TRANSPORTATION SCIENCE
Vol. 45, No. 3, August 2011, pp. 413–434
ISSN 0041-1655 | EISSN 1526-5447 | 11 | 4503 | 0413

DOI 10.1287/trsc.1100.0335

The Mobile Facility Routing Problem

Russell Halper
Applied Math and Scientific Computation Program, University of Maryland, College Park, Maryland 20742,
rdhalper@gmail.com

S. Raghavan
Smith School of Business and Institute for Systems Research, University of Maryland, College Park, Maryland 20742,
raghavan@umd.edu

In many applications, ranging from cellular communications to humanitarian relief logistics, mobile facilities are used to provide a service to a region with temporal and spatially distributed demand. This paper introduces the mobile facility routing problem (MFRP), which seeks to create routes for a fleet of mobile facilities that maximizes the demand serviced by these mobile facilities during a continuous-time planning horizon. In this setting, demand is produced by discrete events at rates that vary over time. Mobile facilities can be positioned at discrete locations to provide service to nearby events. In addition, mobile facilities can be relocated at any time, although the relocation times are significant in relation to the length of the planning horizon. The demand serviced by a mobile facility depends on the arrival and departure times at each location it visits. Although the MFRP is NP-hard, the optimal route for a single mobile facility can be computed in polynomial time. We describe three heuristics for creating routes for the fleet of mobile facilities and evaluate their performance. Our results demonstrate that these heuristics produce high-quality routes for mobile facilities, especially in scenarios where the demand for service changes significantly over time.

Key words: mobile facility location; vehicle routing; heuristics
History: Received: June 2009; revision received: February 2010; accepted: May 2010. Published online in *Articles in Advance* September 24, 2010.

1. Introduction

Mobile facilities are used in many application domains, ranging from cellular telephone coverage to humanitarian relief logistics. For example, cellular telephone service providers deploy portable cellular base stations, such as Cell-Site-on-Light-Trucks (COLTs) and Cell-Site-on-Wheels (COWs), to provide cellular service for events generating demand for service at a higher rate than an established network of fixed base stations can provide for, or when an existing network of fixed base stations is not operational. These portable cellular base stations can be positioned at a location and provide service to cellular customers without any need for existing infrastructure nearby. To provide service, portable cellular base stations must be stationary at a location. They cannot service cellular phone calls while in transit. Over 100 COLTs and COWs were deployed to the Gulf Coast of the United States after Hurricane Katrina disabled the cellular networks in the region (see Jay 2006). COLTs and COWs have also been deployed to provide additional coverage for large events such as Superbowl XL (see Wendland 2006) and the 2009 Presidential Inauguration of Barack Obama (see Moss, Alan, and Farren 2008). Portable cellular base stations may be quickly relocated to provide service where it is most needed. Consequently, in settings where the demand for service changes over time, these mobile facilities can potentially be used to provide service to a large region more effectively than an equal number of fixed facilities with an equivalent capacity. The challenge this creates is how to effectively deploy such mobile facilities.

Similar mobile facilities are also used in other application domains. In some contexts, mobile facilities are used to provide services to dense urban areas where the cost of establishing a permanent fixed facility is prohibitive, or the demand for services is sporadic. For example, the U.S. Postal Service, Royal Mail (see United Kingdom Department of Transportation 2009), and the Hong Kong Post (see Hong Kong Post 2008) deploy mobile post offices in some urban areas to provide services for customers beyond the delivery and pickup of mail. A mobile post office may be sent to a location, allowing customers to purchase services without having to travel to a more distant, traditional post office. Similar to portable cellular base stations, mobile post offices may only provide these services while stationary at a location. No services may be provided while in transit. The demand for these services also varies over time. A decision maker wishing to schedule such mobile facilities faces the challenge of determining a schedule that will provide as much service as possible. Another type of mobile facility deployed in urban areas are trailer-mounted radar speed monitors, which when placed by the side of the road inform passing motorists of their speed (some even photograph speeding violations). Law enforcement agencies use these mobile facilities to encourage motorists to obey speed limits. Over a planning horizon, an operator could wish to deploy such mobile facilities to maximize the number of vehicles observed, the number of vehicles observed that are exceeding the speed limit, or to areas where pedestrians interact with traffic at a high rate.

Mobile facilities are also used in humanitarian relief to give relief organizations the ability to provide aid to populations dispersed in large, remote regions, and in dense urban areas. For example, the Red Cross has mobile blood collection vehicles that it deploys to collect blood donations. Mobile medical clinics are used to provide care to rural areas (see Alexy and Elnitsky 1996) and in developing countries. Mobile vaccination clinics are also used in developing countries. In each of these settings, operators of these mobile facilities ideally would like to service all demand. However, in practice, limited budgets and resources can force operators to instead maximize the amount of service provided.

In all of the above applications, the operational settings in which these mobile facilities are used are quite similar. The rate at which service is demanded at a location changes over time. The mobile facilities provide a service while stationary but can also be transported between locations. No service can be provided by a mobile facility that is in transit between locations. Furthermore, the operational objective is to maximize the service provided.

In this paper we study the problem of effectively deploying a limited fleet of mobile facilities when demand for the service they provide changes over time. We call this problem the mobile facility routing problem (MFRP). The MFRP seeks to determine routes for a fleet of mobile facilities to maximize the amount of demand serviced in a *continuous-time planning horizon*. In the MFRP, there is a discrete set of locations where a mobile facility may be positioned to provide service (these could be, for example, where appropriate permits have been obtained to place the facilities) and a discrete set of event points generating demand (these could be towns, events, or individuals generating demand for service). While at a location, a mobile facility may service demand from a given subset of event points nearby. Mobile facilities may depart one location for another location at any time, although a mobile facility cannot service demand while in transit. Because the rate at which demand is generated by each event point varies over time, the MFRP is novel among facility location problems in that the total demand serviced by the mobile facilities is dependent on their arrival and departure time at each location visited. In the MFRP, the planning horizon, the locations where a mobile facility may visit, the travel times between locations, the event points generating demand, and their rates are known ahead of time. We show that the problem is NP-hard and describe several heuristics for generating effective routes for mobile facilities.

The remainder of this paper is organized as follows. Section 2 gives a formal introduction to the MFRP and shows that it is NP-hard. In §3, we present an algorithm that determines an optimal route for a single mobile facility in polynomial time. Section 4 describes several heuristics for the MFRP. Section 5 discusses a time-discretized integer programming model to calculate high-quality approximate solutions (that are lower bounds on the optimal solution). Section 6 provides a large set of computational experiments evaluating the heuristics and the time-discretized approximation. Section 7 provides concluding remarks.

2. The Mobile Facility Routing Problem

The MFRP is set on a complete graph in a continuous-time planning horizon $[0, T]$ with a set M of mobile facilities. The vertices of the graph are divided into two subsets, E and L. The vertices in E represent event points. The vertices in L represent locations where the mobile facility may stop to service demand from nearby event points. The length of the edge between location l and location l' is denoted by $TT_{ll'}$ and represents the time it takes a mobile facility to travel from location l to location l'. These travel times are assumed to satisfy the triangle inequality. Because the graph is complete and travel times obey the triangle inequality, a mobile facility may always travel from location l to location l' without passing through any intermediate locations. The objective is to determine a route for each mobile facility and an assignment of demand from event points to the stops on the route that maximizes the total demand serviced.

Because the model is set in a continuous-time planning horizon, demand is modeled as being generated at rates that vary over time. For each event point $e \in E$, there is a nonnegative, real-valued *instantaneous demand function* $d_e(t)$ describing the *rate* at which demand is generated by event point e at time t. A mobile facility at a location $l \in L$ is capable of servicing demand from a subset of event points E_l. The subset E_l may consist of all event points within a specified distance of location l or possibly some other subset of event points defined by service constraints particular to an application. For each location $l \in L$, we define $f_l(t)$ to be the cumulative rate of demand generation by all event points in E_l (i.e., $f_l(t) = \sum_{e \in E_l} d_e(t)$). The fleet of mobile facilities is homogeneous, and each facility has rate capacity C (although the MFRP does allow for uncapacitated mobile facilities, i.e., $C = \infty$). The maximum rate at which a mobile facility at location l at time t can service demand is the minimum of the rate capacity of the mobile facility, C, and $f_l(t)$; however, a mobile facility at location l could service demand at a lower rate if demand from one or more of the event points in E_l is serviced by another mobile facility. Because $d_e(t)$ describes the rate at which demand is generated, the total demand generated by event point e between times σ and τ is given by $\int_\sigma^\tau d_e(s)\,ds$. Thus, between time σ and time τ, a mobile facility at location l can service at most $\int_\sigma^\tau \min\{C, f_l(s)\}\,ds$ units of demand.

A route for a mobile facility is a sequence of *stops*, $(l_n, \sigma_n, \tau_n)_{n=0}^N$, where l_n is the location visited during stop n, σ_n is the arrival time at stop n, and τ_n is the departure time from stop n. In other words, service is provided at location l_n during $[\sigma_n, \tau_n]$. We specify that a mobile facility begins providing service immediately upon arrival at a location. Thus, for any route, $\sigma_n = \tau_{n-1} + TT_{l_{n-1}l_n}$ for $n = 1, \ldots N$. We do not assume that the mobile facility must begin and end its route at a depot, but rather that each mobile facility always begins at stop 0 of its route at time 0 and ends at stop N at time T. The requirement that each mobile facility start and end at a depot D is equivalent to adding the constraint that no demand may be serviced at a location l during $[0, TT_{Dl})$ and $(T - TT_{lD}, T]$. This constraint may be enforced by modifying the definition of $f_l(t)$ so that $f_l(t) = \sum_{e \in E_l} d_e(t)$ for $t \in [TT_{Dl}, T - TT_{lD}]$, and $f_l(t) = 0$ otherwise. In this paper, we will assume $d_e(t)$ to be a piecewise-constant function that may only take on nonnegative values during the planning horizon $[0, T]$. Because any continuous function may be approximated arbitrarily closely by a piecewise-constant function, this assumption is not very restrictive.

Figure 1 displays the locations and event points for an instance of the MFRP. This example contains four locations, $L = \{1, 2, 3, 4\}$, and three event points, $E = \{a, b, c\}$. Suppose that each mobile facility in this example has a rate capacity of 10. A route for a single mobile facility, for example, could start at location 1 and service demand from event point a from time 0 until 1.5. At this point the mobile facility could depart location 1 for location 3. It would arrive at location 3 at time 3 and could continue to provide service from there for event point b until time 7. The mobile facility could then immediately depart location 3 and travel to location 4, servicing demand from event point c from time 8 until the end of the planning horizon. In this route, 40.5 units of demand are serviced. The mobile facility services demand from event point a at rate 3 during $[0, 1.5]$, servicing 4.5 (=3×1.5) units of demand. It services demand from event point b

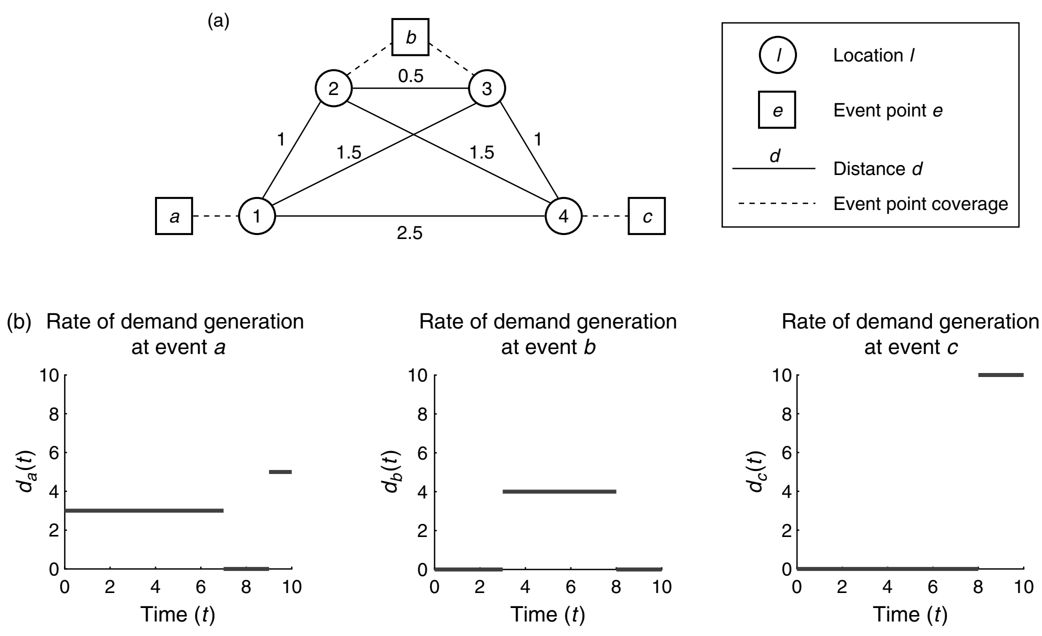

Figure 1 An Instance of the MFRP

Notes. Panel (a) shows a configuration of the locations and event points. The travel times between locations are denoted along the solid line connecting two locations. A dashed line connects each location to each event point it can cover. Panel (b) displays the instantaneous demand function, $d_e(t)$, for each event point in the configuration.

at rate 4 during $[3, 7]$, servicing 16 units of demand. Finally, the mobile facility services demand at rate 10 from event point c during $[8, 10]$, servicing 20 units of demand. Note that this is an optimal route for a single mobile facility with rate capacity 10.

