta U|$%.

Results clearly show that the *No-backtracking* design method (III) gives the best results for lower demands (50, 80 and 100), but deviates for higher demand (240), where *Approximate TSP* curves provide the closest match. Since Feeder services generally operate at lower demand levels, we can recommend using no-backtracking model (III) for most practical situations. Better performance for (III) is also seen for slimmer service areas (Fig. 8 and Table 7) as opposed to square ones (Fig. 6 and Table 5a) to further validate the concept mentioned earlier (Section 3.3) that the *No-backtracking* design method (III) is more suitable for high length-to-width ratio of the service area.

Though we do not report results from the Random design method (IV), we did observe occasional simulation match with the analytical equations of the Random design in the prediction of C_{opt}. However, as compared to other methods, there was a regular lack of consistency and its curves would evidently not follow the behavior of simulation curves. We observe that a Random scheduling strategy is widely, practically and intuitively understood to be inefficient and its few matches occur at extremely low-demand levels (such as one or two demands per cycle) for which no algorithm/strategy would be needed for scheduling purposes anyway.

Note that the inequality (8) is found to be true for all our experiments in the ranges of interest around the C_{opt}. In particular, lower demand and/or more compact area cases (like Case 1), which are less critical, are fully satisfied for all values of C. For example, using N = 80, T = 4 h, C = 9 min (C_m for Case 1), we have that n should be >2.9, which is satisfied by the values in Table 1, where n is a value between 3 and 4 for a C = 9 min. For higher C (less critical scenarios), the inequality is satisfied with increasing margin. For higher demand cases and/or more dispersed area cases (like Case 3), we found that some scenarios do not satisfied (8) for lower C (close to C_m), but all begin to be satisfied at some value of C smaller than the C_{opt} (and continue to be valid from that value up), ensuring that our model assumptions are verified around the critical point of interest (C_{opt}).

Besides verifying (8), which is a deterministic approximation, we also similarly verified that all our simulation experiments do not show customers left un-served after one extra cycle.

5. Extra driving and optimal cycle length

To further explain and validate the intuitive existence of an optimal C discussed in Section 3.2 and shown in Figs. 6–8. The expressions $C(n)$ discussed in Section 3.3 can be split into two terms. The first term consists of the estimated time taken by the shuttle to go from the terminal to the first passenger in the serving sequence and from the last passenger back to the terminal (this term identifies the estimated *extra* driving to/from terminal). And the second terms consists of the estimated

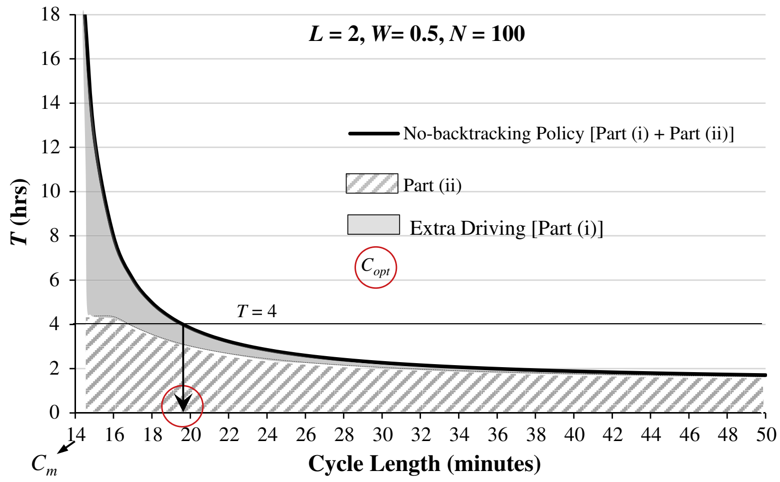

Fig. 9. The concept of extra driving with fixed N.

travel time taken to go from the first passenger to the last one, passing by the remaining $(n-2)$ passengers in the same cycle. For the No-backtracking design method III, Eq. (4) can be rearranged to show these two terms, denoted by (i)/(ii).

