Routing for Relief Efforts

Ann Melissa Campbell

Department of Management Sciences, University of Iowa

108 John Pappajohn Business Building, Iowa City, Iowa 52242-1000

ann-campbell@uiowa.edu

Dieter Vandenbussche

William Hermann

Mechanical and Industrial Engineering, University of Illinois

1206 West Green St., Urbana, IL 61801

dieterv@uiuc.edu

hermann2@uiuc.edu

Abstract

In the aftermath of a large disaster, the routing of vehicles carrying critical supplies can greatly impact the arrival times to those in need. Since it is critical that the deliveries are both fast and fair to those being served, it is not clear that the classic cost-minimizing routing problems properly reflect the priorities relevant in disaster relief. In this paper, we take the first steps in developing new methodologies for these problems. We focus specifically on two alternative objective functions for the TSP and VRP: one that minimizes the maximum arrival time (minmax) and one that minimizes the average arrival time (minavg). To demonstrate the potential impact of using these new objective functions, we bound the worst case performance of optimal TSP solutions with respect to these new variants and extend these bounds to include multiple vehicles and vehicle capacity. Similarly, we examine the potential increase in routing costs that result from using these alternate objectives. We present solution approaches for these two variants of the TSP and VRP which are based on well known insertion and local search techniques. These are used in a series of computational experiments to help identify the types of instances where TSP and VRP solutions can be significantly different from optimal minmax and minavg solutions.

1 Introduction

In recent years, several regions around the world have fallen victim to natural disasters of a massive scale. In 2004, an earthquake in the Indian Ocean spawned a series of tsunamis which caused damage and loss of life as far away as 5000 miles from the epicenter of the earthquake [1]. In 2005, Hurricane Katrina affected a region of approximately 90,000 square miles which is about the same size as Great Britain [2]. The enormous scale of these disasters has brought attention to the need for methodology and technology for effectively managing relief supply chains. A recent report published by the Fritz Institute, a nonprofit agency concerned with the logistics of relief efforts, indicated that most aid organizations involved after the 2004 tsunami were significantly lacking in logistics expertise and technology [26]. A European ambassador attending a UN-sponsored conference for donor nations said: "We don't need a donors' conference; we need a logistics conference" [43]. While a great deal of research and technology is available for commercial supply chains, the challenges associated with managing a humanitarian relief chain following a large-scale emergency are often quite different than in commercial applications. Beamon [9] cites several examples of these differences including the unpredictability of demand for humanitarian aid, where commercial supply chains are designed around a known set of customers with relatively predictable demand patterns. Another key difference is in the strategic goal of each supply chain. Where commercial supply chains are focused on quality and profitability, humanitarian supply chains are usually focused on minimizing loss of life and suffering. Also, when goods are distributed to the public, equity and fairness become much more of a concern than in commercial applications.

With these ideas in mind, we became interested in exploring how such differences in strategic goals could impact the routing of the vehicles delivering emergency aid or if it would impact the routing at all. There are many existing tools for solving vehicle routing problems, and most focus on minimizing the total distance travelled, which serves as a proxy for total cost. After a disaster, the arrival time of relief supplies at the affected communities clearly impacts the survival rate of the citizens and the amount of suffering. In the routes created by optimizing with respect to total distance, some communities may be served significantly later than others in order to save on total cost. Since it is critical that the deliveries are both fast and fair to those being served in a relief context, such observations suggest that using service-based objective functions may better reflect the different priorities and strategic goals found in delivering humanitarian aid. Specifically, we examine two alternative objective functions for the classic traveling salesman problem (TSP) and vehicle routing problem (VRP): one that minimizes the maximum arrival time (minmax routing) and one that minimizes the average arrival time (minavg routing). These objective functions are based on the literature on fairness, which has developed primarily outside of a routing context [33].

Even though these objective functions are by definition different than the traditional objectives, it is not clear how much impact switching to one of these objectives will have on the solutions. Due to the convenience of using existing cost-based routing tools, it is important to verify that such a switch has the potential to make significant improvements in service to those affected by the disaster in order to justify such a change. First, we will bound the worst case performance of optimal TSP solutions for these new objectives and then extend these results to include multiple vehicles and vehicle capacity. Next, we will present solution approaches for these two variants of the TSP and VRP which are based on well known insertion [13] and local search techniques. These tools are used in a series of computational experiments that help identify the types of instances where TSP and VRP solutions can be significantly different from optimal minmax and minavg solutions. Both the bounds and computational experiments will demonstrate that the optimal solutions for traditional routing problems can be significantly different than those found with these alternative objective functions, and that optimal solutions for the two alternatives can also differ. Importantly, both in developing our theoretical bounds and in our computational experiments, we will also examine the impact on total cost as a result of using these alternate objectives. We emphasize that the tools developed in this paper are not intended to be sufficient for solving realistic routing problems in a relief context, rather our intent is to demonstrate the potential impact of using alternative objective measures when developing techniques to be used in practice. This is simply a first step in developing better tools for the delivery of humanitarian aid. Our paper concludes with a discussion of the some of the many issues that remain to be addressed.

2 Literature Review

This paper is part of a recent, emerging effort to apply operations research techniques to facilitate the delivery of humanitarian aid. In addition to [9], examples include a study of the inventory systems for disaster relief [10] and a worldwide facility location model to site warehouses in anticipation of major emergencies [6]. Özdamar et al. [36] describe a model that coordinates deliveries of supplies between different supply depots in the context of a relief operation, and Barbarosoglu et al. [7] look at how to effectively use helicopters in a relief operation. Long [30] discusses the strategic and tactical issues that relief organizations face in preparing for and responding to disasters. Long [30] points out some of the unique issues to relief supply chain management, including clearing supply routes and the critical role of real-time non-computerized communications systems. Another key area of concern in providing relief after a disaster is coordination of the organizations involved. Pettit and Beresford [37], for example, model the relationships between participating bodies including military and non-military organizations. Other efforts to examine the supply chains required by relief efforts have been carried out by various nonprofit agencies, including the Fritz Institute [25].

There are papers that focus specifically on the dispatch of vehicles after a disaster. For example, Shen et al. [44] analyze and develop solution methods for a stochastic vehicle routing problem motivated by strategic planning for large-scale emergencies, where the total quantity of unmet demand is minimized. The model is managed from a two-stage perspective, where pre-planned routes are designed in the first stage, and adjustments to these routes are made in the second stage. Desai et al. [16] propose a model that decides the number of emergency responders, such as police or firefighters, to dispatch to each of several regions affected by a major disaster. The responders sent to a particular region mitigate the risks resulting from the disaster, and the model's objective aims to distribute those risks equitably across regions.

In a relief context, fairness and equity are often important factors. Incorporating equity in operations research models , though, is not new. Mandell [31], for example, provides an overview of modeling equity in public systems, and Larson [29] discusses equity as a factor in the perception of justice in queuing systems.

Equity has been a significant focus in the location of public facilities, as discussed in [33, 34]. The typical assumption in this literature is that each client is interested in minimizing his/her own outcome $f_i(x)$ [34]. This function can measure a variety of outcomes in the location context, but usually represents distance traveled or travel time to the nearest facility. For the weighted location problem, both center and median objectives have been studied extensively, which minimize the largest f_i value and the average f_i value, respectively [23, 20, 32, 40, 39]. Both of these objectives capture the idea of minimizing individual outcomes while controlling inequity but remain very simple to evaluate. They serve as the basis of our minmax and minsum routing objective functions, which we will discuss in further detail in sections 5.1 and 5.2. In location studies, inequality measures are often combined with a more traditional efficiency based objective function in a bi-criteria approach, such as in [34]. This enables the study of location problems to incorporate inequality measures developed originally for problems in the field of economics. These measures were created to evaluate the fairness of a distribution of a particular commodity or resource. The most popular of these measures is the Gini index [17]. For more details on inequality measurement in economics, see the book by Sen [42].

In addition to location problems, equitable objectives are also considered in communication networks, where users of the network are given allocations of bandwidth. In these applications, equitable solutions are often characterized through a concept called proportional fairness, see e.g. [27] for details. In the context of communication networks, Ogryczak et al. [35] discuss an axiomatic approach to the question of fairness, presenting various objective functions that reflect the need for both efficiency and equity.

Several of the objectives we will consider can also be found in the scheduling literature. For instance, the analogue of minimizing the makespan in scheduling (see [38]) is minimizing the arrival time at the last community to receive aid. Similarly, minimizing the sum of weighted completion times is analogous to minimizing a weighted sum of arrival times.

There is little research on alternate objectives for vehicle routing, and we have not found any work that compares using one of these alternate objectives to minimizing total travel time. In [5], the authors consider the 2-TSP where two salesmen must together visit all of the nodes on a tree. The objective is to minimize the length of the longest of the two tours. Even on a tree, this problem is NP-hard, and an approximation algorithm is provided. França et al. [19] present a heuristic for minimizing the length of the longest tour in an m-TSP. Similarly, the problem solved in [3] is a particular instance of a VRP where the length of the longest of four routes is minimized. The emphasis of [3] is on developing specialized cutting planes and a distributed search algorithm. Equity has also appeared in the routing literature as a constraint. In [15], the authors propose a model for an overnight delivery business that includes a constraint to limit the permissible deviation from the average delivery time.

There are some related routing problems, including routing for hazardous materials and school bus routing, that incorporate equity. In designing routes for the delivery of hazardous materials (see e.g. [22]), the focus is on balancing risk to the regions visited on the routes. This focus makes the resulting problems very different than the ones addressed here. There is a dual focus on efficiency and equity in the case of school bus routing. It is important that school bus routing be efficient in terms of total mileage, but it is also important that there are not students who spend an inordinate amount of time on a bus. In [11], the authors optimize both efficiency and equity through a multi-criteria objective function.

3 Assumptions and Definitions

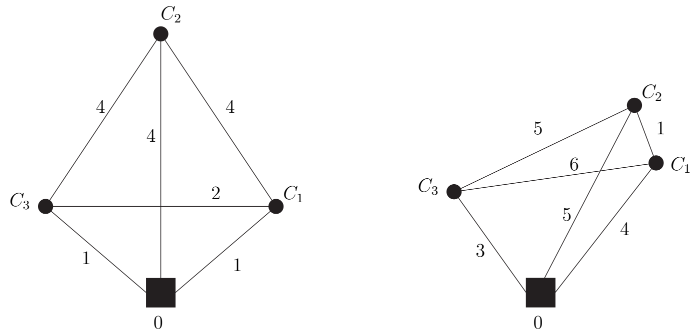
In this paper, we make the following assumptions. We will be developing tours for customers labeled 1 to n. Since we may think of the customers as being located on a network, we will also refer to them as nodes. Because the tours are designed to deliver a particular commodity to the customers, we assume all tours will start from a designated depot, labeled 0. Thus, the arrival time at a customer will be based on travel time from this depot, and the tour will be directed. For simplicity, we will assume all tours start at time 0, and customers do not have time windows limiting delivery times. We assume the travel times are nonnegative and satisfy the triangle inequality. We do not explicitly model service times, but these may easily be included in the travel times. When we consider capacity constraints, we will assume that each customer has unit demand, so that the vehicle capacity indicates how many customers can be visited on one route. For convenience, rather than minimizing the average arrival time, we will use the equivalent objective of minimizing the sum of arrival times. Henceforth, we will refer to the two objectives of interest as minmax and minsum.

The minmax routing problem minimizes the latest time any customer receives service, where the traditional TSP or VRP focuses on the travel time or travel distance for the complete roundtrip from the depot. Similarly, an optimal minsum route minimizes the sum of the arrival times for all of the customers, which, as in the case of minmax, does not include the return trip from the last customer back to the depot. In the delivery of aid supplies, it may be more important to quickly deliver supplies to those in need than it is to get the truck back to the depot quickly.

The advantage of the minsum objective over minmax is that the arrival times at all customers are reflected in the objective function. The minmax problem does not reflect, for example, the second latest arrival time.

In this paper, we will address the performance of an optimal solution for each of the objectives with respect to one or more of the other objectives. To do this, we require some notation. We use $c(TSP)$ to denote the length of an optimal TSP tour. Similarly, we use $c(VRP)_Q^k$ to denote the optimal value of the usual VRP problem with k vehicles and capacity Q on each vehicle. If either k or Q are omitted, we assume their default values to be 1 and ∞, respectively.

