D-STOP

Technical Report 152

Project Title:

Data-driven, Real-time Traffic Signal Optimization: A Distributed Approach

Research Supervisor: Stephen Boyles
Center for Transportation Research

August 2020

Data-Supported Transportation Operations & Planning Center (D-STOP)

A Tier 1 USDOT University Transportation Center at The University of Texas at Austin

D-STOP is a collaborative initiative by researchers at the Center for Transportation Research and the Wireless Networking and Communications Group at The University of Texas at Austin.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
D-STOP/2020/152		

4. Title and Subtitle	5. Report Date
Optimal Placement of Reservation-Based Intersections in Urban Networks	August 2020
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Rahul Patel, Prashanth Venkatraman, and Stephen D. Boyles	Report 152
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Data-Supported Transportation Operations & Planning Center (D-STOP) The University of Texas at Austin School of Architecture 310 Inner Campus Drive, B7500 Austin, TX 78712	11. Contract or Grant No. DTRT13-G-UTC58
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
United States Department of Transportation University Transportation Centers 1200 New Jersey Avenue, SE Washington, DC 20590	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program.
Project Title: Data-driven, Real-time Traffic Signal Optimization: A Distributed Approach

16. Abstract

Reservation-based traffic control is a revolutionary intersection management system which involves the communication of autonomous vehicles and an intersection to request space-time trajectories through the intersection. Although previous studies have found congestion and throughput benefits of reservation-based control that surpass signalized control, other studies have found negative impacts at peak travel times. The main purpose of this paper is to find and characterize favorable mixed configurations of reservation-based controls and signalized controls in a large city network which minimize total system travel times. As this optimization problem is bi-level and challenging, three different methods are proposed to heuristically find effective mixed configurations. The first method is an intersection ranking method that uses simulation to assign a score to each intersection in a network based on localized potential benefit to system travel time under reservation control and then ranks all intersections accordingly. The second is another ranking method; however, it uses linear regression to predict an intersection's localized score. Finally, a genetic algorithm is presented that iteratively approaches high-performing network configurations yielding minimal system travel times. The methods were tested on the downtown Austin network and configurations found that are less than half controlled by reservation intersections that improve travel times beyond an all reservation controlled network. Overall, the results show that the genetic algorithm finds the best performing configurations, with the initial score-assigning ranking method performing similarly but much more efficiently. It was finally find that favorable reservation placement is in consecutive chains along highly trafficked corridors.

17. Key Words	18. Distribution Statement
Reservation-based control; genetic algorithm	No restrictions. This document is available to the public through NTIS (http://www.ntis.gov): National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Acknowledgements

The authors recognize that support for this research was provided by a grant from the U.S. Department of Transportation, University Transportation Centers.

Optimal placement of reservation-based intersections in urban networks

Rahul Patel, Corresponding Author
Graduate Research Assistant
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Phone: 512-471-3548, Fax: 512-475-8744; Email: rahulpatel@utexas.edu

Prashanth Venkatraman
Graduate Research Assistant
Department of Operations Research and Industrial Engineering
The University of Texas at Austin
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Phone: 512-471-3548, Fax: 512-475-8744; Email: prashv92@utexas.edu

Stephen D. Boyles
Associate Professor
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Phone: 512-471-3548, Fax: 512-475-8744; Email: sboyles@mail.utexas.edu

Word count: 6823 words text + 2 table x 250 words (each) = 7323 words

ABSTRACT

Reservation-based traffic control is a revolutionary intersection management system which involves the communication of autonomous vehicles and an intersection to request space-time trajectories through the intersection. Although previous studies have found congestion and throughput benefits of reservation-based control which surpass signalized control, other studies have found negative impacts at peak travel times. The main purpose of this paper is to find and characterize favorable mixed-configurations of reservation-based controls and signalized controls in a large city network which minimize total system travel times. As this optimization problem is bi-level and challenging, we propose three different methods to heuristically find effective mixed-configurations. The first method is an intersection ranking method uses simulation to assign a score to each intersection in a network based on localized potential benefit to system travel time under reservation control and then ranks all intersections accordingly. The second is another ranking method, however uses linear regression to predict an intersection's localized score. Finally, we present a genetic algorithm which iteratively approaches high-performing network configurations yielding minimal system travel times. We test the methods on the downtown Austin network and find configurations which are less than half controlled by reservation intersections that improve travel times beyond an all-reservation controlled network. Overall, our results show that the genetic algorithm finds the best performing configurations with the initial score-assigning ranking method performing similarly but much more efficiently. We finally find that favorable reservation placement is in consecutive chains along highly trafficked corridors.