$$C = \underbrace{\frac{W}{2V} + \frac{4L}{(n+1)V} + 2t_s}_{(i)} + \underbrace{\frac{2L(n-2)}{(n+1)V} + \frac{Wn}{6V} + \frac{W}{6V} + (n-1)t_s}_{(ii)} \quad (15)$$

We show an example to illustrate the influence of the *extra driving* leading to varying C_{opt}. We assume a fixed demand of $N = 100$ for the service duration of the feeder with service area dimensions of $L = 2$ miles and $W = 0.5$ mile. The chart in Fig. 9 shows the increase in Feeder service duration needed with lower cycle lengths. This increase is attributed due to the extra driving portion represented by Part (i) of Eq. (15) which becomes dominating as C continues to decrease below the C_{opt} obtained for $T = 4$ h. This is shown using the shaded area in Fig. 9. In fact, with considerably lower cycle length values close to minimum cycle length value, $C_m = 14$ min, the Feeder shuttle would take more than a day to serve all the demand, which is impractical. An opposite behavior of the extra driving is seen as $C > C_{opt}$, during which, Part (ii) of the Eq. (15) dominates. This portion of the cycle length allows shuttle to serve all the demand within $T = 4$ h of service duration, however, undermining longer waiting and riding times incurred by the customers. Thus, an optimal cycle length C_{opt} for any given T exists which would ensure an equilibrium between the oversaturation effect of *extra driving* and too much of waiting and riding by the customers.

A companion chart shown in Fig. 10 illustrates the variation of N versus cycle length C for a portion of Feeder service duration, $T = 4$ h. A decrease in demand N results in decrease of C_{opt} which allows shuttle to perform more cycles due to greater availability of slack times. This can be visualized clearly with the example for $N = 50$, 80and 100 shown in Fig. 10. For really low values of N, the minimum cycle length ($C_m = 14$ min) becomes the optimal cycle length, C_{opt}, needed to serve a customer, which is 14 min for the assumed service area.

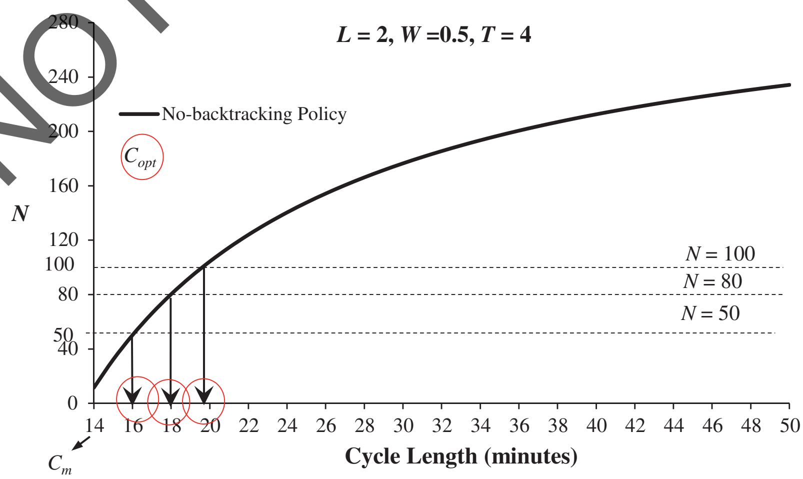

Fig. 10. The concept of extra driving with fixed T.