$la(MM)_Q^k$ denotes the latest arrival time in an optimal minmax routing when k vehicles with capacity Q are available. Given a solution that minimizes the usual VRP objective, i.e. a solution that achieves $c(VRP)_Q^k$, we denote the latest arrival time as $la(VRP)_Q^k$. Note this is a slight



(a) TSP is different from minmax and minsum

(b) minmax is different from minsum

Figure 1: Different objectives yield different solutions

abuse of notation, as alternate optima to the VRP objective may have different latest arrival times. Unless stated otherwise, $la(VRP)_Q^k$ will refer to any of the optimal solutions of the VRP. We define $la(TSP)$ analogously.

We refer to $sa(MS)_Q^k$ as the minimum sum of arrival times when using at most k vehicles, each with capacity Q. We can also consider quantities such as $sa(VRP)_Q^k$ and $sa(TSP)$. Other quantities of interest will be $c(MM)_Q^k$, which represents the value of an optimal minmax solution with respect to the total duration objective. Similarly, we denote $c(MS)_Q^k$ as the total travel time for an optimal minsum solution.

To illustrate our notation and to better illustrate the difference between minmax or minsum routing and the traditional TSP, as well as between minmax and minsum, consider Figures 1(a) and 1(b). Figure 1(a) is a graph where an optimal TSP tour is $0 - C_1 - C_2 - C_3 - 0$. The length of the tour is 10, the latest arrival time is 9, and the sum of arrival times is 15, i.e. $c(TSP) = 10$, $la(TSP) = 9$, and $sa(TSP) = 15$. Optimal minmax and minsum routes traverse $0 - C_1 - C_3 - C_2$ instead, with $c(MM) = c(MS) = 11$, $la(MM) = 7$, and $sa(MS) = 11$.

Optimal minmax and minsum routes can also be different. In Figure 1(b), we show a graph where the optimal minmax route is $0 - C_3 - C_2 - C_1$ ($la(MM) = 9$, $sa(MM) = 20$), but the optimal minsum route is $0 - C_1 - C_2 - C_3$ ($la(MS) = 10$, $sa(MS) = 19$). These examples illustrate that the sets of optimal solutions for the different objectives need not be equal, although they may intersect.

4 Integer Programming Formulations

Like TSP and VRP, the minmax and minsum routing problems can be formulated as mixed integer programming problems. We follow the notation and formulation introduced by Bard et al. [8]. Denote the set of customers with $N := \{1, \ldots, n\}$, the depot with 0, and let $N_0 = N \cup \{0\}$. Define t_{ij} as the travel time between nodes i and j in N_0. The variables x_{ij} are 0-1 variables that indicate if a vehicle travels from node i to j and a_i denotes the arrival time at customer i. The uncapacitated VRP can be formulated as

$$\min \sum_{i,j \in N_0} t_{ij} x_{ij} \tag{1}$$

$$\text{subject to} \sum_{j \in N_0} x_{ij} = 1 \quad \forall i \in N \tag{2}$$

$$\sum_{j \in N_0} x_{ij} - \sum_{j \in N_0} x_{ji} = 0 \quad \forall i \in N_0 \tag{3}$$

$$t_{ij} + a_i \leq a_j + T(1 - x_{ij}) \quad \forall i, j \in N \tag{4}$$

$$a_i \geq t_{0i} x_{0i} \quad \forall i \in N \tag{5}$$

$$x_{ij} \in \{0, 1\} \quad \forall i, j \in N_0, \tag{6}$$

where $T > 0$ is sufficiently large. The constraints (2) ensure that each customer is visited by a vehicle. Equations (3) are standard flow balance constraints and ensure that all routes return to the depot. Inequalities (4) and (5) make sure the variables a represent the appropriate arrival times. Furthermore, inequalities (4) ensure there are no subtours that do not pass through node 0. By adding appropriate variables and constraints, we can easily enforce capacities on the vehicles.

In order to model the minsum objective, we replace the objective (1) with $\sum_{i \in N} a_i$. To solve the minmax problem, we add an auxiliary variable $\bar{a}$ to represent the latest arrival time at a customer. For the minmax problem, we will minimize $\bar{a}$ and add the constraints $a_i \leq \bar{a} \; \forall i \in N$.

We found that these models are very difficult to solve with standard IP solvers. While more advanced methods using cutting planes such as those presented in [8] would certainly improve the performance, the goal of this paper is to obtain a better understanding of the impact of choosing one of these alternative objective functions. Future work may consider developing more sophisticated exact solution approaches for these alternate objectives.

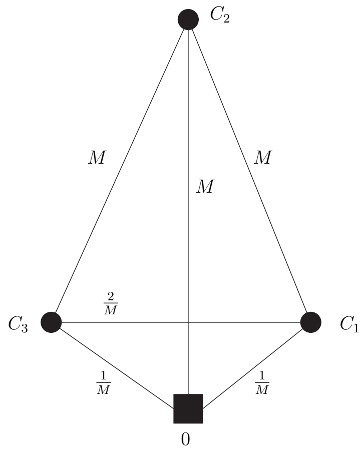


Figure 2: Bound $la(TSP) \leq 2la(MM)$ is tight

5 Bounds

5.1 Minmax

In this section, we examine the relationship between the total duration objective and the minmax objective. We begin with the uncapacitated, one vehicle case.

5.1.1 Latest arrival: TSP versus minmax

Proposition 1. *$la(TSP) \leq 2la(MM)$.*

Proof. Since the maximum arrival time in an optimal minmax tour represents the length of a Hamiltonian path, which is a special case of a spanning tree, we know that $la(MM) \geq MST$, where MST is the total length of a minimum spanning tree. Under the triangle inequality, it is well known that $c(TSP) \leq 2MST$. Therefore, since $la(TSP) \leq c(TSP)$, using an optimal TSP tour to solve a minmax variant will yield a solution no more than 2 times the optimal solution to the minmax problem. □

We can get arbitrarily close to this bound using the example pictured in Figure 2. It is easy to see that the optimal TSP tour in Figure 2 is $0 - C_1 - C_2 - C_3 - 0$. Note that regardless of the orientation of the tour, the last customer will have an arrival time of $\frac{1}{M} + 2M$ whereas the sequence $0 - C_1 - C_3 - C_2$ yields $la(MM) = \frac{3}{M} + M$. Taking the limit as $M \to \infty$, we get

$$\lim_{M \to \infty} \frac{\frac{1}{M} + 2M}{\frac{3}{M} + M} = 2,$$

which shows that one can come arbitrarily close to this ratio of 2 by increasing M.

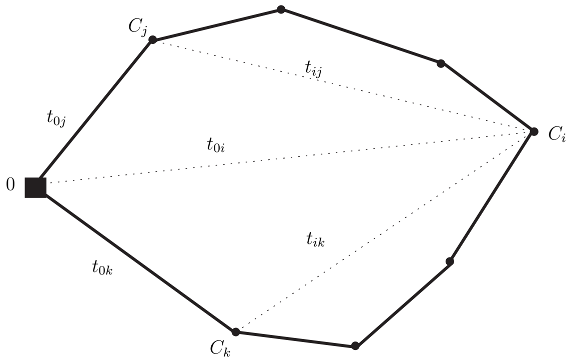


Figure 3: Optimal TSP tour in proof of Proposition 2

When put into the context of vehicle routing for disaster relief, Proposition 1 is significant, indicating that an optimal TSP tour could double the time needed to reach the last group compared to an optimal minmax solution.

5.1.2 Tour length: TSP vs. minmax

Next, we examine the quality of minmax tours with respect to the traditional TSP objective. This will help us determine how much we can lose in overall efficiency, i.e. total travel time, by using minmax as the objective. We show that the worst case ratio is $\frac{3}{2}$, indicating that an optimal minmax tour increases the total roundtrip traveled by at most 50 %.

Proposition 2. *$c(MM) \leq \frac{3}{2}c(TSP)$.*

Proof. Consider an optimal TSP tour as displayed in Figure 3. We denote the first and last customers on this tour as C_k and C_j. Denote the last customer served on the optimal minmax route as C_i, where $c(MM) = la(MM) + t_{0i}$. The minmax route is the shortest Hamiltonian path starting from the depot. Since $c(TSP) - t_{0j}$ is the length of a Hamiltonian path starting from the depot, then $la(MM) \leq c(TSP) - t_{0j}$. Similarly, $la(MM) \leq c(TSP) - t_{0k}$. Now proceed by contradiction and assume $c(MM) > \frac{3}{2}c(TSP)$. This implies $la(MM) + t_{0i} > \frac{3}{2}c(TSP)$. Subtracting $[la(MM) + t_{0j}]$ from both sides and using the triangle inequality, we find:

$$t_{ij} \geq t_{0i} - t_{0j} > \frac{1}{2}c(TSP) + c(TSP) - [la(MM) + t_{0j}] \geq \frac{1}{2}c(TSP).$$

Hence, $t_{ij} > \frac{1}{2}c(TSP)$ and a similar relation establishes $t_{ik} > \frac{1}{2}c(TSP)$. Adding these two inequalities gives $t_{ij} + t_{ik} > c(TSP)$, which is a contradiction, since t_{ij} and t_{ik} are lower bounds on the lengths of two disjoint subpaths of the tour. This establishes the result. $\square$

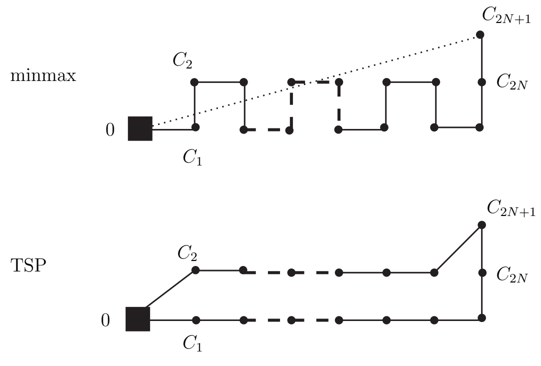


Figure 4: Bound of minmax versus TSP is tight

The optimal TSP and minmax routes given in Figure 4 demonstrate that this bound is tight. In this example, all $n = 2N + 1$ customers lie on a grid in the Euclidean plane where each vertical and horizontal edge has length 1. The dashed segments in each route represent a portion of the route that is repeated as many times as N requires. It is a simple matter to check that $c(TSP) = 2N + 2\sqrt{2}$ and $c(MM) = 2N + 1 + \sqrt{N^2 + 4}$. This gives the following ratio as we increase N:

$$\lim_{N\to\infty} \frac{2N + 1 + \sqrt{N^2 + 4}}{2N + 2\sqrt{2}} = \frac{3}{2}.$$

5.1.3 Multiple Vehicles Latest Arrival: VRP versus minmax

It is easy to see that if there are no capacity constraints, then the usual VRP objective has no incentive to use more than one vehicle. On the other hand, the latest arrival time can be significantly improved with additional vehicles since customers can be served simultaneously. The following proposition bounds how much the minmax objective can improve as a function of the number of vehicles, k.

Proposition 3. $la(MM)^k \geq \max\{\frac{1}{2k-1} la(MM), \max_{i\in N} t_{0i}\}$.