Keywords: Autonomous vehicles, Reservation-based intersection control, Dynamic traffic assignment, Genetic algorithm

1 INTRODUCTION

Connected autonomous vehicles (CAVs) can potentially revolutionize urban traffic operations with new technologies currently being tested on public roads. Cooperative adaptive cruise control (*1),* reduced reaction times, and vehicle-to-vehicle communications could increase road capacity (*2, 3*) and stability by shortening following headways. Patel et al. (*4*) showed that increases in autonomous vehicle (AV) penetration could consistently result in reduced congestion on large-scale networks, assuming current traffic demand conditions. Despite the capacity increases, AVs could counteract these improvements by inducing additional demand (*5*). Additionally, the Braess (*6*) and Daganzo (*7*) paradoxes show that increases in capacity could lead to increased travel times due to rerouting. Nonetheless, wireless vehicle-to-infrastructure (V2I) communication in CAVs can further improve traffic operations with new traffic controls. Reservation-based intersection control (*8, 9*) uses V2I communications and the reduced safety margins of CAVs to increase intersection capacity and throughput, and is the focus of this paper. In some scenarios, reservation-based control using a first-come-first-serve (FCFS) policy reduced intersection delay and improved system travel times beyond optimized signals for a single intersection in microsimulation (*10*, *11*). Patel et al. (*4*) also observed this positive effect in large-scale networks, using mesoscopic dynamic traffic assignment (DTA). They also found that FCFS reservations performed better at lower demand levels due to the low intersection saturation, allowing for more progression.

However, there are scenarios where signals outperform FCFS reservations. Levin et al. (*12*) produces several examples of this phenomenon. One example involves a local road interrupting the progression of an intersecting arterial at high demand, which would be otherwise avoided with a signal. Such paradoxes may not seem apparent in large-scale networks due to alternate route choices, but may still exist and cause congestion. Therefore, FCFS reservations do not dominate signals, and an appropriate combination of the two should be found.

Thus far, no literature has attempted to optimize or evaluate this mixed-control on large-scale networks. Several studies have compared signals to reservation-based control, using microsimulation to model networks of just one or a few intersections (*8*, *10*). However, microsimulation is not tractable for large-scale networks, or may not capture dynamic selfish route choice. The main purpose of this paper is to find and characterize favorable subsets of FCFS reservations and signals in a larger network than has been previously studied (174 intersections). To this end, we develop different heuristic methods and evaluate their performance by solving for dynamic user equilibrium (DUE). It is valuable to know the characteristics of these favorable configurations and methods from a policy and planning standpoint as they can be generalized to other networks either to prioritize deployment of a limited number of reservation-based intersections, or to identify long-term configurations.

Although FCFS might not be the most efficient traffic control policy, it has been the focus of most reservation-based control literature (*8,10,11*) and could be extended to a wide range of other policies due to its generality. FCFS is an inherently fair policy and will likely remain a good candidate to be widely implemented in reservation-based control. Due to the large range of alternative policies extended from FCFS, it is nearly impossible to generalize the network effects of reservations using an arbitrary policy. We therefore assume a FCFS policy in this paper, detailed in Section 2.1. Note that the term *reservation(s)* is used throughout the rest of the paper and refers to FCFS reservation-based intersection control.

The contributions of this paper are as follows. We present and assess the effectiveness of several heuristic methods used to find favorable mixed-configurations of reservations and signals in a network. We then show that the paradoxical effects of FCFS reservations (*12*) exist in a large downtown network by identifying hybrid-configurations which reduce congestion beyond uniform reservation control in DTA. Finally, we develop general reservation intersection deployment strategies based on quantitative and qualitative observations.

These methods include several ranking methods which assign or predict a score for each intersection to encapsulate its potential benefit to system congestion under reservation vs. signal control. Additionally, a genetic algorithm (GA) is used to iteratively find effective mixed-configurations in a network. These methods are evaluated by solving DTA on a city network. Results show mixed-configurations that outperform the all-reservation case, with reservations placed in chains along highly demanded roads. We find that the GA provides the higher performing but slower results, compared to the ranking methods.

The remainder of this paper is organized as follows. Section 2 overviews previous work on reservation-based controls. Section 3 presents the optimization problem statement and assumptions. Section 4 presents methods used to find favorable mixed-configurations of reservations and signals. Section 5 details experimental results on the downtown Austin, TX network, and we conclude in Section 6.

