

Exploring the effects of mixed request schemes for demand-responsive feeder services

R. Kelly ^a, N. Ronald ^b, M. Wallace ^a and S. Winter ^b

^a*Faculty of Information Technology, Monash University, Caulfield 3145, Australia*

^b*Department of Infrastructure Engineering, Melbourne School of Engineering, University of Melbourne
3010, Australia*

Email: nicole.ronald@unimelb.edu.au

Abstract: Modern day communications mean travellers can request transport options, such as taxis and seats on demand-responsive vehicles, on the fly without having to plan ahead. However, last minute requests for transport could be inefficient for the operator. This paper explores the effects of mixed book-ahead/immediate request schemes in the context of a feeder service.

Demand responsive transportation (DRT) combines aspects of both buses and taxis: passengers travel together, but not necessarily to or from the same locations. A simple form of demand-responsive service is that of a feeder bus, permitting passengers to connect to a mass transit service such as a train.

A DRT service can run by getting passengers to book ahead or by allowing requests to arrive at the last minute, however the performance of the system may differ under different mixes of these requests. The ratio of the immediate (last minute) requests to total requests is known as the degree of dynamism. The impact of the degree of dynamism is measured by the total vehicle kilometres travelled (VKTs), the success rate (the ratio of requests that actually got served) and the waiting time for customers.

This paper simulates a DRT service using an event-based model where requests arrive during a simulated day for booking of both book-ahead and immediate requests. A booking system assigns the request to an appropriate time slot if possible and then an optimiser finds efficient routes to pickup and drop-off the customers at the required locations. Finally a dispatching system notifies and updates the driver of the vehicle with the current route.

To optimise the route, the customer demands are modelled as a series of vehicle routing problems that optimise the total distance travelled, with the current solution used as the starting point for the next iteration. A solution method using adaptive large neighbourhood search attempts to fit the new customer request into the existing routes while still allowing the feeder service to meet the scheduled train.

Three scenarios are tested using multiple vehicles. Instances with realistic demand and 250 customers are generated using realistic distributions derived from the Victorian Integrated Survey of Activities and Travel.

The first scenario varies the degree of dynamism across different instances. A second scenario varies the degree of dynamism with a spatial distribution of immediate requests reflecting distance to the train station. We also investigate the impact of train frequencies on the feeder service by altering the headway between trains. More frequent trains leaves less scope for optimisation and the longer headway has a higher success rate as the vehicle has time to pick everyone up. Overall we find that the most inefficient service occurs around 80% degree of dynamism.

Keywords: *Demand-responsive transportation, optimisation, degree of dynamism*

1 INTRODUCTION

Traditional public transportation services generally run to fixed timetables and routes, only being reviewed occasionally. This is efficient for moving large numbers of passengers around, for example, to the city centre in the morning peak for work-related activities. However, some public transport services, such as those laid on for travellers with restricted mobility, are offered for individuals rather than crowds, and at much shorter timescales of days or hours. Now, due to advances in communications, modern day transport can be made increasingly adaptable for all travellers, to the point of requesting rides immediately. However, not knowing demand in advance could lead to inefficiencies for operators, such as increased vehicle-kilometres and excessive detouring to collect last-minute passengers. Unknowns like these are a serious implementation impediment for large-scale (crowd-serving) demand-responsive public transport systems.

The objective of this paper is to explore the effects of not knowing demand in advance on a demand-responsive service, from both the operator's and passengers' points of view. This paper studies these dynamics under varying proportions of book-ahead and immediate ride requests, or varying degrees of dynamism in the demand. This forms part of a larger project using simulation to investigate the viability of demand-responsive transportation (DRT) services. DRT can be defined as a transportation service with properties similar to both a bus and a taxi: passengers travel together, but not necessarily to or from the same locations.