A second mobile facility with rate capacity 10 could then be sent on a route starting at location 1, providing service for event point a during $[1.5, 6]$. At time 6, it could travel to location 2, arriving at time 7 and servicing demand from event point b during $[7, 8]$. It could then be sent back to location 1 to provide service during $[9, 10]$ to event point a. In total, this route would capture 22.5 units of demand.

These two routes together define an optimal solution to this instance of the MFRP with two mobile facilities, each with rate capacity 10. In this scenario, it is easy to see that two mobile facilities cannot service the demand from event point a during $[6, 7]$, event point b during $[7, 8]$, and event point c during $[8, 9]$. At best, two out of the three can be serviced. This solution is optimal because it services the two of these three periods that generate the most demand, and all other demand in the scenario. An equal amount of demand would be serviced if the first mobile facility departs location 1 for location 3 at any time $t \in [0, 1.5]$ and the second mobile facility begins servicing demand at location 1 at time t instead of time 1.5. In this example, the optimal route for a single mobile facility is also one of the two routes in an optimal solution for two mobile facilities. However, in general this need not be the case.

2.1. Formulating the MFRP as an Infinite-Dimensional Mixed-Integer Program

The mobile facility routing problem may be formulated as an infinite-dimensional mixed-integer program (IDMIP). The program seeks to find functions $\tilde{h}_{mle}(t)$ that describe the rate that demand from event point e is serviced by mobile facility m at location l at time t. Let $\tilde{x}_{ml}(t)$ be 1 if mobile facility m is at location l at time t, and 0 otherwise. Let $L_e = \{l: e \in E_l\}$ be the set of locations from which service may be provided to event point e. The MFRP can be formulated as follows:

$$\text{Maximize} \quad \sum_{m\in M}\sum_{l\in L}\sum_{e\in E_l}\int_0^T \tilde{h}_{mle}(s)\,ds \tag{1}$$

$$\text{subject to:} \quad \sum_{e\in E_l}\tilde{h}_{mle}(t) \le C\tilde{x}_{ml}(t) \quad \text{for each } m\in M,\ l\in L,\ t\in[0,T]; \tag{2}$$

$$\sum_{m\in M}\sum_{l\in L_e}\tilde{h}_{mle}(t) \le d_e(t) \quad \text{for each } e\in E,\ t\in[0,T]; \tag{3}$$

$$\tilde{x}_{ml}(t) + \tilde{x}_{ml'}(t') \le 1 \quad \text{for each } m\in M,\ l\in L,\ t\in[0,T]; \quad \max\{0, t - TT_{l'l}\} < t' < \min\{T, t + TT_{ll'}\}. \tag{4}$$

$$\tilde{h}_{mle}(t) \ge 0 \quad \text{for each } m\in M,\ l\in L,\ e\in E_l,\ t\in[0,T]; \tag{5}$$

$$\tilde{x}_{ml}(t) \in \{0,1\} \quad \text{for each } m\in M,\ l\in L,\ t\in[0,T]. \tag{6}$$

The objective function maximizes the amount of demand serviced. If mobile facility m is at location l at time t, constraint (2) ensures that the rate capacity of mobile facility m is not violated. If mobile facility m is not at location l at time t, then $\tilde{x}_{ml}(t) = 0$, which in turn forces $\tilde{h}_{mle}(t) = 0$ for each event point $e \in E_l$. Constraint (3) ensures that for each time t, the rate that demand is serviced from event point e is no greater than the rate that demand is generated by event point e. Constraint (4) enforces the travel times. If mobile facility m is at location l at time t, then this constraint says mobile facility m may not service demand from any other location l' at time t' if either mobile facility m could not leave location l' at time t' and arrive at location l by time t, or mobile facility m could not leave location l at time t and arrive at location l' by time t'. There are an *infinite number of such constraints*, because these constraints are defined for a continuum of times.

Although there has been considerable work on infinite-dimensional linear programs, to our knowledge, not much is known in terms of solution methods for IDMIPs. Consequently, the focus of this paper is on approximate solution methods for the MFRP. In particular, we discuss three heuristics and a discretization of the IDMIP.

2.2. Related Work

Facility location is a well-studied field within operations research. For background on facility location theory, see the excellent monographs by Mirchandani and Francis (1990) and Daskin (1995). To the best of our knowledge, the MFRP has not been studied before.

In a certain sense, the MFRP may be viewed as resembling a dynamic and continuous version of the maximum covering location problem (MCLP). Given a set of demand points and a set of potential facility sites, the MCLP seeks to site a fixed number of facilities to cover the maximum possible amount of demand. Typically, demand is considered to be covered if it is within a given service radius of a facility. Daskin (1995, Ch. 4) is a good survey for background on the MCLP. Other work on the MCLP includes Current and Storbeck (1998), Pirkul and Schilling (1991), Church, Stoms, and Davis (1996), and Karasakal and Karasakal (2004).

A unique feature of the MFRP is that the amount of demand serviced from an event point is variable and dependent on the arrival and departure times of mobile facilities at nearby locations. In addition, these arrival and departure times can occur at any time during the planning horizon $[0, T]$, provided that the travel times between locations along a route are not violated. As a result, an explicit decision must be made to determine when a mobile facility should depart each location along its route. This distinguishes the MFRP from the MCLP, because the amount of demand serviced from an event point in a solution to the MFRP varies depending on the arrival and departure times of mobile facilities at nearby event points. On the other hand, the MCLP, which seeks to find a static placement of facilities, specifies the demand at each demand point as either covered or not.

Covering path problems also seek to determine one or more paths for vehicles to cover demand. These problems can be separated into two classes. The first class contains those problems where all demand must be covered, such as Current, ReVelle, and Cohon (1984), Current and Schilling (1989), and Current, Pirkul, and Rolland (1994). The second class contains those problems that seek to maximize coverage, sometimes as one objective in a multiobjective optimization problem, such as Current and Schilling (1994) and Current, ReVelle, and Cohon (1985). The problems in this second class resemble the MFRP in that they seek to determine a collection of paths that maximize demand coverage. However, unlike the MFRP, demand in these problems only needs be covered at some point along the path. In the MFRP, the demand serviced from each event point is variable because it is dependent on the arrival and departure times of the mobile facilities at nearby locations.

The MFRP is also distinct from real-time vehicle routing problems. Generally, in real-time vehicle routing problems, requests for service arrive stochastically during a planning horizon $[0, T]$, and a policy must be determined on how to service these customers. Often the objective is to determine a policy that minimizes costs or minimizes the time until the next available vehicle (see Ghiani et al. 2003 for a survey on the subject). Although difficult in their own regard, real-time vehicle routing problems are often designed to make a series of discrete decisions, either when a new customer request arrives or when sufficiently many customer requests arrive to warrant dispatching another vehicle. There may be minimum service times associated with each customer request; however, the demand serviced is otherwise not dependent on the length of time spent at the customer. Conversely, a solution to the MFRP must specify each route for the entire planning horizon, including the times when a mobile facility departs from and arrives at each stop in its route. In addition, the MFRP seeks to maximize coverage in a setting where the demand serviced by a mobile facility at a location is dependent on the arrival and departure times.

Gendreau, Laporte, and Semet (2001) consider the problem of relocating ambulances to maintain coverage constraints when ambulances respond to calls that occur stochastically during a planning horizon $[0, T]$. The objective was to maximize the amount of demand covered by at least *two* ambulances minus a penalty for each ambulance relocated. (For a survey of research in ambulance location models, see Brotcorne, Laporte, and Semet 2003.) In this problem, the demand for coverage is constant, and relocation of ambulances is driven by the need to maintain coverage when ambulances cease to provide coverage while responding to calls. These calls arrive stochastically during the planning horizon. Herein lies the distinction from the MFRP, where mobile facilities provide service during the entire planning horizon and the relocation of mobile facilities is driven by the changing rate at which demand for service is being generated at each event point and by the objective of maximizing the demand serviced. In addition, although demand in the MFRP varies over time, it is deterministic.

The MFRP is also distinct from other problems in the literature that seek to locate facilities over a planning horizon. One well-studied problem is the dynamic facility location problem that seeks to locate capacitated facilities to service all demand during a multiperiod planning horizon at minimum cost. (For references, see Ballou 1968, Wesolowsky 1973, Wesolowsky and Truscott 1975, Erlenkotter 1981, Van Roy and Erlenkotter 1982, and Canel et al. 2001.) The multiperiod planning horizon consists of a discrete set of periods. Facilities with sufficient capacity must be sited to service all demand during each of these periods. Typically, a cost is incurred to relocate, add, or resize a facility; however, the problem does not give specific consideration to the time it takes to perform these actions. This may be well suited for long planning horizons when the facilities are infrequently moved, or for scenarios where the time it takes to relocate a facility is relatively short compared to the planning horizon. The continuous-time planning horizon in the MFRP provides the ability to consider problems when the time to relocate a mobile facility may constitute a significant portion of the planning horizon, as is the case, for example, in the telecommunications setting.

Bespamyatnikh et al. (2000) and Durocher and Kirkpatrick (2006) studied a problem related to mobile facility location. Similar to the MFRP, this problem is set in a continuous-time planning horizon. Given a collection of continuously moving points in the plane, the objective is to find paths for each of k continuously moving facilities so that at any time t, the distances from the points to the closest facility minimizes a given metric. Because facilities move continuously and are always providing service, the facilities in this problem differ from those used in the MFRP, where mobile facilities may only provide service while stationary. Additionally, demand in the MFRP does not originate from moving customers and varies over time.

2.3. Computational Complexity of the MFRP

We now show that the MCLP may be considered as a special case of the MFRP where demands are generated at a constant rate and mobile facilities cannot be relocated because of travel times that are longer than the planning horizon. Consequently, the MFRP is NP-hard. Given a set of demand points J, each with demand d_j, a set of potential facility sites I, and the distances between each facility site and each demand point, the MCLP seeks to site p or fewer facilities to cover the maximum possible amount of demand. A facility at a site i is capable of providing coverage to a subset of demand points N_i. Megiddo, Zemel, and Hakimi (1983) prove that the MCLP is NP-hard. To realize an instance of the MCLP as an instance of the MFRP, let the planning horizon end at time 1. For each potential facility site i create a location $l(i)$, and for each demand point j create an event point $e(j)$ generating demand for service at constant rate d_j. Let the travel time between any two locations be greater than one, so a mobile facility can visit at most one location during the planning horizon (and stays at that location for the entire planning horizon, $[0, 1]$). Finally, create p mobile facilities, each with rate capacity $C = \sum_{j\in J} d_j$ (this rate capacity ensures that a mobile facility's capacity is never exceeded).

3. The Single Mobile Facility Routing Problem

In this section we show that the MFRP with a single mobile facility (i.e., $|M| = 1$) is polynomially solvable. We refer to the MFRP with a single mobile facility as the single mobile facility routing problem (SMFRP). Although the SMFRP is posed in a continuous-time setting, our result that allows for a polynomial-time solution is that the SMFRP can be equivalently considered in a discrete-time setting. To show this, we first introduce some notation that we can use to describe the piecewise-constant functions $f_l(t)$.

For each location $l \in L$, let K_l be two plus the number of discontinuities of $f_l(t)$ during $(0, T)$. (We add two to account for times 0 and T.) Let $q_1^l = 0$, $q_{K_l}^l = T$, and for $2 \le i \le K_l - 1$, let q_i^l be the ith discontinuity of $f_l(t)$. For $2 \le i \le K_l - 1$, we refer to such a time q_i^l as a *critical time* of location l. Let g_i^l be the value of $f_l(t)$ for $t \in [q_i^l, q_{i+1}^l)$ for each $1 \le i \le K_l - 1$. Our main result is that for the SMFRP, it suffices to restrict our attention to critical times of the locations in L.

Theorem 3.1. *There exists an optimal route $(l_n, \sigma_n, \tau_n)_{n=0}^N$ where either the departure time τ_n is a critical time of $f_{l_n}(t)$, or the arrival time σ_{n+1} is a critical time of $f_{l_{n+1}}(t)$, for each $n = 0, 1, \ldots, N-1$.*

Proof. Define $f_l^C(t) = \min\{C, f_l(t)\}$. The function $f_l^C(t)$ describes the maximum rate at which demand may be serviced by a single mobile facility with rate capacity C at location l at time t. As we will show, any stop with $\sigma_n = \tau_n$ may be removed from the route without reducing the amount of demand serviced. Consequently, we may assume that $\sigma_n < \tau_n$ for $n = 0, 1, \ldots, N$. The amount of demand serviced as a function of the departure times may be written as

$$\begin{aligned}\Delta(\tau_0, \ldots, \tau_N) &= \int_0^{\tau_0} f_{l_0}^C(t)\,dt + \cdots + \int_{\tau_{n-1}+TT_{l_{n-1}l_n}}^{\tau_n} f_{l_n}^C(t)\,dt \\ &\quad + \int_{\tau_n+TT_{l_nl_{n+1}}}^{\tau_{n+1}} f_{l_{n+1}}^C(t)\,dt + \cdots + \int_{\tau_{N-1}+TT_{l_{N-1}l_N}}^{\tau_N} f_{l_N}^C(t)\,dt.\end{aligned}$$

Suppose the sequence $(l_n, \sigma_n, \tau_n)_{n=0}^N$ defines an optimal route and there exists an n such that τ_n is not a critical time of $f_{l_n}(t)$ and σ_{n+1} is not a critical time of $f_{l_{n+1}}(t)$. It suffices to show that this route may be modified to create a new route servicing an equal amount of demand that either leaves location l_n at a critical time of $f_{l_n}(t)$, arrives at location l_{n+1} at a critical time of $f_{l_{n+1}}(t)$, or skips stop n entirely. With this in hand, a finite sequence of such modifications will produce a route that satisfies the theorem.