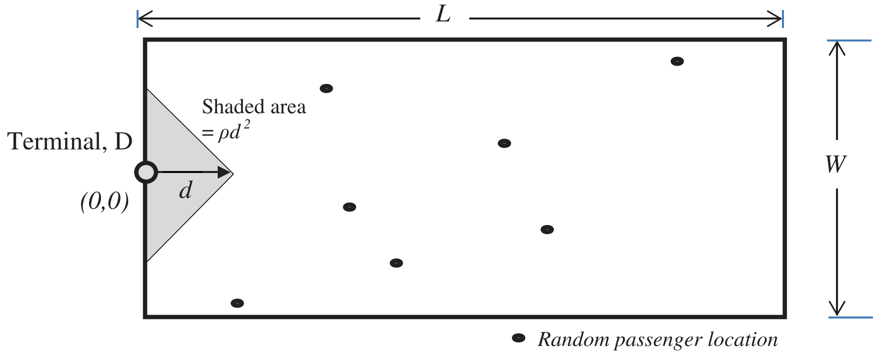

Fig. A.1. Illustration of expected closest distance between terminal and a passenger.

Table 6
Performance and service information data for RTD Call-n-Ride.[a]

Performance data		Service information							
Route	Boardings per service hour	Area coverage (sq. mile)	Approx. *Length*[A] (miles)	Approx. *Width*[A] (miles)	Peak headway in practice (minutes)	Disutility (U) for peak headway in practice	Peak headway analytical (minutes)[B]	Disutility (U) for peak headway analytical	Potential for disutility improvements?
N Inverness	9.5	2.12	1.5	1.5	10	8.0	15.7	0.6	NC[C]
Meridian	8.7	1.07	1.6	0.9	15	0.6	13.8	0.5	Almost same
Interlocken	7.7	5.39	3.0	1.7	30	1.2	26.7	1.0	Almost same
S Inverness	7.4	1.11	1.5	0.7	10	7.0	12.1	0.5	YES
Broomfield	4.5	7.13	3.2	2.2	30	1.1	27.4	1.0	Almost same
Louisville	4.1	8.20	3.3	2.5	30	1.1	28.6	1.1	Almost same
Dry Creek	3.8	4.74	2.5	2.0	30	1.0	21.6	0.8	Almost same

[a] Selected Call-n-ride services for which peak headway data were available (Potts et al., 2010).
[A] Approximate values obtained using Google Earth.
[B] Using Eq. (14) and design method (III).
[C] NC = No-conclusion with design method (III) since service area is a square.

6. Application example

The *No-backtracking* design method III appears to be the best for most scenarios in Tables 5a–5c, especially for lower demands. Thus, we compared our model with some of the existing real case transit system examples that experience similar conditions. Comparisons are made with RTD Call-n-Ride (CnR) that function similar to demand responsive transit with the data from the month of October 2008 (see Table 6) and the peak headway used by the operating transit agencies (Potts et al., 2010). The spatial distribution of the passenger requests is assumed to be uniform over the rectangular service area for the real CnR services which is the same assumption as for deriving our analytical framework in (14). This is an approximate assumption of demand distribution as for most of the CnR Routes; the Google Earth image shows almost a uniform spread of buildings and houses over the service area. The peak hours are assumed to be from 5 am to 9 am as assumed in the simulation.

Analytical peak headways/cycle lengths computed for these Routes show close similarities with the peak hour headways used in practice with no-backtracking design method. It is quite interesting to note that the operators have almost optimized efficiency of the Routes without any analytical aid. However, DRT systems having dimensions and demand settings that are suitable for the no-backtracking design (or, any of the three design methods mentioned in the paper), the efficiency gains can be increased by using the analytically calculated peak headway to the peak headway used in practice (see Table 6). For example, the route of S Inverness can improve its efficiency by adjusting the peak headway according to the proposed analytical formula of this paper. Thus, this model presented in this paper could serve as quite a good reference for the operators.

7. Conclusions

Feeder services are a potential solution to the first/last mile transit connectivity problem faced by modern society. An improvement of these services would ensure a more pleasant experience for passengers and eventually increase transit ridership, reduce congestion and pollution and increase the livability of residential areas. As noted by Potts et al. (2010), scheduling problems are a major concern for transit operators for these types of services. This paper proposes the development of an analytical queuing model to estimate the best duration of the cycle length from terminal to terminal using continuous approximation and inputs from demand data and geometrical parameters of the service area. An optimal cycle length C_{opt} must exist to balance two opposite effects: too long cycles would result in excessive riding and waiting time for passengers; too short cycles would cause an oversaturation of the system unable to serve all demand for excessive driving to/from the terminal.

