

# Election Day Routing of Rapid Response Attorneys

Justin C. Goodson

July 29, 2014

## Abstract

Since the United States presidential election of 2000, political campaigns have deployed volunteer lawyers to provide real-time information and assistance to poll observers. To assist with the routing of these *rapid response attorneys* (RRAs) among polling locations, we formulate two optimization models to equitably and efficiently allocate volunteer resources. We propose heuristic solution procedures based on simulated and compressed annealing and describe how our efforts to route RRAs in one state benefited election day operations during the 2012 election.

**Keywords:** election operations, equitable routing, simulated and compressed annealing.

## 1 Introduction

Election operations give rise to a variety of operational challenges including routing problems for voting machine delivery (Fry and Ohlmann, 2009), equitable allocation of voting machines to precincts (Yang et al., 2013), politician routing (Cook, 2011), political districting (Mehrotra et al., 1998; King et al., 2012), and excessive voter wait times (Allen and Bernshteyn, 2006). Since the controversial United States presidential election of 2000, campaigns have devoted significant resources to ensure polling places follow established procedures. In this paper, we focus on the management of these resources, specifically the operational task faced by United States campaign operations teams to efficiently and equitably manage volunteer lawyers on election days. We discuss the implementation of our methods in one state on election day 2012.

During United States presidential campaigns, two types of volunteers provide crucial election day support: *poll observers* and *rapid response attorneys* (RRAs). A poll observer assists with election day activities at polling places by helping election officials identify problems such as illegal electioneering, voting without proper identification, voting at the wrong polling location, early opening or late closing of polling locations, intimidation of voters, and insufficient ballots. RRAs rove among multiple polling locations, providing real-time information, legal expertise, and assistance to poll observers.

During the 2012 election, one state wanted to determine the most effective and efficient way to route attorneys among polling locations. The operational task facing the state-level campaign team required the assignment of each RRA to multiple polling locations and the routing of RRAs among their assigned locations. Given a sequence of polling locations, a single RRA repeatedly visits locations throughout the day in the order prescribed by their route. If an urgent issue arises at a particular location, the RRA travels directly to that polling place, interrupting the assigned visit sequence. After the issue is resolved, the RRA continues their assigned route from where they left off.

The operational problem faced by the campaign operations team is characterized by three considerations. First, to enable quick response times to issues requiring immediate attention, the geographic dispersion of polling locations assigned to each RRA should be tight. Second, prescribed RRA routes should minimize travel time, affording RRAs more time at polling locations. Third, with the aim of treating volunteers fairly, the workload should be distributed equitably among RRAs.

To address these considerations, we develop two optimization models and heuristic solution procedures to equitably and efficiently route RRAs. Our optimization models differ in their treatment of equitable workload distribution across RRAs. In the first formulation, we follow the guidelines imposed by the campaign operations team: an equal number of polling locations is assigned to each RRA route. Results of this model were implemented in one state on election day 2012. In the second model we consider an alternative definition of equity based on travel and service time. In a post-election-day analysis, we explore how these two approaches to modeling equity can lead

to different routing schemes.

Our solution approach for the first model consists of two stages. In the first stage, we identify geographically tight clusters of polling locations, operating under the assumption that close proximity will lead to short RRA travel times among locations in a given cluster. The heuristic solution procedure employs simulated annealing to guide the search of *swap* neighborhoods. We show how to quickly evaluate solutions in the swap neighborhood of a current solution. For small problem instances where optimal solution values can be verified, our computational results show that the simulated annealing procedure identifies optimal solutions. Although we tailor our local search to the clustering of polling locations, it should be possible to adapt the procedure to other types of clustering problems. In the second stage, we determine the travel-time-minimizing visit sequence among polling locations in each cluster by optimally solving small instances of the classical traveling salesman problem (TSP).

Our solution approach for the second model simultaneously treats assignments of polling locations and routing decisions. We take as our objective the minimization of total travel and service time such that the difference between the largest and smallest route travel times satisfies a given tolerance level. Our compressed annealing heuristic punishes violations of the equity constraint via a penalty term in the objective function. Computational experiments demonstrate that as the tolerance for inequity decreases, the total routing cost increases, but only marginally for reasonable tolerance levels.

The need for quick solution development and easy-to-explain routes were primary drivers behind our choice of the first model and solution procedure. Because of the short time span allotted to create the solution, the typical phases of problem definition, problem formulation, solution development, and implementation proceeded rapidly. Consequently, future implementations may benefit from deeper exploration of some issues. Our second problem formulation further explores the issue of equitable work distribution. In the conclusion, we discuss additional issues that may merit more investigation.

Overall, our statewide routing solution for the 2012 election significantly accelerated the task of routing RRAs and more efficiently allocated the time of volunteer attorneys. Prior to our involve-

ment, RRA routes were developed “by hand” and without the guidance of optimization models and methods. Subsequent applications of our optimization routines are likely to realize even larger benefits in states where poll observers are placed at larger numbers of polling locations and “by-hand” routing methods are even more likely to yield suboptimal RRA routes. In particular, as campaigns rely more and more on person-to-person contact and “grass-roots” politics, the need for effective workforce management increases. Our work may serve as a baseline for coordinating microtargeting efforts such as “persuadables” targeting and “get-out-the-vote” targeting. These and related campaign activities that depend on volunteer work may realize significant productivity increases by utilizing optimization techniques to guide operational decision making.

The remainder of the paper proceeds as follows. We present our models and solution procedures in §2 and §3. In §4, we report on the application of our methods in one state on election day 2012 and explore how different treatments of equity lead to different solution structures. We conclude the paper in §5.

## **2 Location-Based Equity Model and Solution Procedure**

In this section, we model equitable workloads by requiring polling locations to be distributed equally among RRAs. We present a two-stage solution procedure that first assigns polling locations to RRAs and second routes RRAs among their assigned polling locations. We formulate the first stage as a clustering problem in §2.1 and present a heuristic solution method in §2.2. We treat the second-stage routing problem as a classical TSP. Because the TSP instances we encountered during the 2012 election were small (no more than nine polling locations), optimal solutions were quickly identified via enumeration. Alternative solution techniques could be employed for larger problem instances. We refer the reader to Gutin and Punnen (2002) for a review of exact and heuristic TSP solution procedures. The remainder of this section focuses on the first-stage solution procedure.