Proof. Consider an optimal set of tours that achieves $la(MM)^k$. No matter how many vehicles there are, the optimal latest arrival time must be larger than the longest direct trip from the depot, i.e. $la(MM)^k \geq t_{0i} \forall i \in N$. Note that the minmax objective has no incentive to use fewer than k tours, so we can assume that $1, \ldots, k$ are the last nodes on each of the k tours. Recall that a_i is the arrival time at node i and assume that $a_k = la(MM)^k$, that is $k \in \text{argmax}_i\{a_i\}$. We can construct a single tour whose minmax value is less than or equal to $(2k - 1)a_k$, which proves the desired result. We traverse each tour in sequence: starting at the depot, traveling to its endpoint, and return to the depot to start the next tour, leaving the tour with endpoint a_k last. By the

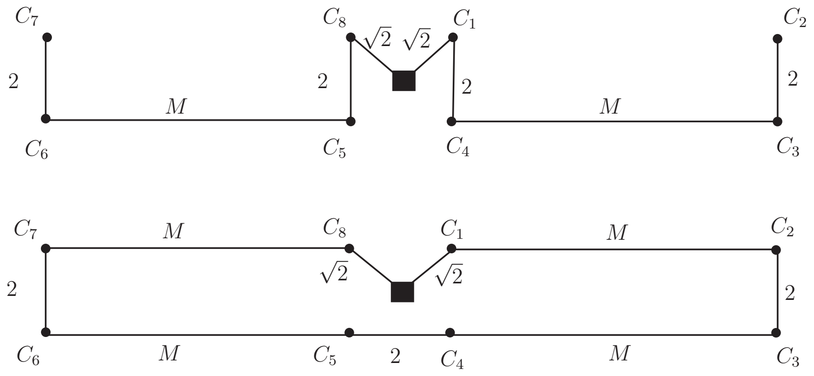


Figure 5: VRP versus minmax with 2 vehicles

triangle inequality and since $a_j \leq a_k\ \forall j \neq k$, we have that the length of this single tour is less than or equal to $\sum_{j=1}^{k-1} 2a_j + a_k \leq (2(k-1)+1)a_k = (2k-1)la(MM)^k$. $\square$

It is not difficult to see that this bound is tight. Consider an instance with k customers where each customer is distance M from the depot, and the distance between any pair of customers is $2M$. Clearly, using k vehicles, the maximum arrival time is M. If we have only one vehicle, the maximum arrival time is $2M(k-1)+M = (2k-1)la(MM)^k$.

Proposition 4. *$la(VRP)^k = la(TSP) \leq 2k\, la(MM)^k$.*

Proof. Suppose the optimal minmax solution consists of k routes and N customers with $k \leq N$. Using the same notation as given in Proposition 3, let $a_k = la(MM)^k$ and a_i be the arrival time at the last node on tour i for $i \in \{1, \ldots k\}$. Since a_k is the maximum arrival time, then $a_i \leq a_k$ for all i. Summing over the arrival times of the last customers on each tour gives: $\sum_{i=1}^{k} a_i \leq ka_k = k\, la(MM)^k$. Multiplying $\sum_{i=1}^{k} a_i$ by 2 gives the total length of a tour that visits all customers and returns to the depot k times by retracing each route. Clearly such a tour must have total length bounded below by the optimal TSP solution. Hence, $c(VRP)^k \leq 2\sum_{i=1}^{k} a_i \leq 2k\, la(MM)^k$. When the vehicles are uncapacitated, the total cost objective provides no incentive to use more than one vehicle; hence we may assume that $c(TSP) = c(VRP)^k$. Since $la(VRP)^k \leq c(VRP)^k$ and $la(TSP) \leq c(TSP)$, the result follows. $\square$

The two graphs in Figure 5 demonstrate that this bound is tight for two uncapacitated vehicles. The upper graph shows the two routes of the optimal minmax tour $(0-C_1-C_4-C_3-C_2)$ and $(0-C_8-C_5-C_6-C_7)$. Here $la(MM)^2 = M+4+\sqrt{2}$. The lower graph gives the optimal VRP solution $(0-C_1-C_2-C_3-C_4-C_5-C_6-C_7-C_8-0)$ with $la(VRP)^2 = la(VRP) = 4M+6+\sqrt{2}$.

As the distances M grow to ∞, we have

$$\frac{la(VRP)^2}{la(MM)^2} = \frac{4M + 6 + \sqrt{2}}{M + 4 + \sqrt{2}} = 4 = 2k.$$

Propositions 3 and 4 demonstrate that the impact of optimizing with respect to alternative objectives increases with the number of vehicles. If additional vehicles are available, the costs associated with them may be justified by the huge potential reductions in arrival time. It is important to carefully evaluate these tradeoffs when determining the appropriate fleet size. Next, we evaluate the potential changes in route duration (cost) that may be required to create these reductions in arrival time.

5.1.4 Multiple Vehicles Tour Length: VRP versus minmax

Proposition 5. *$c(MM)^k \leq k\, c(VRP)^k = k\, c(TSP)$.*

Proof. We will take an optimal TSP tour and use it to construct two paths from the depot. Consider the midpoint of an optimal TSP. If this halfway point occurs along an edge, then delete this edge from the tour. We are left with two connected paths from the depot. If the midpoint occurs at a node, then arbitrarily assign the first half of the TSP tour to be the first path with the second half becoming the second path from the depot. Dividing the TSP tour in this fashion corresponds to creating two paths starting from the depot where the length of each path is less than or equal to $\frac{c(TSP)}{2}$. For $k \geq 2$, $la(MM)^k$ is a decomposition of nodes into k paths which start from the depot such that the longest path is minimized. Hence, $la(MM)^k \leq \frac{c(TSP)}{2}$ for $k \geq 2$. Following the notation as given in Proposition 4, a_k is the maximal arrival time of the optimal minmax solution with k routes. By use of the triangle inequality and the division of the TSP tour given above, we find the relation: $c(MM)^k = \sum_{i=1}^{k}(a_i + t_{0i}) \leq \sum_{i=1}^{k} 2a_i \leq 2\sum_{i=1}^{k} a_k \leq 2k\, la(MM)^k \leq k\, c(TSP) = k\, c(VRP)^k$. $\square$

It is easy to see this bound is tight. Consider an instance with k customers, all a large distance M from the depot and the distance between any two customers is some small $\epsilon > 0$. The VRP solution will use one vehicle, hence we have that $c(VRP)^k = 2M + (k-1)\epsilon$, while the minmax solution will send one vehicle to each customer, which results in $c(MM)^k = 2k\, M$. Hence as M grows large, $c(MM)^k$ approaches $k\, c(VRP)^k$.

Thus, using k vehicles may reduce arrival times by a factor of $2k$ but increase total duration by a factor of k.

5.1.5 Capacity

Next we will consider what happens when we explicitly consider capacity limitations for the vehicles. Recall that we have assumed each customer has uniform demand, so capacity here will reflect the maximum number of customers that can be served by each vehicle.

Since the proof of Proposition 3 does not change with the introduction of capacity, we have

Proposition 6. $la(MM)^k_Q \geq \max\{\frac{1}{2k-1} la(MM), \max_{i \in N} t_{0i}\}$.

This bound is tight with groups of Q customers co-located at a distance M from the depot and $2M$ from each other.

However, this does not mean that capacity has no impact on minmax solutions. When there are k vehicles, solutions achieving $la(MM)^k$ and $la(MM)^k_Q$ will use all k vehicles, but the customers may be allocated differently due to the capacity limitations. The following proposition provides a bound on the difference between $la(MM)^k$ and $la(MM)^k_Q$:

Proposition 7. $la(MM)^k_Q \leq (2k-1) la(MM)^k$.

Proof. The value for $la(MM)^k_Q$ must be less than $la(MM)$ since there are k tours, each visiting a subset of the customers in the one route solution. With triangle inequality, each tour will have a lower latest arrival time than the one complete tour. We know from Proposition 3 than $la(MM) \leq (2k-1) la(MM)^k$ which gives us our desired result. □

Consider the following example with $k = 2$, $Q = 2$, and $n = 4$. Three of the four customers are co-located at distance M from the depot, and the fourth customer is M from the depot and $2M$ from the other three. With two uncapacitated trucks, we can visit the three with one truck and the fourth with the other truck, yielding a latest arrival time of $la(MM)^2 = M$. Whereas with capacitated trucks, we can only visit two of the three co-located points with a single truck. To reach the other two, it will take $la(MM)^2_2 = M + 2M = 3M = 3la(MM)^2 = (2k-1) la(MM)^k$.

We note that Proposition 7 does not explicitly involve the value for Q. We expect the bound to be tighter, though, when Q is close to its minimum value of $\lceil \frac{n}{k} \rceil$. This is something we will evaluate in our computational experiments.

5.2 Minsum

5.2.1 Sum Arrival: TSP vs. minsum

In this section, we examine the performance of TSP solutions with respect to the minsum objective. Recall that a minsum objective reflects the sum of arrival times, so service times to all customers are considered. We will show that TSP gives a factor n approximation for the minsum objective.

Proposition 8. *$sa(TSP) \leq n\, sa(MS)$.*

Proof. Suppose that $0-1-2-\ldots-n-0$ is an optimal TSP tour. Since the orientation in which we traverse the tour affects the latest arrival time, the best minsum objective that can be obtained with this tour is

$$\begin{aligned} sa(TSP) &= \min\left\{nt_{01} + (n-1)t_{12} + \ldots + t_{n-1,n},\ nt_{0,n} + (n-1)t_{n-1,n} + \ldots + t_{12}\right\} \\ &\leq \frac{n}{2}\left(t_{01} + t_{12} + \ldots + t_{n-1,n} + t_{n0}\right) \\ &= \frac{n}{2}c(TSP) \leq nMST. \end{aligned}$$

Clearly, MST is less than or equal to the optimal minsum objective. Thus, we can conclude that using an optimal TSP tour to solve a minsum variant will yield a solution no more than n times the optimal value of the minsum problem. □

In Figure 6, we present an example of a graph that achieves a ratio of $\frac{n}{2}$. It is easy to check that the optimal TSP tour is $0 - C_{N+1} - C_1 - \ldots - C_N - C_{N+2} - 0$ and the optimal minsum tour is $0 - C_1 - \ldots - C_N - C_{N+2} - C_{N+1}$. Hence, the performance ratio with respect to the minsum objective is

$$\frac{(N+2)M + (N+1)M + N\epsilon + \ldots + 2\epsilon + M}{(N+2)\epsilon + (N+1)\epsilon + \ldots + 3\epsilon + 2M + 2M}$$

Taking the limit as $M \to \infty$, we get $\frac{N+2}{2} = \frac{n}{2}$.

Proposition 8 indicates that the sum of the arrival times in an optimal TSP solution can be significantly worse than in an optimal minsum route. This again confirms that a TSP tour may not be a very equitable solution and will likely yield much higher average service times to those in need.

We can also present an alternative bound to Proposition 8 that depends on the spread of the edge lengths, t_{ij}.

Proposition 9. *$sa(TSP) \leq \frac{\max_{i,j} t_{ij}}{\min_{i,j} t_{ij}} sa(MS)$.*

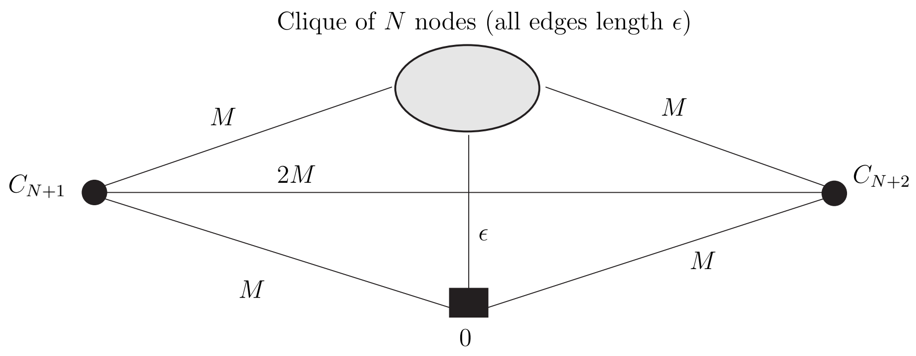


Figure 6: Factor $n/2$ ratio for TSP versus minsum

Proof. Suppose $0 - 1 - 2 - \ldots - n$ is an optimal minsum tour. We have

$$sa(MS) = \sum_{i=0}^{n-1}(n-i)t_{i,i+1} \geq \left(\sum_{i=0}^{n-1}(n-i)\right)\min_{i,j} t_{ij}.$$

Similarly, we can show that $sa(TSP) \leq \left(\sum_{i=0}^{n-1}(n-i)\right)\max_{i,j} t_{ij}$. The result follows from combining the two inequalities. $\square$

Note that an analogous result can be shown for the minmax objective. It is particularly relevant here because it gives us a condition when $sa(TSP)$ could be much closer to $sa(MS)$ than a factor of n. Specifically, if the difference between the longest and shortest edge is small, the two objectives can be close in value.