Because neither is τ_n a critical time of $f_{l_n}(t)$ nor is σ_{n+1} a critical time of $f_{l_{n+1}}(t)$, the piecewise-constant function $f_{l_n}(t)$, and thus $f_{l_n}^C(t)$, is constant in a neighborhood of τ_n and, similarly, $f_{l_{n+1}}^C(t)$ is constant in a neighborhood of σ_{n+1}. Consequently, $\Delta(\tau_0, \ldots, \tau_N)$ is differentiable in a neighborhood of τ_n and

$$\begin{aligned}\frac{\partial}{\partial \tau_n}\Delta(\tau_0, \ldots, \tau_N) &= f_{l_n}^C(\tau_n) - f_{l_{n+1}}^C(\tau_n + TT_{l_nl_{n+1}}) \\ &= f_{l_n}^C(\tau_n) - f_{l_{n+1}}^C(\sigma_{n+1}). \qquad (7)\end{aligned}$$

Because by assumption the route is optimal, it follows that

$$f_{l_n}^C(\tau_n) - f_{l_{n+1}}^C(\sigma_{n+1}) = 0. \qquad (8)$$

Because $f_{l_n}^C(t)$ and $f_{l_{n+1}}^C(t)$ are piecewise-constant functions, Equation (8) holds true for the interval of time around τ_n where both $f_{l_n}^C(t)$ and $f_{l_n}^C(t + TT_{l_nl_{n+1}})$ remain constant.

Let $\tilde{p}$ be the latest critical time of $f_{l_n}(t)$ before time τ_n and let $\tilde{q}$ be the latest critical time of $f_{l_{n+1}}(t)$ before time σ_n. Because both $f^C_{l_n}(t)$ and $f^C_{l_{n+1}}(t + TT_{l_n l_{n+1}})$ are piecewise-constant functions, it follows from Equation (8) that $f^C_{l_n}(t) = f^C_{l_{n+1}}(t + TT_{l_n l_{n+1}})$ for all t in the interval $\max\{\tilde{p}, \tilde{q} - TT_{l_n l_{n+1}}\} \le t \le \tau_n$. Thus, if the mobile facility departs from stop n at time $\tilde{\tau}_n = \max\{\tilde{p}, \tilde{q} - TT_{l_n l_{n+1}}, \sigma_n\}$ instead of time τ_n, the mobile facility will service an equal amount of demand. Consequently, this modified route remains optimal.

If $\tilde{\tau}_n = \tilde{p}$ or $\tilde{\tau}_n = \tilde{q} - TT_{l_n l_{n+1}}$, then either the mobile facility departs stop n or arrives at stop $n+1$ at a critical time. If $n \ge 1$ and $\tilde{\tau}_n = \sigma_n$, then the mobile facility may depart stop n the instant it arrives and service an equal amount of demand. In this new optimal route, zero demand is serviced at stop n. Because travel times satisfy the triangle inequality, if the mobile facility were to leave stop $n-1$ at time τ_{n-1} and travel directly to location l_{n+1}, skipping stop n entirely, it would arrive at location l_{n+1} at time $\tilde{\tau}_n + TT_{l_n l_{n+1}}$ or earlier, and therefore services no less demand. Thus, a new optimal route may be followed by leaving stop $n-1$ at time τ_{n-1} and traveling directly to stop $n+1$, arriving at time $\tau_{n-1} + TT_{l_{n-1} l_{n+1}}$. Finally, if $\tilde{\tau}_n = 0$, the mobile facility must be at the first stop in the route, so $n = 0$. In this case, a new optimal route may be defined by starting at location l_1 at time 0 and following the remainder of the route. In either case, a new optimal route may be defined that skips stop n, proving the theorem. □

An immediate consequence of Theorem 3.1 is that there is a discrete set of times when the lone mobile facility can be expected to leave from or arrive at each location. Furthermore, the size of this set is polynomial in the number of critical times of each location. In particular, if S_l is the set of times when the mobile facility can be assumed to leave from or arrive at location l in an optimal solution, then

$$S_l = \left(\bigcup_{l' \ne l} \{ q^{l'}_j + TT_{l'l} \mid j = 0, 1, \ldots K_{l'},\ q^{l'}_j + TT_{l'l} < T \} \right) \cup \left(\bigcup_{l' \ne l} \{ q^{l'}_j - TT_{ll'} \mid j = 0, 1, \ldots K_{l'},\ q^{l'}_j + TT_{ll'} > 0 \} \right) \cup \{ q^l_j \mid j = 0, 1, \ldots, K_l \}, \tag{9}$$

and because S_l only contains times in $[0, T]$,

$$|S_l| \le K_l + 1 + 2 \sum_{l' \ne l} K_{l'}. \tag{10}$$

Using this discrete sets of times, it is possible to find the optimal solution for a single mobile facility in polynomial time. To do so, we define a directed acyclic graph (DAG) and find the longest path between a designated source and sink node in the graph. Let the set S_l be ordered by time and denoted by $\{s^l_0 < s^l_1 < \cdots\}$. For each location l and each time $s^l_k \in S_l$, we create a vertex v^l_k in the graph. If $v^l_k < T$, an arc is created from vertex v^l_k to vertex v^l_{k+1} and given a weight equal to the amount of demand that the mobile facility could service if it was at that location l during the time period $[s^l_k, s^l_{k+1}]$. For each vertex $v^l_k > 0$, an arc with weight 0 is created to vertex $v^{l'}_{k'}$ if $s^l_k + TT_{ll'} = s^{l'}_{k'}$ and $s^{l'}_{k'} < T$. A source vertex v_{start} is added along with arcs with weight 0 from v_{start} to each v^l_0 for each $l \in L$. Finally, a sink node v_{end} is added along with arcs with weight 0 from each node $v^l_{|S_l|}$ to v_{end} for each $l \in L$.

Note that because each arc originates from a vertex representing an earlier time than that of the vertex at which the arc terminates, no cycles exist in the graph. Thus, a path through this network from v_{start} to v_{end} can be thought of as representing a route for the mobile facility. Each arc on the path from a vertex v^l_k to vertex v^l_{k+1} represents the mobile facility providing service at location l from time s^l_k to s^l_{k+1}. Each arc on the path from a vertex v^l_k to a vertex $v^{l'}_{k'}$ represents the mobile facility traveling from location l to location l'. The length of the path equals the amount of demand the mobile facility can capture along this route.

Finding the longest path through the network from v_{start} to v_{end} can easily be accomplished in linear time using a dynamic program that for the purposes of this paper we will refer to as the *single mobile facility algorithm* (SMFA). (Note that the problem is identical to the well-known PERT problem in project management, where one also seeks to find the longest (or so-called critical) path in a DAG.)

Single Mobile Facility Algorithm:

Step 0: Set length(v_{end}) := 0. Create a sorted list of the remaining vertices in the network in decreasing order of the time associated with each vertex and with v_{start} at the end of the list.

Step 1: Remove the vertex v from the top of the list. Set length(v) = $\max_{(v,v')}\{\text{weight}(v, v') + \text{length}(v')\}$ and set next(v) to equal the vertex v' for which this maximum is achieved. Ties are broken arbitrarily.

Step 2: If the list is empty, terminate. Otherwise, repeat Step 1.

At the conclusion of the SFMA, length(v_{start}) provides the length of the longest path, and the actual path is obtained by following the nodes stored in the next(v) array (starting from v_{start}). The SMFA can be viewed as being identical to running a single-sink shortest-path algorithm on a DAG after the arc lengths have been negated, and it is well-known that the running time is linear in the number of vertices and arcs of the DAG.

The DAG constructed contains $2 + \sum_{l \in L} |S_l|$ vertices. Vertex v_{end} has no outgoing arcs, and each other vertex has at most $|L|$ outgoing arcs. Therefore the DAG has at most $|L| + |L| \sum_{l \in L} |S_l|$ arcs. Thus the SMFA has running time $O(|L|^2 \sum_{l \in L} K_l)$, which is linear in the number of arcs in the DAG and polynomial in the number of locations and critical times of the functions $f_l(t)$.

4. Heuristics for the MFRP

The previous section describes how to quickly find an optimal route for a single mobile facility. In this section we consider the MFRP with multiple mobile facilities, which is NP-hard. Unfortunately, Theorem 3.1 may not be extended to the case with multiple mobile facilities. Recall the example problem in Figure 1 and the optimal routes for two mobile facilities. In any optimal solution the second mobile facility leaves location 1 at time 6 and arrives at location 2 at time 7. However, neither is time 6 a critical time of location 1, nor is time 7 a critical time of location 2. This occurs because the first mobile facility departs location 3 at time 7. If one were to create a function representing the amount of demand generated by event b that is not serviced by the first mobile facility, then it may be noted that there is a discontinuity at time 7 caused by the departure of the first mobile facility from location 3 at time 7. This discontinuity, which is not a critical time of $f_2(t)$, causes the route of the second mobile facility to depart location 1 at time 6 and arrive at location 2 at time 7. Here, time 6 is produced by subtracting $TT_{l_3 l_4}$ and $TT_{l_1 l_2}$ from critical time 8 of location 4. Similar examples can be constructed where, given $|M|$ mobile facilities, the time a facility departs from or arrives at a location is not a critical time of that location but rather is a critical time of another location plus or minus a sequence of travel times between $|M|$ pairs of locations.

In general, the potential departure times for a mobile facility from a location l are dependent on the critical times of the locations in L and on the demand from the events in E_l that is serviced by the other mobile facilities. Each time a mobile facility relocates from one location to another in an optimal solution, it either departs or arrives at a critical time or at a time when another mobile facility is arriving or departing from some other location. If not, by a similar argument to the proof of Theorem 3.1, the mobile facility could either leave earlier or later and service more demand. For the MFRP, there seems to be no obvious way to generate a polynomially bounded set of discrete times when a mobile facility may depart from or arrive at a location. It remains an open question whether an optimal solution to the MFRP always exists in which each time a mobile facility travels, it departs a location at a time that is a critical time of that location or of some other location plus or minus a sequence of travel times between pairs of locations. Even if this were true, and the sequence had length at most $|M|$, the number of potential departure times would still be exponential in $|M|$. Furthermore, even if there was a method to reduce this set of potential departure times so that its size is polynomially bounded, the problem of determining the optimal routes for a fleet of mobile facilities is still extremely difficult. For example, in perhaps the simplest case where each mobile facility could arrive at each location only at time 0 and depart only at time T, the end of the planning horizon, we have observed that solving the MFRP is no easier than the MCLP. In fact, in our experiments the MFRP appears to be significantly more difficult!

Consequently, our strategy has been to develop several heuristics for generating routes for the MFRP, which we describe below. We then apply a local search method to improve the routes. As we stated earlier, we utilize the SMFA when generating routes. As each heuristic executes, a record is kept of the demand that is not serviced in the instance of the MFRP. We define the variable $\tilde{d}_e(t)$, for each $e \in E$, to be the rate at which demand that is not serviced is generated at event point e at time t. Similarly, for each location $l \in L$, we define the variable $\tilde{f}_l(t)$ to be the cumulative rate at which demand that is not serviced is generated by the event points in E_l at time t. Thus, $\tilde{f}_l(t) = \sum_{e \in E_l} \tilde{d}_e(t)$ for each $t \in [0, T]$. Furthermore, we define $\tilde{d}_{me}(t)$ to be the rate that demand is being serviced by mobile facility m from event point e at time t and $\tilde{f}_{ml}(t)$ to be the rate that mobile facility m is servicing demand from location l at time t.

Given a subset of locations L_m that mobile facility m visits, the SMFA can be used to generate a route that captures the maximum possible amount of demand not being serviced from these locations. In other words, the SMFA can be seen as a function that takes in the remaining serviceable demand at each location in L_m, the travel times between those locations, and the rate capacity of a mobile facility and returns a sequence of locations for the mobile facility to visit, the arrival and departure time at each location in the sequence, and functions $\tilde{f}_{ml}(t)$ that describe the amount of demand serviced by mobile facility m from location l at each time t.

4.1. Demand Assignment

After a route is created for a mobile facility with the SMFA, the rate that demand is serviced at each time t from each event point must still be determined. In applications where the mobile facilities have no rate capacity or have a sufficiently high rate capacity such that it cannot be exceeded (e.g., $C > \max_{l \in L,\, t \in [0,T]} f_l(t)$), a mobile facility is capable of servicing all nearby demand, in which case all maximal assignments of demand in the model service the same amount of demand as in practice. Contexts where mobile facilities have small service areas can potentially be modeled by creating a single event that can be serviced for each location. Overlapping coverage in this setting corresponds to two or more mobile facilities at the same location. In this case, the choice of which mobile facility in the model services demand does not affect the amount of demand serviced in practice.

However, in applications where the rate capacity of a mobile facility is a limiting factor, for each stop made by a mobile facility the rate at which demand is serviced from each nearby event point must be determined. In many contexts where mobile facilities at distinct locations can provide service to a single event, operators of mobile facilities can explicitly control how much demand will be serviced by each mobile facility. For example, a cellular telephone service provider can instruct individual cell phones to use specific base stations. When providing humanitarian relief using mobile facilities after a disaster, the population must be notified of the locations of the mobile facilities. In such a case, instructions could also be sent about which mobile facility the population should go to in order to receive aid. These instructions can be broadcast over the radio or communicated by relief workers sent out to each event point in the model. Implementing such measures to control demand assignment allows operators to maximize the service provided by their mobile facilities. Additionally, in contexts where the operator has limited ability to control demand assignment, the MFRP can potentially be used to provide a reasonable approximation of the routing of mobile facilities.