Of particular relevance to this paper is the study of a *partially* dynamic demand-responsive transport service. The requests are made partly in advance and partly dynamically, which reflects a more realistic use pattern of such services: some people plan and book ahead their travels, some have regular trips, and some only come up with ad-hoc requests. A partially dynamic demand-responsive service has also been studied before by Wong *et al.* (2014). Unlike the current paper, in Wong *et al.* (2014) demands have different origins *and destinations*, and these are randomly distributed over the environment, with no better knowledge of demand patterns available to them. Different bus-scheduling strategies, as well as the ratio of advanced bookings to dynamic requests, are investigated with respect to the impact on the quality and cost of the service. They found that a degree of dynamism of 70% provided the worst performance.

In contrast, our paper takes an applied perspective and focuses on the system effects of dynamism, considering spatial patterns within the population. Demand is generated based on household travel survey data from Victoria, Australia. Furthermore, in this paper demand is constrained to different origins but same destinations – the feeder bus principle – in order to minimise the effect of random assumptions and stay true to a likely real world application of a demand-responsive transport system. Differing headways of the mass transit service are also explored. Our hypothesis is there is a particular value of dynamism, for a feeder bus system, that provides the optimal performance for both the operator and passengers.

Section 2 provides more background on the dynamic Dial-A-Ride Problem (DARP) and related research. Section 3 describes the model, while Section 4 outlines the experiments, data and results. Section 5 presents conclusions and further work.

2 RELATED LITERATURE

Many forms of DRT exist in real world, which can be categorised by assessing the route type (fixed, deviation), scheduling type, vehicle type, and the nature of the origins and destinations (whether it is door to door or based at stops, and whether the service has a single origin or destination, or many different trip combinations) (Enoch *et al.*, 2004).

Feeder buses are a common form of demand-responsive service; according to the DRT categorisation they are a many-to-one service. They have been often used in optimisation and simulation studies, including algorithms for feeder buses servicing many stations (Lee and Savelsbergh, 2014) and determining the critical level of demand to switch between fixed and flexible services (Li and Quadrifoglio, 2010). Chandra and Quadrifoglio (2013) developed a street connectivity index to determine performance of flexible feeder services.

The booking scheme is one part of how customers interact with a DRT service. If the scheme requires booking in advance, then they need to plan ahead to ensure they are picked up. If it is immediate, then they need to consider their wait time, or the risk of not getting a ride.

Several demand-responsive services run in Victoria, Australia: Telebus (Public Transport Victoria, 2013b), Flexiride (Public Transport Victoria, 2013c), Gowanbrae (Public Transport Victoria, 2013a). In these cases, the bookings are mostly made in advance. There is some level of dynamism though: in the Flexiride scheme, only the pickup point is known, not the dropoff point. For Telebus and Gowanbrae, the requests must be

booked in advance for a limited number of stops, and therefore it is not too much of an imposition for the driver to work out the new route, even if passengers board at stops other than the terminus.

The optimisation of demand-responsive transportation schemes has led to many publications. The optimisation community use the term Dial-A-Ride Problem (DARP) to refer to this family of algorithms, which is a generalisation of the Pickup and Delivery Vehicle Routing Problem (PDVRP) and the Vehicle Routing Problem with Time Windows (VRPTW) (Cordeau and Laporte, 2007). The aim of the DARP is to minimise the cost of the routes required to service a number of requests with a fixed number of identical vehicles (Cordeau and Laporte, 2007).

Cordeau and Laporte (2007) published a detailed review covering approaches to DARP, including static and dynamic algorithms. The authors conclude that “After more than twenty years of research, it is fair to say that excellent heuristics exist for the static case...more emphasis should now be put on the dynamic version of the problem”. Berbeglia *et al.* (2010) reviewed dynamic pickup and delivery problems, with a section devoted to DARP approaches. Several different variations have been addressed: the taxi scheduling problem, the uncapacitated problem, the hospital patient problem, with request priorities, and an executive aircraft scheduling problem.

More recently Pillac *et al.* (2013) reviewed dynamic vehicle routing problems. They pointed out that “researchers have mainly focussed on the routing aspect of dynamic fleet management”. The requirement to deliver multiple passengers to a certain location at a certain time – to meet a train – requires a different emphasis. Indeed the frequency of the train service, has an important impact on the best solution approach, as will be evident from the current paper.