2 RESERVATION-BASED INTERSECTION MODEL

This section first reviews the tile-based reservation control mechanism proposed by Dresner and Stone (*8, 9*) under a FCFS policy with possible paradoxical properties (*12*). Finally, we detail the simplified conflict region intersection model (*13*) used in this study's simulations to tractably model reservation-based control in large DTA networks.

2.1 FCFS Tile-Based Reservations

Dresner and Stone's (*8, 9*) proposed tile-based reservation mechanism relies on the V2I communication between CAVs and an intersection manager (IM) agent. Basically, the IM divides the intersection into a grid of space-time tiles. As CAVs enter the detection radius (CAVs must know their intersection arrival time), they make requests with the IM for a reservation to move through the intersection. The IM then simulates the vehicle's desired path through the tile grid. If there is no conflict with another vehicle's reserved path, the reservation is approved. Else, the reservation is rejected.

A control policy determines priority during conflicting requests. In this paper, we use a FCFS policy for reservation control, as do most other previous studies. FCFS is a fairness-based priority which grants reservations according to intersection arrival times. If a vehicle's request for a reservation is rejected due to conflict, the vehicle is delayed and the IM suggests a later time for safe traversal.

Although simple, FCFS properties can lead to paradoxes in reservation-based control. Levin et al. (*12*) produces three theoretical examples of signals outperforming reservations. The first shows that vehicles with lower priority and fewer conflict limitations could move before higher priority vehicles with more conflict limitations. For example, a vehicle on a small and empty local road approaching an intersection with a large arterial and long queue will move before someone farther back in the arterial queue, whereas a signal would give more green time to the arterial. The second exploits the property that vehicles cannot request a reservation unless they can execute it so a vehicle at the back of a platoon can't make a request until it can enter the intersection. The third shows once a request is reserved, any request that does not conflict with it can move as well, which can lead to vehicles moving in a different order than their requests.

Simulation on smaller arterial and freeway networks demonstrate this phenomenon. However, for the downtown DTA network tested in this paper, previous results suggest improvements to travel times for all demand scenarios tested, however at a decreasing rate of improvement with demand increase (*4*, *12*).

2.2 Conflict Region Model

To make large-scale DTA simulations tractable when modeling tile-based reservation control, we use a simplified conflict region model. The conflict region model (*13*) aggregates tiles into larger conflict regions, each limited by a capacity, and is able to capture the long-run behavior of reservation control. An example is shown in Figure 1. Additionally, we use the cell transmission model (*14*, *15*) to dynamically propagate flow on links.

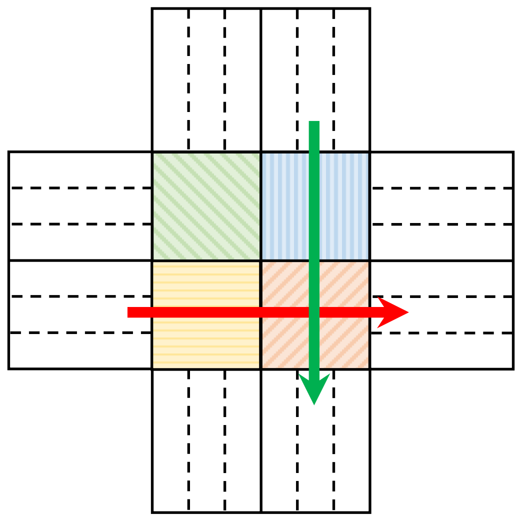

FIGURE 1 Conflict region representation of a four-way intersection, showing two conflicting turning movements (*13*)

3 PROBLEM STATEMENT AND ASSUMPTIONS

This section presents the bi-level optimization problem that is the focus of this paper and which we attempt to solve using heuristic methods presented in Section 4. We then state the major assumptions made for this paper.

Simply put, we want to achieve the lowest total system travel time ($TSTT$) in a network with the best configuration of reservations and signals. Equations 1 – 4 present the general bi-level optimization problem. The direct decision variable $\vec{z}$ defines a network-control configuration in which each z_i is an *eligible* intersection i's control (reservation or signal), and the indirect decision variable $\vec{x}$ is a DUE link flow mapping. E is the set of *eligible* intersections defined in the assumptions below, and is a subset of the set of all intersections in the network.

$$\min_{\vec{x},\vec{z}} \quad TSTT(\vec{x},\vec{z}) \tag{1}$$

$$s.t. \quad \vec{x} = \mathrm{F}(\vec{z}) \tag{2}$$

$$z_i \in \{0,1\} \quad \forall i \tag{3}$$

$$i \in E \tag{4}$$

1. The higher level is to minimize the $TSTT$ of a city network by assigning each *eligible* intersection's z_i as reservation-based $\{1\}$, or signalized $\{0\}$ control;
2. The lower level is to solve for DUE to obtain the link flows $\vec{x}$ for $TSTT(\cdot)$ as shown in Equation 2, where the function $\mathrm{F}(\vec{z})$ finds $\vec{x}$ by using DTA.

