

# Optimizing dial-a-ride services in Maryland: Benefits of computerized routing and scheduling <sup>1</sup>

Nikola Marković<sup>a</sup>, Rahul Nair<sup>b</sup>, Paul Schonfeld<sup>a</sup>, Elise Miller-Hooks<sup>a</sup>, Matthew Mohebbi<sup>c</sup>

<sup>a</sup>*Department of Civil and Environmental Engineering, University of Maryland, College Park, USA*

<sup>b</sup>*IBM Research, Dublin, Ireland*

<sup>c</sup>*IT Curves, Gaithersburg, USA*

---

## Abstract

This paper reports on a system developed to address the dial-a-ride problem and an implementation for Maryland where real-world practical constraints are considered in providing customized vehicle routing and scheduling for about 450 trip requests daily. The system, called Mobile Resource Management System (MRMS), allows for dispatch operators to quickly study different operational scenarios and, in a strategic setting, explore tradeoffs between level-of-service and various system characteristics, including fleet composition, fleet size and dispatch rules. Such insights play a key role in making long-term investment decisions or estimating cost of servicing contracts that have service level agreements. Test comparison of manual and MRMS-based routes indicated an estimated annual operational expense reduction of \$0.82 million, or about 18% of the total annual expense. In addition to the cost benefits, improved quality of service and the reduction in total vehicle-kilometers traveled leading to environmental benefits are demonstrated.

*Keywords:* dial-a-ride, fleet management, scheduling, software, applications

---

## 1. Introduction

Better accessibility to public transportation has become an important objective for many transit systems across the world. This objective has been partially achieved by introducing low-floor buses complemented with “kneeling devices” and specialized ramps, installing elevators in rail stations, and introducing higher platforms. Never-

---

<sup>1</sup>An earlier version of this paper (Marković et al., 2014) was presented at the Transportation Research Board meeting in 2014.

theless, many handicapped and elderly people still find it difficult to use public transit despite these recent enhancements. Some handicapped or elderly people need additional help, while others may find the closest stop to be too far away, or the wait too long (Borndörfer et al., 1999). Transit systems across the world offer these people a special demand-responsive door-to-door transportation service, which is often called dial-a-ride (DAR) service. The major difference between DAR and taxi service is that DAR allows ridesharing; thus, multiple unrelated passengers, with different origins and destinations, may be served simultaneously with the same vehicles.

The average number of paratransit trips provided by a transit agency has increased by 7% from 2007 to 2010 (GAO, 2012), and such trend will likely continue because of the aging population. Due to increasing ridership, the DAR services have been among the fastest growing budget fractions of many transit agencies. These trends pressed DAR service providers to improve the efficiency of their operations and thereby limit some of the budget increases. To better control costs and manage their growing operations more efficiently, DAR agencies have started to implement communication and computer technology. These technologies also serve as enablers for optimization algorithms to generate better routes and schedules for drivers and vehicles.

This paper addresses the vehicle routing and scheduling problem in DAR operations and makes the following contributions. First, key practical considerations in real-world dispatching operations, such as “locked-blocks” and the heterogeneous nature of demand and supply, are identified and modeled. A “locked-block” refers to a series of passenger demands that do not change from one day to the next, and as such, must be served in the same order every day. Heterogeneity of supply manifests in two ways. Trip requests that have associated with it particularly egregious costs due to spatial or time criteria can be outsourced to a taxi. Additionally, the operator may have a fleet of vehicles with varying capacity. This capacity should be allocated based on the heterogeneity of demand for particular routes. For example, such demand variation is likely to occur when there are wheelchair-bound passengers who require additional space onboard. Such practical constraints have been considered only in a limited manner (for example Madsen et al. (1995) considers demand and fleet heterogeneity).

The second contribution of the work is in deploying the proposed algorithm to a suite of test instances, and for a real-world deployment in Maryland. The paper examines the benefits of improved efficiency and results from deployment in the field, including the improvement in quality of service. Most of all, the results reported in

this paper indicate that considerable savings can be achieved in providing DAR services through implementation of custom-designed vehicle routing and scheduling algorithms. These results show a potential for reducing the governmental expenditures provided proper investments in computer-aided decision making in DAR operations.

The remainder of the paper is organized in five sections. Section 2 describes the overall system architecture and its components. In Section 3 we present the model, solution algorithm, and its application on benchmark problems. Section 4 describes a real-world deployment of MRMS in Maryland, and emphasizes benefits of computerized routing and scheduling. In Section 5 we discuss components of the MRMS used to dynamically optimize DAR systems. Finally, Section 6 draws conclusions and recommends further implementations.

## **2. The MRMS system**

A system to manage DAR services, from trip request management to dispatch and operations was implemented. At the core of the system is the algorithm for efficient scheduling and dispatching of vehicles satisfying various operational and level-of-service (LOS) constraints. The MRMS was designed to provide operators with many parameters that could be set for different runs allowing the operators to experiment and find the optimal tradeoff between LOS and operating and capital expenses. Current wide availability of wireless data, tablets and wireless devices, built-in navigation tools, and powerful back office computers enable the role of optimization algorithms to improve operations.

The MRMS development and implementation included three distinct parts:

1. In-vehicle navigation and communication equipment that work over standard 3G or 4G wireless data services. This equipment allows monitoring and controlling the location and status of the vehicles at all times. It also allows dynamic modifications of driver task lists, which include the sequence of pick-up and drop-off locations and time windows.
2. Cloud-based monitoring, control and rule-based decision-making that provides the intelligence for routing and performance optimization based on provided LOS and system constraints.
3. A presentation layer that enables operators to view in real-time the performance of the operation. It also provides the operators the ability to set up performance

parameters, and modify the constraints and optimization rules.