5.2.2 Tour Length: TSP vs minsum

We next show how much the tour length using the minsum solution may differ from the tour length for the TSP solution. Again, we are interested in determining how much total efficiency, in terms of total travel time, is lost by considering an alternative objective such as minsum. In particular, we show that the ratio $\frac{c(MS)}{c(TSP)}$ cannot be bounded by a constant ratio. To do this, we use the graph pictured in Figure 7. In this graph, we find the depot and K nodes located in the plane. Each node C_i will refer to a clique of size N_i, where the edge lengths within the clique are assumed to be 0. All other distances will be determined by assuming the cliques lie in the Euclidean plane. We assume that K is even, so that clique C_1 is on the right side. For $M > 2$ and integer and $N_i = M^{i-1}$, we will show that the optimal minsum route will traverse the solid line depicted in Figure 7. Observe that the cliques on the solid line are labeled in descending order, so that the first clique, C_K, will contain the largest number of nodes. Given these assumptions, we have the following lemma, which we prove in the appendix:

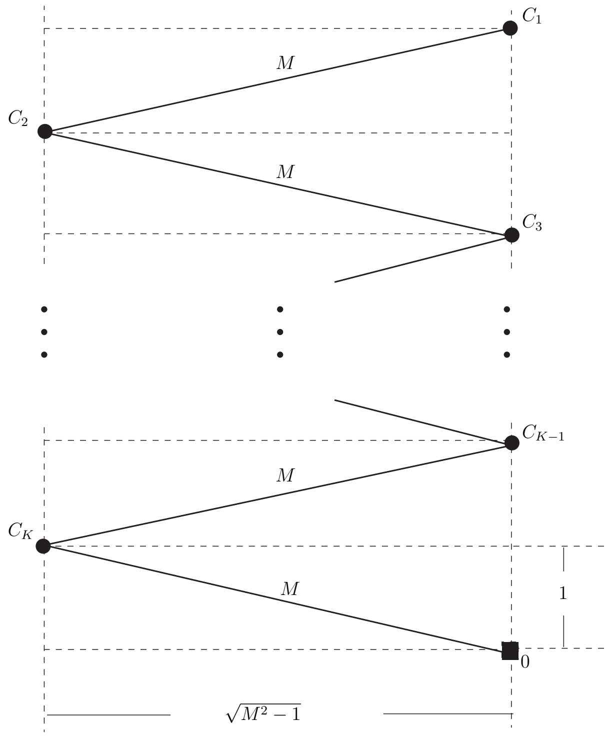


Figure 7: Illustration for tour length of minsum versus TSP

Lemma 1. *The optimal minsum tour in Figure 7 is* $0 - C_K - C_{K-1} - \ldots - C_3 - C_2 - C_1$*, where nodes within a clique are visited consecutively in any order.*

Assuming this is the case, we have that $c(MS) = MK + K$. Furthermore, it is easy to see that the optimal TSP tour is $0 - C_{K-1} - C_{K-3} - \ldots - C_1 - C_2 - C_4 - \ldots - C_k$ and has length $c(TSP) = 2\frac{K}{2} + M + 2\frac{K-2}{2} + M = 2M + 2K - 2$. This gives the ratio $\frac{c(MS)}{c(TSP)} = \frac{MK+K}{2M+2K-2}$ which goes to $\frac{K}{2}$ as $M \to \infty$. By choosing a large enough number of cliques and a large enough integer M, we can make the ratio $\frac{c(MS)}{c(TSP)}$ arbitrarily large.

This indicates that if cost is a concern or if we want to be able to use a vehicles to make multiple trips in a day, we will need to be very careful in choosing an appropriate objective function. We may need to consider a combination of objectives to balance the needs for equity and efficiency. This will be discussed further in section 8.

5.2.3 Multiple Vehicles

Just as with minmax, we can expect to reduce the minsum objective by increasing the number of available vehicles. The following proposition shows how much better the minsum objective can be as a result of being able to simultaneously serve customers.

Proposition 10. $sa(MS) \leq (2n(k-1)+1)sa(MS)^k$.

Proof. Consider an optimal set of tours that achieves $sa(MS)^k = \sum_{i=1}^n a_i$. Note that like the minmax objective, there is no incentive to use fewer than k tours. Next let $l_m = \max_j\{l_j\}$ where l_j represents the arrival time at the end of the jth of the k tours. Note also that $l_m = \max_{i\in N}\{a_i\}$. We now construct a single tour from these k tours. We visit each tour in sequence: starting at the depot, traveling to its endpoint, and returning to the depot to start the next tour, until we reach the last customer on the last tour. The arrival time to the customers on the original first (chosen arbitrarily) tour will be unchanged. The arrival time to the customers on the kth tour will be increased by at most $2l_1 + 2l_2 + \cdots + 2l_{k-1}$.

Since $l_m \geq l_j \forall j \in 1, \ldots, k$, the new arrival time at customer i is clearly less than or equal to $2(k-1)l_m + a_i$. Hence, we have that

$$sa(MS) \leq \sum_{i\in N}(2(k-1)l_m + a_i) = 2n(k-1)l_m + sa(MS)^k.$$

Since $l_m < sa(MS^k)$, we have the desired result. □

Combining Propositions 8 and 10, we get

Proposition 11. $sa(TSP) \leq (n^2(k-1)2+n)sa(MS)^k$.

As discussed in section 5.1.3, it is important to evaluate the tradeoffs between the costs associated with additional vehicles and the potential improvements in service.

5.2.4 Capacity

The impact of having additional vehicles is affected by the capacity of the vehicle. Unlike with minmax, the following bounds explicitly includes Q.

Proposition 12. $sa(MS) \leq (k(k-1)Q+1)sa(MS)^k_Q$.

Proof. We proceed as in the proof of Proposition 10 and concatenate all the routes in the solution that achieves $sa(MS)^k_Q$. Since there are at most Q customers on the second route, the sum of arrival times for those customers increases by at most $2l_1Q$. The sum of arrival times for the original third tour increases by a total of $2(l_1 + l_2)Q$, etc. Thus

$$\begin{aligned} sa(MS) &\leq sa(MS)^k_Q + 2l_1Q + \ldots + 2(l_1 + l_2 + \ldots l_{k-1})Q \\ &\leq sa(MS)^k_Q + 2l_mQ + \ldots + 2(k-1)l_mQ \\ &= sa(MS)^k_Q + (k-1)(k)l_mQ \\ &\leq sa(MS)^k_Q + (k-1)(k)Qsa(MS)^k_Q, \end{aligned}$$

which is our desired result. □

6 Heuristic Algorithms

In this section, we will discuss how to modify the popular insertion algorithm to handle these new objective functions as well as discuss the use of improvement heuristics. It is nontrivial to modify construction and improvement heuristics to achieve good solutions for these problems while keeping them computationally efficient and simple to implement. Both of these factors are important for these tools to be successfully used in practice.

Insertion heuristics have proven to be popular methods for solving a variety of vehicle routing and scheduling problems. They were first introduced and analyzed, as many other popular optimization techniques, for the TSP [41]. Insertion heuristics construct a feasible solution, i.e. a set of feasible routes, by repeatedly and greedily inserting an unrouted customer into a partially constructed feasible solution. Different variants of the insertion heuristic arise as a result of choices in how routes are initially created and the criteria for selecting unrouted customers to insert and where to insert them in the partial solution. In some variants, the heuristic starts with a series of null routes, while others use seeds to initiate routes. Typically, the customer and insertion point are selected that yield the least increase in the current objective function, but this procedure can be modified to include randomization in the selection process [18]. For a review of insertion heuristics, see Campbell and Savelsbergh [13].

6.1 Minmax

As indicated above, the basic version of the insertion algorithm repeatedly inserts the customer creating the least expensive increase in total distance. For the minmax objective, it would appear that the best insertion customer and insertion point would be the combination that leads to the least increase in the latest arrival time. This is true for one vehicle, but not when multiple vehicles are available. In this case, there may be several tours where some of the customers can be inserted without creating an increase in the latest arrival time over all tours. Not all insertions that do not increase the latest arrival time at some iteration will result in the same minmax objective at the end of the algorithm. Some insertion points may lead to larger or smaller increases in the arrival times to the last customer on a particular tour. We take this into consideration in our modification of the insertion algorithm. Assume there exists at least one customer who can be inserted on some tour without causing an increase in the maximum arrival time. We then choose the customer and insertion point creating the least increase in the arrival time at the last customer on *any* of the tours that do not determine the latest arrival time. Otherwise, the customer and insertion point yielding the least increase in the latest arrival over all tours is selected. Algorithm 1 details this approach.

To give the algorithms the best chance to "succeed" in terms of finding good solutions to these problems, we will experiment both with initiating the algorithm with k null (empty) routes and k routes started from seed customers. This choice should not impact the complexity of the algorithm (it depends on the complexity of the seed selection procedure), but will likely impact the final solution. If we begin with null routes, the $LATEST_r$ value will initially be 0 for all r. When using seeds, this value will be the arrival time at the seed customer for each route.

In Algorithm 1, for each potential insertion point for each uninserted customer, G_1 reflects any change in the latest arrival time over all routes as a result of the insertion. G_2 reflects the change in the latest arrival time for the route under consideration. At each iteration, all of the uninserted customers (N) are considered. If $G_2^\star < \infty$ at the end of an iteration, then this indicates there are customers that can be inserted without increasing the objective function. The cheapest of these to insert is customer $j_2^\star$ in position $i_2^\star$ on route $r_2^\star$. Similarly, if all customers create an increase in the latest arrival time, the best customer is stored in $j_1^\star$, and the best insertion point is $i_1^\star$ on route $r_1^\star$.

Algorithm 1 runs in $O(n^3)$, like the traditional insertion algorithm. For each of the n insertions, n^2 customer/insertion point combinations are considered. The updating procedure is in constant time and thus does not impact the complexity.

6.2 Minsum

In adapting the basic insertion algorithm for minsum, we start by examining how a single insertion impacts the objective function. An insertion that makes a vehicle arrive x minutes later to the following customer also makes the arrival x minutes later to each subsequent customer on the tour.

Algorithm 1 Insertion Algorithm for Minmax Objective

N = set of unassigned nodes, R = set of routes;
$LATEST_r$ = initial latest arrival time for route $r \in R$.
$ROUTEMAX = \max_{r \in R} LATEST_r$;
while $N \neq \emptyset$ **do**
  LET $G_1^\star = \infty$, $G_2^\star = \infty$;
  for $j \in N$ **do**
    for $r \in R$ **do**
      for $(i-1, i) \in r$ **do**
        $LATEST = LATEST_r + t_{i-1j} + t_{ji} - t_{i-1i}$
        $G_1 = \max(ROUTEMAX, LATEST) - ROUTEMAX$
        $G_2 = LATEST - LATEST_r$
        if $G_1 < G_1^\star$ **then**
          update($G_1^\star$, $j_1^\star$,$i_1^\star$, $r_1^\star$);
        end if
        if $(G_1 = 0)$**and**$(G_2 < G_2^\star)$ **then**
          update($G_2^\star$, $j_2^\star$,$i_2^\star$, $r_2^\star$);
        end if
      end for
    end for
  end for
  if $(G_2^\star < \infty)$ **then**
    insert($j_2^\star$,$i_2^\star$, $r_2^\star$);
    $LATEST_{r_2^\star} = LATEST_{r_2^\star} + G_2^\star$;
    $N = N - j_2^\star$;
  else
    insert($j_1^\star$, $i_1^\star$, $r_1^\star$);
    $LATEST_{r_1^\star} = LATEST_{r_1^\star} + G_1^\star$
    $ROUTEMAX = ROUTEMAX + G_1^\star$;
    $N = N - j_1^\star$;
  end if
end while

This is key in the development of Algorithm 2.

In Algorithm 2, the A variables represent arrival times and will have initial values other than 0 if we start from seeded routes. G will represent the change in the sum of arrival times as a result of each proposed insertion, where $\mid r \mid$ represents the number of customers on route r. The difference in sums will include the arrival time for the added customer plus the lateness incurred at all of the subsequent customers. At the end of each iteration, customer $j^\star$ will be inserted in position $i^\star$ on route $r^\star$.