To clarify, a feasible solution to this problem $\vec{z}$ is a network with a subset of reservation intersections and a remaining subset of signalized intersections. The $TSTT$ of the configured network is the solution's performance measure and is used to evaluate network congestion effects. Because solving for DUE is itself a difficult problem, we propose methods to heuristically solve for optimal $\vec{z}$'s.

To compare and evaluate the methods, experiments are conducted with different proportions of reservation intersections α (alpha) on the same city network. For example, if our test network contains 100 *eligible* intersections, an $\alpha = 0.2$ requires 20 reservations and 80 signals. Here, E would be the set of 100 eligible intersections and $\sum_{i=1}^{100} z_i = 20$. This restriction also resembles application in practice as transportation authorities have budgets and will most likely deploy a limited number of reservation intersections.

Below, we list the assumptions made in this paper.

- The set of *eligible* intersections E whose controls can be switched is the set of currently signalized intersections in the real network. The City of Austin uses pre-timed signals downtown during the peak period. The model does not consider the set of merges, diverges, or stop sign controlled intersections because reservations provide little system-wide benefit when applied there (*4*). The signals in our model reflect the timings and phase patterns (including offsets for progression) currently used by the city;
- Only CAVs can use the reservation intersections, so all simulations are composed of 100% CAV demand;
- All reservation-based control uses a FCFS policy.

4 METHODS

This section details several ranking methods and a meta-heuristic method used to obtain solutions to the optimization problem (Equations 1 – 4). In addition, we identify differential measures which generalize an intersection's performance under reservation vs. signalized control.

The first two methods assign a score to each intersection which allows them to be ranked in order of the best reservation-control candidates. The scores are representative of potential benefit to $TSTT$ under reservation vs. signalized control, relative to other intersections in the network. To assign a score, the first method uses local intersection simulation results and the second uses weighted sums of intersection characteristics obtained from system simulation. The third is a more sophisticated ranking method that uses multilinear regression to predict an intersection's score found from the first method. It chooses a feasible $\vec{z}$ which maximizes the total predicted score using easily obtainable intersection characteristics. Finally, we propose a meta-heuristic genetic algorithm which iteratively moves toward higher performing $\vec{z}$ solutions. This method finds nice solutions, however provides few "pro-reservation" generalizations and is slowed by long computation times due to the required fitness-calculation of a DUE solution. On the other hand, ranking methods are easier to execute and can offer quantitative selection criteria, however may not guarantee good solutions.

4.1 Intersection Ranking Methods

This section details two methods which assign a score to each intersection i and rank them in order of their differential potential benefit to $TSTT$ under reservation control compared to signal control. The standalone scores may have limited realistic interpretations, however are used to compare intersections with each other and capture inefficient reservation behavior. These inefficiencies are typically seen in smaller networks with limited route choice, as this exploits FCFS paradoxes, and motivates more localized scores.

The first ranking method approximates an "effective sub-system travel time" ($\Delta SSTT$) score by locally simulating each intersection under reservation and signal control. The $\Delta SSTT$ score is shown in Equation 5as the difference between two sub-system travel times ($SSTT$). For each i, an $SSTT_{sig}$ and $SSTT_{FCFS}$ are obtained by solving DUE on a subnetwork consisting of only i and its incoming and outgoing links with $z_i = 0$ and $z_i = 1$, respectively. For all subnetworks, origin-destination (OD) demands are obtained by solving DUE on the whole parent network with $z_i = 0 \; \forall i$ and extracting the individual intersection's flows from the parent $\vec{x}$. OD demand is gathered from the all-signal case as it is representative of current real-world conditions, before any reservation-control has been implemented. This score estimation process is illustrated in Figure 2 for a single intersection. Though this method involves the DTA simulation of every eligible intersection subnetwork at least twice, the subnetworks are very small and have little demand compared to their parent city network. Single-intersection subnetworks, however, assume no interdependencies between intersections and, as presented in the following section, may be difficult to predict with a linear regression trend.