The key element of the MRMS is an algorithm used to efficiently solve the vehicle routing and scheduling problem in DAR operations. Since a vast majority of trips are scheduled between one and seven days in advance, the main algorithm was based on the static DAR problem. For large-scale instances exact solutions to DAR problem are computationally prohibitive. So, a specialized solution heuristic was developed, incorporating the practical considerations such as, outsourcing of outlier requests, locked-blocks, and heterogeneity in fleet and requests. Elements of the aforementioned algorithm were also used to tackle requests arising in the real-time, and to continuously reoptimize the DAR system throughout its operations.

### 3. Static dial-a-ride problem

When the DAR requests are made long in advance, which is the case with 90% of requests, the problem of assigning multiple trip requests to be served concurrently on multiple vehicles is called a static DAR Problem (DARP) (Cordeau and Laporte, 2007). The static DARP is usually formulated as a generalized vehicle routing problem with pick-up and delivery (Cordeau, 2006). Since DARP involves transportation of people, tighter time windows and an additional maximum ride time constraint are considered (Luo and Schonfeld, 2007). The latter ensures that a passenger does not spend too much time onboard while other passengers are picked up and dropped off. The most commonly defined objective is to design  $m$  least cost vehicle routes to accommodate  $n$  requests under a set of constraints (Cordeau and Laporte, 2007; Cordeau, 2006; Cordeau and Laporte, 2003*a,b*). The mathematical formulations assume that  $m$  is a given integer. The problem is repeatedly solved by imputing different values for  $m$  to find the minimum number of vehicles that can accommodate all the requests (Cordeau and Laporte, 2003*b*). For a more detailed literature survey, readers are referred to (Berbeglia et al., 2007).

In the following section, we provide the mathematical formulation of the static DARP that explains the practical decisions and issues arising in real-world DAR operations. We formulate it as a fleet size and mix vehicle routing problem (Golden et al., 1984) with pick-up and delivery and trip outsourcing. This formulation includes features like fleet size, trip outsourcing to taxis, and no idling of loaded vehicles that are very important in practical applications and are commonly considered in heuristic solution methods (Luo and Schonfeld, 2007; Jaw et al., 1986). Despite the practical

importance of the aforementioned features, the mathematical formulations found in the literature typically assume a fixed fleet (i.e., treated  $m$  as a constant), and few consider trip outsourcing (Heilporn et al., 2011) or idling for loaded vehicles (Parragh, 2011). The proposed formulation is built upon Cordeau (2006) and follows most of its notation.

### 3.1. Formulation

Let  $n$  denote the number of users (or requests) to be served. The DARP may be defined on a complete directed graph  $G = (N, A)$ , where  $N = P \cup D \cup \{0, 2n + 1\}$ ,  $P = \{1, \dots, n\}$ , and  $D = \{n + 1, \dots, 2n\}$ . Subsets  $P$  and  $D$  contain pick-up and drop-off nodes, respectively, while nodes 0 and  $2n + 1$  represent the depot. With each user  $i$  are thus associated an origin node  $i$  and a destination node  $n + i$ .

Let  $K$  be the set of vehicles. Each vehicle  $k \in K$  has a capacity  $Q_k$ , and the total duration of its route cannot exceed  $T_k$ . With each node  $i \in N$  are associated a load  $q_i$  and a nonnegative service duration  $d_i$  such that  $q_0 = q_{2n+1} = 0$ ,  $q_i = -q_{n+i}$  ( $i = 1, \dots, n$ ), and  $d_0 = d_{2n+1} = 0$ . A time window  $[e_i, l_i]$  is also associated with node  $i \in N$ , where  $e_i$  and  $l_i$  represent the earliest and latest time, respectively, at which service may begin at node  $i$ . With each arc  $(i, j) \in A$  and vehicle  $k \in K$  are associated a routing cost  $c_{ij}^k$  and a travel time  $t_{ij}$  which is assumed to be identical for all the vehicles. Let  $f_k$  denote the fixed cost for using vehicle  $k \in K$ , and let  $g$  and  $v$  denote the fixed and variable taxi cost, respectively. Finally, let  $r$  denote the maximum ride time coefficient such that  $r \cdot t_{i,n+i}$  represents the maximum ride time for user  $i \in P$ .

For each arc  $(i, j) \in A$  and each vehicle  $k \in K$ , let  $x_{ij}^k = 1$  if vehicle  $k$  travels from node  $i$  to node  $j$ . For each node  $i \in N$  and each vehicle  $k \in K$ , let  $B_i^k$  be the time at which vehicle  $k$  begins service at node  $i$ , and  $Q_i^k$  be the load of vehicle  $k$  after visiting node  $i$ . Let  $z_i^k = 1$  if vehicle  $k$  is non-empty after visiting node  $i$ . For each user  $i$ ,  $L_i^k$  be the ride time of user  $i$  on vehicle  $k$ , and  $y_i = 1$  if the customer is outsourced to a taxi.