For each of the n insertions, we again evaluate n^2 possible insertion combinations. After each insertion is selected, $O(n)$ operations are required to do the necessary updating. This makes the algorithm $O(n \times (n^2 + n)) = O(n^3)$, so there is no increase in complexity.

Algorithm 2 Insertion Algorithm for Minsum Objective

N = set of unassigned nodes, R = set of routes;
$A(i, r)$ = initial arrival time at ith customer on route r;
while $N \neq \emptyset$ **do**
    LET $G^\star = \infty$
    for $j \in N$ **do**
        for $r \in R$ **do**
            for $(i-1, i) \in r$ **do**
                $ARR = A(i-1, r) + t_{i-1j}$
                $CHANGE = t_{i-1j} + t_{ji} - t_{i-1i}$
                $G = ARR + CHANGE \times (\mid r \mid - i + 1)$
                if $G < G^\star$ **then**
                    update($G^\star$, $ARR^\star$, $CHANGE^\star$, $j^\star$, $i^\star$, $r^\star$);
                end if
            end for
        end for
    end for
    insert($j^\star$, $i^\star$, $r^\star$); $A(i^\star, r^\star) = ARR^\star$;
    for $k \in r^\star$ after $i^\star$ **do**
        $A(k, r^\star) = A(k-1, r^\star) + CHANGE^\star$;
    end for
    $N = N - j^\star$
end while

6.3 Seeding the Routes

The choice of seeds can have a large impact on the final routes. With the alternate objective functions, we found the following algorithm successful in determining seeds for the routes. In Algorithm 3, we iteratively determine the seeds. In each iteration, any unrouted customer is considered a potential seed. For each potential seed, we evaluate the cost to insert that customer on any of the already seeded routes. We retain the cheapest insertion point for each potential seed with $i_2^\star$ and $r_2^\star$, and its cost is $G_2^\star$. The customer with the highest $G_2^\star$ value is selected as the new seed since this is the customer who can be served least efficiently by the existing routes.

6.4 Improvement heuristics

Insertion heuristics are often followed by iterative improvement heuristics. Various papers have been written on how to efficiently implement iterative improvement heuristics in the presence of complicating constraints. For a survey of these techniques, see Kindervater and Savelsbergh [28].

Before we describe the improvement heuristics used here, we will describe an interesting observation that illustrates the difference in the objective functions, even from a problem-solving perspective. In the case of Euclidean problems, where the distances represent Euclidean distances between points in the plane, exchange heuristics help eliminate edges in the tours that cross since

Algorithm 3 Partitioning Insertion Algorithm

N = set of unassigned nodes; AR = set of assigned routes; R =set of routes;
while $|AR| < |R|$ **do**
  LET $G^\star = \infty$
  for $j \in N$ **do**
    LET $G_2^\star = \infty$
    for $r \in AR$ **do**
      for $(i-1, i) \in r$ **do**
        G = insertion cost for minmax or minsum objective;
        if $G < G_2^\star$ **then**
          update($G_2^\star, i_2^\star, r_2^\star$);
        end if
      end for
    end for
    if $G_2^\star > G^\star$ **then**
      update($G^\star, j^\star, i^\star$,$r^\star$);
    end if;
  end for
  $N = N - j^\star$;
end while

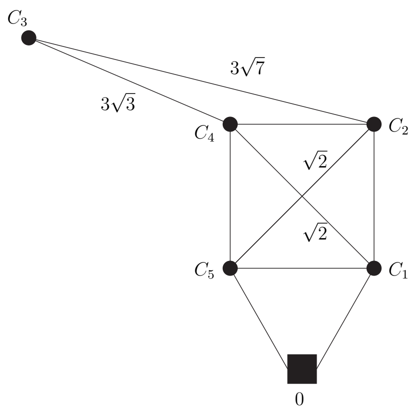


Figure 8: Uncrossing may not improve the solution for minsum

this has been shown to be suboptimal for the standard VRP objective. For the minsum objective, however, it is not always better to uncross two edges. Consider the graph given in Figure 8. All unmarked edges have length one. One can verify that this problem can be embedded in the plane, and distances are Euclidean. The sequence $0 - C_1 - C_4 - C_3 - C_2 - C_5$ yields a smaller minsum objective than uncrossing this sequence to get $0 - C_1 - C_2 - C_3 - C_4 - C_5$. This is because in the second sequence, the order of the edges with length $3\sqrt{3}$ and $3\sqrt{7}$ is exchanged, putting the longer edge earlier, which more than offsets the cost of taking the longer diagonal edges. Note also that an optimal minmax tour may cross itself with the last edge of the tour since its length is not included in the objective (see, for instance, the optimal minmax tour in Figure 1(a)).

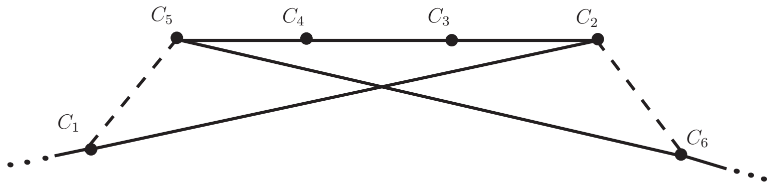


Figure 9: 2-Edge exchange example.

In our computational experiments, we used two forms of local improvement: 2-edge exchange and 2 node-relocation. The first of these seeks to find an improved solution by taking any two edges within a single tour and replacing them. The 2 node-relocation removes any two nodes from any tours and checks if inserting them elsewhere would create an improvement. These are described in greater detail in [28]. We selected these two improvement schemes since they involve different search spaces, and thus should work well together. In our experiments, we run the 2 node-relocation improvement heuristic first until no improving relocations can be found, then run the 2-edge exchange improvement heuristic. These two heuristics are alternated until there are no improving changes of either type.

Implementing the improvement heuristics is straightforward for the VRP objective, and only minor modifications are necessary when dealing with the minmax objective. For the minsum objective, it is necessary to maintain extra information in order to efficiently test the possible improvements to the tours. Consider, for instance, applying a 2-edge exchange in the example pictured in Figure 9, where edges (C_5, C_6) and (C_1, C_2) are replaced by (C_1, C_5) and (C_2, C_6). This exchange reverses the order of customers $C_2, \ldots, C_5$ on the tour. In the case of the VRP objective or the minmax objective, this has no impact since the change in objective is determined only by the exchanged edges. To recalculate the minsum objective, however, a myopic approach may iterate through the nodes whose order was reversed, increasing the overall complexity of the heuristic. By storing and maintaining the arrival time at each node in the current solution, we are able to calculate the change in the minsum value more efficiently.

7 Computational Results

We have implemented the above algorithms and will next explore the impact of different data characteristics on the resulting solutions using these algorithms.

7.1 Data

Our computational experiments used data representing several different geographical distributions. The first dataset, hereafter referred to as Augerat-A, is available through `http://branchandcut.org/VRP/data/` and was introduced in [4]. Augerat-A consists of 27 different instances with 31-79 customers. In each instance, customers are distributed rather uniformly, but the depot is not necessarily in the center. For example, in certain instances, all of the customers may be to the lower left of the depot on a graph. The second dataset, hereafter referred to as Augerat-B, is also available through `http://branchandcut.org/VRP/data/` and was introduced in [4]. Augerat-B consists of 23 datasets with between 30 and 77 customers each. In Augerat-B, unlike A, customers are grouped in clusters. The use of the two datasets should help us understand the impact of the clustering on the relative performance of the different objectives. The third dataset, hereafter referred to as Golden is available through `http://neo.lcc.uma.es/radi-aeb/WebVRP/` and included in [21]. These instances all are based on special structures. The special structures include concentric circles, diamonds, squares, and stars. We have selected 11 of the smaller instances to test here. We further eliminated customers from these selected instances so that the structure of the instance was maintained but the number of customers is less than 100.

7.2 Heuristics Quality

In an effort to verify that our heuristics provide reasonably good solutions, we generated small test instances and compared the results from our heuristics with the solutions obtained from the MIP formulations presented in section 4. A total of 16 instances were derived from the Augerat-A and -B collection, by removing customers such that the instances were of size ranging between 15 and 20 customers, but such that the overall distribution pattern of the customers was maintained. For each instance, the MIP solver [24] was given 11,000 CPU seconds, and the best solution found during the branch-and-bound process is compared to the best solution found using the heuristics discussed in section 6. For each instance, we solved two routing problems: the first allows exactly one vehicle, the second problem allows two. To keep the MIP model simple, the vehicles were uncapacitated.

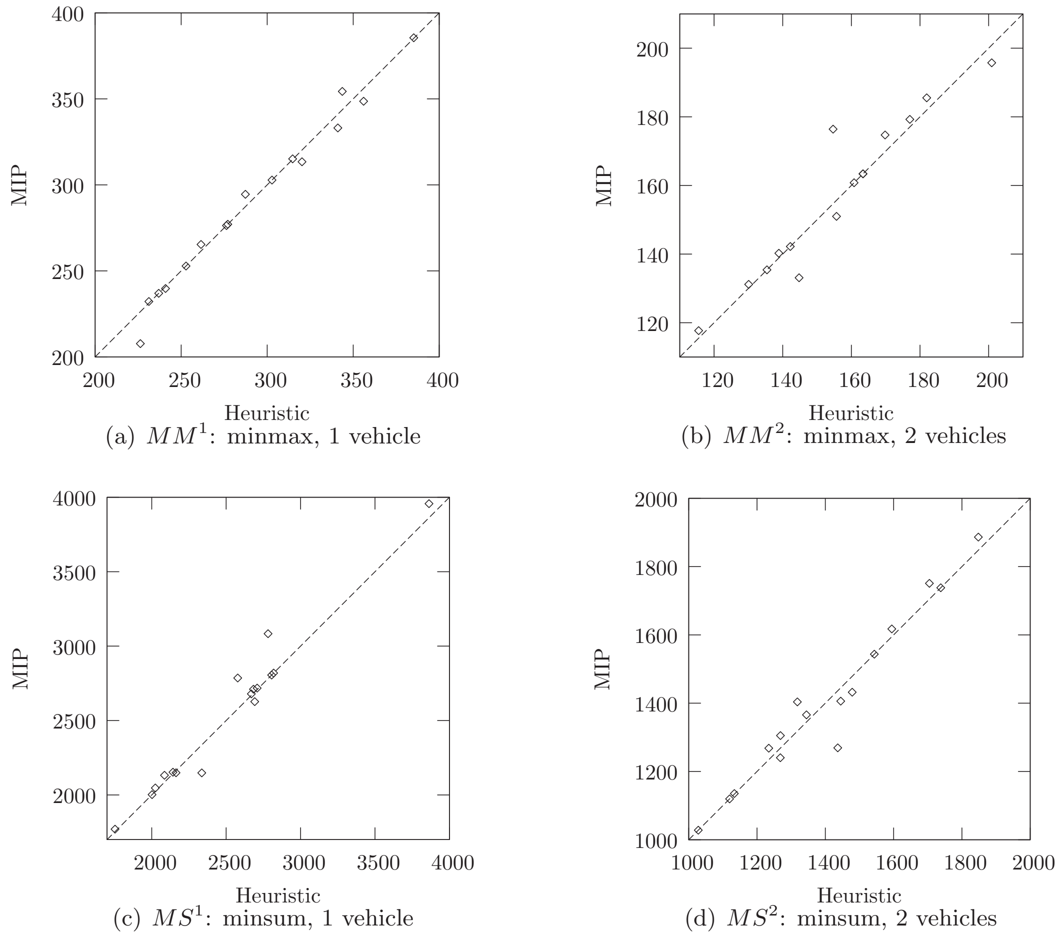
We decided to use a time limit of 11,000 CPU seconds for a number of reasons. One was because of the number of tests we were doing, but second, and more importantly, was the fact that experiments with longer time limits either created no change in the solution, or at best, matched the solution found by our heuristics. We did experiments with longer time limits for six different instances including two where the MIP objective was worse than the heuristics after 11,000 CPU seconds, two where the MIP performed better than the heuristics after 11,000 CPU seconds, and two where the MIP and the heuristics reached similar objective values after 11,000 CPU seconds. For each of these six instances, we let the solver run for 10 hours. Running them longer is infeasible because after 10 hours the branch-and-bound tree requires more than 1 Gb of memory. For the two instances where the MIP objective was worse after 3 hours (11,000 CPU seconds), the solutions after 10 hours of CPU time were the same solutions as created by our heuristics. For the other four instances, none of the objective values improved from the solution found after 3 hours, and optimality gaps barely changed as well. Thus, even these relatively small problems are extremely difficult to solve, but 11,000 CPU seconds gives us a good picture of the solution quality that is possible with an MIP approach.