Results give us a handy closed form expression which can be readily used to decide the best dispatch policy to operate a demand responsive feeder transit system. For square service area side length $L = W = 1$ mile, and total demands $N = 50–100$ over a $T = 4$ h period, the approximate C_{opt} are found to be 11–15 min for the peak hours of commuting. For service areas of length $L = 2$ miles, width $W = 0.5$ mile the estimated optimal cycle lengths C_{opt} vary between 16–20 min for the peak hours of commuting. Further, increasing the service area dimensions to a very high length ($L = 3$ miles) to width ($W = 0.33$ mile) ratio, the estimated optimal cycle lengths are found to be 22–27 min. Simulation experiments and some comparison with real services validate our approach well.

In conclusion, the developed model suggests and encourages transit planners and operators to make use of this methodological approach in selecting the correct operating policy for feeders, whose proper design and operations are becoming increasingly important to enhance the performance of the public transportation system network, within modern sprawled urban and suburban areas. We are aware of the limitations of the model, due to the needed approximation and assumptions used, but we believe our results give a contribution to the research in this area and are useful for practitioners too. Our future work consists of developing models for changes in fleet size and shuttle capacity as an extension of this current work. We are also investigating optimal cycle lengths for non-uniform demand distributions and for appropriate design methods that describe relationships between C and n, other than those mentioned in the paper.

Appendix A

A.1. Derivation of expected distance E[D] to the closest passenger demand from the terminal

Within a rectangular service area with length, L and width, W, spatial–temporal 'uniform' demand distribution of passenger requests follow a Poisson distribution having expected value of ρA_R, with the probability of finding a given number of points (A_n) within an area A_R as

$$P(A_n = q) = \frac{(\rho A_R)^q}{q!} e^{-\rho A_R} \quad \text{where, } \rho = \text{demand density and} \quad q = 1, 2, 3, \ldots \tag{A.1}$$

The average closest distance between the terminal and a passenger can be obtained for $q = 0$ and $A_R = d^2$ (see Fig. A.1), where d is the variable rectilinear distance between the terminal and the closest passenger.

$$E[D] = \int_0^\infty e^{-\rho d^2} \mathrm{d}d = \frac{1}{2}\sqrt{\frac{\pi}{\rho}} \approx \frac{0.89}{\sqrt{\rho}} \tag{A.2}$$

Using the average closest passenger-to-passenger distance as $\frac{0.63}{V\sqrt{\rho}}$ (by Quadrifoglio et al. (2006)), the cycle length C can be expressed as,

$$C = 2\left(\frac{0.89}{V\sqrt{\rho}}\right) + (n-1)\frac{0.63}{V\sqrt{\rho}} + (n+1)t_s = \left(\frac{1.15}{V\sqrt{\rho}}\right) + \frac{0.63n}{V\sqrt{\rho}} + (n+1)t_s \tag{A.3}$$

References

Ahouissoussi, N.B.C., Wetzstein, M.E., 1998. A comparative cost analysis of biodiesel, compressed natural gas, methanol, and diesel for transit bus systems. Resource and Energy Economics 20 (1), 1–15.

Aldaihani, M.M., Quadrifoglio, L., Dessouky, M.M., Hall, R., 2004. Network design for a grid hybrid transit service. Transportation Research Part A 38 (7), 511–530.

Alp, O., Erkip, N.K., Güllü, R., 2003. Optimal lot-sizing/vehicle-dispatching policies under stochastic lead times and stepwise fixed costs. Operations Research 51 (1), 160–166.

Badami, M.G., Haider, M., 2007. An analysis of public bus transit performance in Indian cities. Transportation Research Part A 41 (10), 961–981.