The static DARP arising in the real-world operations can be formulated as the following mixed integer program:

$$\min \sum_{k \in K} f_k \sum_{j \in N} x_{0,j}^k + \sum_{k \in K} \sum_{i \in N} \sum_{j \in N} c_{ij}^k x_{ij}^k + \sum_i (g + v \cdot t_{i,i+n}) y_i \quad (1)$$

$$\sum_{k \in K} \sum_{j \in N} x_{ij}^k = 1 - y_i, \quad \forall i \in P \quad (2)$$

$$\sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{n+i,j}^k = 0, \quad \forall i \in P, k \in K \quad (3)$$

$$\sum_{j \in N} x_{0j}^k \leq 1, \quad \forall k \in K \quad (4)$$

$$\sum_{j \in N} x_{ji}^k - \sum_{j \in N} x_{ij}^k = 0, \quad \forall i \in P \cup D, k \in K \quad (5)$$

$$\sum_{i \in N} x_{i,2n+1}^k \leq 1, \quad \forall k \in K, \quad (6)$$

$$B_j^k \geq B_i^k + d_i + t_{ij} - M(1 - x_{ij}^k), \quad \forall i \in N, j \in N, k \in K \quad (7)$$

$$Q_i^k \leq M \cdot z_i^k, \quad \forall i \in N, k \in K \quad (8)$$

$$Q_i^k > M(z_i^k - 1), \quad \forall i \in N, k \in K \quad (9)$$

$$B_j^k \leq B_i^k + d_i + t_{ij} + M(2 - x_{ij}^k - z_i^k), \quad \forall i \in N, j \in N, k \in K \quad (10)$$

$$Q_j^k \geq Q_i^k + q_j - M(1 - x_{ij}^k), \quad \forall i \in N, j \in N, k \in K \quad (11)$$

$$L_i^k = B_{n+i}^k - (B_i^k + d_i), \quad \forall i \in N, j \in N, k \in K \quad (12)$$

$$B_{2n+1}^k - B_0^k \leq T_k, \quad \forall k \in K \quad (13)$$

$$e_i \leq B_i^k \leq l_i, \quad \forall i \in N, k \in K \quad (14)$$

$$t_{i,n+i} \leq L_i^k \leq r \cdot t_{i,n+i}, \quad \forall i \in P, k \in K \quad (15)$$

$$Q_i^k \leq Q_k, \quad \forall i \in N, k \in K \quad (16)$$

$$x_{ij}^k, z_i^k, y_i \in \{0, 1\}; B_i^k, L_i^k, s_i^k \in R^+; Q_i^k \in Z^+, \quad \forall i \in N, j \in N, k \in K \quad (17)$$

The objective function (1) minimizes fixed and variable vehicle routing and trip outsourcing costs. Constraints (2) ensure that each request is served exactly once or is outsourced to a taxi. Constraints (3) ensure that the origin and destination nodes are visited by the same vehicle. Constraints (4)-(6) guarantee that the route of each used vehicle starts and ends at the depot. Constraints (7) ensure consistency of the time windows, while (8)-(10) prevent a loaded vehicle from idling. Consistency of the load variable is ensured by constraints (11). Equalities (12) define the ride time of each user, which is bounded by constraints (15). The latter also act as precedence constraints, because the non-negativity of the  $L_i^k$  variables ensures that node  $i$  will be visited before node  $n+i$  for every user  $i$ . Inequalities (13) bound the duration of each route. Finally, constraints (14) and (16) impose time windows and capacity limits, respectively.

This formulation builds on the math program in Cordeau (2006), but differs in that it includes outsourcing of trips ((1) and (2)), variable fleet size ((1) and (4)), and

Table 1: Overview of solution methods for the DARP

Group	Specific Solution Methods	Pros	Cons
Exact	<ol style="list-style-type: none"> <li>1. Dynamic programming</li> <li>2. Mathematical optimization using: Branch &amp; Cut, Branch &amp; Price or integer L-shaped method</li> </ol>	Optimality is guaranteed	Cannot solve real-world problem instances (max about 50 requests)
Heuristic	<ol style="list-style-type: none"> <li>1. Cluster first, chain second</li> <li>2. Insertion heuristic</li> </ol>	Quickly builds a good feasible solution	Optimality is NOT guaranteed
Metaheuristic	<ol style="list-style-type: none"> <li>1. Tabu Search</li> <li>2. Simulated annealing</li> <li>3. Genetic algorithms</li> <li>4. Variable neighborhood search</li> <li>5. Ant colony optimization</li> </ol>	Can find a near-optimal solutions if run long enough	Optimality is NOT guaranteed; typically takes much longer than heuristic

precludes idling of loaded vehicles ((7)-(11)). The locked blocks can be considered via preprocessing. Namely, an optimal itinerary for a group of passengers traveling together can be found separately, and inputed in the above formulation as one long request with a load equivalent to capacity of the vehicle.

### 3.2. Solution method implemented in MRMS

The optimization model presented above can be solved optimally with branch-and-bound-based algorithms, but only for relatively small instances, e.g.  $n < 50$  (Cordeau, 2006). For somewhat larger instances of the static DARP, several metaheuristics proved successful in finding very good or near-optimal solutions (Cordeau and Laporte, 2003b; Parragh et al., 2010). The downside of these latter techniques is that they often require significant time to tackle practically small instances, e.g. over an hour with 144 requests (Parragh et al., 2010). Since run times have been noted to increase nonlinearly with problem size, a practical solution technique for instances with several hundred or thousand requests is a greedy insertion heuristic. Such a heuristic can construct good feasible solutions to very large problems in reasonable time. In Table 1 we summarize several of these potential solution techniques as well as their pros and cons.

Due to its efficiency relative to other methods, an insertion-based heuristic (Algorithm 1) was implemented in the MRMS. In particular, the parallel insertion heuristic (Jaw et al., 1986) with improvement operators (Toth and Vigo, 1996; Luo and Schonfeld, 2007) was implemented and modified to incorporate several operational requirements specified by potential users. These requests included working with blocks of trips that may or may not be modified, considering different passenger needs, and different

specifications of the objective function. The insertion algorithm outlined next differs from other insertion heuristics mentioned previously as follows.