The results for the full set of instances can be found in figure 10, where we plot for each instance the value found by the heuristic against the best MIP solution found. One can see that the heuristic solutions are comparable to, and often better than, the MIP solutions, both for minsum and minmax and with 1 or 2 vehicles. The MIP obviously can eventually find a better solution than the heuristics, but often not before the computer runs out of memory, as indicated by our tests. Here, the heuristics required only fractions of a second, while the MIP solver was terminated at the time limit on each instance. The MIP solver did not terminate prematurely on any of the 32 problems.

7.3 Results

In the experiments, we solve each instance using two different methods and select the best solution for our tables. For the minmax and minsum objectives, initial solutions are constructed using the basic insertion algorithm without seeded routes as well as the insertion algorithm seeded using the partitioning algorithm described earlier. For the traditional TSP and VRP objectives, initial solutions were constructed using the basic insertion algorithm without seeded routes and the well known Clarke-Wright algorithm [14]. For fairness, we wanted to create the best solutions possible for each objective, and Clarke-Wright proved to be more successful than partitioning in our experiments with the traditional TSP and VRP objectives. After each initial solution was constructed, improvement heuristics were applied as described earlier. The best solution after improvement was the solution used in the tables.

In reporting the results, we have used two methods to compare arrival times in different solutions. The first method evaluates the objective value of a routing solution with respect to the other



(a) MM^1: minmax, 1 vehicle

(b) MM^2: minmax, 2 vehicles

(c) MS^1: minsum, 1 vehicle

(d) MS^2: minsum, 2 vehicles

Figure 10: Objective values for heuristic solutions versus best MIP solutions

Table 1: Ratios for $k = 1$ route comparing objective functions and upper semideviations

		$\frac{la(TSP)}{la(MM)}$	$\frac{la(MS)}{la(MM)}$	$\frac{sa(TSP)}{sa(MS)}$	$\frac{sa(MM)}{sa(MS)}$	$\frac{c(MM)}{c(TSP)}$	$\frac{c(MS)}{c(TSP)}$	$\frac{us(TSP)}{us(MM)}$	$\frac{us(TSP)}{us(MS)}$	$\frac{us(MM)}{us(MS)}$
Augerat-A	MIN	0.948	0.970	0.974	0.973	1.004	1.027	0.926	0.902	0.850
	MAX	1.051	1.257	1.229	1.270	1.246	1.431	1.114	1.138	1.234
	AVG	1.013	1.114	1.089	1.124	1.075	1.184	1.016	1.016	1.003
Augerat-B	MIN	0.912	0.922	0.852	0.938	0.995	0.963	0.822	0.730	0.728
	MAX	1.146	1.275	1.424	1.321	1.233	1.384	1.391	1.565	1.315
	AVG	1.020	1.092	1.061	1.074	1.047	1.114	1.041	1.068	0.987
Golden	MIN	0.977	0.966	1.006	1.020	1.012	1.007	0.951	0.955	0.901
	MAX	1.049	1.263	1.185	1.210	1.095	1.303	1.089	1.080	1.052
	AVG	1.014	1.115	1.105	1.101	1.055	1.167	1.003	1.017	0.987

objective functions. For example, in the notation from section 5, $la(TSP)$ represents the latest arrival time for the best TSP routing solution. Note this is a slight abuse of notation, as these values now correspond to heuristic rather than optimal solutions. We then compute the ratio of the appropriate routing solutions using the same objective measure. For example, $\frac{la(TSP)}{la(MM)}$ is the ratio of the latest arrival time for a TSP solution as compared to a minmax solution. The further this ratio deviates from 1 is in an indicator of how much the objective function changes the solutions. We average these ratios over all of the instances within a particular dataset. This weights the instances within a dataset identically and allows easy comparison among different datasets. We also provide the minimum and maximum ratios within a given dataset to capture the variance in these values.

The second method for comparing routing solutions is intended to measure the inequity in the arrival times. These comparisons are based on the mean absolute upper semideviation (see e.g. [33]), which is a measure of the deviation of arrival times that have a higher value than the average arrival time. In particular, for a given solution using n nodes, denote the arrival times as a_i for $i \in \{1, \ldots, n\}$, and denote the mean arrival time as $\mu = \frac{1}{n}\sum_{i=1}^{n} a_i$. The mean absolute upper semideviation is computed as: $\frac{1}{n}\sum_{a_i \geq \mu}(a_i - \mu)$. Solutions with large upper semideviations indicate that the arrival times induced by the solution may not be equitable. As an example of notation, the mean absolute upper semideviation of a TSP route will be denoted as $us(TSP)$. Further, the mean absolute upper semideviation of a VRP solution using k routes and capacity Q is denoted as $us(VRP)_Q^k$. Similar to the first method of comparing routes, ratios of mean absolute upper semideviations have been computed for all instances and objectives on a dataset. The minimum, maximum, and average of these ratios are reported alongside the appropriate objective ratios in the tables.

In Table 1, we compare the relative performance of the different objective functions when one vehicle is available to serve the customers. For Augerat-B, we see that, on average, the latest arrival time in a TSP solution is 2% higher than in a minmax solution, while the upper semideviation is 4.1% higher. This comes at the cost of a 4.7 % increase in total route duration. There is approximately a 6 % difference between the sum of arrival times in a TSP solution and in a minsum solution, with a 6.8% increase in upper semideviation. These improvements in sum of arrival times and upper semideviation come at cost of an 11% increase in total route duration.

We note that the minimum values for the minmax and minsum ratios are less than one, but this is due to the fact the problems are solved using heuristics. For the Augerat-A and Golden datasets, we find similar results with both yielding larger improvements in the minsum objective than with minmax when compared to TSP solutions. Overall, we do not observe the significant differences indicated by the worst-case results presented earlier.

In Tables 2 and 3, we see how the results change as additional vehicles are considered, in particular for the number of vehicles $k \in \{1, 5, 10\}$. In these tables, all of the vehicles have the minimum capacity such that there is a feasible solution, i.e. $Q = \lceil \frac{n}{k} \rceil$. For all of the datasets, we find that the $la(VRP)_Q^k / la(MM)_Q^k$ ratio is significantly larger with multiple vehicles, while the cost increases are not as significant. For instance, for Augerat-A, having 5 vehicles provides a 36.5% improvement in latest arrival time, with an average increase in total cost of 12%. The results for minsum are similar. For Augerat-A, we obtained an average increase of 20% in sum of arrival times, with a 13 % increase in total cost. Under tight capacity, the additional vehicles did not seem to provide significant improvements in upper semideviation.

It is interesting to note that for Augerat-A and B, the increase to 5 vehicles creates a significant difference in the average latest arrival ratio (e.g. 1% to 36.5% for Augerat-A), but there is not much of a jump when 10 vehicles are considered (e.g. 36.5% to 39% for Augerat-A). The results for the upper semideviation measure are similar. For the specially structured Golden dataset, the impact of choosing a minmax objective appears to grow more steadily with the increasing number of vehicles with increases from 1.5% with one vehicle to 36.4% with 5 vehicles to 68.3% with 10 vehicles. This behavior is similar with the minsum objective.

Table 2: Effect of multiple vehicles for minmax routing using tight capacity, $Q = \lceil \frac{n}{k} \rceil$

		$\frac{la(VRP)^k_Q}{la(MM)^k_Q}$			$\frac{c(MM)^k_Q}{c(VRP)^k_Q}$			$\frac{us(VRP)^k_Q}{us(MM)^k_Q}$		
		$k=1$	$k=5$	$k=10$	$k=1$	$k=5$	$k=10$	$k=1$	$k=5$	$k=10$
Augerat-A	MIN	0.948	1.050	1.170	1.004	1.044	1.067	0.926	0.942	0.993
	MAX	1.051	1.618	1.627	1.246	1.220	1.370	1.114	1.376	1.305
	AVG	1.013	1.365	1.390	1.075	1.122	1.212	1.016	1.163	1.150
Augerat-B	MIN	0.912	0.890	1.066	0.995	1.013	0.994	0.822	0.622	0.771
	MAX	1.146	2.129	2.047	1.233	1.340	1.403	1.391	1.880	1.629
	AVG	1.020	1.390	1.390	1.047	1.129	1.180	1.041	1.180	1.126
Golden	MIN	0.977	1.032	1.233	1.012	0.972	0.925	0.951	0.968	1.086
	MAX	1.049	1.723	2.880	1.095	1.234	1.288	1.089	1.324	2.178
	AVG	1.014	1.364	1.683	1.055	1.120	1.138	1.003	1.141	1.340

Table 3: Effect of multiple vehicles for minsum routing using tight capacity, $Q = \lceil \frac{n}{k} \rceil$

		$\frac{sa(VRP)^k_Q}{sa(MS)^k_Q}$			$\frac{c(MS)^k_Q}{c(VRP)^k_Q}$			$\frac{us(VRP)^k_Q}{us(MS)^k_Q}$		
		$k=1$	$k=5$	$k=10$	$k=1$	$k=5$	$k=10$	$k=1$	$k=5$	$k=10$
Augerat-A	MIN	0.974	1.015	1.137	1.027	1.008	1.061	0.902	0.949	0.999
	MAX	1.229	1.351	1.287	1.431	1.242	1.420	1.138	1.430	1.382
	AVG	1.089	1.202	1.203	1.184	1.132	1.190	1.016	1.184	1.188
Augerat-B	MIN	0.852	1.003	1.026	0.963	0.974	1.029	0.730	0.777	0.891
	MAX	1.424	1.304	1.188	1.384	1.229	1.269	1.565	2.038	1.440
	AVG	1.061	1.136	1.099	1.114	1.107	1.136	1.068	1.256	1.156
Golden	MIN	1.006	1.068	1.179	1.007	1.015	0.925	0.955	0.994	1.170
	MAX	1.185	1.317	1.671	1.303	1.223	1.206	1.080	1.339	2.178
	AVG	1.105	1.183	1.312	1.167	1.142	1.116	1.017	1.186	1.405

Table 4: Effect of capacity on minmax routing using 5 vehicles

		$\frac{la(VRP)^5_Q}{la(MM)^5_Q}$			$\frac{c(MM)^5_Q}{c(VRP)^5_Q}$			$\frac{us(VRP)^5_Q}{us(MM)^5_Q}$		
		$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$
Augerat-A	MIN	1.050	2.175	3.914	1.044	1.329	1.355	0.942	2.081	4.080
	MAX	1.618	3.769	4.953	1.220	1.816	2.083	1.376	3.507	5.361
	AVG	1.365	2.796	4.347	1.122	1.593	1.717	1.163	2.709	4.714
Augerat-B	MIN	0.890	1.686	2.067	1.013	1.395	1.509	0.622	1.392	2.363
	MAX	2.129	3.313	4.475	1.340	2.245	2.602	1.880	3.849	8.658
	AVG	1.390	2.453	3.630	1.129	1.708	2.072	1.180	2.637	4.462
Golden	MIN	1.032	2.337	4.350	0.972	1.203	1.242	0.968	2.097	3.802
	MAX	1.723	2.820	4.928	1.234	1.697	1.806	1.324	2.716	5.135
	AVG	1.364	2.573	4.587	1.120	1.410	1.477	1.141	2.488	4.706

Next, Tables 4 and 5 address what happens if capacity is not as “tight” and 5 vehicles are available. The columns identified with $C = 1$ represent tight capacity ($Q = \lceil \frac{n}{k} \rceil$), $C = 3$ represents uncapacitated ($Q = n$), and $C = 2$ represents the results when capacity is halfway between tight and uncapacitated ($Q = \lceil \frac{1}{2} \lceil \frac{n}{k} \rceil + \frac{1}{2}n \rceil$). We see the anticipated substantial impact by considering additional capacity. For example, with $C = 2$, using the traditional VRP objective can yield a latest arrival time on average almost 2.8 times larger than when latest arrival time is explicitly considered for Augerat-A. Similar improvements occur for the upper semideviation measure. It is interesting to note that the total length ratios are not nearly as large, indicating that significant improvements in latest arrival do not necessarily come at equally significant increases in total route duration. The results are also similar with the Augerat-B and Golden datasets. For the minsum objective, we also find significant improvements in the sum of arrival times, with smaller increases in total route duration.