## 2.1 Clustering Problem Formulation

The problem of identifying geographically tight groups of polling locations is similar to the multi-source Weber problem (MWP), a type of clustering problem and a model for many location-allocation problems. The MWP is characterized by a set of customers (polling places) with given locations and a set of facilities (RRAs) to be located. The MWP requires facilities to be located and customers assigned to facilities such that the sum of the Euclidian distances from each customer to its assigned facility is minimized. Applications of the MWP include partitioning of sales territories (Fleischmann and Paraschis, 1988) and placement of warehouses (Bhaskaran, 1992), airports (Saaty, 1972), and emergency services (Dokmeci, 1977). Rosing (1992) and Krau (1997) examine exact solution methods for the MWP and Brimberg et al. (2000) compare and contrast heuristic schemes.

The model we design in this section to assign polling locations to RRAs differs from the MWP in two ways. First, per the guidelines imposed by the campaign operations team, we require RRAs to service equal numbers of polling locations. Second, because a pairing of a RRA to a group of polling locations is not entirely analogous to locating a facility, we use geographic centroids to evaluate solutions. Specifically, we evaluate a feasible solution by summing Euclidian distances between each polling location and the centroid of that polling location's group.

Denote the set of polling locations by  $\mathcal{L} = \{1, 2, \dots, L\}$ . The horizontal and vertical coordinates of polling location  $l$  in  $\mathcal{L}$  are  $x_l$  and  $y_l$ , respectively. We desire to partition the set  $\mathcal{L}$  among clusters  $\mathcal{C} = \{1, 2, \dots, C\}$ , where we assume  $C \leq L$ , such that the sum of Euclidian distances from each location to its cluster's centroid is minimized. To equitably distribute the workload among RRAs, we require the number of locations assigned to each cluster to be at least  $\lfloor L/C \rfloor$  and no more than  $\lceil L/C \rceil$ , where  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  are the floor and ceiling operators, respectively.

We formulate the optimization problem as a non-linear integer program, with decision variable  $z_{lc}$  set to one if location  $l$  is assigned to cluster  $c$  and set to zero otherwise. We denote a solution vector by  $z = (z_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$ .

$$\min \left\{ f(z) = \sum_{l \in \mathcal{L}} \sum_{c \in \mathcal{C}} z_{lc} \left( (x_l - \mu_c^x)^2 + (y_l - \mu_c^y)^2 \right)^{\frac{1}{2}} : \right. \quad (1)$$

$$\mu_c^x = \frac{\sum_{l \in \mathcal{L}} z_{lc} x_l}{\sum_{l \in \mathcal{L}} z_{lc}}, c \in \mathcal{C}, \quad (2)$$

$$\mu_c^y = \frac{\sum_{l \in \mathcal{L}} z_{lc} y_l}{\sum_{l \in \mathcal{L}} z_{lc}}, c \in \mathcal{C}, \quad (3)$$

$$\sum_{c \in \mathcal{C}} z_{lc} = 1, l \in \mathcal{L}, \quad (4)$$

$$\max_{c \in \mathcal{C}} \left\{ \sum_{l \in \mathcal{L}} z_{lc} \right\} - \min_{c \in \mathcal{C}} \left\{ \sum_{l \in \mathcal{L}} z_{lc} \right\} \leq \delta \quad (5)$$

$$z_{lc} \in \{0, 1\}, \{(l, c) : l \in \mathcal{L}, c \in \mathcal{C}\}. \quad (6)$$

Equation (1) is the objective function, with centroid coordinates defined by equations (2) and (3). Equation (4) requires each location be assigned to exactly one cluster. Inequity tolerance parameter  $\delta$  in equation (5) is a user-supplied integer in the range  $[1, L]$  indicating the decision maker's tolerance for the difference between the largest number of polling locations assigned to a route and the smallest. Setting  $\delta = 1$  enforces the campaign team's requirement that each RRA be assigned at least  $\lfloor L/C \rfloor$  polling locations and no more than  $\lceil L/C \rceil$ . When  $\delta = 1$ , equation (5) implies  $L \bmod C$  clusters each have  $\lceil L/C \rceil$  locations assigned to them and the remaining  $C - L \bmod C$  clusters are each assigned  $\lfloor L/C \rfloor$  locations. The simulated annealing heuristic of §2.2 is designed for the case  $\delta = 1$ .

In §3, we consider an alternative to constraint (5) and define equity in terms of travel and service time assigned to each RRA. The computational experiments of §4.3 explore differences between these two definitions of equity and examine the effects of increasing  $\delta$  beyond 1.

The number of solution vectors satisfying equation (6) is  $2^{C \times L}$ . Constraints (4) reduce this figure to  $C^L$ . To identify an optimal solution, one can enumerate these solutions, discarding solutions not in compliance with constraint (5). In §4.2, we implement this procedure to determine optimal solution values for small problem instances. For many problems of practical interest, the enumeration procedure becomes computationally prohibitive, unable to identify an optimal solution after days of computing time. In these cases, we rely on the simulated annealing heuristic described in

the next section.

## 2.2 Simulated Annealing Heuristic

To solve clustering problem (1)-(6), we use simulated annealing to guide the search of *swap* neighborhoods. Simulated annealing is a local search algorithm in which non-improving moves are probabilistically accepted in an attempt to avoid becoming trapped in a low-quality, locally-optimal solution (Kirkpatrick et al., 1983; Johnson et al., 1989, 1991). We choose simulated annealing to direct the search because of its wide success across a variety of problem types and for its straightforward implementation. Further, in contrast to most metaheuristics, simulated annealing converges in probability to the set of global minima (Hajek, 1988).

Given a feasible assignment of locations to clusters, the swap neighborhood consists of all solutions that can be obtained by swapping location  $l$  assigned to cluster  $c$  with location  $l'$  assigned to a different cluster  $c'$ . More formally, to construct a solution vector  $\bar{z}$  in the swap neighborhood of  $z$ , first select location  $l$  from  $\{\hat{l} : z_{\hat{l}c} = 1\}$  and location  $l'$  from  $\{\hat{l}' : z_{\hat{l}'c'} = 1\}$ . Then, set  $\bar{z}$  equal to  $z$  and perform the swap by subsequently setting  $\bar{z}_{lc} = 0$ ,  $\bar{z}_{l'c'} = 0$ ,  $\bar{z}_{lc'} = 1$ , and  $\bar{z}_{l'c} = 1$ . Solutions in the swap neighborhood of a feasible solution are also feasible because they satisfy constraints (4)-(6).