Barr, R.S., Golden, B.L., Kelly, J.P., Resende, M.G.C., Stewart, W.R., 1995. Designing and reporting on computational experiments with heuristic methods. Journal of Heuristics 1 (1), 9–32.

Beardwood, J., Halton, J.H., Hammersley, J.M., 1959. The Shortest Path through Many Points. Mathematical Proceedings of the Cambridge Philosophical Society, vol. 55(6). Cambridge Univ. Press, pp. 299–327.

Bertini, R.L., El-Geneidy, A.M., 2004. Modeling transit trip time using archived bus dispatch system data. Journal of Transportation Engineering 130 (1), 56–67.

Campbell, A.M., Savelsbergh, M., 2004. Efficient insertion heuristics for vehicle routing and scheduling problems. Transportation Science 38 (3), 369–378.

Cayford, R., Yim, Y.B., 2004. Personalized Demand-responsive Transit Service. California PATH Research Report UCB-PRR-2004-12, Berkeley, California.

Chandra, S., Shen, C.W., Quadrifoglio, L., 2011. A simulation study to derive the optimal cycle length for feeder transit services. International Journal of Information Systems in the Service Sector 3 (3), 27–46.

Chang, S.K.J., Hsu, S.C., 2004. Modeling of passenger waiting time in intermodal station with constrained capacity on intercity transit. International Journal of Urban Sciences 8 (1), 51–60.

Chowdhury, M.S., Chien, S.I.J., 2011. Joint optimization of bus size, headway, and slack time for efficient timed transfer. Transportation Research Record: Journal of the Transportation Research Board 2218, 48–58.