Describing the degree of dynamism, both Berbeglia *et al.* (2010) and Pillac *et al.* (2013) note that the degree of dynamism was presented by Lund *et al.* (1996), which is the “ratio between the number of dynamic requests and the total number of requests” (Berbeglia *et al.*, 2010). Pillac *et al.* (2013) also notes other sources of dynamism, such as the geographical distribution of requests.

Degree of dynamism has been used in experimentation. Wong *et al.* (2014) explored “a partially dynamic DRT problem and its operational efficiency under different levels of degree of dynamism, which can represent a form of information regarding future request arrivals.” They used random demand over a 20 x 20km area over 8 hours. Three strategies for the dispatching of vehicles were trialled: Drive First (DF), Wait First (WF) and Modified Dynamic Wait (MDW). These strategies are less important in approaches such as those used by this paper that allow full reordering of requests based on the position of the vehicle. They found that the system was most inefficient with a scenario of 70% dynamic requests.

3 METHODOLOGY

A number of feeder network scenarios are tested using a model that simulates a dispatching system and handles both advance and immediate requests. The model is characterised by three components - event simulation, vehicle routing and vehicle tracking.

3.1 Definitions

Requests: A request i consists of a location l_i and a request time r_i . The destination for all requests is the train station depot given by location 0. Advance requests also specify a desired train service and are known at the start of the planning horizon. Immediate requests are served as soon as possible. If a request cannot be served in the desired cycle or the next cycle it is rejected. Once accepted, requests are guaranteed to occur in the specified cycle and will not be moved to a different cycle or rejected.

Vehicles: The fleet for serving the requests are given by the set K . Vehicles are homogeneous and have a given capacity q . Vehicles routes start and end at the train station depot. Each vehicle has a start time a_k associated with it.

Cycles: A cycle is a tour for a given period that begins and ends at a train station (unlike taxis and fixed buses) and visits a set of request locations exactly once. It has a duration constraint that ensures that the tour finishes in time to meet the next departing train. All advance requests are known at the beginning of the cycle, while the tour may be adjusted to handle immediate requests while the vehicle is en route. Changes to a tour to include immediate requests only take place when the vehicle has arrived at its next location. The length of the cycle is the parameter L .

A travel time c_{ij} and distance d_{ij} are associated with every pair of locations i and j in the network. The set of

locations consists of the train station depot l_0 and the set of request locations $\{l_1, \dots, l_n\}$. The distance between each location is calculated using the Manhattan distance and the travel time is calculated by multiplying the distance and a constant velocity of 30km/h.

A pickup time p_i and dropoff time d_i are associated with each request i . The excess travel time E_i is the amount of time the passenger for request i spends on the DRT transport less the direct travel time between location i and the station (location 0) denoted as c_{i0} . It is defined as:

$$E_i = d_i - p_i - c_{i0}$$

The waiting time W_i is defined only for immediate customers and the amount of time the passenger for request i spends waiting for the transport to arrive. It is defined as:

$$W_i = p_i - r_i$$

Degree of dynamism: The degree of dynamism is the percentage of requests that are immediate and follows the definition of Lund et al Lund et al. (1996):

$$\text{degree of dynamism} = \frac{\# \text{ of immediate requests}}{\# \text{ of requests}}$$

3.2 Event Simulation

The aim of the event simulation component is to simulate the dispatching service for both advance and immediate requests. The event simulation first takes a set of requests from an instance and generates an appropriate set of events. These events are then placed on a priority queue which is sorted in ascending order of start time. The start time represents the time the event takes place (booking time for requests, start time for a pickup/delivery cycle). Three types of events are placed on the priority queue and are handled in the following fashion:

1. *Advance Requests*: An attempt to serve the advance request in the most appropriate cycle is made. If a quick insertion heuristic rules this out, an attempt to insert it into the next cycle is made. If these both fail, the request is rejected.
2. *Immediate Requests*: A quick insertion heuristic determines if the immediate request can be inserted into the current cycle. If it is possible, it is inserted using a complete reoptimisation across vehicles. If not, an attempt is made to add it to the next cycle. If this fails, the request is rejected.
3. *Cycle Starts*: The start of a cycle triggers a vehicle routing calculation resulting in a set of moves for the next cycle. After finding the route, the vehicle proceeds with the moves in order, unless advised otherwise.