The phase of our heuristics that determines the demand from each event point serviced by each mobile facility is called *demand assignment*. Given a route, this phase assigns demand from event points to each stop on the route individually. For stop n in the route, each of our heuristics initially set $\tilde{d}_{me}(t) = 0$ for all $e \in E_{l_n}$ and all times t between the arrival time σ_n and the departure time τ_n. The event points in E_{l_n} are sorted according to one of four criteria that we specify below. Taking the first event point e in the list, we define $\tilde{d}_{me}(t)$ at each time $t \in [\sigma_n, \tau_n]$ to be the minimum of the current rate demand that is not serviced that is being generated at event point e (i.e., $\tilde{d}_e(t)$), and the unused rate capacity of the mobile facility. Accordingly, we then subtract $\tilde{d}_{me}(t)$ from $\tilde{d}_e(t)$. This process is then repeated for each event point in the sorted list until either the list has been exhausted or the rate capacity of the mobile facility has been reached for the entire duration of the stop. The four sorting criteria we consider are as follows.

Sort 1: Sort the event points in increasing order of the number of locations from which a mobile facility could provide service to the event point (i.e., $|L_e|$). For event points that are covered by the same number of locations, we sort them in increasing order of the amount of demand not serviced during Stop n.

Sort 2: Sort the event points in increasing order of the number of locations from which a mobile facility could provide service to the event point (i.e., $|L_e|$). For event points that are covered by the same number of locations, we sort them in decreasing order of the amount of demand not serviced during Stop n.

Sort 3: Sort the event points in increasing order of the amount of demand not serviced during Stop n.

Sort 4: Sort the event points in decreasing order of the amount of demand not serviced during Stop n.

Intuitively, Sort 1 and Sort 2 first assign demand from those event points that can be serviced from fewer locations, making it less likely that the route of another mobile facility will be able to service demand from those event points. Sort 1 and Sort 3 attempt to first assign demand from those event points with small amounts of demand not serviced by sorting the event points in increasing order of the amount of demand not serviced. This would leave more event points with large amounts of demand that is not serviced, which would hopefully allow other routes to be created having stops where a larger amount of demand is serviced. Alternatively, sorting the event points in decreasing order of the amount of demand not serviced in Sort 2 and Sort 4 would leave more event points with demand that is not serviced, hopefully allowing any route created to minimize travel time, during which a mobile facility services no demand.

4.2. Sequential Routing for the MFRP

The sequential routing heuristic generates a route for one mobile facility at a time. This heuristic initializes by setting $\tilde{d}_e(t) = d_e(t)$ for each $e \in E$ and $\tilde{f}_l(t) = f_l(t)$ for each $l \in L$. Each subsequent stage of the algorithm generates a route for a single mobile facility considering all locations (i.e., $L_m = L$) using the SMFA. Thus, each time a route is generated, it will service the maximum possible amount of demand that was previously not serviced. Because the SMFA only determines locations for the mobile facility to visit and a departure time for each location in the route, after each iteration a decision must be made on how demand should be assigned from event points to the stops in the route of the mobile facility. Finally, the variables $\tilde{d}_e(t)$ and $\tilde{f}_l(t)$ are updated to reflect the demand serviced by the new route. The steps of the sequential routing heuristic are outlined below.

Step 0 **(Initialization)**: For each $e \in E$, initialize $\tilde{d}_e(t) := d_e(t)$. For each $l \in L$, initialize $\tilde{f}_l(t) := f_l(t)$.

Step 1 **(Route Determination)**: Choose a mobile facility without a route. (Because all mobile facilities are identical, this selection may be made arbitrarily.) Run the SMFA, using the capacitated remaining serviceable demand functions $\tilde{f}_l^C(t) = \min\{C, \tilde{f}_l(t)\}$ for each location l, to determine each location in the route, and the arrival and departure time for each stop in the route.

Step 2 **(Demand Assignment)**: Compute the demand serviced from each event point for each stop in the route following the method for demand assignment described in §4.1.

Step 3 **(Demand Update)**: Update $\tilde{d}_e(t)$ for each $e \in E$ and $\tilde{f}_l(t)$ for each $l \in L$ to reflect the demand that is serviced by this route.

Step 4: If all mobile facilities have been assigned a route or all demand is serviced, terminate. Otherwise, return to Step 1.

4.3. Generating Routes with an Insertion Heuristic

The second heuristic we present is an insertion heuristic. The heuristic associates a set of locations L_m with each mobile facility m. Each mobile facility m will only visit locations in L_m on its route. Each set L_m is initially empty. At each stage of the heuristic, a location $\tilde{l}$ is considered for insertion into each set L_m. When location $\tilde{l}$ is being considered for insertion into L_m, the SMFA is used to calculate the demand serviced along the best route for mobile facility m, assuming that it may only visit the locations in L_m and location $\tilde{l}$, and assuming that mobile facility m may only service demand that it is already servicing or demand that is currently not serviced. If the route followed by at least one mobile facility may be improved by allowing it to visit location $\tilde{l}$, then location $\tilde{l}$ is added to the set L_m for the mobile facility m whose route may be most improved by the addition of location $\tilde{l}$. The routes of the remaining mobile facilities remain unchanged. The process is repeated until no more improvements may be found. The details of the insertion heuristic are outlined below.

Step 0 **(Initialization)**: For each $e \in E$, $l \in L$, and $m \in M$, initialize $\tilde{d}_e(t) := d_e(t)$, initialize $\tilde{f}_l(t) := f_l(t)$, initialize the set L_m to be empty, initialize $\tilde{d}_{me}(t) := 0$, and initialize $\tilde{f}_{ml}(t) := 0$. Furthermore, initialize the set $\tilde{L} := L$.

Step 1 **(Insertion Selection)**: Select the location $\tilde{l} \in \tilde{L}$ with the largest total amount of demand that is not serviced during the planning horizon. For each mobile facility m with $\tilde{l} \notin L_m$, do the following. For each location $l \in L_m \cup \{\tilde{l}\}$, define the function $\delta_l(t) = \min\{C, \tilde{f}_l(t) + \tilde{f}_{ml}(t)\}$. Calculate how much demand could be serviced by mobile facility m by executing the SMFA using the functions $\delta_l(t)$ as the rate that demand can be serviced at time t from location l by this mobile facility and the travel times between the locations in $L_m \cup \{\tilde{l}\}$.

Step 2 **(Route Selection)**: If it was found in Step 1 that the route of no mobile facility can be improved by allowing it to visit location $\tilde{l}$, then remove location $\tilde{l}$ from the set $\tilde{L}$ and go to Step 5. Otherwise, the route that may be most improved by visiting location $\tilde{l}$ is changed to follow the new route calculated in Step 1. In the case where this maximum improvement may be realized by inserting location l into several routes, we choose the route that was first considered. Location $\tilde{l}$ is added to L_m.

Step 3 **(Demand Assignment)**: Demand from event points is assigned to each stop in the route following the method for demand assignment described in §4.1. If no demand remains that may be serviced from location $\tilde{l}$, or $\tilde{l} \in L_m$ for every mobile facility m, then location $\tilde{l}$ is removed from $\tilde{L}$.

Step 4 **(Demand Update)**: Update $\tilde{d}_e(t)$ for each $e \in E$ and $\tilde{f}_l(t)$ for each $l \in L_m$ to reflect the demand that is now serviced or no longer serviced by this new route.

Step 5: If $\tilde{L}$ is empty, terminate. Otherwise, return to Step 1.

4.4. Local Search for the MFRP

We now describe a local search algorithm that looks for improvements to routes generated with our heuristics. A naive method for looking for improvements could examine solutions in a local search neighborhood defined by the following two types of exchanges:

- Choose a stop n in a route r. Try changing the location visited in stop n to a different location, maintaining either the arrival and departure times for stop n, or the departure time from stop $n-1$ and the arrival time at stop $n+1$.
- Choose a pair of routes, r_1 and r_2, and a stop n_1 in route r_1. Delete stop n_1 so that the demand serviced at stop n_1 by route r_1 is no longer serviced. Next, delete every stop of route r_2 and recalculate route r_2 to service the maximum amount of demand that is currently not serviced using the SMFA. (This could improve route r_2 by allowing it to service some of the demand previously serviced during stop n_1 of route r_1.) Finally, delete all the remaining stops in route r_1 and recalculate route r_1 using the SMFA.

Implementing such a local search procedure has two distinct problems. The first type of exchange never improves the routes generated by either heuristic. The insertion heuristic associates with each route a set of locations that the route may visit and then uses the SMFA to calculate a route that services the maximum amount of demand that is currently not serviced while visiting only the stops in that set. It terminates when, for each route, no improvement may be found by allowing that route to visit another location. Consequently, each route generated by the insertion heuristic cannot be improved by changing the location visited in a stop to any other location. In the sequential heuristic, each route was created to service the maximum amount of demand that was not serviced immediately prior to the generation of that route. Because the demand being serviced at some stage of the sequential routing heuristic will remain serviced during execution of the subsequent steps of the sequential routing heuristic, the demand not serviced immediately prior to the generation of any route includes all demand that is not serviced by the final routes produced by the sequential heuristic. Consequently, after the sequential heuristic, if a stop in a route is replaced with a stop at a different location, the mobile facility will service no more demand after this exchange than it did on its route before the exchange.

The second (and more severe) problem is that the running time of a local search algorithm that considers every exchange of the second type blows up wildly. This is because the SMFA is executed each time the second type of exchange is considered. In our experience, searching over every exchange of the second type could take an unreasonably long time because most exchanges in the neighborhood are fruitless. Motivated by identifying types of neighborhoods that can yield improvements, we look at a similar (but smaller) special type of local search neighborhood where a lower bound on the improvement may be computed rapidly.

Suppose we have two routes, route r_1, $(l_n^1, \sigma_n^1, \tau_n^1)_{n=0}^{N_1}$, and route r_2, $(l_n^2, \sigma_n^2, \tau_n^2)_{n=0}^{N_2}$, and that route r_1 was generated prior to route r_2 in the sequential routing heuristic. Furthermore, suppose that route r_1 and route r_2 were allowed to visit the same set of locations and were generated using the SMFA. All demand serviced in route r_2 was not being serviced when route r_1 was generated. Consequently, route r_1 cannot be improved by servicing demand that is being serviced in route r_2. However, it may be possible to improve route r_2 if the mobile facility in route r_2 was allowed to capture some of the demand serviced in route r_1.

For each pair of routes in a solution to the MFRP generated by one of our heuristics, our local search algorithm looks for opportunities to improve the solution by removing a stop from route r_1, inserting it into route r_2, and filling in the gap in route r_1 with demand that is not serviced. Specifically, if stop $(l_{n_0}^1, \sigma_{n_0}^1, \tau_{n_0}^1)$ is removed from route r_1 and inserted into route r_2, we temporarily assume that the arrival time and departure time of stop $(l_{n_0}^1, \sigma_{n_0}^1, \tau_{n_0}^1)$ are preserved. Because travel times are assumed to obey the triangle inequality, this produces a unique time t_0 when the mobile facility following route r_2 must deviate from route r_2 to arrive at location $l_{n_0}^1$ at time $\sigma_{n_0}^1$ and a unique time t_1 when the mobile facility returns to route r_2 after departing location $l_{n_0}^1$ at time $\tau_{n_0}^1$. The demand

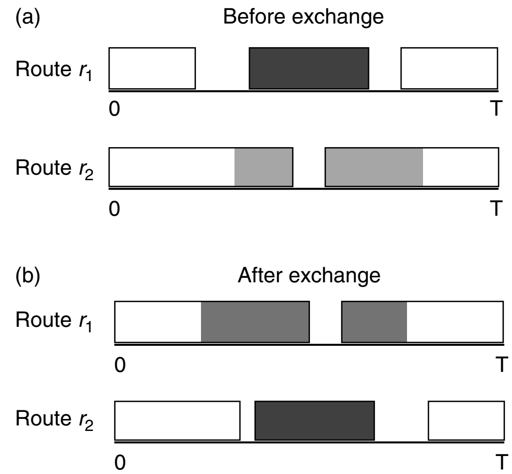

Figure 2 An Example of an Exchange in a Local Search Algorithm
Notes. Boxes above each timeline each represent a stop along a route. The solid black box represents the stop that is moved from route r_1 to route r_2. The demand lost in the swap is shaded by light gray. The demand added in the exchange is shaded with dark gray.

covered along route r_2 between time t_1 and time t_2 is no longer serviced by this mobile facility.