1. Several changes in trip insertion are made to account for groups of trips that are often referred to as “blocks” and “locked-blocks” in the DAR industry. The MRMS allows a user to specify a sequence of trips that may or may not be changed with subsequent insertions. Blocks and locked-blocks are often formed for repetitive trips that are requested from day-to-day or week-to-week. This is also the case for groups of people with common origins or destinations. Locked-blocks refer to groups of trips that are not only required to travel together, but also for contractual reasons, operators are not allowed to add additional trips to these groups.
2. Early morning, late afternoon, or long trips are filtered and exempt from the insertion. The operator is notified of these trips and may decide whether to outsource them to taxis. Outsourcing requests to taxis is common practice in DAR operations. It occurs when dispatchers assess that some trips have particularly egregious costs due to spatial or time criteria demand and, thus, would be cheaper to outsource to a third party rather than force into ridesharing tours.

Similar to Jaw et al. (1986) and Madsen et al. (1995) the developed method considers heterogeneity of demand and supply. Different weights are be associated with passengers depending on the required capacity (e.g., passengers in wheelchairs need more space than others). The vehicles in the depot are heterogeneous in their capacity and number of wheelchair passengers they can carry, and each route has a pre-specified start time.

The above modifications enabled an algorithm with academic origins to be implemented in the field, allowing its application in the daily operations of several DAR companies. Others, including Madsen et al. (1995) and Luo (2006), also extended Jaw’s parallel insertion algorithm to account for practical aspects arising in the applications they considered.

### *3.3. Testing the solution method on benchmark problems*

To demonstrate the validity of the implemented solution, the algorithm was run on several benchmark instances where solutions from exact methods were known. The benchmark problems including 24 instances were taken from Cordeau (2006). He used



---

**Algorithm 1** Outline of the algorithm implemented in MRMS

---

**Initialize:** Sort requests according to desired pick-up time and introduce the first vehicle.

**Step 0:** Notify operator of early, late, and long trips.

**Step 1:** Consider the next unassigned request.

**Step 2:** For each introduced vehicle:  
Generate all feasible insertions of the request into the schedule and compute the change in the objective function

**Step 3:** If there exists a feasible insertion of the request, then the insertion with the minimum increase in the objective function is selected and the request is inserted while updating the schedule. If a feasible insertion does not exist, a new vehicle is introduced and the request is assigned to it.

**Step 4:** If there are unassigned requests, then go to step 1, otherwise go to step 5.

**Step 5:** Apply the following two operators through desired number of iterations:

*Trip reinsertion operator:* remove trip  $i$  from its current route and try reinserting it into all the vehicle routes. Make the reinsertion with the minimum associated cost (the final route may be the same as the current one);

*Trip exchange operator:* remove trip  $i$  from its route  $r$  and remove trip  $j$  from its route  $s$ , ( $r \neq s$ ); insert the two stops of trip  $i$  in the best positions of route  $s$  and insert the two stops of trip  $j$  in the best positions of route  $r$ . If the exchange yields a cost reduction, make the swap.

---

a branch-and-cut algorithm to find  $m$  least cost vehicle routes to accommodate  $n$  requests, assuming that  $m$  is given. The benchmark problems include up to 48 trips. The largest instance solved optimally has 32 requests. Table 2 compares Cordeau’s solutions and lower bounds with solutions obtained by the implemented heuristic method. Note that the implemented heuristic ensures higher LOS since it prevents loaded vehicles from idling. Here, we do not consider outsourcing of requests to taxi in order to make the heuristic comparable with Cordeau’s setting. The heuristic was implemented in Matlab 2010a on a PC with an AMD Athlon 3300 GHz processor and the computation times are compared with those reported in Cordeau (2006).

The implemented heuristic provides competitive results within no more than 35 seconds computation time. In 75% of the cases, the implemented heuristic provided results with the same or fewer vehicles. The driving time was within 25 to 55% of the optimal solution or bound reported in Cordeau 2006. Moreover, computation times needed to solve optimally the instances in Table 2 justify the use of the heuristic. For example, instance b4-32 includes 32 requests and is solved optimally in about 3 hours with the branch-and-cut algorithm. Thus, solution of the problem instances with several hundred requests that face midsize or large DAR companies would be computationally intractable. Even with recent developments in computer technology, finding exact solutions would be formidable and heuristics or metaheuristics must be applied.

#### 4. Regency Taxi case study

Regency Taxi (RT) is a taxi company and a DAR service provider located in Maryland, US. Currently, RT has 40 employees and 150 drivers who are fully trained in handling transportation needs of disabled and senior citizens. The company’s fleet consists of 150 vehicles, some of which are equipped with ramps or lifts that facilitate access for disabled passengers traveling in wheelchairs. Moreover, the vehicles are equipped with 3G or 4G wireless smart devices, i.e., tablets, which allow regular communication with the dispatch center.