3.3 Vehicle routing

The vehicle routes are solved as a series of static vehicle routing problems (VRPs) with the objective of minimising the total distance travelled. An additional constraint is added to the capacitated vehicle routing problem and is defined for each vehicle k as:

$$a_k + \sum_{i,j \in C} t_{ij}^k \leq L$$

This ensures that each vehicle route does not exceed the cycle length. The start time a_k for each vehicle is an offset from the start of the cycle and is either the current time if the vehicle has not departed or the time of arrival at the next customer. The feasibility of this constraint can be checked in constant time if the finish time of the route is stored and updated after each insertion.

For each cycle, a set of requests to serve during this period is stored. The initial optimisation process solves a VRP using the fleet, the set of requests and the train station depot as the start and end of the route. Each time a request is picked up it is removed from the set of requests and the location of the vehicle is updated.

When the dispatching system receives an immediate request it first attempts to determine if the request can be inserted into any of the current routes and still allow the vehicles to arrive at the train station in time. This is

tested using a quick greedy insertion heuristic. If it is feasible to insert the request, the route is again optimised as a VRP with the following changes. If the vehicle is moving, the change takes place after the vehicle has arrived at the next location. The starting location of the route is then set to the current location of the vehicle. A VRP is solved for the remaining requests in the current cycle.

The VRP is optimised using an adaptive large neighbourhood search (ALNS) heuristic as described in Pisinger and Ropke (2007). The ALNS algorithm first constructs an initial solution using a greedy approach and then attempts to improve the solution by applying a series of moves. Each move removes a relatively large neighbourhood of requests from the solution using a destroy operator and attempts to insert them at better positions using a repair operator. Simulated annealing is used to determine whether the modified solution is accepted or not.

3.4 Vehicle tracking

The vehicle tracking component updates the statistics of the customers as they are visited. After each customer i is picked up, the pickup time p_i is updated appropriately. When the vehicle drops all the customers off at the train station, the drop off times d_i are updated for all customers onboard the vehicle. The wait time and excess travel times are generated using these statistics. The distance travelled by the vehicle is updated as each move on the route is completed and the success rate is updated as the requests are processed in the simulation.

4 COMPUTATIONAL EXPERIMENTS AND RESULTS

This section reports the experimental results for each scenario. The experiments were run on an Intel i5 1.8GHz processor with 4GB of memory.

The vehicle routing component uses 100 iterations of LNS for improvement for each request that is inserted. This attempts to achieve a balance between CPU time and accuracy. A warm start method is used where the previous results of the route optimisation are used as the starting point for the next insertion.

Requests are inserted into either the requested cycle or the following cycle. The additional cycle is to allow for requests that occur late in the end of a cycle to be served. If the request cannot fit in either cycle it is rejected. A fleet of ten vehicles is used each with a capacity of 10 customers.

4.1 Generation of Demands

The sample input data was derived using data from the Victorian Integrated Survey of Activities and Travel (Department of Transport, Planning and Local Infrastructure, State Government of Victoria, 2014), a household travel survey carried out in Melbourne and various cities in Victoria in 2009/10.

Trips from homes to train stations between 5am and 10am were extracted from the data. The distance travelled to the station was divided into segments (0-1km, 1-3km, 3-5km, 5-10km, and 10-15km, and the time between leaving home and departing on a train (in 5 minute intervals) was measured. These were used to derive a probability matrix based on distance and lead time. The distribution is skewed towards closer locations as approximately two-thirds of customer requests are located within 3km of the station. A distribution of departure times from stations was also created, i.e., the demand was not uniform over the time period.

From the input data, different levels of demand can be generated. For all runs, we used 250 customers. Over 5 hours, this corresponds to 50 customers per cycle.

4.2 Computational Results

The experiments consist of three scenarios which test the degree of dynamism (dod) under a range of conditions. The results reported are the average over five instances of demand (that is, with randomly selected departure time and origin) and thirty iterations. Each scenario was tested with a cycle length of 60 minutes.