Route r_1 no longer contains stop n_0, which allows the mobile facility following that route to stay longer at stops $n_0 - 1$ and $n_0 + 1$, possibly servicing more demand. To calculate the amount of additional demand serviced, we define the set S of potential departure times from stop $n_0 - 1$. By a similar argument to the proof of Theorem 3.1, one can show that the mobile facility will either depart stop $n_0 - 1$ at a discontinuity of $\tilde{f}_{l_{n_0-1}}(t)$ or at $\tau_{n_0-1}^1$, or arrive at stop $n_0 + 1$ at a discontinuity of $\tilde{f}_{l_{n_0+1}}(t)$ or at σ_{n_0+1}. Specifically, S will contain the discontinuities of $\tilde{f}_{l_{n_0-1}}(t)$ and $\tilde{f}_{l_{n_0+1}}(t + TT_{l_{n_0-1} l_{n_0+1}})$ during the interval of time between the times $\tau_{n_0-1}^1$ and $\sigma_{n_0+1}^1 - TT_{l_{n_0-1} l_{n_0+1}}$. Sort S, and denote the sorted order as $S = \{s_0 < s_1 < \cdots < s_K\}$. The following procedure (illustrated in Figure 2) gives an efficient method for computing the locally optimal time the mobile facility should depart stop $n_0 - 1$ for stop $n_0 + 1$.

Step 0: Initialize $k := 1$, $D^- := 0$, $D^+ := \int_{s_0 + TT_{l_{n_0-1} l_{n_0+1}}}^{\sigma_{n_0+1}^1} \tilde{f}_{l_{n_0}+1}(s)\, ds$. Then initialize $k^{\max} := 0$ and $D^{\max} := D^+$.

Step 1: Set $D^+ := D^+ - \int_{s_{k-1} + TT_{l_{n_0-1} l_{n_0+1}}}^{s_k + TT_{l_{n_0-1} l_{n_0+1}}} \tilde{f}_{l_{n_0+1}}(s)\, ds$. Set $D^- = D^- + \int_{s_{k-1}}^{s_k} \tilde{f}_{l_{n_0-1}}(s)\, ds$.

Step 2: If $D^- + D^+ > D^{\max}$, set $D^{\max} := D^- + D^+$ and $k^{\max} := k$.

Step 3: If $k = K$, terminate. Otherwise, set $k := k + 1$.

At the end of this procedure, $s_{k^{\max}}$ gives the locally optimal time that the mobile facility following should depart stop $n_0 - 1$ for stop $n_0 + 1$.

To see if this exchange produced an improvement, one may wish to see if the demand serviced in route r_1 and route r_2 after the exchange is greater than the demand serviced by the two routes before the exchange. This is equivalent to computing the difference in the amount of demand that is serviced after the exchange but was not serviced before the exchange and the amount of demand that was serviced before the exchange but is not serviced after the exchange. The additional amount of demand serviced after the exchange is the amount of demand serviced in route r_1 between times $\tau^1_{n_0-1}$ and $\sigma^1_{n_0+1}$. The amount of demand that is no longer serviced after the exchange is the amount of demand previously serviced along route r_2 between times t_0 and t_1. (We note that this may count the demand from an event point twice if it was covered by route r_2 before the exchange and route r_1 after the exchange. However, this demand will be canceled out in the final calculation.) The difference in these two is the net change in the amount of demand serviced. If this net change is positive, this is an improvement in the routes. If an improvement is found, we could then delete route r_2 and delete stop n_0 from route r_1, leaving the remainder of route r_1 intact, and recompute route r_2 using the SMFA to service the maximum amount of remaining demand. After demand assignment to this new route (to see if we could further increase the amount of demand serviced), we could delete the remaining stops in route r_1 and recompute route r_1 using the SMFA. However, this operation is computationally expensive. Thus, we will only perform such an operation after considering all possible exchanges and selecting the one that could potentially yield the most improvement.

Notice that because the arrival time and departure time of stop n_0 of route r_1 are preserved when it is inserted into route r_2, even if the difference in the demand serviced after the exchange is negative, it is possible that an overall increase in the amount of demand serviced could be found when we recompute route r_1 and route r_2 using the SMFA. For this reason, we define a negative threshold value Δ and will consider all exchanges where the change in the amount of demand serviced along the two routes after the exchange is greater that Δ.

To start the local search procedure, we define an empty tabu list $\mathcal{T}$ that will keep track of exchanges of a stop between routes that are not to be considered again. Each iteration of the local search algorithm searches over all such exchanges of a stop between two routes that are not in the tabu list and selects the exchange that provides the greatest improvement. If the greatest difference in the amount of demand serviced is less than Δ, we terminate. Otherwise, we recalculate the two routes as follows. We first delete the stop n_0 that was moved from route r_1 to route r_2 and delete all stops originally in route r_2 (i.e., $(l_n, \sigma^2_n, \tau^2_n)^{N_1}_{n=0}$), temporarily leaving the remainder or route r_1 intact. Next, we recompute route r_2 using the SMFA to generate a new route r_2 that services the maximum amount of demand currently not serviced. Demand assignment is performed for the new route r_2 using the procedure described in §4.1. Now, to recalculate route r_1 we delete the remaining stops in route r_1 and then recompute route r_1 using the SMFA. Demand assignment is performed for the new route r_1 using the procedure described in §4.1. If the total demand serviced along the new route r_1 and along route r_2 is not more than before the exchange, we revert to the two routes before the exchange and add the exchange to the tabu list $\mathcal{T}$. Otherwise, the new routes are kept. The local search algorithm continues to iterate until no further improvement can be found. The details of the local search procedure are outlined below.

Step 0 **(Initialization)**: Initialize route pointers r_1 and r_2, set $n_0 := 0$, $\delta := \Delta$, and $\mathcal{T} = \phi$.

Step 1 **(Explore All Solutions in Neighborhood)**: For each pair of routes, route a and route b, and for each stop n in route a, run Step 1a.

Step 1a **(Compute Particular Neighborhood Solution)**: Consider moving stop n from route a to route b, preserving the arrival and departure time at stop n. Compute the demand that is no longer serviced in route b if stop n must be visited and store it as δ_1. Compute the locally optimal time that the mobile facility following route a must leave stop $n-1$ for stop $n+1$. Store the additional demand serviced along route a at stops $n-1$ and $n+1$ as δ_2. If $\delta < \delta_2 - \delta_1$, set $r_1 := a$, $r_2 := b$, $n_0 := n$, and $\delta := \delta_2 - \delta_1$.

Step 2 **(Attempt to Implement Improvement)**: If $\delta \leq \Delta$, terminate. Otherwise, delete stop n_0 from route r_1 and route r_2, leaving the remainder of route r_1 intact. Recalculate route r_2 using the SMFA and allowing the mobile facility to visit any location. Run the demand assignment phase to assign demand to each stop in the new route r_2. Next, delete the remaining stops in route r_1 and recalculate route r_1 using the SMFA. Run demand assignment to assign demand to each stop in the new route r_1. If this does not produce an improvement, add the exchange of stop $(l_{n_0}, \sigma^a_{n_0}, \tau^a_{n_0})$ from route r_1 to route r_2 to $\mathcal{T}$ and revert to the previous routes. Return to Step 1.

Given a pair of routes generated by the sequential routing heuristic, during Step 1a of this local search procedure we only considered a pair of routes a and b if route a was generated prior to route b in the sequential routing heuristic.

5. Calculating Lower Bounds for the MFRP with Integer Programming

In this section, we discuss how mixed-integer programming can be used to provide informative lower bounds for the MFRP that are useful in evaluating the performance of our heuristics. These lower bounds provide feasible solutions to the MFRP and are therefore heuristics in their own respect. However, because calculating the optimal solutions to these mixed-integer programs (MIP) is computationally very challenging, these can only be used to compute solutions to very small problems.

5.1. Fixing Facility Locations

Mobile facilities in the MFRP can be relocated over time. However, keeping each facility at one location for the entire planning horizon $[0, T]$ is a feasible solution to the MFRP and thus provides a lower bound. We refer to such a solution as a static placement of facilities. The optimal static placement of facilities that services the most demand may be found by solving an MIP, which we name the Static MIP (SMIP). The optimal static solution may be used to get a sense of the value of mobile facilities (in the sense of how much more demand may be serviced with mobile facilities) and identify when mobile facilities provide the most value.

Let F be the set of fixed facilities. Given a scenario, let $S = \{0 = s_0 < s_1 < \cdots < s_K = T\}$ be the collection of all critical times of all locations in L. For each $f \in F$ and $l \in L$, define the binary variable x_{fl} to be 1 if fixed facility f is positioned at location l and 0 otherwise. For each $f \in F$, $l \in L$, $e \in E$, and $k = 0, 1, \ldots, K-1$, define the nonnegative continuous variable d_{fek} to be the amount of demand serviced by fixed facility f from event point e during $[s_k, s_{k+1})$. Let D_{ek} be the total amount of demand generated by event point e during $[s_k, s_{k+1})$. The SMIP is formulated as follows:

$$\text{(SMIP)} \quad \text{Maximize} \ \sum_{f\in F}\sum_{e\in E}\sum_{k=0}^{K-1} d_{fek} \tag{11}$$

subject to:

$$\sum_{l\in L} x_{fl} = 1 \quad \text{for each } f \in F, \tag{12}$$

$$\sum_{e\in E} d_{fek} \le C(s_{k+1} - s_k) \quad \text{for each } f \in F,\ k = 0, 1, \ldots, K-1; \tag{13}$$

$$\sum_{f\in F} d_{fek} \le D_{ek} \quad \text{for each } e \in E,\ k = 0, 1, \ldots, K-1; \tag{14}$$

$$d_{fek} \le D_{ek} \sum_{l\in L_e} x_{fl} \quad \text{for each } f \in F,\ e \in E,\ k = 0, 1, \ldots, K-1; \tag{15}$$

$$d_{fek} \ge 0 \quad \text{for each } f \in F,\ e \in E,\ k = 0, 1, \ldots, K-1; \tag{16}$$

$$x_{fl} \in \{0, 1\} \quad \text{for each } f \in F,\ l \in L. \tag{17}$$

The objective of the SMIP, given by (11), is to maximize the amount of demand serviced. Constraint (12) assigns each facility to a single location l. Because the rate at which each event point generates demand is constant during each interval of the form $[s_k, s_{k+1})$, constraint (13) preserves the rate capacity of the facility by dictating that the amount of demand serviced during $[s_k, s_{k+1})$ by the facility cannot exceed the maximum amount of demand it could serve during that time. Constraint (14) says that the amount of demand serviced by the facilities during period k from event point e is less than or equal to the amount of demand generated by event point e during period k. Constraint (15) says that a facility may service demand from event point e only if the facility is positioned at a location $l \in L_e$. Constraints (16) and (17) specify the values the variables may take on.

5.2. Time Discretization of the IDMIP

An approximation of the IDMIP formulation of the MFRP may be obtained by discretizing time. This provides an approximate solution and a lower bound for the MFRP. The finer the discretization, the better the approximation. On the other hand, the size of this time-discretized formulation is extremely large and grows rapidly with the level of discretization.

For the time discretization, let $\{t_0, t_1, \ldots, t_K\}$ be the set of times when any mobile facility can be assumed to either arrive at or depart from any location. Furthermore, to discretize the IDMIP, the rate at which each event point generates demand must be constant during each interval of time $[t_k, t_{k+1})$. Notice that the set $\bigcup_{l\in L} S_l$ satisfies this condition. We use this time discretization in our computational experiments.

To formulate this time-discretized MIP (TDMIP), let P be the set of time periods between these times, $\{(t_k, t_{k+1})\colon k = 0, \ldots, K-1\}$. For each $p \in P$, define $\theta_p = t_{p+1} - t_p$, the length of period p. By design, during each period p, a mobile facility will either be traveling from one location to another or be providing service from a location for the entire period. Let D_e^p be the total demand generated at event point e during period p. Define the binary decision variable x_{ml}^p to be 1 if mobile facility m is providing service from location l in period p, and 0 otherwise. For each $e \in E$, $l \in L_e$, and $p \in P$, define the nonnegative real-valued decision variable d_{el}^p to be the amount of demand generated from event point e during period p that is serviced by mobile facilities at location l. The TDMIP can be formulated as follows:

$$\text{(TDMIP)} \quad \text{Maximize} \ \sum_{p\in P}\sum_{e\in E}\sum_{l\in L_e} d_{el}^p, \tag{18}$$

subject to:

$$\sum_{l\in L_e} d_{el}^p \le D_e^p \quad \text{for each } p \in P,\ e \in E, \tag{19}$$

$$\sum_{e\in E_l} d_{el}^p \le (C\theta_p) \sum_{m\in M} x_{ml}^p \quad \text{for each } p\in P,\, l\in L, \tag{20}$$

$$x_{ml}^p + x_{ml'}^{p'} \le 1 \quad \text{for each } m\in M,\ p,p'\in P,\ l,l'\in L,$$

such that

$$(t_{p'}, t_{p'+1}) \cap (t_p - TT_{l'l}, t_{p+1} + TT_{ll'}) \ne \varnothing, \tag{21}$$

$$d_{el}^p \ge 0 \quad \text{for each } p\in P,\ e\in E,\ l\in L_e, \tag{22}$$

$$x_{ml}^p \in \{0,1\} \quad \text{for each } p\in P,\ m\in M, l\in L. \tag{23}$$

The objective function (18) seeks to maximize the amount of demand serviced. Constraint (19) ensures that the amount of demand serviced during period p from event point e does not exceed the amount of demand generated by event point e during period p. Constraint (20) ensures that the amount of demand serviced by mobile facilities at location l during period p does not exceed the amount of demand that those mobile facilities may capture during period p, which is $C\theta_p$ for each mobile facility. Because the instantaneous demand functions are constant during each period, the rate capacity of a mobile facility will not be exceeded if this constraint is met. Constraint (21) ensures that the route of each mobile facility is feasible, respecting the travel times between locations. Constraints (22) and (23) specify the ranges of the variables.