A solution vector  $\bar{z}$  obtained from  $z$  by swapping two polling locations in clusters  $c$  and  $c'$  can be evaluated by directly calculating  $f(\bar{z})$  via equation (1). However, because  $\bar{z}$  differs from  $z$  in the assignment of only two polling locations, the computation required to calculate  $f(\bar{z})$  can be reduced by modifying  $f(z)$  to reflect the changes in clusters  $c$  and  $c'$ . Algorithm 1 details the procedure, where we assume centroid coordinates for each cluster in solution vector  $z$  have been stored. Line 1 calculates the new centroid coordinates for clusters  $c$  and  $c'$  of solution vector  $\bar{z}$ . Line 2 initializes  $f(\bar{z})$  with  $f(z)$ . Lines 3 and 4 subtract from  $f(\bar{z})$  the objective function contributions of clusters  $c$  and  $c'$  in solution vector  $z$ . Finally, lines 5 and 6 add to  $f(\bar{z})$  the objective function contributions of updated clusters  $c$  and  $c'$  in solution vector  $\bar{z}$ . Our computational experience suggests Algorithm 1 significantly reduces the computation required to evaluate a solution.

Our simulated annealing procedure is displayed in Algorithm 2. Three solution vectors are

---

**Algorithm 1** Evaluation of  $\bar{z}$  in the Swap Neighborhood of  $z$ 

---

- 1: For  $\bar{z}$  calculate centroid coordinates  $\mu_c^x, \mu_c^y, \mu_{c'}^x,$  and  $\mu_{c'}^y$  via equations (2) and (3)
  - 2:  $f(\bar{z}) \leftarrow f(z)$
  - 3:  $f(\bar{z}) \leftarrow f(\bar{z}) - \sum_{l \in \{\hat{l}: z_{i_c}=1\}} ((x_l - \mu_c^x)^2 + (y_l - \mu_c^y)^2)^{\frac{1}{2}}$
  - 4:  $f(\bar{z}) \leftarrow f(\bar{z}) - \sum_{l \in \{\hat{l}: z_{i_{c'}}=1\}} ((x_l - \mu_{c'}^x)^2 + (y_l - \mu_{c'}^y)^2)^{\frac{1}{2}}$
  - 5:  $f(\bar{z}) \leftarrow f(\bar{z}) + \sum_{l \in \{\hat{l}: \bar{z}_{i_c}=1\}} ((x_l - \mu_c^x)^2 + (y_l - \mu_c^y)^2)^{\frac{1}{2}}$
  - 6:  $f(\bar{z}) \leftarrow f(\bar{z}) + \sum_{l \in \{\hat{l}: \bar{z}_{i_{c'}}=1\}} ((x_l - \mu_{c'}^x)^2 + (y_l - \mu_{c'}^y)^2)^{\frac{1}{2}}$
- 

maintained –  $z_{\text{best}}$ ,  $z_{\text{curr}}$ , and  $z_{\text{neigh}}$  – corresponding to the best-found, the current, and a neighbor solution vector, respectively. An iteration of the inner loop begins on line 3 by randomly selecting a solution  $z_{\text{neigh}}$  from the swap neighborhood of current solution  $z_{\text{curr}}$ . Line 4 probabilistically updates  $z_{\text{curr}}$  with  $z_{\text{neigh}}$ , where  $(a)^+ = a$  if  $a > 0$  and 0 otherwise and  $\tau$  is the current temperature (a control parameter in simulated annealing). Lines 5 and 6 update the best-found solution vector. The inner loop terminates after 10,000 iterations. The outer loop implements a geometric cooling schedule with a temperature multiplier of 0.99 and terminates after at least 1000 iterations and without having updated the best-found solution for 500 successive iterations. We set the initial temperature to  $\tau = 5$ . To obtain an initial feasible solution, we randomly assign polling locations to clusters such that  $L \bmod C$  clusters each have  $\lceil L/C \rceil$  locations assigned to them and the remaining  $C - L \bmod C$  clusters are each assigned  $\lfloor L/C \rfloor$  locations. Our computational experience suggests these parameters yield high-quality solutions for the problem instances we consider in this work.

### 3 Time-Based Equity Model and Solution Procedure

Per the request of the campaign operations team, we modeled equitable workloads in §2 by assigning equal numbers of polling locations to RRAs. This requirement, combined with a very short timetable for development and implementation, prevented the exploration of alternative models for equitable work distribution. In this section, we present a post-election-day analysis that treats equity in terms of route duration and service time. Our model allows the decision maker to determine the level of equity he or she deems necessary.



---

**Algorithm 2** Simulated Annealing Procedure

---

```
1: repeat
2:   repeat
3:     Randomly select  $z_{\text{neigh}}$  in the swap neighborhood of  $z_{\text{curr}}$ 
4:      $z_{\text{curr}} \leftarrow z_{\text{neigh}}$  with probability  $\exp\{-(f(z_{\text{neigh}}) - f(z_{\text{curr}}))^+/\tau\}$ 
5:     if  $f(z_{\text{curr}}) < f(z_{\text{best}})$  then
6:        $z_{\text{best}} \leftarrow z_{\text{curr}}$ 
7:   until 10,000 iterations
8:    $\tau \leftarrow \tau \times 0.99$ 
9: until At least 1000 iterations and not update  $z_{\text{best}}$  for 500 successive iterations
```

---

A time-based approach to equity was motivated by the route structures obtained via solution method one: due to the geographic spread of polling locations in some regions, travel times for some routes were much longer than travel times for other routes. Consideration of service time in addition to route duration balances travel time with the number of polling locations assigned to each RRA. We present our problem formulation in §3.1 and a heuristic solution procedure based on compressed annealing in §3.2.

### 3.1 Problem Formulation

In contrast to the two-stage approach of §2, the problem formulation in this section simultaneously treats assignments of polling locations and routing decisions. As in §2.1, let  $\mathcal{C} = \{1, 2, \dots, C\}$  be a set of  $C$  RRAs. Denote by  $v^c = (v_1^c, v_2^c, \dots, v_{I^c}^c)$  a sequence of  $I^c$  locations assigned to RRA  $c$  in  $\mathcal{C}$ , where each  $v_i^c$  on route  $v^c$  belongs to the set of polling locations  $\mathcal{L} = \{1, 2, \dots, L\}$ . Denote by  $v = (v^c)_{c \in \mathcal{C}}$  a collection of RRA routes such that each polling location in  $\mathcal{L}$  appears exactly once on exactly one route in  $v$  and at least one polling location is assigned to each route. Denote by  $t(l, l')$  the travel time from polling location  $l$  to polling location  $l'$  and by  $s_l$  the average service time at polling location  $l$ . Let  $g(v^c) = t(v_{I^c}^c, v_1^c) + s_1 + \sum_{i=2}^{I^c} (t(v_{i-1}^c, v_i^c) + s_i)$  be the total time required to traverse and service route  $v^c$  and let  $h(v) = \sum_{c \in \mathcal{C}} g(v^c)$  be the total time required to traverse and service all routes. Denoting by  $\mathcal{V}$  the set of all route collections, the problem we seek

to solve is

$$\min \left\{ h(v) : \right. \quad (7)$$

$$\left. \max_{c \in \mathcal{C}} \{g(v^c)\} - \min_{c \in \mathcal{C}} \{g(v^c)\} \leq \gamma, \quad (8)$$

$$\left. v \in \mathcal{V} \right\}, \quad (9)$$

where  $\gamma$  is a parameter supplied by the user to control tolerance of inequity, as defined by the difference between the largest and smallest route times. Large values of  $\gamma$  permit large differences in route times while small values require route times to be similar. Thus, by selecting appropriate values for  $\gamma$ , the decision maker can enforce strict equity among RRA assignments, can disregard equity altogether, or can select an intermediate alternative.