Varying degree of dynamism. The first scenario measures the impact of different degrees of dynamism. The degree of dynamism ranges from 0% to 100% in 10% increments and the impact is measured by comparing total distances travelled. The immediate requests are selected at random from the full set of requests with probability based only on the dod; no weight is given to the spatial locations of the customers.

All customers are picked up for all dod's when using multiple vehicles. The least efficient scenario is around 50% dod, with the highest VKTs. The system is most efficient when all customers are doing the same thing, either all immediate requests or all advance requests.

Varying degree of dynamism with a spatial component. The second scenario explores spatially distributed dynamism. In a real-world situation it is more likely that a customer who must travel far to reach the station will book in advance, while more local customers will book dynamically, knowing they can normally be fitted in at the last minute. We explore the impact of such a distribution of requests on the performance of the service. For each degree of dynamism between 0% and 100% with 10% steps, immediate requests are selected from customers closest to the station. The efficiency of the different dod is then compared using distances travelled. Using multiple vehicles (Figure 1), the least efficient scenario is around 80% dod with all customers picked up.

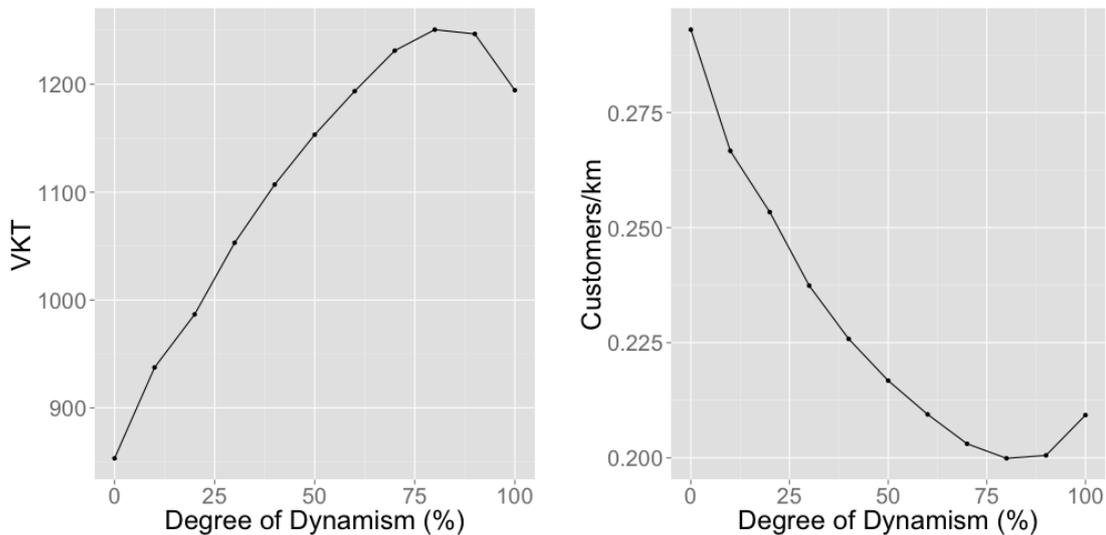


Figure 1. Impact of spatially distributed degree of dynamism on VKTs and customers/km for multiple vehicles.

Impact of train frequency. The final scenario investigates the impact of train frequency on the bus service cost and quality. The scenario uses three train frequencies at 15, 30 and 60 minutes and measures the percentage of customers that can be successfully picked up for differing levels of dod. More frequent trains could be matched by shorter bus service cycles, less customers per cycle and ultimately less scope for optimisation. On the other hand if trains are high frequency it may not be necessary to tie bus services to train departure times, since the bus travel is longer in comparison to the waiting times at the train station.

The success rate is reasonably constant for each cycle length, however it declines slightly with increased dod. The best performance occurs when all customers make their bookings in the same way.