6. Computational Results

In §4, we described two heuristics to generate routes for mobile facilities and a local search procedure that attempts to improve a set of routes. During execution, both of these heuristics employ one of four sorting orders that are defined in §4.1 to assign demand coverage. To test these heuristics, we have developed a variety of simulated data sets. These data sets are generated with varying parameter values and demand profiles to represent a wide range of data. The one parameter we keep fixed is the rate capacity of each mobile facility, which we fix at 10. Fixing this parameter value does not restrict the scope of the data sets. In every data set, all locations and events are positioned at random points in a rectangular region of the plane. The travel time between each pair of locations is given by the Euclidean distance between them. A mobile facility at a location may provide service to events within a given distance of that location.

It is important to note the scope of demand profiles for the MFRP when simulating the demand in data sets. In particular, we are interested in studying the use of mobile facilities to provide service over a large geographic region with a dynamic demand profile. When demand is generated at a relatively constant rate for a large portion of the planning horizon, it is unlikely that mobile facilities could be used more efficiently than fixed facilities. In such a case, a static model may be more appropriate for determining fixed locations for the facilities. Thus, we do not wish to study scenarios where there are many event points that each generate demand at a relatively constant rate for a large percentage of the planning horizon. Similarly, we also do not wish to study scenarios with many locations where demand could be serviced at a relatively constant rate for the entire planning horizon. Keeping this in mind, we have created two distinct types of simulated scenarios for the MFRP.

The first type of scenario simulates, at a high level, demand profiles that we might expect to occur in practice, such as when routing a fleet of portable cellular base stations over a day. We refer to these scenarios as *"realistic" scenarios*. Each realistic scenario simulates part of a single day, beginning at 7 A.M. and ending at midnight in a 10 by 15 rectangular region of the plane. The demand profile of each event point in the scenario simulates the forecasted demand from a single event that might occur in practice, such as a sporting event, rush hour traffic, a county fair, a convention, etc. For each type of event point, we define a time window during which that type of event point may generate demand. For each event point e of that type, we choose at random a time t_{start} when demand generation begins and a later time t_{end} when demand generation ends inside that window, although we specify that each event point must generate demand for at least one hour (i.e., $t_{end} - t_{start} > 1$). During that time period, $d_e(t)$ gradually increases to a maximum level, remains at that maximum level for a period of time, and then gradually decreases to 0. To ensure that each event point provides a significant maximum level of demand, the maximum level is selected from a uniform distribution on $[0.5 * C, C]$. The rates of increase and decrease are constant and are also chosen randomly but allow the maximum level to be reached at least momentarily and demand to decrease back to 0 during the time period t_{start} and t_{end}. Outside of the time period between t_{start} and t_{end}, no demand will be generated by event point e.

The demand profiles at the event points in a realistic scenario share a specific structure. Namely, the demand is generated during a single interval of time, and during that interval demand gradually increases to, remains at, and then decreases from a maximum rate of demand generation. To see how our heuristics perform in less structured data sets, we created a second type of scenario. We refer to these scenarios as *mathematically challenging scenarios*. As the name suggests, the mathematically challenging data sets are designed to be less structured, and therefore more difficult, than the realistic scenarios. The planning horizon of each of these scenarios begins at time 0 and ends at time 100. Each of these scenarios is set in a

25-unit square region of the plane. The demand profile for an event point is generated as follows. The number of pieces in the piecewise-constant function $d_e(t)$ for an event point e is chosen at random from a specified range. To generate a wide range of demand profiles, we specified that each event point would have at least 5 and no more than 35 pieces. To find the length of each piece, we assign each piece a random number between 0 and 1. These numbers are then normalized so that their sum equals the length of the planning horizon. Each normalized number is taken to be the length of the corresponding piece. The rate that demand is generated by that event point (i.e., the height of $d_e(t)$) during that time is 0 with probability $p = 0.7$, and otherwise chosen randomly from a lognormal distribution with a standard deviation of either 1, 1.5, 2, or 2.5.

In our results, the names of realistic scenarios will begin with the letter "R" and the mathematically challenging scenarios will begin with the letters "MC." This is followed by a unique index for the scenario. Table 1 gives the number of mobile facilities, locations, event points, and coverage radius in each scenario. Computational results were compiled on a Dell Optiplex 740 with an AMD Athlon 64 X2 5000+ dual core processor with 3GB of RAM running Microsoft Windows XP. The heuristics were coded and compiled in Microsoft Visual C++ 2005.

Table 1 **The Number of Mobile Facilities, Locations, and Event Points and the Coverage Radius of a Mobile Facility at a Location for Each Type of Scenario**

Data set	Facilities	Locations	Event points	Coverage distance
R0	3	25	75	1
R1	5	25	75	1
R2	7	25	75	1
R3	10	25	75	1
R4	3	25	35	1
R5	5	25	35	1
R6	7	25	35	1
R7	10	25	35	1
R8	3	15	35	1.5
R9	5	15	35	1.5
R10	7	15	35	1.5
R11	10	15	35	1.5
R12	3	15	20	1.5
R13	5	15	20	1.5
MC0	3	25	75	3.5
MC1	5	25	75	3.5
MC2	7	25	75	3.5
MC3	10	25	75	3.5
MC4	5	10	20	4.5
MC5	10	25	50	4.5
MC6	5	15	20	4.5
MC7	10	15	20	4.5
MC8	5	10	20	1
MC9	5	10	15	4.5

6.1. Sequential Routing vs. Insertion Heuristic

Table 2 displays the performance of the two route generation heuristics when different sorting orders are used with each heuristic on a variety of data sets. Each row displays the averaged results from 25 realistic data sets or 40 mathematically challenging sets generated with varying demand parameters. We make two key observations. First, notice that regardless of the sorting order used in the demand assignment step, the sequential routing heuristic typically outperforms the insertion heuristic. Second, it appears that in both heuristics, using Sort 2 produces routes that on average service the most total demand. Because the difference in the average demand serviced is small, the Wilcoxon signed rank test (see Wilcoxon 1945) was used to see if Sort 2 outperforms the other sorting methods in a statistically significant manner. For the sequential routing heuristic, the hypothesis that Sort 2 outperforms Sort 3 and Sort 4 was validated at a 0.00066 significance level for the realistic data sets and at a 0.0000002 significance level for the mathematically challenging data sets. For the insertion heuristic, the same hypothesis was validated at a 0.0000101 significance level for the mathematically challenging data sets. The Wilcoxon signed rank test on the realistic data sets shows that for the insertion heuristic, Sort 2 outperforms Sort 3 at a 0.002 significance level, but the test was inconclusive when comparing Sort 2 and Sort 4. In addition, the Wilcoxon test was inconclusive when comparing Sort 1 to Sort 2. This is possibly because Sort 1 and Sort 2 only differ by the secondary sorting criteria. (Intuitively, the reason Sorts 2 and 1 outperform Sorts 3 and 4 may be because they first assign demand from event points that can be serviced from fewer locations.) Consequently, the demand assignment phase of our heuristics is unlikely to produce significantly different results using either Sort 1 or Sort 2. (For example, when used with the sequential routing heuristic, Sort 1 and Sort 2 produce the same solution for 279 of the 350 realistic instances and 201 of the 400 mathematically challenging instances.) Because on the remaining instances Sort 2 on average outperforms Sort 1, we chose to use Sort 2 in the demand assignment phase of our heuristics for the remainder of our computational experiments.

Implementing the local search procedure with the routes generated with either the sequential routing heuristic or the insertion heuristic provided improvement. Tables 3 and 4 display the demand serviced in solutions created with the two route generation heuristics before and after the local search procedure, as well as the runtimes of the heuristics in seconds. The sequential routing heuristic typically finds higher-quality solutions than the insertion heuristic but has a longer runtime. Although the insertion heuristic runs the SMFA more frequently than

Table 2 Performance of the Sequential Routing Heuristic and the Insertion Heuristic with Different Sorting Orders

	Sequential heuristic				Insertion heuristic			
Data set	Sort 1	Sort 2	Sort 3	Sort 4	Sort 1	Sort 2	Sort 3	Sort 4
R0	134.55	134.55	134.46	134.55	130.17	130.17	130.18	129.96
R1	197.39	197.57	197.22	197.52	193.70	193.65	193.44	193.56
R2	260.67	260.77	260.33	260.71	254.71	254.71	255.04	254.89
R3	316.49	316.77	316.29	315.22	311.42	311.92	310.12	310.55
R4	87.12	87.12	87.11	87.12	83.28	83.61	83.28	83.61
R5	133.25	133.34	133.19	133.04	129.45	129.57	129.22	130.11
R6	150.51	150.55	150.56	150.51	148.44	148.34	148.46	148.70
R7	169.80	169.77	169.78	169.77	169.49	169.38	169.61	169.06
R8	91.64	91.64	91.48	91.37	90.46	90.46	90.01	90.19
R9	134.19	134.21	134.11	134.17	131.88	131.81	131.40	131.80
R10	156.57	156.50	156.51	156.33	155.10	155.18	154.80	154.98
R11	170.55	170.55	170.23	170.42	170.13	170.20	169.87	170.36
R12	65.61	65.61	65.59	65.61	63.68	63.68	63.68	63.68
R13	88.10	88.10	88.10	88.00	86.44	86.49	86.69	86.92
Average (R)	154.03	154.07	153.93	153.88	151.31	151.37	151.13	151.27
MC0	2,494.71	2,494.96	2,492.52	2,492.00	2,436.39	2,436.55	2,433.20	2,427.83
MC1	4,042.53	4,040.43	4,031.52	4,024.26	3,936.63	3,934.95	3,921.62	3,907.47
MC2	5,390.01	5,394.17	5,369.33	5,374.52	5,236.50	5,234.01	5,218.15	5,217.11
MC3	7,117.66	7,122.76	7,079.51	7,086.65	6,853.98	6,855.66	6,826.91	6,832.50
MC4	2,360.31	2,360.58	2,357.73	2,357.74	2,322.06	2,322.04	2,322.19	2,319.02
MC5	6,014.99	6,019.46	5,993.12	5,996.95	5,806.29	5,804.62	5,791.46	5,792.32
MC6	2,248.01	2,247.33	2,244.27	2,243.67	2,205.04	2,204.13	2,201.72	2,198.37
MC7	3,471.89	3,469.71	3,470.63	3,466.91	3,414.45	3,415.97	3,408.82	3,410.38
MC8	2,430.20	2,430.29	2,425.40	2,426.83	2,405.90	2,405.41	2,402.18	2,397.19
MC9	1,947.57	1,947.70	1,946.06	1,946.75	1,929.34	1,933.73	1,928.95	1,932.74
Average (MC)	3,751.79	3,752.74	3,741.01	3,741.63	3,654.66	3,654.71	3,645.52	3,643.49

Note. Each row displays the averaged results from either 25 realistic data sets or 40 mathematically challenging data sets.

the sequential routing heuristic, the insertion heuristic found no improvement the vast majority of the time when trying to add a location to the collection of locations that may be visited on a route. In fact, routes generated by the insertion heuristic typically contained just one or two stops. Thus, each time the SMFA was run in the insertion heuristic, it was run on a substantially smaller network and therefore executed much more quickly.

The sequential routing heuristic produces higher-quality routes than the insertion heuristic. When the sequential routing heuristic generates a route, it considers the demand that is not serviced at every location. As a result, it can produce routes where the mobile facility makes a number of stops visiting locations during periods where demand can be serviced at a high rate. In addition, by considering all locations in the SMFA, the sequential routing heuristic also implicitly considers the spatial and temporal configuration of demand that is not serviced. On the other hand, the insertion heuristic initially generates routes where each mobile facility makes one stop at a location where a large amount of demand for service is generated over the planning horizon, regardless of the proximity of other locations or temporal configuration of other demand. This makes it unlikely during the execution of the insertion heuristic that a route will be improved by insertion of another location into the set of locations a mobile facility can visit. As a result, the insertion heuristic has trouble deviating from these initial routes.