Problem (7)-(9) is most similar to the no-depot multiple TSP of Na (2006), which takes as the objective to minimize the longest route. Although a minimax objective is an often-employed inequity metric, for routing problems, minimizing the longest route does not guarantee efficient visit sequences for routes with shorter travel and service times. The inequity measure we employ in the left-hand side of constraint (8) is referred to by Yang et al. (2013) as a *range* metric. For a class of resource allocation problems, Yang et al. (2013) take as their objective to minimize the range of inequity. In the routing context, because a wide array of route collections might achieve similar range values, we incorporate the range metric into the constraint set and keep as our objective the minimization of total travel and service time. This scheme gives preference to efficient routings that satisfy the given tolerance for inequity.

### 3.2 Compressed Annealing Heuristic

Problem (7)-(9) is composed of two main parts, a no-depot multiple TSP and an inequity constraint. The no-depot TSP is a NP-hard optimization problem and the inequity constraint presents additional feasibility difficulties. Using a penalty-based approach, we partially decompose these two components and conduct a heuristic search. We consider infeasible solutions by relaxing inequity constraint (8) into the objective function with penalty

$$p(v) = \max \left\{ 0, \max_{c \in \mathcal{C}} \{g(v^c)\} - \min_{c \in \mathcal{C}} \{g(c^c)\} - \gamma \right\}, \quad (10)$$

a function returning the portion of the range inequity exceeding tolerance parameter  $\gamma$ . Letting  $\lambda$  be a non-negative penalty multiplier, the relaxed problem we seek to solve is

$$\min \{h(v) + \lambda p(v) : v \in \mathcal{V}\}. \quad (11)$$

Any route collection  $v$  with penalty  $p(v) = 0$  is a feasible solution to both problems (11) and (7)-(9). Further, Hadj-Alouane and Bean (1997) show that for a sufficiently large  $\lambda$ , relaxed problem (11) enjoys strong duality with original problem (7)-(9).

We employ compressed annealing to solve problem (11) in search of high-quality solutions to problem (7)-(9). Compressed annealing varies the penalty multiplier  $\lambda$ , referred to as “pressure,” within the framework of traditional simulated annealing (Ohlmann et al., 2004). Over the course of the heuristic search, pressure is increased, thereby biasing the solution landscape toward inequity-feasible solutions satisfying constraint (8). We utilize compressed annealing to direct the search because of its success in other problem domains and because it provides a straightforward method for penalty-based local search. Further, unlike the vast majority of metaheuristic methods, compressed annealing converges in probability to the set of global minima (Ohlmann et al., 2004).

We use compressed annealing to guide the search of *relocation* neighborhoods. The relocation neighborhood of a route collection in  $\mathcal{V}$  consists of all route collections that can be obtained by moving a single customer to a different position on the same route or on a different route. Prohibiting relocations that result in empty routes, all route collections in the neighborhood of a given solution remain feasible for problem (11). We implement relocation neighborhoods as discussed in Kindervater and Savelsbergh (2003), calculating solution costs via standard “delta” evaluations.

Our compressed annealing procedure is detailed in Algorithm 3. As with the simulated annealing procedure described in §2.2, three solution vectors are maintained –  $v_{\text{best}}$ ,  $v_{\text{curr}}$ , and  $v_{\text{neigh}}$  – corresponding to the best-found, the current, and a neighbor route collection, respectively. Line 1 initializes iteration counter  $k$  to zero. An iteration of the inner loop begins on line 4 by randomly selecting a route collection  $v_{\text{neigh}}$  from the relocation neighborhood of current collection  $v_{\text{curr}}$ . Line

5 probabilistically updates  $v_{\text{curr}}$  with  $v_{\text{neigh}}$ , where parameter  $\tau$  is the current temperature. Lines 6 and 7 update the best-found solution to be the least infeasible route collection with the smallest total travel and service time. The inner loop terminates after 80,000 iterations. The outer loop increments  $k$ , updates the temperature via a geometric cooling schedule with a temperature multiplier of 0.99, and updates the pressure via a limited exponential compression schedule with a pressure cap of 900,000 and a compression coefficient of 0.06. The outer loop terminates after at least 800 iterations and without having updated the best-found route collection for 100 successive iterations. We set the initial temperature to  $\tau = 350$  and the initial pressure to  $\lambda = 0$ . We obtain an initial route collection by randomly assigning polling locations to RRAs such that each RRA is assigned to at least one polling location. Our selected parameter values are tailored to our largest problem instance and are obtained via the insights of Ohlmann and Thomas (2007).

---

**Algorithm 3** Compressed Annealing Procedure

---

```

1:  $k \leftarrow 0$ 
2: repeat
3:   repeat
4:     Randomly select  $v_{\text{neigh}}$  in the relocation neighborhood of  $v_{\text{curr}}$ 
5:      $v_{\text{curr}} \leftarrow v_{\text{neigh}}$  with probability  $\exp\{-(h(v_{\text{neigh}}) + \lambda p(v_{\text{neigh}}) - h(v_{\text{curr}}) - \lambda p(v_{\text{curr}}))^+ / \tau\}$ 
6:     if  $p(v_{\text{curr}}) \leq p(v_{\text{best}})$  and  $h(v_{\text{curr}}) < h(v_{\text{best}})$  then
7:        $v_{\text{best}} \leftarrow v_{\text{curr}}$ 
8:     until 80,000 iterations
9:    $\tau \leftarrow \tau \times 0.99$ 
10:   $k \leftarrow k + 1$ 
11:   $\lambda \leftarrow 900,000(1 - \exp\{-0.06k\})$ 
12: until  $k \geq 800$  iterations and not update  $v_{\text{best}}$  for 100 successive iterations

```

---

## 4 Computational Experience

In §4.1, we describe the problem instances encountered in one state during the 2012 United States presidential election. In §4.2, we report on the application of the location-based equity model and solution procedure to route RRAs on election day 2012. In §4.3, we use the location- and time-based equity models and solution procedures to explore the tradeoffs associated with inequity tolerance.