In RT’s operations, about 90% of DAR requests are made in advance. A passenger schedules a trip by specifying pick-up and delivery locations, as well as the desired pick-up or drop-off time. After the trip requests are gathered, the dispatchers use the afternoon to design routes and schedules for the following day. A route manifest is electronically sent to drivers who can start their assignments on the following morning

Table 2: Testing the implemented heuristic on the benchmark problems

Instance	Input				Heuristic			Cordeau 2006			Heuristic vs. Cordeau 2006	
	n	T	Q	L	Veh	Driving (hr)	CPU (sec)	Veh	Driving (hr)	CPU (sec)	$\Delta$ Veh	Gap
a2-16	16	480	3	30	2	7.33	3	2	*4.90	1	0	0.33
a2-20	20	600	3	30	3	8.67	4	2	*5.75	5	+1	0.34
a2-24	24	720	3	30	2	10.24	8	2	*7.19	16	0	0.30
a3-18	18	360	3	30	3	7.60	3	3	*5.01	12	0	0.34
a3-24	24	480	3	30	3	8.86	6	3	*5.75	250	0	0.35
a3-30	30	600	3	30	4	12.37	9	3	*8.25	2543	+1	0.33
a3-36	36	720	3	30	4	15.09	15	3	*9.72	748	+1	0.36
a4-16	16	240	3	30	3	8.29	2	4	*4.71	152	-1	0.43
a4-24	24	360	3	30	3	10.48	5	4	*6.25	1511	-1	0.40
a4-32	32	480	3	30	4	12.61	10	4	7.46	14400	0	0.41
a4-40	40	600	3	30	4	16.07	20	4	7.92	14400	0	0.51
a4-48	48	720	3	30	4	18.05	35	4	8.10	14400	0	0.55
b2-16	16	480	6	45	2	6.94	2	2	*5.16	9	0	0.26
b2-20	20	600	6	45	3	7.35	4	2	*5.54	1	+1	0.25
b2-24	24	720	6	45	3	10.27	6	2	*7.41	8	+1	0.28
b3-18	18	360	6	45	4	7.31	2	3	*5.03	42	+1	0.31
b3-24	24	480	6	45	4	9.29	5	3	*6.58	217	+1	0.29
b3-30	30	600	6	45	3	12.54	12	3	*8.86	409	0	0.29
b3-36	36	720	6	45	3	14.30	19	3	*10.06	3724	0	0.30
b4-16	16	240	6	45	4	6.82	2	4	*4.95	47	0	0.27
b4-24	24	360	6	45	4	8.87	5	4	*6.19	351	0	0.30
b4-32	32	480	6	45	4	12.92	13	4	*8.25	10609	0	0.36
b4-40	40	600	6	45	4	15.57	20	4	9.86	14400	0	0.37
b4-48	48	720	6	45	5	17.58	32	4	9.78	14400	+1	0.44

Note: n, T, Q, L stand for the number of requests, route duration, capacity, and maximum ride time for each user, respectively. Values marked with ‘\*’ represent optimal total driving time for m vehicles. Other values for driving time represent lower bounds. The last column shows the percentage gap in driving time between the solution obtained through heuristic and the lower bound.

with the schedules available on their tablets. The routes are given in the form of a sequence of pick-up and drop-off locations with requested arrival times, and drivers are asked to serve the routes in exactly the same sequence that is provided to them. Another 10% of requests are made for the same day and those trips are quickly inserted into previously built routes while considering the GPS positions of the vehicles. When dispatchers determine that a request is an “outlier,” i.e., does not fit well with the rest of the demand, they dispatch these trips as a single taxi trip.

As RT’s DAR operations have grown over the past several years, its management determined that manual dispatch not only produced inefficient routes but also became impractical. Therefore, they decided to computerize the dispatching process. The management envisioned that computerized routing will considerably reduce dispatching effort and provide an even higher LOS to their passengers than through manual routing. LOS is an essential part of any DAR contract. Considering its daily routine, the company sought a solution method capable of finding good feasible solutions to problems of forming routes/manifests within less than an hour. The management studied the specifications of several commercially available software solutions and decided that the MRMS would be the best fit for their operation. In the following section we evaluate the benefits of implementing the MRMS in the described operations.

#### *4.1. Benefits of implementing the MRMS: Operational level*

The immediate benefits of implementing the MRMS system were observed in reducing dispatching, transportation, and external costs. RT receives about 450 requests on a daily basis. The MRMS was able to route this number of trips within 7 minutes and thereby save the dispatch staff about 12 dispatcher-hr/day that were previously needed to build the manifests manually. Moreover, the transportation cost was reduced by decreasing the number of vehicles and drivers needed to serve all requests. The MRMS-based routes were compared with the manually built route manifest for one day of RT’s operations (Table 3). Test comparisons showed that the number of routes and vehicle-km traveled decreased by 12.8% and 22.4%, respectively. Moreover, the number of routes and vehicle-km traveled decreased despite the imposition of tighter LOS constraints than those observed in manually built routes. Table 3 indicates that the MRMS-based routes yield estimated total annual savings of approximately \$0.82 million. This represents roughly 18% of total operating cost. Finally, negative external costs, such as emissions and traffic congestion, were similarly substantially reduced as a result of significant anticipated reductions in the number of vehicle-hours driven.

Table 3: Operational-level savings achieved through implementation of MRMS

Operational Level Costs	Manual Routing Scheduling	MRMS Routing Scheduling	Daily Savings	Annual Savings
Dispatching cost	12 dispatcher-hr/day	7 computer-min/day	\$300	\$105,000
Fixed cost	39 drivers & vehicles	34 drivers & vehicles	\$950	\$332,500
Variable cost	17,900 vehicle-km	13,890 vehicle-km	\$1,083	\$379,050
<b>total savings:</b>				\$816,550

Note: The savings are computed by the authors of this manuscript based on field data and assuming: \$25/dispatcher-hr, \$170/driver-day, \$20/vehicle-day (vehicle depreciation), \$0.27/vehicle-km (variable cost), and 350 working days/year. LOS and operator’s constraints for computer-based routing correspond to those observed in manual routes.

In addition to evident cost reductions shown in Table 3, the MRMS allowed computation of various performance measures and visualization of schedules and routes. These features were especially important for gaining trust in the methodologies and the ultimate adoption of the tool and its operations research-based techniques.