After local search, the gap in the runtimes and solution quality between the sequential routing heuristic and the insertion heuristic narrows. The sequential routing heuristic with local search significantly outperforms the insertion heuristic with local search on the mathematically challenging data sets, whereas the difference in the performance of the two heuristics with local search on the realistic data sets is minimal. More specifically, the hypothesis that the sequential heuristic with local search produces higher-quality solutions than the insertion heuristic with local search was verified by the Wilcoxon signed rank test at a 0.0005 significance level for the mathematically challenging data sets but only at a 0.329 significance level for the realistic data sets. The local search procedure also seems to provide proportionally more improvement for realistic data sets than it does for mathematically challenging data sets. This may be because the demand at each event point in a realistic data set is generated during a distinct interval of time, whereas the demand at an event point in a mathematically

Table 3 Performance of Sequential Heuristic with and Without Local Search

	Sequential routing heuristic						
	Sequential heuristic		With local search		Improvement (%)		
Data set type	Demand serviced	Run time (s)	Demand serviced	Run time (s)	Maximum	Minimum	Median
R0	134.55	50.18	135.30	199.24	3.41	0.00	0.00
R1	197.57	18.85	199.11	135.60	4.43	0.00	0.31
R2	260.77	38.03	264.37	369.58	5.15	0.00	1.16
R3	316.77	116.88	320.88	1,144.15	4.29	0.00	1.35
R4	87.12	5.61	87.57	39.18	3.93	0.00	0.00
R5	133.34	3.64	134.62	27.35	5.37	0.00	0.68
R6	150.55	12.49	152.56	87.25	4.80	0.00	0.81
R7	169.77	33.84	171.85	105.23	2.96	0.00	1.12
R8	91.64	4.68	92.22	32.08	4.39	0.00	0.00
R9	134.21	0.91	135.71	6.27	4.86	0.00	0.46
R10	156.50	2.36	158.44	14.03	6.67	0.00	0.98
R11	170.55	6.42	171.98	21.68	3.50	0.00	0.47
R12	65.61	0.10	66.21	0.61	13.10	0.00	0.00
R13	88.10	0.20	89.03	1.25	3.71	0.00	0.41
Average (R)	154.07	21.01	155.70	155.96	5.04	0.00	0.55
MC0	2,494.96	142.61	2,495.18	164.06	0.93	0.00	0.00
MC1	4,040.43	24.13	4,046.92	27.93	1.53	0.00	0.00
MC2	5,394.17	166.73	5,409.85	229.41	2.42	0.00	0.01
MC3	7,122.76	116.72	7,140.36	191.80	2.18	0.00	0.12
MC4	2,360.58	3.51	2,377.57	5.48	3.70	0.00	0.59
MC5	6,019.46	128.57	6,042.36	304.20	2.62	0.00	0.21
MC6	2,247.33	8.48	2,259.29	20.11	3.77	0.00	0.36
MC7	3,469.71	8.14	3,493.89	39.36	3.13	0.00	0.73
MC8	2,430.29	6.06	2,450.00	11.58	5.12	0.00	0.53
MC9	1,947.70	3.33	1,962.74	9.51	5.94	0.00	0.39
Average (MC)	3,752.74	60.83	3,767.82	100.34	3.13	0.00	0.29

Notes. Each row displays the averaged results from either 25 realistic or 40 mathematically challenging data sets. The maximum, minimum, and median improvement found by local search is given for each row.

challenging data set could be generated at many different periods in the planning horizon. This could produce more opportunities to find improvements by moving stops between routes for realistic data sets than between routes for mathematically challenging data sets.

6.2. The Static Placement of Facilities

As discussed in §5.1, when evaluating the quality of solutions produced by our heuristic, it is natural to ask the following question. How much more demand could be serviced by the solutions produced by our heuristics than in the optimal static solution? If our heuristics outperform the optimal static solution, it gives an indication of their effectiveness of utilizing the capabilities of the mobile facilities to relocate over time. To answer this question, we compare the demand serviced in our heuristic solutions to the demand serviced in an optimal static placement of the same number of fixed facilities with equal rate capacity. Solving the SMIP given by (11)–(17) can be computationally challenging because K can be quite large. We implemented the SMIP in ILOG OPL 5.2. For many of the mathematically challenging data sets, the SMIP was computationally intractable. For the data sets where we were able to solve the SMIP to find the optimal static solution, Table 5 presents a comparison of solutions to MFRP instances generated by the sequential routing heuristic with local search and the optimal static solution. On average, using the sequential heuristic with local search to route mobile facilities allows 6.4% more demand to be serviced than the optimal static solution for the mathematically challenging data sets. This improvement increases to 16.6% when comparing results of the realistic scenarios. We observed improvement as high as 61.83%. In a few instances we did observe the optimal static solution outperform our heuristic solution, although by less than 3% in all but six instances and typically by less than 1%.

To further evaluate the trade-off between the solutions generated by our heuristics and the optimal static solution, we developed a third type of scenario. We begin the name of each of the third type of scenario with the letter "T." These scenarios are designed to test the performance of the heuristics relative to the percentage of time each event point is generating demand. Each of these scenarios is defined together

Table 4 Performance of the Insertion Heuristic with and Without Local Search

	Insertion heuristic						
	Insertion heuristic		With local search		Improvement (%)		
Data set type	Demand serviced	Run time (s)	Demand serviced	Run time (s)	Maximum	Minimum	Median
R0	130.17	0.17	135.08	307.34	12.56	0.00	3.94
R1	193.65	0.29	199.11	191.78	7.24	0.00	1.99
R2	254.71	0.31	264.20	677.30	9.62	0.00	3.24
R3	311.92	0.30	320.82	1,283.30	6.55	0.23	2.27
R4	83.61	0.05	87.73	61.25	16.26	0.00	4.60
R5	129.57	0.08	134.43	16.12	12.27	0.00	3.15
R6	148.34	0.07	153.11	41.26	8.70	0.00	2.95
R7	169.38	0.05	171.53	32.36	5.24	0.00	0.93
R8	90.46	0.04	92.26	10.36	9.75	0.00	0.16
R9	131.81	0.05	135.62	28.83	8.38	0.00	2.32
R10	155.18	0.04	158.32	10.24	10.90	0.00	1.12
R11	170.20	0.04	172.11	14.18	3.61	0.00	0.85
R12	63.68	0.01	66.43	8.83	31.88	0.00	2.35
R13	86.49	0.01	88.22	10.62	12.29	0.00	0.95
Average (R)	151.37	0.11	155.64	185.81	11.09	0.02	2.20
MC0	2,436.55	0.49	2,461.06	47.36	8.25	0.00	0.00
MC1	3,934.95	1.06	4,015.27	14.29	7.27	0.00	1.84
MC2	5,234.01	1.73	5,380.02	161.76	9.72	0.00	2.17
MC3	6,855.66	2.80	7,119.38	182.76	11.34	0.00	3.76
MC4	2,322.04	0.13	2,370.97	8.52	7.57	0.00	1.66
MC5	5,804.62	2.00	6,031.09	360.15	17.62	0.04	2.67
MC6	2,204.13	0.17	2,259.30	21.31	9.61	0.00	1.62
MC7	3,415.97	0.35	3,502.45	51.77	5.51	0.00	1.57
MC8	2,405.41	0.10	2,457.52	9.23	6.33	0.00	1.18
MC9	1,933.73	0.10	1,966.64	8.44	7.87	0.00	0.93
Average (MC)	3,654.71	0.89	3,756.37	86.55	9.11	0.00	1.74

Notes. Each row displays the averaged results from 25 realistic or 40 mathematically challenging data sets. The maximum, minimum, and median improvement found by local search is given for each row.

Table 5 A Comparison of the Total Demand Serviced in the Optimal Static Solution and in the Solution to the MFRP Generated by the Sequential Routing Heuristic with Local Search

Data set type	Static solution	Sequential heuristic with local search	Improvement (%) Maximum	Minimum	Median	Instances improved
R0	117.19	133.41	32.46	4.56	15.05	25/25
R1	169.76	199.11	32.57	7.90	17.06	25/25
R2	232.78	264.30	29.07	1.94	12.20	25/25
R3	283.08	320.64	20.90	4.44	13.80	25/25
R4	71.08	87.57	49.84	7.83	22.70	25/25
R5	111.81	134.62	39.74	9.02	20.27	25/25
R6	129.06	152.60	40.93	6.65	17.50	25/25
R7	152.43	171.85	30.66	2.67	12.08	25/25
R8	78.84	92.22	47.62	0.99	15.53	25/25
R9	119.03	135.71	33.69	0.93	13.46	25/25
R10	144.78	158.44	21.76	1.57	8.70	25/25
R11	162.67	171.99	13.23	1.17	5.32	25/25
R12	53.96	66.21	61.83	1.99	26.54	25/25
R13	77.41	89.03	37.13	2.94	13.92	25/25
MC0	2,473.36	2,495.69	11.32	−4.03	0.66	30/40
MC1	3,996.97	4,065.89	5.94	−2.87	1.50	35/40
MC4	2,202.41	2,377.22	21.14	−9.85	7.76	38/40
MC6	2,109.78	2,259.30	28.01	−13.92	6.07	35/40

Notes. Each row contains the averaged results of either 40 mathematically challenging data sets or 25 realistic data sets. The maximum, minimum, and median improvement for the data sets in each row is displayed, as well as the proportion of data sets for which an improvement was found.

with a parameter $\lambda \in (0, 1]$. Upon creation, an event point e in one of these scenarios is assigned a total amount of demand that it will generate during the planning horizon, D_e, and a randomly chosen time in the planning horizon, t_e. For a fixed λ, event point e will generate demand at rate $D_e/(\lambda T)$ during the interval of time $[(1-\lambda)t_e, t_e + \lambda(T - t_e))$ and at rate 0 outside of that interval. Thus, the total amount of demand generated by a particular event point e is equal for every value of λ. Figure 3 displays an example of the instantaneous demand function, $d_e(t)$, for a single event point e and several values of the parameter λ. Notice that as λ approaches 0, the demand becomes more "spiky."

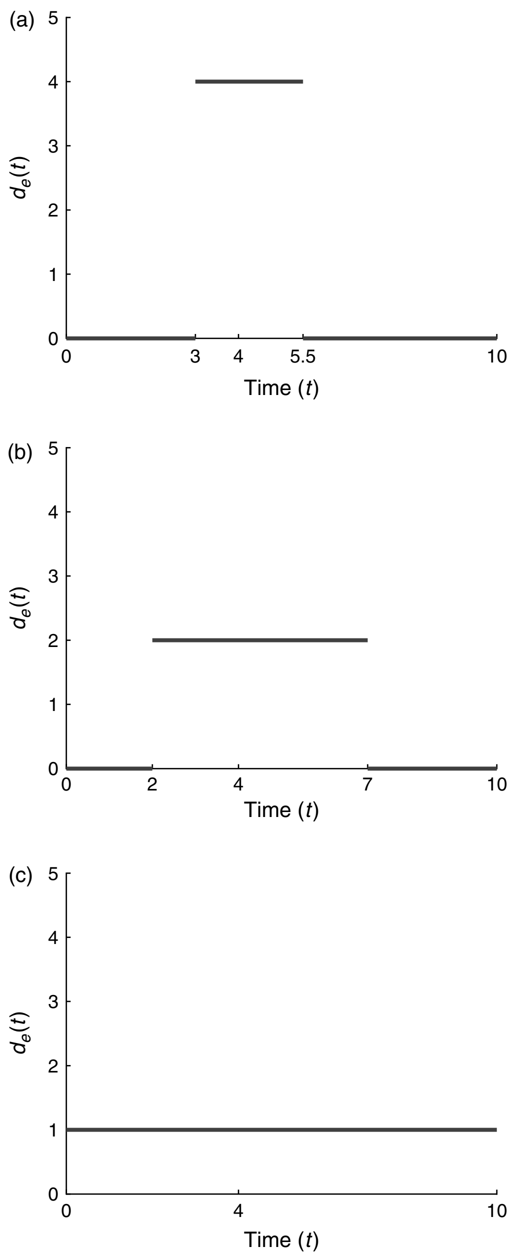

Figure 3 **An Example of the Demand Profile of a Single Event Point from the Third Types of Scenarios for Three Different Values of λ**

Notes. In all three graphs, $D_e = 10$, $T = 10$, and $t_e = 4$. Panel 3(a) displays the rate demand that is generated at the event point for $\lambda = 0.25$. Panel 3(b) displays the rate demand that is generated at the same event point for $\lambda = 0.50$. Panel 3(c) displays the rate demand that is generated at the same event point for $\lambda = 1$.

Each of these scenarios is generated in a 25 by 25 region of the plane with a planning horizon of length 100. Each location is generated at a random position where at least one event point can be covered. We chose a moderately wide coverage radius of 3.5 to encourage the existence of locations that can cover multiple events. The rate capacity of each mobile facility is $(\sum_{e\in E} D_e/(\lambda T))$. By definition, this rate capacity will never be exceeded for any value of λ. Furthermore, because any particular event point generates the same amount of total demand for every value of λ, this rate capacity guarantees that the optimal static solution is the same for every $\lambda \in (0, 1]$.

Figure 4 shows four graphs comparing the demand captured in solutions generated by the sequential heuristic with local search and in the optimal static solution for several instances of this type of scenario as λ takes on values between 0 and 1. These four were chosen as being qualitatively representative of the types of behavior displayed by these data sets. These plots demonstrate that the amount of demand the mobile facilities are able to service relative to the amount of demand serviced by the optimal static placement of facilities increases as the length of time each event point generates demand decreases, and the overall demand profile becomes more "spiky." It should also be noted that as the coverage radius decreases, each facility is able to cover less nearby demand while at a location, and the quality of our heuristic solution for the MFRP relative to the static solution should be expected to increase.