We implement our procedures in C++ and execute all experiments on 2.8GHz Intel Xeon processors with 12-48GB of RAM and the CentOS 5.3 operating system. The annealing procedures outlined in §2.2 and §3.2 execute quickly, requiring less than one CPU minute for the largest problem instance.

### 4.1 Problem Instances

We were told poll observers would be located in 208 polling locations among nine regions across the state. The first three columns of Table 1 display the number of polling locations and RRAs assigned to each region. Regions 1-6 were each assigned two or more RRAs and regions 7-9 were each assigned one RRA. We discuss the remainder of Table 1 in §4.2.

We use the latitude and longitude of each polling location as the horizontal and vertical coordinates required as input to the clustering problem described in §2.1. To convert physical addresses into latitude and longitude coordinates we use a batch geocoder publicly available at [www.findlatitudeandlongitude.com/batch-geocode](http://www.findlatitudeandlongitude.com/batch-geocode). An important feature of this geocoder was the ability to check the accuracy of the conversion process. The accuracy report led us to correct several address entry errors.

As input for routing optimization, we estimate the travel time between each pair of polling locations in each region. To do this, we first modify the standard Euclidian distance matrix to approximate city driving. Given polling locations  $l$  and  $l'$  with latitude and longitude coordinates  $(\text{lat}_l, \text{long}_l)$  and  $(\text{lat}_{l'}, \text{long}_{l'})$ , we approximate the driving distance  $d_{ll'}$  between  $l$  and  $l'$  as suggested by Simchi-Levi et al. (2003):  $d_{ll'} = \rho 69((\text{long}_l - \text{long}_{l'})^2 + (\text{lat}_l - \text{lat}_{l'})^2)^{\frac{1}{2}}$  miles, where 69 is

Table 1: Location-Based Equity First-Stage Results

Region	RRAs ( $C$ )	Polling Locations ( $L$ )	Optimal	Annealing
1	14	119	–	2.205
2	3	24	–	0.395
3	2	16	0.562	0.562
4	2	12	0.206	0.206
5	2	11	0.159	0.159
6	2	10	0.165	0.165
7	1	9	N/A	N/A
8	1	5	N/A	N/A
9	1	2	N/A	N/A

the approximate number of miles per degree of latitude in the continental United States and  $\rho$  is a “circuitry factor,” suggested by Simchi-Levi et al. (2003) to be 1.3 for the regions of the state included in our study. We estimate the travel time between locations  $l$  and  $l'$  as  $t_{ll'} = (60/r_{ll'})d_{ll'}$  minutes, where  $r_{ll'}$  is the average driving speed in miles per hour from location  $l$  to location  $l'$ . Given the location of most polling places, we set  $r_{ll'} = 25$  for all  $l, l'$  pairs. We set the average service time for each polling location  $l$  to  $s_l = 15$  minutes, the average time RRAs were expected to spend at a polling location. Although our method for approximating driving distances and travel times has been proven useful in practice, improved estimates could be obtained via specialized mapping software.

## 4.2 Application on Election Day 2012

For regions 1-6, we assign polling locations to RRAs via the two-stage optimization routine outlined in §2. Per the campaign team’s requirement, inequity tolerance parameter  $\delta$  is set to 1 in equation (5). Because regions 7-9 were each assigned one RRA, we perform only the second-stage routing optimization as clustering is unnecessary.

The small sizes of the clustering problems for regions 3-6 allow the feasible solution spaces

to be enumerated and the optimal solutions to be quickly identified via the enumeration method discussed in §2.1. The fourth column of Table 1 displays the optimal solution values and the fifth column shows the best solution values returned by 10 executions of our simulated annealing heuristic on each problem instance with two or more RRAs. For regions 3-6, the annealing heuristic identifies an optimal solution. To provide a point of comparison for region 2, we benchmark against the performance of the Microsoft Excel non-linear solver, the spreadsheet-based solver that would have been available to the campaign operations team. The Excel solver identifies a solution value of 0.397, a 0.5 percent increase over the solution value identified by the simulated annealing algorithm. The annealing procedure and the Excel-based method both terminate in less than one minute. Because the size of the optimization problem for region 1 exceeds the limits of even the premium version of the Excel non-linear solver, a similar comparison for region 1 is not possible. However, the performance of the annealing procedure in regions 2-6 suggest the solution obtained for region 1 is of high quality.

The optimization-based solutions also improve upon manually-obtained solutions. This was particularly true in region 1 where a “by-hand” clustering method achieved an objective value of 2.409, a 9.25 percent increase over simulated annealing’s clustering solution for the minimization problem. Converting objective values to miles, the annealing-based solution decreases total distance to cluster centroids from 216 to 197 miles. The degree to which this level of improvement is meaningful depends on how crucial a given area is to the outcome of the election. In swing states, for example, even incremental increases in RRA efficiency can be impactful. A further benefit is the ability to automate the creation of RRA routes, a task that, prior to our involvement, required weeks of volunteer time.

We used the Google Maps API (publicly available at <https://developers.google.com/maps/>) to develop an interactive, visual implementation of the RRA routes resulting from our optimization routines. Figure 1a displays the routes developed for region 1.<sup>1</sup> In Figure 1a, markers of the same color denote polling locations assigned to a RRA and numbers in markers

---

<sup>1</sup>To maintain confidentiality, Figure 1 translates the locations of polling places in region 1 to a different area in the United States.