Clarens, G.C., Hurdle, V., 1975. An operating strategy for a commuter bus system. Transportation Science 9 (1), 1–20.
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., 2007. The dial-a-ride problem: models and algorithms. Annals of Operations Research 153 (1), 29–46.
Cortés, C.E., Jayakrishnan, R., 2002. Design and operational concepts of a high coverage point-to-point transit system. Transportation Research Record 1783, 178–187.
Daganzo, C.F., 1990. Some properties of polling systems. Queueing Systems 6 (1), 137–154.
Daganzo, C.F., 1996. Logistics Systems Analysis. Springer.
Desaulniers, G., Desrosiers, J., Erdmann, A., Solomon, M.M., Soumis, F., 2002. VRP with pickup and delivery. In: Toth, P., Vigo, D. (Eds.), The Vehicle Routing Problem. SIAM Monographs on Discrete Mathematics and Applications, Philadelphia.
Dessouky, M., Hall, R., Nowroozi, A., Mourikas, K., 1999. Bus dispatching at timed transfer transit stations using bus tracking technology. Transportation Research Part C 7 (4), 187–208.
Diana, M., Dessouky, M.M., 2004. A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. Transportation Research Part B 38 (6), 539–557.
Diana, M., Dessouky, M.M., Xia, N., 2006. A model for the fleet sizing of demand responsive transportation services with time windows. Transportation Research Part B 40 (8), 651–666.
Fielding, G.J., Babitsky, T.T., Brenner, M.E., 1985. Performance evaluation for bus transit. Transportation Research Part A 19 (1), 73–82.
Gaboune, B., Laporte, G., Soumis, F., 1993. Expected distances between two uniformly distributed random points in rectangles and rectangular parallelpipeds. Journal of the Operational Research Society 44 (5), 513–519.
Gleason, J.M., Barnum, D.T., 1982. Toward valid measures of public sector productivity: performance measures in urban transit. Management Science 28 (4), 379–386.
Horn, M.E.T., 2002. Fleet scheduling and dispatching for demand-responsive passenger services. Transportation Research Part C 10 (1), 35–63.
Ignall, E., Kolesar, P., 1974. Optimal dispatching of an infinite-capacity shuttle: control at a single terminal. Operations Research 22 (5), 1008–1024.
Jaillet, P., 1988. A priori solution of a traveling salesman problem in which a random subset of the customers are visited. Operations Research 36 (6), 929–936.
Jaw, J.J., Odoni, A.R., Psaraftis, H.N., Wilson, N.H.M., 1986. A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. Transportation Research Part B 20 (3), 243–257.
Khattak, A.J., Yim, Y., 2004. Traveler response to innovative personalized demand-responsive transit in the San Francisco Bay Area. Journal of Urban Planning and Development 130 (1), 42–55.
Koffman, D., 2004. Operational Experiences with Flexible Transit Services. Transportation Research Board, Washington, DC.
Langevin, A., Mbaraga, P., Campbell, J.F., 1996. Continuous approximation models in freight distribution: an overview. Transportation Research Part B 30 (3), 163–188.
Li, X., Quadrifoglio, L., 2010. Feeder transit services: choosing between fixed and demand responsive policy. Transportation Research Part C 18 (5), 770–780.
Melachrinoudis, E., Ilhan, A.B., Min, H., 2007. A dial-a-ride problem for client transportation in a health-care organization. Computers and Operations Research 34 (3), 742–759.
Özekici, S., 1987. Average waiting time in queues with scheduled batch services. Transportation Science 21 (1), 55–61.
Pagès, L., Jayakrishnan, R., Cortes, C.E., 2006. Real-time mass passenger transport network optimization problems. Transportation Research Record 1964, 229–237.
Potts, J.F., Marshall, M.A., Crockett, E.C., Washington, J., 2010. A Guide for Planning and Operating Flexible Public Transportation Services. Transportation Research Board, Washington, DC.
Psaraftis, H.N., 1980. A dynamic programming solution to the single vehicle many-to-many immediate request dial-a-ride problem. Transportation Science 14 (2), 130–154.
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.W., Dessouky, M.M., 2006. Performance and design of mobility allowance shuttle transit services: bounds on the maximum longitudinal velocity. Transportation Science 40 (3), 351–363.
Quadrifoglio, L., Dessouky, M.M., Palmer, K., 2007. An insertion heuristic for scheduling mobility allowance shuttle transit (MAST) services. Journal of Scheduling 10 (1), 25–40.
Quadrifoglio, L., Dessouky, M.M., Ordóñez, F., 2008. A simulation study of demand responsive transit system design. Transportation Research Part A 42 (4), 718–737.
Quadrifoglio, L., Chandra, S., Shen, C.W., 2009. Transit Services for Sprawling Areas with Relatively Low Demand Density: A Pilot Study in the Texas Border's Colonias. Texas Transportation Institute, Texas A&M University System.
Roos, D., 1971. Summary Report: The Dial-a-ride Transportation System. MIT Urban Systems Laboratory, Cambridge, Massachusetts, USL TR-70-10.
Ropke, S., Cordeau, J.F., Laporte, G., 2007. Models and branch-and-cut algorithms for pickup and delivery problems with time windows. Networks 49 (4), 258–272.
Santos, A., McGuckin, N., Nakamoto, H.Y., Gray, D., Liss, S., 2011. Summary of Travel Trends: 2009 National Household Travel Survey. Federal Highway Administration, Washington, DC.
Savelsbergh, M.W.P., Sol, M., 1995. The general pickup and delivery problem. Transportation Science 29 (1), 17–29.
SuperShuttle. <http://www.supershuttle.com/> (accessed on 07.01.12).
Szplett, D.B., 1984. Approximate procedures for planning public transit systems: a review and some examples. Journal of Advanced Transportation 18 (3), 245–257.
TriMet Paratransit Program. <http://trimet.org/lift/index.htm> (last accessed on 05.03.12).
Wardman, M., 2004. Public transport values of time. Transport Policy 11 (4), 363–377.
Zhao, F., Zeng, X., 2008. Optimization of transit route network, vehicle headways and timetables for large-scale transit networks. European Journal of Operational Research 186 (2), 841–855.