6.3. Time-Discretized IDMIP

In general, the TDMIP given by (18)–(23) is computationally intractable for any problem of reasonable size. As mentioned previously, we solved the TDMIP using $\bigcup_{l\in L} S_l$ as the set of times defining the periods in P in ILOG OPL 5.2. Even for small problems, $|P|$ can be very large. Table 6 displays a comparison of the demand serviced in the routes generated with the TDMIP and with the sequential routing heuristic with local search for small instances of the problem with two mobile facilities and four or five event points and locations. We specified in each of these scenarios that for location l, E_l must contain the two closest event points, in addition to any event points within the coverage range from location l. In implementing the TDMIP, we found that problems with more than two

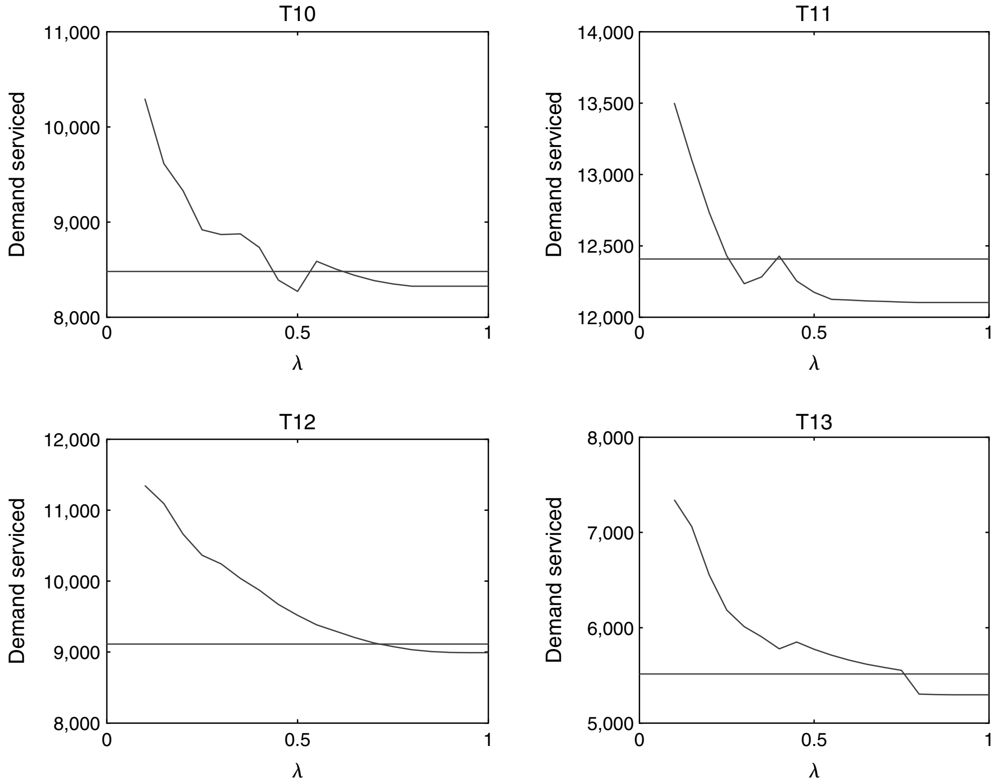

Figure 4 Results from Four Scenarios of the Third Type

Notes. The horizontal line displays the amount of demand serviced in the optimal static solution, which is equal for every value of λ. The second curve displays the demand serviced in the solution to the MFRP generated by the sequential routing heuristic with local search for several scenarios as λ takes on values between 0 and 1. Each of these data sets has 25 events and 75 locations. T10 has 5 mobile facilities, T11 has 10 mobile facilities, T12 has 7 mobile facilities, and T13 has 3 mobile facilities.

Table 6 A Comparison of the Demand Serviced by Solutions Obtained by Solving the TDMIP Given by (18)–(23) and the Demand Serviced by the Solutions Obtained by the Sequential Routing Heuristic with Local Search

Data set	$\|M\|$	$\|L\|$	$\|E\|$	TDMIP solution	TDMIP runtime (s)	Solution gap (%)	Sequential heuristic with local search	Percentage of TDMIP (%)
MC10.0	2	4	4	1,310.40	31.29	0.00	1,306.73	99.72
MC10.1	2	4	4	1,148.87	2.32	0.00	1,148.87	100.00
MC10.2	2	4	4	711.53	6.90	0.00	701.97	98.66
MC10.3	2	4	4	1,365.70	1.60	0.00	1,334.22	97.69
MC10.4	2	4	4	1,826.99	2.31	0.00	1,826.99	100.00
MC10.5	2	4	4	1,642.82	30.59	0.00	1,642.82	100.00
MC10.6	2	4	4	641.21	3.07	0.00	623.48	97.23
MC10.7	2	4	4	749.18	2.57	0.00	749.18	100.00
MC10.8	2	4	4	1,251.38	3.00	0.00	1,189.16	95.03
MC10.9	2	4	4	515.86	3.26	0.00	515.86	100.00
MC11.0	2	5	5	1,409.38	14,400.00	1.54	1,391.01	98.70
MC11.1	2	5	5	1,318.38	2,522.28	0.00	1,315.68	99.79
MC11.2	2	5	5	1,300.05	14,400.00	0.19	1,300.05	100.00
MC11.3	2	5	5	1,436.13	14,400.00	3.32	1,387.75	97.23
MC11.4	2	5	5	1,396.10	1,053.80	0.00	1,389.24	99.51
MC11.5	2	5	5	1,211.44	2,999.03	0.00	1,211.44	100.00
MC11.6	2	5	5	711.00	19.79	0.00	711.00	100.00
MC11.7	2	5	5	998.24	14,400.00	2.32	998.24	100.00
MC11.8	2	5	5	690.14	7,483.68	0.00	658.63	95.43
MC11.9	2	5	5	1,329.55	8,656.25	0.00	1,329.55	100.00

mobile facilities, more than five event points or locations, and more than five pieces in each piecewise-constant demand profile were too large to load into memory. Furthermore, some MFRP instances with five event points and five locations were difficult to solve to optimality. Because each event point in a realistic scenario only produces demand during a single interval of time, a realistic data set with only four or five event points could easily have no more than two events generating demand at any given time. Consequently, we present results for mathematically challenging problems generated having two mobile facilities, four or five locations, and four or five event points, with each having between two and five pieces in their piecewise-constant demand profile.

Even when considering problems of such a small size, the TDMIP can take a long time to solve. Consequently, the TDMIP was terminated if it had not finished running after four hours and the best solution found was taken. The solution gap produced by the TDMIP at termination is also displayed in Table 6. The table does not display the runtimes of the sequential heuristic with local search because it was under 0.01 seconds in all instances. The solutions to the MFRP given by the sequential heuristic with local search are competitive with the solutions found by solving the TDMIP. The heuristic solutions to these instances were computed almost instantly, whereas the TDMIP sometimes could not be solved within four hours. In only four of the larger test MFRP instances was the optimal solution to the TDMIP found in under one hour, running in 20 seconds, 18 minutes, 43 minutes, and 50 minutes, respectively.

The inability to load the TDMIP into memory for problems when restricting mobile facilities to depart from or arrive at a location l at times in S_l and the large amount of time it takes to solve small data sets with the TDMIP suggest that in a practical setting, heuristics are the best choice to determine routes for mobile facilities in the MFRP. We used our sequential routing heuristic with local search to generate routes for these problems. For each of these data sets, our heuristic executed almost instantly, returning a runtime of under 0.01 seconds. The solutions generated by our heuristics were competitive with the solutions produced by the TDMIP, on average servicing 98.95% of the amount of demand serviced in the routes computed with the TDMIP. Because there are only two mobile facilities, we believe that the TDMIP solution is near optimal for these small instances. The very fact that the heuristics were on average able to service 98.95% of the demand serviced by the TDMIP is a strong endorsement of our heuristic methods. We note that the data sets that the TDMIP could solve are in some sense outside the scope of the problem. Because these data sets are very small in size, the rate that demand is generated by each event point varies less dramatically than in the larger data sets, which are more representative of the type of demand profiles a mobile facility would encounter in practice, and for which our heuristics are designed.

7. Conclusion

In this paper, we introduced the MFRP. We showed that the optimal route for a single mobile facility may be found in polynomial time, although in general the MFRP is NP-hard. We presented several heuristics for routing mobile facilities to maximize the amount of demand serviced. Our computational results on a variety of data sets confirm the effectiveness of our heuristics. In comparison to the optimal placement of fixed facilities, we generated routes for a fleet of mobile facilities that service larger amounts of demand when demand levels are more "spiky." In addition, the heuristics performed competitively with a time-discretizated MIP formulation of the MFRP, which could only be solved for small instances of the problem.

A number of related problems have the potential to be fruitful avenues of future research. For example, no exact method is known for solving the MFRP. Such a method would be useful in evaluating the quality of heuristic solutions. Practical considerations of routing mobile facilities may introduce additional constraints that would affect the heuristics proposed in this paper. For example, to increase customer satisfaction and ensure that customers have an adequate opportunity to use the services provided by a mobile facility, operators of mobile facilities may wish to ensure that minimum levels of service are provided at each stop. To do so, operators may wish to specify a minimum amount of time that a mobile facility must spend at each location or a minimum amount of demand that must be serviced at each location. Alternatively, an operator may wish to specify a maximum amount of time that a mobile facility can spend at each location in an attempt to maximize the number of event points serviced over the planning horizon.

Opportunities exist to study applications where the model for operating mobile facilities may differ slightly from the MFRP. For example, the MFRP assumes that the instantaneous demand function of each event point is known for the entire planning horizon. In practice, the rate that demand is generated at an event point may deviate from a forecast due to unforeseen circumstances. Event points may also appear stochastically during the planning horizon if the population unexpectedly relocates, or for other reasons, such as the failing of a fixed base station in a cellular network. The addition of stochastic elements to the MFRP may provide an improved decision-making tool in such environments. The MFRP assumes that if demand is not serviced, then it is lost. However, there may be applications where demand for services at an event point accumulates when not being serviced by a mobile facility.

References

Alexy, B. B., C. A. Elnitsky. 1996. Community outreach: Rural mobile health unit. *J. Nursing Admin.* **26**(12) 38–42.

Ballou, R. H. 1968. Dynamic warehouse location analysis. *J. Marketing Res.* **5**(3) 271–276.

Bespamyatnikh, S., B. Bhattacharya, D. Kirkpatrick, M. Segal. 2000. Mobile facility location. *Proc. 4th Internat. Workshop Discrete Algorithms Methods Mobile Comput. Comm.* ACM, New York, 46–53.

Brotcorne, L., G. Laporte, F. Semet. 2003. Ambulance location and relocation models. *Eur. J. Oper. Res.* **147**(3) 451–463.

Canel, C., B. M. Khumawala, J. Law, A. Loh. 2001. An algorithm for the capacitated, multi-commodity multi-period facility location problem. *Comput. Oper. Res.* **28**(5) 411–427.

Church, R. L., D. M. Stoms, F. W. Davis. 1996. Reserve selection as a maximal covering location problem. *Biol. Conservation* **76**(2) 105–112.

Current, J. R., D. A. Schilling. 1989. The covering salesman problem. *Transportation Sci.* **23**(3) 208–213.

Current, J. R., D. A. Schilling. 1994. The median tour and maximal covering tour problems: Formulations and heuristics. *Eur. J. Oper. Res.* **73** 114–126.

Current, J. R., J. E. Storbeck. 1988. Capacitated covering models. *Environ. Planning B: Planning Design* **15**(2) 153–163.

Current, J., H. Pirkul, E. Rolland. 1994. Efficient algorithms for solving the shortest covering path problem. *Transportation Sci.* **28**(4) 317–327.

Current, J. R., C. ReVelle, J. Cohon. 1984. The shortest covering path problem: An application of locational constraints to network design. *J. Regional Sci.* **24**(2) 161–183.

Current, J. R., C. S. ReVelle, J. L. Cohon. 1985. The maximum covering/shortest path problem: A multiobjective network design and routing formulation. *Eur. J. Oper. Res.* **21** 189–199.

Daskin, M. 1995. *Network and Discrete Location*. John Wiley & Sons, Inc., New York.

Durocher, S., D. Kirkpatrick. 2006. The Steiner centre of a set of points: Stability, eccentricity, and applications to mobile facility location. *Internat. J. Comput. Geometry Appl.* **16**(4) 345–371.

Erlenkotter, D. 1981. A comparative study of approaches to dynamic location problems. *Eur. J. Oper. Res.* **6**(2) 133–143.

Gendreau, M., G. Laporte, F. Semet. 2001. A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. *Parallel Comput.* **27**(12) 1641–1653.

Ghiani, G., F. Guerriero, G. Laporte, R. Musmanno. 2003. Real-time vehicle routing: Solution concepts, algorithms and parallel computing strategies. *Eur. J. Oper. Res.* **151**(1) 1–11.

Hong Kong Post. 2008. Mobile post office locations. Accessed January 21, 2009, http://www.hongkongpost.com/eng/locations/mobile/index.htm.

Jay, M. L. 2006. Katrina leaves wireless industry better prepared for next disaster. http://www.ctia.org/content/index.cfm/AID/10410.

Karasakal, O., E. K. Karasakal. 2004. A maximal covering location model in the presence of partial coverage. *Comput. Oper. Res.* **31**(9) 1515–1526.

Megiddo, N., E. Zemel, S. L. Hakimi. 1983. The maximum coverage location problem. *SIAM J. Algebraic Discrete Methods* **4**(2) 253–261.

Mirchandani, P. B., R. L. Francis. 1990. *Discrete Location Theory*. Wiley-Interscience, New York.

Moss, M., B. Alan, J. Farren. 2008. COLTs and COWs to help with cell phone demand. *WTOP Radio*. Accessed December 10, http://www.wtop.com/?nid=25&sid=1544462.

Pirkul, H., D. A. Schilling. 1991. The maximum covering location problem with capacities on total workload. *Management Sci.* **37**(2) 233–248.

United Kingdom Department of Transportation. 2009. Mobile post office. Accessed January 21, 2009, http://www.dft.gov.uk/pgr/regional/ltp/accessibility/knowledgepool/aos/mobilepostoffice.

Van Roy, T. J., D. Erlenkotter. 1982. A dual-based procedure for dynamic facility location. *Management Sci.* **28**(10) 1091–1105.

Wendland, M. 2006. Cell capacity will be super during the big game in Detroit. *Detroit Free Press* (February 1).

Wesolowsky, G. O. 1973. Dynamic facility location. *Management Sci.* **19**(11) 1241–1248.

Wesolowsky, G. O., W. G. Truscott. 1975. The multiperiod location-allocation problem with relocation of facilties. *Management Sci.* **22**(1) 57–65.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics* **1** 80–83.