Tables 6 and 7 present the effect of increases in capacity when 10 vehicles are available. For example, with the Golden dataset and 10 uncapacitated vehicles, explicitly minimizing latest arrival times yields solutions that differ on average by a factor of 8. For the minsum objective, we find very similar results as capacity loosens. We note that Augerat-B has the least impact with both objectives when $C = 3$. This indicates that clustering may limit the contribution of loosening capacity at some point. Just as with 5 vehicles, Tables 6 and 7 show that the magnitudes of the improvements in latest arrival or sum of arrival times are significantly larger than the increases in total route length when the alternate objectives are used.

Table 5: Effect of capacity on minsum routing using 5 vehicles

		$\frac{sa(VRP)^5_Q}{sa(MS)^5_Q}$			$\frac{c(MS)^5_Q}{c(VRP)^5_Q}$			$\frac{us(VRP)^5_Q}{us(MS)^5_Q}$		
		$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$
Augerat-A	MIN	1.015	1.869	3.035	1.008	1.370	1.410	0.949	2.181	4.179
	MAX	1.351	2.665	4.613	1.242	1.789	2.052	1.430	3.553	5.366
	AVG	1.202	2.273	4.036	1.132	1.561	1.683	1.184	2.741	4.778
Augerat-B	MIN	1.003	1.251	1.582	0.974	1.383	1.497	0.777	1.455	2.378
	MAX	1.304	2.344	4.619	1.229	1.951	2.630	2.038	4.197	6.590
	AVG	1.136	1.839	2.942	1.107	1.651	2.005	1.256	2.574	4.306
Golden	MIN	1.068	2.083	3.994	1.015	1.197	1.208	0.994	2.218	4.022
	MAX	1.317	2.860	5.729	1.223	1.678	1.786	1.339	3.080	5.729
	AVG	1.183	2.499	4.846	1.142	1.393	1.460	1.186	2.597	4.905

Table 6: Effect of capacity on minmax routing using 10 vehicles

		$\frac{la(VRP)^{10}_Q}{la(MM)^{10}_Q}$			$\frac{c(MM)^{10}_Q}{c(VRP)^{10}_Q}$			$\frac{us(VRP)^{10}_Q}{us(MM)^{10}_Q}$		
		$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$
Augerat-A	MIN	1.170	2.989	4.340	1.067	1.791	1.825	0.993	2.579	4.688
	MAX	1.627	4.680	8.318	1.370	2.776	3.185	1.305	4.563	9.237
	AVG	1.390	2.796	4.347	1.212	2.382	2.609	1.150	3.668	6.742
Augerat-B	MIN	1.066	1.746	2.076	0.994	1.663	1.809	0.771	1.654	2.514
	MAX	2.047	3.715	5.595	1.403	3.531	4.683	1.629	6.724	13.951
	AVG	1.390	2.819	4.178	1.180	2.647	3.269	1.126	3.193	5.267
Golden	MIN	1.233	3.148	5.946	0.925	1.497	1.572	1.086	3.592	6.850
	MAX	2.880	6.846	11.181	1.288	2.688	2.888	2.178	7.136	12.379
	AVG	1.683	4.770	8.163	1.138	1.929	2.036	1.340	4.903	9.067

Table 7: Effect of capacity on minsum routing using 10 vehicles

		$\frac{sa(VRP)^{10}_Q}{sa(MS)^{10}_Q}$			$\frac{c(MS)^{10}_Q}{c(VRP)^{10}_Q}$			$\frac{us(VRP)^{10}_Q}{us(MS)^{10}_Q}$		
		$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$	$C=1$	$C=2$	$C=3$
Augerat-A	MIN	1.137	2.141	3.477	1.061	1.802	1.836	0.999	2.713	4.932
	MAX	1.287	3.682	7.261	1.420	2.969	3.405	1.382	4.784	9.308
	AVG	1.203	2.273	4.036	1.190	2.316	2.540	1.188	3.838	7.043
Augerat-B	MIN	1.026	1.332	1.639	1.029	1.946	2.193	0.891	1.494	2.473
	MAX	1.188	2.786	5.256	1.269	3.432	4.704	1.440	5.669	11.761
	AVG	1.099	2.093	3.342	1.136	2.714	3.352	1.156	3.211	5.294
Golden	MIN	1.179	2.668	5.075	0.925	1.489	1.554	1.170	3.749	7.149
	MAX	1.671	5.766	11.044	1.206	2.662	2.860	2.178	7.136	12.379
	AVG	1.312	4.169	7.902	1.116	1.889	1.994	1.405	5.189	9.598

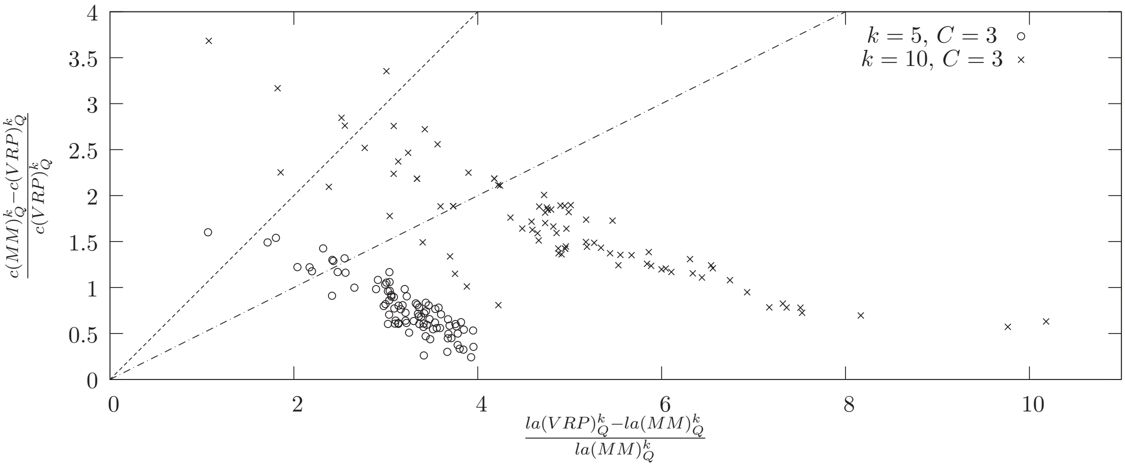


Figure 11: Improvements in latest arrival versus total duration

Figures 11 and 12 further illustrate the trade-off between improving latest arrival time or sum of arrival times and the total route duration. In Figure 11, for example, we plot the point $\left(\frac{la(VRP)_Q^k - la(MM)_Q^k}{la(MM)_Q^k}, \frac{c(MM)_Q^k - c(VRP)_Q^k}{c(VRP)_Q^k}\right)$ for each instance. This plots the percentage increase in latest arrival time by using the VRP objective against the percentage increase in total duration by using the minmax objective. We do this for $k = 5$ and 10, uncapacitated vehicles. Since almost all points lie below the 45 degree line and the majority lie below the line $y = \frac{1}{2}x$, we can remark that the relative increase in latest arrival is consistently and often significantly larger than the relative increase in total route duration. Similar conclusions can be drawn for minsum from Figure 12.

In looking at the full set of tables, it appears that geography of the datasets does have some impact on our results. For both Augerat-A (randomly distributed) and Augerat-B (clustered), almost all of the improvement in service from having additional vehicles is felt by expanding to 5 vehicles, but steady improvements in service are possible with increasing capacity for both objectives. Interestingly, increasing capacity seems to have as much or more impact on improving service for Augerat-A as for Augerat-B, but it usually accompanies lower increases in cost for Augerat-A for both minmax and minsum objectives. The Golden datasets (special structures) yield quite different results than Augerat-A and B. For both objectives, the improvement in service found with increasing capacity is much more dramatic for Golden than found with the other datasets, and the cost increases are often less. Also, the Golden sets exhibit steady increases in service with the increase in vehicles.

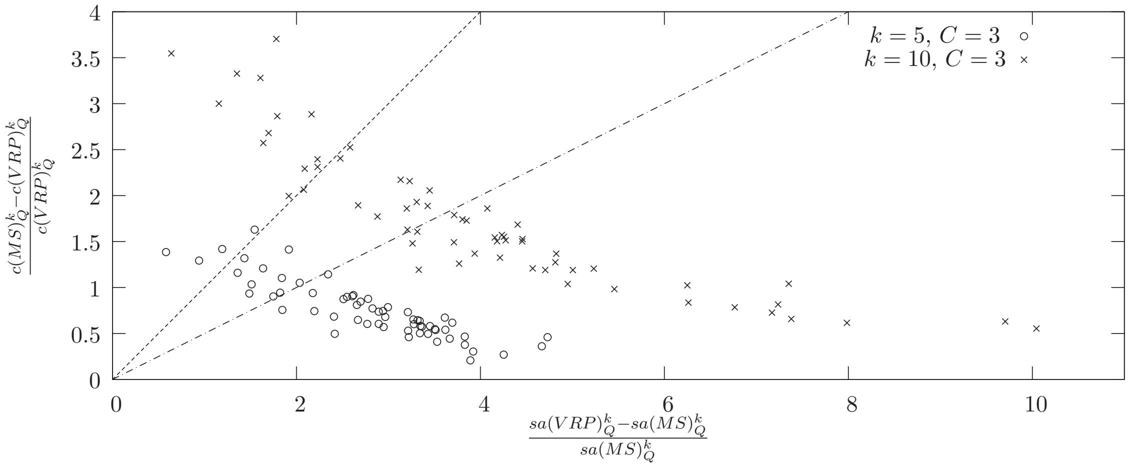


Figure 12: Improvements in sum of arrival times versus total duration

8 Conclusions and Next Steps

This paper provides strong evidence that much better service times to customers are possible than those created by traditional routing problems and algorithms. In a situation where service time equates to survival, such as after a disaster, it is important to have appropriate routing technology that addresses these concerns. We realize that this effort represents only a first step in developing new methodology for routing humanitarian aid. There are many other issues in a relief context that need to be considered, including additional factors in the objective function and additional constraints. In terms of objectives, it is clearly important to consider combinations of cost and service, since both are indeed relevant in practice. Future work may consider an objective function that combines a traditional cost-based objective with the service-based objectives considered here. It is also important to consider the fact that the different customers in our datasets represent different size groups, so customer weighting schemes may be worth examining.

Key constraints to be considered involve the reliability of arcs in the network after the disaster. After a disaster, existing roads may suddenly become closed or the travel time may become drastically increased, so routes may need to be designed to incorporate detours or alternate paths. For some roads, their availability or travel time in the network may be dynamically changing during the days following the disaster, so it may be necessary to dynamically change the routes as well. These arc reliability issues similarly impact the location decisions for the distribution centers of relief supplies to ensure reliable and efficient routes can be created. We are currently studying how arc reliability can be incorporated into these location decisions [12].

Another area of future research is to consider the complications that result from delivering multiple commodities. Consider what happens if the vehicles delivering relief are aid-specific, such a water truck and a food truck. If we know a tour that is efficient and equitable for one commodity, the natural choice is to send both trucks on the same route. If they arrive at the same distribution point at the same time, there may be a large delay before the second truck can be unloaded due to limited staff or security concerns. Such delays are not good from an efficiency perspective. Also, the distribution point that is last on the route would receive both types of supply last, which does not seem equitable. Thus, we would like to consider the issue of multiple arrival times in conjunction with multiple products/arrival times. There are many questions yet to be addressed in developing appropriate routing tools for disaster relief.