4.3 Discussion

The results correspond with those presented by Wong *et al.* (2014) using random demand, who found that 70% dod was the most inefficient. With realistic demand in our paper the critical point occurs around 80%. At first glance, this seems somewhat puzzling, as for any given set of requests if we know them in advance, we can only improve the performance of the system. The inefficient behaviour seen at 80% dod is a result of different success rates as the degree of dynamism increases. Very high percentages of immediate requests result in increased numbers of ‘hard’ requests being rejected, thus decreasing the total overall distance travelled by the vehicles.

Comparing the first two scenarios (random and spatial distributions of dod), we see that the wait time for customers decreases significantly in the non-random scenario. The excess time is shorter for higher dods as vehicles are often making trips for one customer only, rather than forcing customers to travel together; this is also demonstrated by the increased VKTs.

R. Kelly *et al.*, Exploring the Effects of Mixed Request Schemes for Demand-Responsive Feeder Services.

In terms of implications for real-world systems, it is clear that a mixed set of immediate and book-ahead requests is inefficient. However, there could be an option to separate the area into zones designated as book-ahead-only or request-only to improve the efficiency of the system.

5 CONCLUSIONS AND FURTHER WORK

This paper has presented experimental results from a simulation of a demand-responsive feeder service. The aim was to understand the effect of booking ahead for DRT services grounded with real-world data. Using realistic demand for home to train trips from Victoria, Australia, the simulation experimented with different degrees of dynamism, vehicle numbers, train frequencies and the spatial distribution of requests.

The results show that the most inefficient option when optimising for VKTS was when requests were mixed, which is in line with previous research (Wong *et al.*, 2014).

The model has some limitations. The lack of a real-world network and the effects of other vehicles are not included; in high traffic areas this could affect the performance of the service. User choice is also not included in this version of the model. This could involve customers assessing whether the feeder bus provides a good service and choosing to use it or not, or adapting in other ways such as changing their departure or request time.

Further work includes experimenting with real-world street networks and configurations of vehicles. More complex systems could also be tested, for example, incorporating transfers between vehicles.

ACKNOWLEDGEMENTS

All authors have been supported by a grant from the Australian Research Council (LP120200130).

REFERENCES

- Berbeglia, G., J.-F. Cordeau, and G. Laporte (2010). Dynamic pickup and delivery problems. *European Journal of Operational Research* 202(1), 8–15.
- Chandra, S. and L. Quadrioglio (2013). A new street connectivity indicator to predict performance for feeder transit services. *Transportation Research Part C: Emerging Technologies* 30, 67–80.
- Cordeau, J.-F. and G. Laporte (2007). The dial-a-ride problem: models and algorithms. *Annals of Operations Research* 153(1), 29–46.
- Department of Transport, Planning and Local Infrastructure, State Government of Victoria (2014). Victorian Integrated Survey of Travel and Activity (VISTA).
- Enoch, M., S. Potter, G. Parkhurst, and M. Smith (2004). INTERMODE: Innovations in Demand Responsive Transport. Technical report, London, UK.
- Lee, A. and M. Savelsbergh (2014). An extended demand responsive connector. *EURO Journal on Transportation and Logistics*.
- Li, X. and L. Quadrioglio (2010). Feeder transit services: Choosing between fixed and demand responsive policy. *Transportation Research Part C: Emerging Technologies* 18(5), 770–780.
- Lund, K., O. B. G. Madsen, and J. M. Rygaard (1996). Vehicle routing problems with varying degrees of dynamism. Technical report, IMM Institute of Mathematical Modelling.
- Pillac, V., M. Gendreau, C. Guéret, and A. L. Medaglia (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research* 225(1), 1–11.
- Pisinger, D. and S. Ropke (2007). A general heuristic for vehicle routing problems. *Computers & Operations Research* 34(8), 2403–2435.
- Public Transport Victoria (2013a). 490 - Airport West to Gowanbrae via Melrose Dr, Gowanbrae Dr.
- Public Transport Victoria (2013b). Mooroolbark Station - Chirnside Park (Telebus Area 2).
- Public Transport Victoria (2013c). Yarrowonga and Mulwala FlexiRide Service.
- Wong, K. I., A. F. Han, and C. W. Yuen (2014). On dynamic demand responsive transport services with degree of dynamism. *Transportmetrica A: Transport Science* 10(1), 55–73.