In manual routing and scheduling it is difficult to keep track of performance measures, because they require tedious computations. In computer-based routing, on the other hand, a variety of statistics can be readily generated. These statistics provide increased insight into overall system performance, which is of particular interest to system managers. Moreover, statistics allow better control of vehicle utilization and, thus, discourage drivers from using the vehicles for their own needs. Some of the performance measures that the management found useful are shown in Table 4. Additionally, the tool’s user-friendly interface (Figure 1) produced recognizable savings in time and effort for the dispatchers and drivers, thus also aiding the tool’s acceptance.

*4.2. Benefits of implementing the MRMS: Tactical level*

Besides cost reductions achieved at the operational level, the MRMS provided its user with a valuable tool that can be used for decision making at the tactical level. The MRMS allows operators to quickly study different operational scenarios and explore tradeoffs between LOS and various system characteristics. Such insights play a key role in making long-term investment decisions in the company’s fleet, or when signing a contract which promises a certain LOS. Some of the insights that operators using MRMS are finding useful are provided next.

One of the factors influencing LOS is the time window promised to customers. Intuitively, wider time windows provide more flexibility to the operator, but lower LOS for

Table 4: Performance measures provide management with many insights and better control over vehicle usage

Trips	452	Rte	Pass	PasKm	VehHr	Utilz	PaVHr	PKmVHr
Vehicles	34	1	18	703.66	11.04	0.905	1.63	63.73
Total Passenger-Km	16740	2	17	538.51	11.72	0.772	1.45	45.95
Total Vehicle-Km	13890	3	19	573.00	11.83	1.000	1.61	48.42
Total Driving Hours	347.25	4	12	462.34	11.27	0.764	1.07	41.04
Total Vehicle Hours	407.59	5	11	386.89	13.21	0.696	0.83	29.29
Average Utilization	0.85	6	10	402.95	12.05	0.748	0.83	33.43
Average Vehicle Out-Depot Hours	11.96	7	14	503.91	13.77	0.823	1.02	36.60
Average Pass on Vehicle	1.39	...	...	...	...	...	...	...
Average of Pass on Loaded Vehicle	1.87	31	2	130.21	3.72	1.000	0.54	34.96
Min Deviation from DPT (min)	30	32	3	136.76	3.18	1.000	0.94	42.96
Average Deviation from DPT (min)	18.01	33	3	107.47	3.27	1.000	0.92	32.82
Max Deviation from DPT (min)	30	34	2	111.83	3.27	1.000	0.61	34.16

Note: DPT = desired pick-up time, PaVHr = pass/veh-hr, and PKmVhr = pass-km/veh-hr.

The screenshot shows the 'Service Request' form in the MRMS system. The form is divided into several sections:

- Client Info:** Includes fields for Funding S., PU Person, Address (City and State), Passengers (set to 1), and Wheel Chairs (set to 0).
- Pick up Information:** Includes POI Cat and POI fields.
- Drop Information:** Includes POI Cat and POI fields.
- Trip Estimates:** A table with columns for Go, Return, and Total, and rows for Distance, Time, and Cost, all showing 0.
- Call Out Info:** Includes Call Out Type (Call Phone), Call Out Info (301-301-5555), Payment Type (Corporate-V), Trip (1248321), Job (N/A), and Authorized By (MIKE).
- Additional Information:** A text area containing 'FLIGHT INFO: NW456 / PERSON NEEDS SIDE WHEELCHAIR VAN'.

Buttons for 'Save Request' and 'End Call' are visible at the bottom right. A cartoon character is also present in the bottom right corner of the interface.

Figure 1: Compact menu for entering new requests by specifying origin and destination locations, number of passengers traveling, and whether the passengers are traveling in wheel chairs

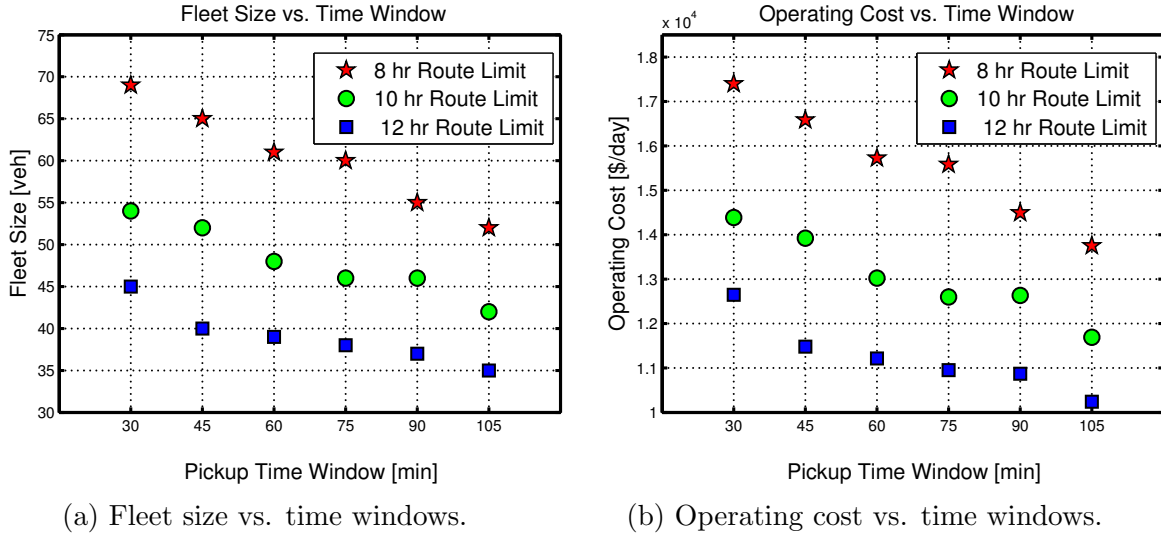


Figure 2: Fewer vehicles are needed as the time windows and route duration limits are relaxed. Consequently, the operating costs are decreased. Note: The assumed costs are as in Table 3