Acknowledgments

This work was partially supported by the National Science Foundation through grant number 0237726(Campbell).

References

[1] Death toll drops in worst tsunamis in memory. `http://www.wavescape.co.za/top_bar/tidings/Oneoffs/Tsunami04/tsunami04.htm`, April 2005.

[2] Red Cross response to Katrina update. `http://www.udel.edu/PR/UDaily/2005/mar/redcross090805.html`, September 2005. UDaily.

[3] D. Applegate, W. Cook, S. Dash, and A. Rohe. Solution of a min-max vehicle routing problem. *INFORMS Journal on Computing*, 14(2):132–143, 2002.

[4] P. Augerat, J. M. Belenguer, E. Benavent, A. Corbern, D. Naddef, and G. Rinaldi. Computational results with a branch and cut code for the capacitated vehicle routing problem. Technical Report 949-M, Universite Joseph Fourier, Grenoble, France, 1995.

[5] I. Averbakh and O. Berman. A heuristic with worst-case analysis for minimax routing of two travelling salesman on a tree. *Discrete Applied Mathematics*, 68:17–32, 1996.

[6] B. Balcik and B. M. Beamon. Distribution network design for humanitarian relief chains. INFORMS Annual Meeting, 2005.

[7] G. Barbarosoglu, L. Ozdamar, and A. Cevik. An interactive approach for hierarchical analysis of helicopter logistics in disaster relief operations. *European Journal of Operational Research*, 140(1):118–133, July 2002.

[8] J. F. Bard, G. Kontoravdis, and G. Yu. A branch-and-cut procedure for the vehicle routing problem with time windows. *Transportation Science*, 2002.

[9] B. M. Beamon. Humanitarian relief chains: Issues and challenges. In *The 34th International Conference on Computers and Industrial Engineering*, 2004. `http://faculty.washington.edu/benita/sfpaper.pdf`.

[10] B. M. Beamon and S. A. Kotleba. Inventory modeling for complex emergencies in humanitarian relief operations. Technical report, University of Washington, 2006.

[11] R. Bowerman, B. Hall, and P. Calamai. A multi-objective optimization approach to urban school bus routing: Formulation and solution method. *Transportation Research A*, 29A(2): 107–123, 1995.

[12] A. Campbell, N. Leifker, and B. Thomas. Location problems considering arc failure. Working paper, 2006.

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

[14] G. Clarke and J. W. Wright. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12:568–581, 1964.

[15] R.F. Dell, R. Batta, and M.H. Karwan. The multiple vehicle TSP with time windows and equity contraints over a multiple day horizon. *Transportation Science*, 30(2):120–133, 1996.

[16] J. Desai, H.D. Sherali, and T.S. Glickman. Allocating emergency response resources to minimize risk under equity considerations. *American Journal of Mathematical and Management Sciences*, 24:367–410, 2004.

[17] P.M. Dixon, J. Weiner, T. Mitchell-Olds, and R. Woodley. Bootstrapping the gini coefficient of inequality. *Ecology*, 68:1548–1551, 1987.

[18] T.A. Feo and M.G.C. Resende. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6:109–133, 1995.

[19] E. D. França, M. Gendreau, G. Laporta, and F. M. Müller. The m-traveling salesman problem with minmax objective. *Transportation Science*, 29:267–275, 1995.

[20] R.S. Garfinkel, W. Neebe, and M.R. Rao. An algorithm for the m-median location problem. *Transportation Science*, 8:217–236, 1974.

[21] B. Golden, E. Wasil, J. Kelly, and I. M. Chao. The impact of metaheuristics on solving the vehicle routing problem: Algorithms, problem sets, and computational results. In T. Crainic and G. Laporte, editors, *Fleet Management and Logistics*, pages 33–56. Kluwer Academic Publishers, 1998.

[22] R. Gopalan, K.S. Kolluri, R. Batta, and M. H. Karwan. Modeling equity of risk in the transportation of hazardous material. *Operations Research*, 38:961–973, 1990.

[23] S.L. Hakimi. Optimum location of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12:450–459, 1964.

[24] ILOG, Inc. *ILOG CPLEX 10.0, User Manual*, 2005.

[25] Fritz Institute. `www.fritzinstitute.org`.

[26] The Fritz Institute. Logistics and the effective delivery of humanitarian relief. `http://www.fritzinstitute.org/PDFs/Programs/TsunamiLogistics0605.pdf`, May 2005.

[27] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control for communication networks: shadow prices, proportional fairness, and stability. *Journal of the Operational Research Society*, 49: 237–252, 1998.

[28] G. A. P. Kindervater and M. W. P. Savelsbergh. Vehicle routing: Handling edge exchanges. In E. H. L. Aarts amd J. K. Lenstra, editor, *Local Search in Combinatorial Optimization*, pages 337–360. Wiley, 1997.

[29] R.C. Larson. Perspectives on queues: Social justice and the psychology of queuing. *Operations Research*, 35:895–905, 1987.

[30] D. Long. Logistics for disaster relief. *IIE Solutions*, 29(6):1085–1259, June 1997.

[31] M.B. Mandell. Modelling effectiveness equity trade-offs in public service delivery systems. *Management Science*, 37:467–482, 1991.

[32] N. Megiddo and A. Tamir. New results on p-center problems. *SIAM Journal of Computing*, 12:751–758, 1983.

[33] W. Ogryczak. Inequality measures and equitable approaches to location problems. *European Journal of Operational Research*, 122:374–391, 2000.

[34] W. Ogryczak. Inequality measures and equitable locations. Technical report, Institute of Control and Computation Engineering, Warsaw University of Technology, 2005.

[35] W. Ogryczak, T. Sliwinski, and A. Wierzbicki. Fair resource allocation schemes and network dimensioning problems. *Journal of Telecommunications and Information Technology*, pages 34–42, 2003.

[36] L. Özdamar, E. Ekinci, and B. Küçükyazici. Emergency logistics planning in natural disasters. *Annals of Operations Research*, 129:217–245, 2004.

[37] S. J. Pettit and A. K. C. Beresford. Emergency relief logistics: an evaluation of military, non-military and composite response models. *International Journal of Logistics*, 8(4):313 – 331, December 2005.

[38] M. Pinedo. *Scheduling: theory, algorithms, and systems.* Prentice Hall, 2002.

[39] C.S. ReVelle. Urban public facility location. In E.S. Mills, editor, *Handbook of Regional and Urban Economics*, volume 2. Elsevier Science, Amsterdam, 1987.

[40] C.S. ReVelle and R. Swain. Central facilities location. *Geographical Analysis*, 2:30–42, 1970.

[41] D. J. Rosenkrantz, R. E. Stearns, and P. M. Lewis. An analysis of several heuristics for the traveling salesman problem. *SIAM Journal on Computing*, 6(3):563–581, 1977.

[42] A. Sen. *On Economic Inequality.* Clarendon Press, Oxford, 1973.

[43] S. Shane and R. Bonner. Annan nudges donors to make good on full pledges. *New York Times*, page 12, January 7 2005.

[44] Z. Shen, M. Dessouky, and F. Ordóñez. Stochastic vehicle routing problem for large-scale emergencies. Technical Report 2005-06, Department of Industrial and Systems Engineering, University of Southern California, 2005.

APPENDIX

We now complete the proof of Lemma 1 from section 5.2.2. Following Figure 7, it is required to show that the optimal minsum route will select cliques of nodes in descending order. Note that the order of nodes *within* a clique is arbitrary as long as the nodes are visited consecutively.

Proof. Assume that a route is given $R : 0 - C_{i_1} - C_{i_2} - \ldots - C_{i_P} - C_K - C_{i_{P+1}} - \ldots$ where C_K is the highest indexed clique in the graph and is not first in route R. We will show that route $R' : 0 - C_K - C_{i_1} - C_{i_2} - \ldots - C_{i_P} - C_{i_{P+1}} - \ldots$ obtained by placing C_K first has a strictly better minsum objective. Let t_{ij} represent the distance between C_i and C_j. The sum of arrivals for R and R' are:

$$sa(R) = (N)t_{0i_1} + (N - N_{i_1})t_{i_1 i_2} + \ldots + (N - \sum_{j=1}^{P} N_{i_j})t_{i_P K} + (N - N_K - \sum_{j=1}^{P} N_{i_j})t_{K i_{P+1}} + \ldots$$

$$sa(R') = (N)t_{0K} + (N - N_K)t_{K i_1} + \ldots + (N - N_K - \sum_{j=1}^{P-1} N_{i_j})t_{i_{P-1} i_P} + (N - N_K - \sum_{j=1}^{P} N_{i_j})t_{i_P i_{P+1}} + \ldots$$

where $N = \sum_{j=1}^{K} N_j$. Note that the final portion of these sums is the same for both routes. We wish to show that the difference $sa(R) - sa(R')$ is strictly positive. This difference can be simplified by adding and subtracting the quantity $(N_K)t_{i_P K}$ to obtain:

$$\begin{aligned} sa(R) - sa(R') &= sa(R) - (N_K)t_{i_P K} + (N_K)t_{i_P K} - sa(R') \\ &= (N - N_K - \sum_{j=1}^{P} N_{i_j})(t_{i_P K} + t_{K i_{P+1}} - t_{i_P i_{P+1}}) + N(t_{0i_1} - t_{0K}) + N_K(t_{i_P K} + \sum_{j=1}^{P-1} t_{i_j i_{j+1}}) - (N - N_K)t_{i_1 K} \\ &\geq N(t_{0i_1} - t_{0K}) + N_K(t_{i_P K} + \sum_{j=1}^{P-1} t_{i_j i_{j+1}}) - (N - N_K)t_{i_1 K} \end{aligned}$$

The inequality follows from dropping the first term, which is positive since $N \geq N_K + \sum_{j=1}^{P} N_{i_j}$ and $t_{i_P K} + t_{K i_{P+1}} \geq t_{i_P i_{P+1}}$ by the triangle inequality. This holds for any route where C_K is not first. If C_K is last, then omitting terms involving $C_{i_{P+1}}$ shows this inequality to still hold. By using the fact that $t_{0K} = M$ in the graph, we arrive at:

$$sa(R)-sa(R') \geq N(t_{0i_1}-M)+N_K(t_{i_P K}+\sum_{j=i}^{P-1} t_{i_j i_{j+1}})-(N-N_K)t_{i_1 K} \geq Nt_{0i_1}-NM+(2N_K-N)t_{i_1 K} \tag{A-1}$$

This last inequality in (A-1) holds because $t_{i_P K} + \sum_{j=1}^{P-1} t_{i_j i_{j+1}} \geq t_{i_1 K}$ due to the triangle inequality. Recall that the number of nodes in clique C_K is given by $N_K = M^{K-1}$ and the total number of nodes in the graph is $N = \frac{M^K-1}{M-1}$. Hence for $M > 2$ the last coefficient, $(2N_K - N)$, in (A-1) is positive. The smallest possible value of $Nt_{0i_1} - NM + (2N_K - N)t_{i_1 K}$ occurs when $C_{i_1} = C_{K-1}$, where $t_{0i_1} = 2$ and $t_{i_1 K} = M$. These values show $sa(R) - sa(R') \geq 2$.

Thus if C_K is not first on a route, then the route can be improved by placing clique C_K first. This argument serves as the basis for an induction proof. We assume that a route is given where the first j cliques are in descending order as $R_1 : 0 - C_K - C_{K-1} - \ldots - C_{K-j+1} - \ldots$, but clique C_{K-j} is not the successor of C_{K-j+1}. Similar to before, let route R'_1 have clique C_{K-j} following C_{K-j+1}. The difference $sa(R_1) - sa(R'_1)$ will eliminate the portions of the two routes that are the same, which now include the descending cliques $C_K, \ldots, C_{K-j+1}$. We can remove nodes $C_K, \ldots, C_{K-j+2}$ from the graph and let C_{K-j+1} represent the depot instead. By symmetry of the graph, the distances involved will be the same and the remaining graph is a smaller instance of the original. This allows the basis argument to be used to establish the result that route R'_1 is an improvement. Hence, the only route that cannot be improved for the minsum objective on this graph is the route that visits all cliques in descending order. $\square$