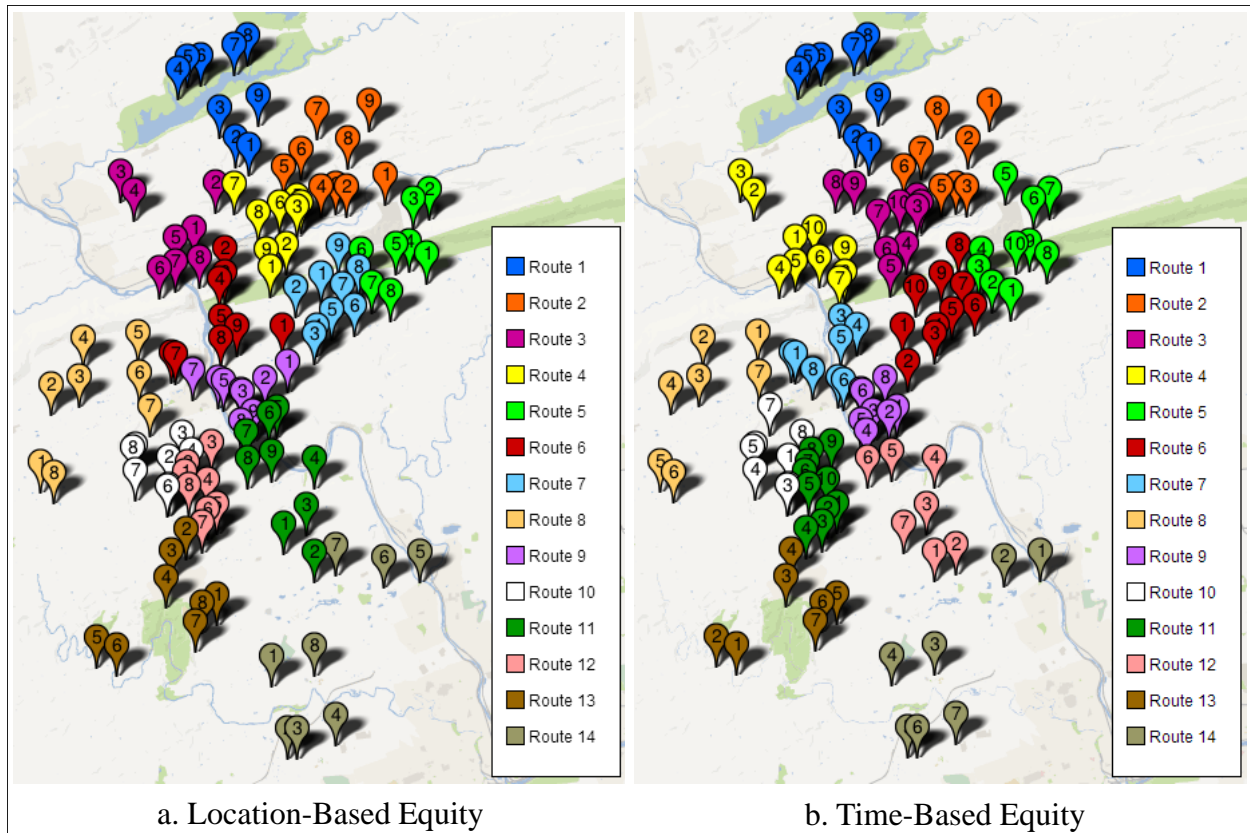


Figure 1: Routing Comparison for Region One

of the same color indicate the visit sequence. Similar maps were developed for each region and each RRA received a map displaying only their assigned route. The visual depiction of the routes produced by our procedure increased solution transparency and facilitated a seamless transition from the output of optimization routines to implementation on election day. Further, the maps and optimization routines were major contributors to a successful operation, significantly accelerating the process of routing RRAs.

### 4.3 The Cost of Equity

In this section, we explore the tradeoffs associated with parameters  $\delta$  and  $\gamma$ , the tolerance level for inequity across polling location assignments in our location-based equity model and the tolerance level for inequity across route times in our time-based equity model, respectively. In this post-election-day analysis, we examine the effects of varying  $\gamma$  via the compressed annealing procedure



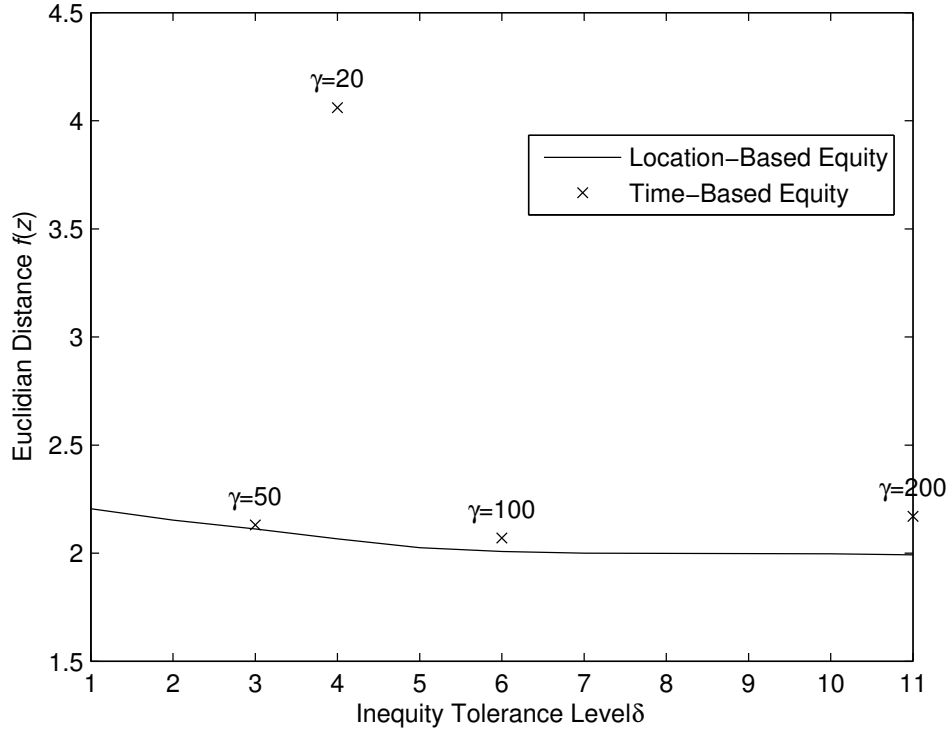


Figure 2: Effect of Inequity Tolerance Level  $\delta$  on Euclidian Distance

of §3.2 and explore the impact of varying  $\delta$  via a similar compressed annealing algorithm that penalizes violations of equity constraint (5). Throughout this section, we focus on region 1, the largest problem instance encountered during our work with the 2012 election.

Figures 2 and 3 facilitate the analysis. The line series in Figure 2 displays the total Euclidian distance to cluster centroids  $f(z)$  achieved for values of  $\delta$  ranging from 1 to 11. Increasing  $\delta$  beyond 11 renders inequity constraint (5) non-binding. Similarly, the line series in Figure 3 shows the total travel and service times  $h(v)$  achieved for values of  $\gamma$  ranging from 20 to 700 minutes. For values of  $\gamma$  less than 20, the compressed annealing procedure identifies few route collections satisfying inequity constraint (8). For values of  $\gamma$  greater than or equal to 10, feasible solutions are discovered, but increasing  $\gamma$  beyond 700 minutes renders inequity constraint (8) non-binding. The series marked by a cross in Figure 2 depicts four time-based equity solutions evaluated via the objective value of the location-based equity model. Likewise, the series marked by a cross in Figure 3 displays four location-based equity solutions evaluated via the objective of the time-based equity model. Above each marker is the value of  $\delta$  or  $\gamma$  leading to the corresponding solution.