the customers. The MRMS allowed its users to quantify the cost of providing different time windows and explore these tradeoffs. The required fleet size ranged between 45 and 35 vehicles as the time windows were relaxed from 30 to 105 min, respectively, while keeping a fixed 12 hour limit on route duration (pentagrams in Figure 2a). This insight is especially important in signing contracts with transit agencies that typically specify minimum LOS to passengers in terms of time windows and maximum system response time. The corresponding operating daily cost may be reduced from roughly \$12,650 to \$10,200 as time windows are relaxed from 30 to 105 minutes (pentagrams in Figure 2b).

Operators using MRMS were also able to quantify the change in fleet needed to meet its demand when the route duration limit was relaxed from 8 to 10 and 12 hours (Figure 2a). This insight is useful in negotiating general agreement contracts that may impose a limit on driver working hours. It also allows operators to better estimate the value of driver overtime hours by assessing the corresponding operating cost (Figure 2b).

A factor that considerably influences operator costs is the distribution of demand over time. It is theoretically expected that steadier demand yields better utilization of transportation assets (Fitzsimmons and Fitzsimmons, 2006). The MRMS allows its users to quantify effects of potential changes in demand distribution. Figure 3 shows the original and two slightly modified demand distributions (Original, Modified 1, and

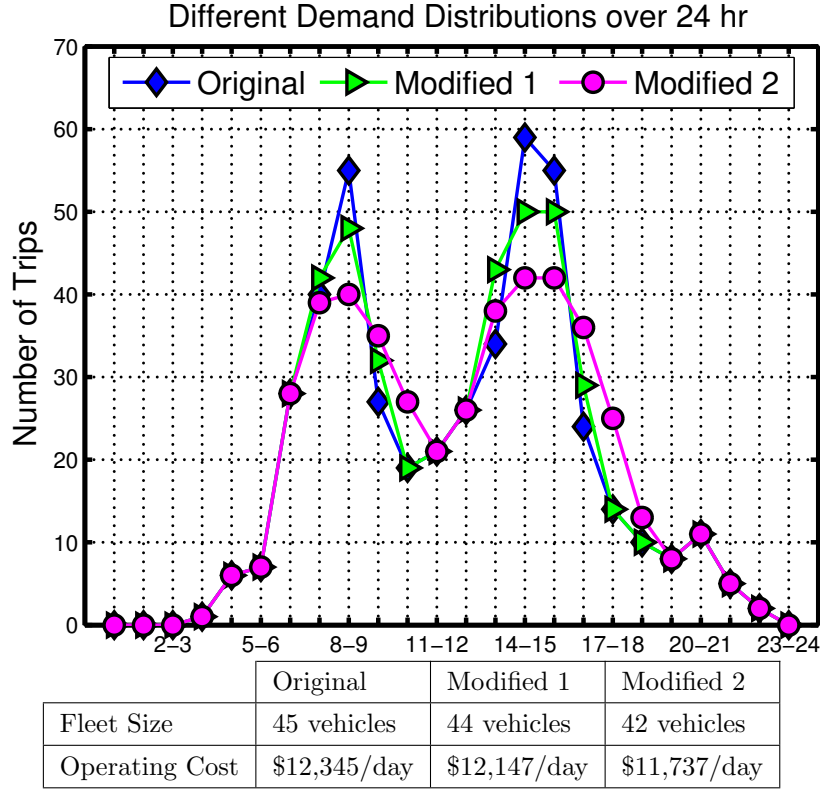


Figure 3: More uniform demand results in smaller fleet and operator’s cost needed to serve all the requests. Note: The assumed costs are as in Table 3

Modified 2). The two modified demand distributions were obtained by redistributing some of the requests originally scheduled during the peak hours. Results in Figure 3 indicate that fewer vehicles are needed to satisfy steadier demand, while keeping the same number of trips and their origin-destination locations, 30 minute time windows, and 12 hour route limit. This insight is important in anticipating potential savings that could be achieved by spreading some of the peak demand. To reach this goal, the DAR operators could provide incentives that encourage their customers to travel during off-peak hours, such as discounts in co-pay paid by the customer and more direct service due to fewer on-board passengers.

Through sensitivity analysis, it was determined that increasing vehicle capacity from a base of 7 passengers does not improve the solutions. This occurs because other operational constraints, including a time window of between 30 and 105 minutes, route durations of between 8 and 12 hours and maximum ride times per passenger prevent the algorithm from taking full advantage of vehicle capacity. This finding is valuable in determining long-term fleet composition.



## 5. Dynamic optimization within the MRMS

The majority of DAR requests are generally pre-scheduled (e.g., 90% RT's demand), which makes the static DARP the key component in managing DAR operations. Solving static DARP is also expected to yield the greatest savings for a DAR company. A similar concept can be applied to tackle requests that need to be served on a short notice (i.e., dynamic requests which should be served on the same day). The MRMS employs Steps 2-3 outlined in Algorithm 1 to sequentially insert dynamic requests as they arrive in the real-time. Every time a new request is received, the system collects locations of all the vehicles as well as information about their remaining schedules. The MRMS then examines all feasible trip insertions, and opts for the insertion with the smallest incremental cost. This automated handling of dynamic requests reduces human effort, and also yields more efficient routes due to ability to examine all feasible insertions.