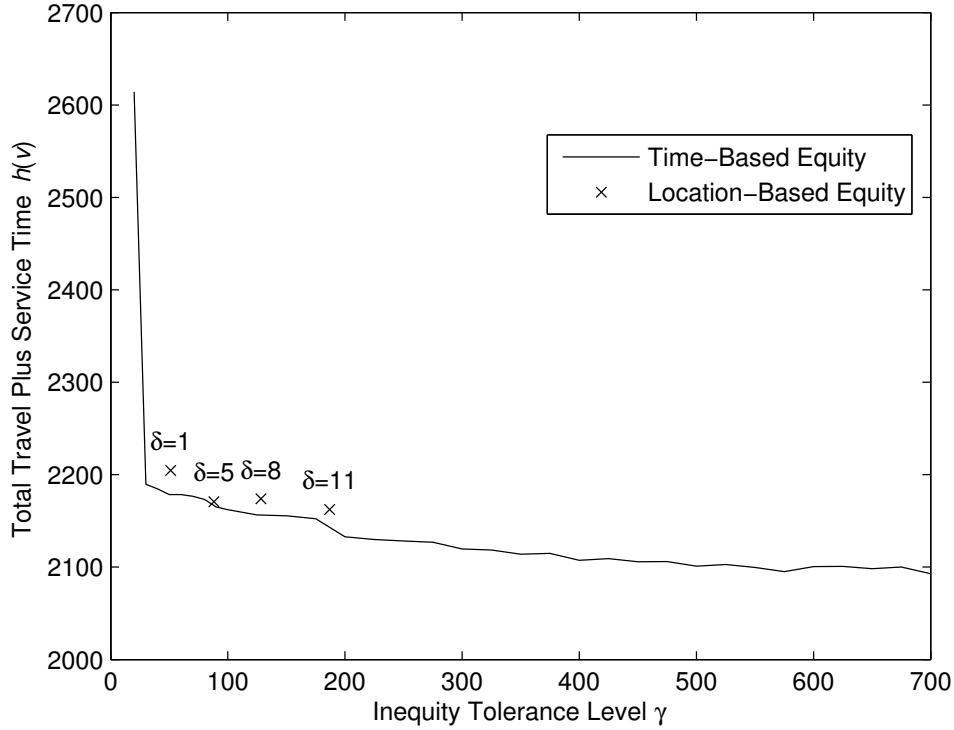


Figure 3: Effect of Inequity Tolerance Level  $\gamma$  on Total Travel and Service Time

Figure 2 indicates total Euclidian distances to cluster centroids increases as inequity tolerance decreases. Although growth occurs more rapidly as  $\delta$  is decreased below five, the overall rise in cost is not too drastic. When  $\delta$  is set to 11,  $f(z)$  is 1.992, whereas when  $\delta$  is set to one,  $f(z)$  is 2.205, representing a 10.69 percent increase. Converting these figures to miles, decreasing  $\delta$  from 11 to one reduces total mileage to cluster centroids from 198 to 179.

Examining the second series in Figure 2, we see a decrease in time-based inequity tolerance  $\gamma$  does not necessarily lead to a reduction in location-based inequity. Specifically, when  $\gamma$  is reduced from 50 minutes to 20 minutes, the spread in the number of polling locations assigned to routes increases from three to four. Further, the location-based cost of the time-based solution is nearly double the cost achieved by the location-based solution method. Thus, using time-based equity may not be a valid substitute for location-based equity.

Figure 3 indicates total travel and service time increases as inequity tolerance decreases. Although sharp growth ensues when  $\gamma$  approaches 20 minutes, the rise in routing cost is otherwise moderate. When  $\gamma$  is set to 700, routing cost  $h(v)$  is 2092.84, whereas when  $\gamma$  is set to 30,  $h(v)$

is 2189.69, representing only a 4.63 percent increase in total travel and service time. Thus, for region 1, reasonable levels of inequity tolerance are possible in exchange for incremental increases in routing cost.

Although routing costs are similar at very different values of inequity tolerance, route structures are markedly different. For larger values of  $\gamma$ , the bulk of the polling locations is assigned to only a few RRAs. As  $\gamma$  is decreased, the number of polling locations is spread across RRAs more evenly, more closely resembling the original equity requirement of the campaign operations team. In fact, when the inequity tolerance is set to 30 minutes, the largest number of polling locations assigned to a single route is 10 and the fewest is seven. This difference of three is only two shy of the campaign's equity condition  $\delta = 1$ .

The second series in Figure 3 suggests that decreasing  $\delta$  in the location-based equity model generally leads to routing schemes with less time-based inequity. The point marked by  $\delta = 1$  is the location-based equity solution implemented on Election Day 2012. The total travel and service time achieved by this route collection is 2204.50 and the difference between the largest and smallest route times is 51.26 minutes. Taking the latter value to be  $\gamma$ , the time-based equity routing cost is 2178.57, a 1.18 percent decrease over the location-based equity routing scheme. Excluding the total service time, which remains fixed for all solutions, the time-based equity routing cost decreases the location-based equity routing cost by 6.18 percent. Figures 1a and 1b place these two routing structures side-by-side. The second route collection, which measures inequity based on differences in route and service times, assigns fewer polling locations to RRA routes in less-geographically-dense areas relative to the first route collection, which measures equity via the spread of polling locations among RRA routes. This difference is particularly evident when comparing routes from each solution method in the lower portion of region 1.

The time-based equity model and solution procedure also improves upon a manually-obtained routing solution. For example, in region 1, a "by-hand" solution procedure resulted in a total travel and service time of 2217.94 minutes and a difference between the largest and smallest route times of 43.27 minutes. Setting  $\gamma$  equal to 43.27, the compressed annealing algorithm identifies a routing solution with total travel and service time of 2183.50 minutes, a 1.56 percent decrease. Excluding

the total service time, a constant for all solutions, the manually-obtained solution increases by 8.64 percent the travel time of the optimization-based routes.

The primary managerial takeaway of this discussion is how one might manage the tradeoff between cost and equity. Having selected a suitable definition of equity (location- or time-based) and given a target level of inequity tolerance, Figure 2 or Figure 3 can be used to inform the decision maker as to how overall efficiency might improve or decrease if the tolerance for inequity is adjusted. Although the discussion in this section focuses on region 1, a similar analysis could be performed for other geographic areas in future elections. While one can generally expect cost to decrease as inequity tolerance increases, the change in cost may depend on specific attributes of the region in question, such as the geographic spread of polling locations and the number of RRAs.