Moreover, the route improvement procedure (Step 5, Algorithm 1) is constantly applied to reoptimize routes by reinserting or swapping requests. This reoptimization is particularly useful when deviations from the schedule occur either due to congestion or vehicle break downs. These components further increase the efficiency of the DAR operations, as well as the competitiveness of the proposed management system. The whole concept of the MRMS which tackles both static and dynamic DARP is currently patent pending (Mohebbi et al., 2011).

It should be noted that concrete savings from dynamically managing a DAR system are much harder to evaluate than in the case of static DARP. For static DARP, a simple off-line comparison of manual and MRMS-based routes suffice to estimate the savings. However, a similar comparison for dynamic DARP would require vast data in order to realistically simulate a day of operations with and without computerized management. Since dynamic components of the MRMS were gradually introduced in the operations of RT, we were unable to quantify the savings. This will be done in future implementations of the MRMS.

## 6. Conclusions

A computerized system, the MRMS, for managing DAR operations was developed. The developed management system was implemented in a number of midsize transportation companies yielding considerable cost reductions. Test comparisons of manual

and MRMS-based routes indicated that a typical mid-size operator performing approximately 450 trips per day could reduce its annual dispatching and transportation costs by an estimated 0.82 million dollars, roughly 18% of the total operational expense. The MRMS also provided the operators with a powerful tool for quickly exploring system performance under different operational scenarios. Insights gleaned from studying a variety of operational scenarios were found important for tactical planning where an operator needs to determine the cost of providing a certain level-of-service, long-term fleet composition, and utility of driver overtime hours. These findings encourage further implementation of optimization methods with realistic constraints in DAR operations, because they can considerably reduce the costs of providing these services. Since DAR operations are heavily subsidized, computerized systems like MRMS could also help reduce some of these governmental expenditures.

## Acknowledgments

This work was funded by the Maryland Industrial Partnership and the IT Curves company. This support is gratefully acknowledged, but implies no endorsement of the findings.

## References

- Berbeglia, G., Cordeau, J.-F., Gribkovskaia, I. and Laporte, G. (2007), ‘Static pickup and delivery problems: a classification scheme and survey’, *Top* **15**(1), 1–31.
- Borndörfer, R., Grötschel, M., Klostermeier, F. and Küttner, C. (1999), *Telebus Berlin: Vehicle scheduling in a dial-a-ride system*, Springer.
- Cordeau, J.-F. (2006), ‘A branch-and-cut algorithm for the dial-a-ride problem’, *Operations Research* **54**(3), 573–586.
- Cordeau, J.-F. and Laporte, G. (2003a), ‘The dial-a-ride problem (DARP): Variants, modeling issues and algorithms’, *Quarterly Journal of the Belgian, French and Italian Operations Research Societies* **1**(2), 89–101.
- Cordeau, J.-F. and Laporte, G. (2003b), ‘A tabu search heuristic for the static multi-vehicle dial-a-ride problem’, *Transportation Research Part B: Methodological* **37**(6), 579–594.

- Cordeau, J.-F. and Laporte, G. (2007), ‘The dial-a-ride problem: models and algorithms’, *Annals of Operations Research* **153**(1), 29–46.
- Fitzsimmons, J. A. and Fitzsimmons, M. J. (2006), *Service management: operations, strategy, and information technology*, McGraw-Hill New York.
- GAO (2012), Ada paratransit services: Demand has increased, but little is known about compliance, Report, United States Government Accountability Office.
- Golden, B., Assad, A., Levy, L. and Gheysens, F. (1984), ‘The fleet size and mix vehicle routing problem’, *Computers & Operations Research* **11**(1), 49–66.
- Heilporn, G., Cordeau, J.-F. and Laporte, G. (2011), ‘An integer L-shaped algorithm for the dial-a-ride problem with stochastic customer delays’, *Discrete Applied Mathematics* **159**(9), 883–895.
- Jaw, J.-J., Odoni, A. R., Psaraftis, H. N. and Wilson, N. H. (1986), ‘A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows’, *Transportation Research Part B: Methodological* **20**(3), 243–257.
- Luo, Y. (2006), Heuristics and performance metamodels for the dynamic dial-a-ride problem, PhD thesis.
- Luo, Y. and Schonfeld, P. (2007), ‘A rejected-reinsertion heuristic for the static dial-a-ride problem’, *Transportation Research Part B: Methodological* **41**(7), 736–755.
- Madsen, O. B., Ravn, H. F. and Rygaard, J. M. (1995), ‘A heuristic algorithm for a dial-a-ride problem with time windows, multiple capacities, and multiple objectives’, *Annals of operations Research* **60**(1), 193–208.
- Marković, N., Nair, R., Schonfeld, P., Miller-Hooks, E. and Mohebbi, M. (2014), Optimizing dial-a-ride services in maryland, in ‘Transportation Research Board 93rd Annual Meeting’, number 14-1705.
- Mohebbi, M., Ditu, C. V. and Siddiqui, M. I. Y. (2011), ‘Efficient automated ride sharing system’. US Patent Application 13/247,446.
- Parragh, S. N. (2011), ‘Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem’, *Transportation Research Part C: Emerging Technologies* **19**(5), 912–930.

Parragh, S. N., Doerner, K. F. and Hartl, R. F. (2010), ‘Variable neighborhood search for the dial-a-ride problem’, *Computers & Operations Research* **37**(6), 1129–1138.

Toth, P. and Vigo, D. (1996), Fast local search algorithms for the handicapped persons transportation problem, *in* ‘Meta-Heuristics’, Springer, pp. 677–690.