## 5 Conclusion and Future Research Directions

We introduce the problem of routing rapid response attorneys (RRAs) among multiple polling locations to provide real-time assistance and information to poll observers on election days. We consider two optimization models to equitably and efficiently route RRAs, one model that treats equity in terms of the assignment distribution of polling locations and another that measures equity via differences in travel and service time. We develop heuristic solution procedures based on simulated and compressed annealing. During the 2012 United States presidential election, an implementation of our methods in one state significantly accelerated the task of routing RRAs and more efficiently allocated the time of volunteer attorneys. We also explore the tradeoffs associated with our treatments of routing equity.

We suggest several avenues for future research. First, operational efficiency may be increased by explicitly accounting for potential detours from the planned RRA route. As described in §1, RRAs may break from the visit sequence to provide assistance at a polling location. Developing routes that are robust to such detours may increase operational efficiency. The notion of anticipatory route selection in vehicle routing (Thomas and White, 2004) may provide a starting point.

Second, rather than direct attorneys to rove among polling locations as required by the cam-

paign operations team, it may be more beneficial to station attorneys at fixed locations (e.g., the geographic centers of their assigned routes) that provide quick response times to the assigned polling locations. A drawback of this approach is a potential decrease in interaction among RRAs and poll observers, who may benefit from RRAs' suggestions throughout the election day. Treating the problem in this fashion may benefit from work on ambulance location (Henderson and Mason, 2005).

Third, it may be possible to augment our models to capture uncertainty in RRA service times across polling places and variation in travel times over an election day. Although the addition of stochastic service and travel times is likely to complicate the optimization model and solution procedure, accounting for such variability may provide a more realistic model. A deterministic alternative to modeling travel times as random variables is to set different travel durations as a function of time of day. The work of Figliozzi (2012) may serve as a baseline for incorporation of time-dependent travel times.

## Acknowledgements

The author wishes to express appreciation to anonymous referees and to Barrett Thomas for feedback on the paper.

## References

Allen, T. and M. Bernshteyn (2006). Mitigating voter waiting times. *Chance* 19(4), 25–36.

Bhaskaran, S. (1992). Identification of transshipment center locations. *European Journal of Operational Research* 63(2), 141–150.

Brimberg, J., P. Hansen, N. Mladenovic, and E. Taillard (2000). Improvements and comparison of heuristics for solving the multisource weber problem. *Operations Research* 48(1), 129–135.

Cook, W. (2011, December 21). The problem of the traveling politician. In *The New York*

- Times*. Retrieved from <http://campaignstops.blogs.nytimes.com/2011/12/21/the-problem-of-the-traveling-politician/>.
- Dokmeci, V. (1977). A quantitative model to plan regional health facility systems. *Management Science* 24, 411–419.
- Figliozzi, M. (2012). The time dependent vehicle routing problem with time windows: Benchmark problems, an efficient solution algorithm, and solution characteristics. *Transportation Research Part E: Logistics and Transportation Review* 48(3), 616–636.
- Fleischmann, B. and J. Paraschis (1988). Solving a large scale districting problem: a case report. *Computers & Operations Research* 15(6), 521–533.
- Fry, M. and J. Ohlmann (2009). Route design for delivery of voting machines in hamilton county, ohio. *Interfaces* 39(5), 443–459.
- Gutin, G. and A. Punnen (2002). *The Traveling Salesman Problem and Its Variations*. New York: Springer-Verlag.
- Hadj-Alouane, A. and J. Bean (1997). A genetic algorithm for the multiple-choice integer program. *Operations Research* 45(1), 92–101.
- Hajek, B. (1988). Cooling schedules for optimal annealing. *Mathematics of Operations Research* 13, 311–329.
- Henderson, S. and A. Mason (2005). Ambulance service planning: simulation and data visualisation. *Operations Research and Health Care* 70, 77–102.
- Johnson, D., C. Aragon, L. McGeoch, and C. Schevon (1989). Optimization by simulated annealing: an experimental evaluation; part I, graph partitioning. *Operations Research* 37(6), 865–892.
- Johnson, D., C. Aragon, L. McGeoch, and C. Schevon (1991). Optimization by simulated annealing: an experimental evaluation; part II, graph coloring and number partitioning. *Operations Research* 39(3), 378–406.

- Kindervater, G. and M. Savelsbergh (2003). Vehicle routing: handling edge exchanges. In E. Aarts and L. J. (Eds.), *Local search in combinatorial optimization*. Princeton, NJ: Princeton University Press.
- King, D., S. Jacobson, E. Sewell, and W. Tam Chao (2012). Geo-graphs: an efficient model for enforcing contiguity and hole constraints in planar graph partitioning. *Operations Research* 60(5), 1213–1228.
- Kirkpatrick, S., J. Gelatt, C., and M. Vecchi (1983). Optimization by simulated annealing. *Science* 220(4598), 671–680.
- Krau, S. (1997). *Extensions du problem de weber*. Ph. D. thesis, Departement de Mathematiques et de Genie Industriel, Universite de Montreal.
- Mehrotra, A., E. Johnson, and G. Nemhauser (1998). An optimization based heuristic for political districting. *Management Science* 44(8), 1100–1114.
- Na, B. (2006). *Heuristic approaches for the no-depot k-traveling salesman problem with a minmax objective*. Ph. D. thesis, Texas A&M University.
- Ohlmann, J., J. Bean, and S. Henderson (2004). Convergence in probability of compressed annealing. *Mathematics of Operations Research* 29(4), 837–860.
- Ohlmann, J. and B. Thomas (2007). A compressed-annealing heuristic for the traveling salesman problem with time windows. *INFORMS Journal on Computing* 19(1), 80–90.
- Rosing, K. (1992). An optimal method for solving the (generalized) multi-weber problem. *European Journal of Operational Research* 58, 414–426.
- Saaty, T. (1972). Optimum positions for airports. *Naval Research Logistics Quarterly* 19, 101–109.
- Simchi-Levi, D., P. Kaminsky, and E. Simchi-Levi (2003). *Designing and Managing the Supply Chain*. New York: McGraw-Hill Irwin.

Thomas, B. and C. White (2004). Anticipatory route selection. *Transportation Science* 38(4), 473–487.

Yang, M., T. Allen, M. Fry, and W. Kelton (2013). The call for equity: simulation optimization models to minimize the range of waiting times. *IIE Transactions* 45(7), 781–795.

Yang, M., M. Fry, D. Kelton, and T. Allen (2013). Improving voting systems through service-operations management. *To Appear in Production and Operations Management*.