$$\Delta SSTT = SSTT_{sig} - SSTT_{FCFS} \tag{5}$$

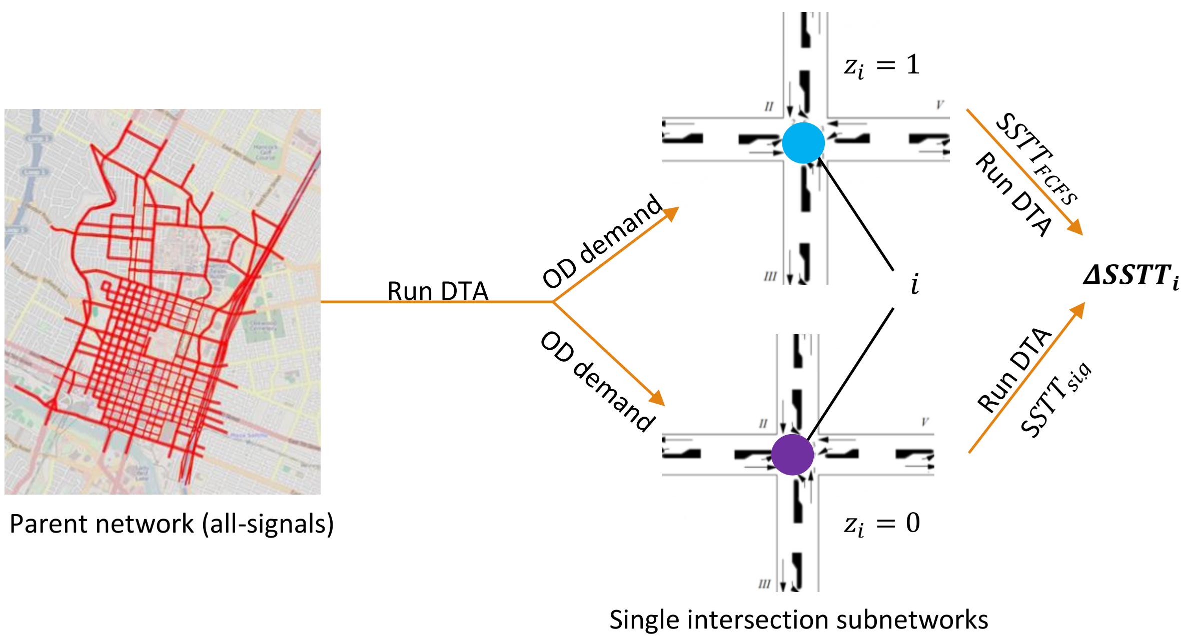

FIGURE 2 Method of data collection for one intersection's effective sub-system travel time

The second ranking method uses a simplified score composed of a weighted combination of intersection turning demands. Several results presented in Section 5 show that through, left, and right demands were the most significant predictors of reservation-based performance for any intersection. If combined effectively according to relative importance, demand-based scores (d_{score}) can be calculated for intersections. However, finding these initial weights may require an existing favorable $\vec{z}$ solution. Equation 6 below defines a possible d_{score}, where $\mu_{t_{FCFS}}$ is the average through, left, or right turning demand of all $z_i = 1$ from an existing $\vec{z}$ solution. Similarly, $\mu_{t_{sig}}$ is the same average, but of all $z_i = 0$. These two averages form a constant weight which is applied to each i's turning demands d_t to get a score for each i. As will be shown, favorably selected reservation intersections tend to have much higher turning demand than signalized intersections. Although this ranking method requires an existing $\vec{z}$ solution, it can prove powerful once effective weights are obtained as it only requires intersection turning demands extracted from the parent network's $\vec{x}$ DUE solution. As mentioned in Section 5.2, this method offers less than average performing configurations, but is very time efficient using the most significant predictors of reservation control benefit.

$$d_{score} = \sum_{t \in \{through,\ left, right\}} \frac{\mu_{t_{FCFS}}}{\mu_{t_{sig}}} * d_t \tag{6}$$

Other weighted scores may be formulated using other intersection characteristics. If an effective score which captures the differential performance of an intersection under reservation vs. signal control can be found, the ranking of intersections is intuitive and can allow for simple deployment strategies. Finding an efficient and accurate ranking method can be difficult however, due to few readily available intersection metrics which may also ignore the interdependent complexities that intersections may share with each other.

4.2 Multilinear Regression for Scoring

This section presents another, more complex intersection ranking method which uses a multilinear regression to predict an intersection's $\Delta SSTT$ score. If effective and extensible, this method may allow "pro-reservation" intersections to be generalized and make for easy reservation deployment strategies as the regression can be used on any signalized intersection with easily obtainable characteristics.

As shown in Section 5, regression rankings performed worse than all other methods in terms of minimizing $TSTT$, however the original $\Delta SSTT$ data performed well. Despite this, we are still able to draw generalizations from significant predictor variables.

4.2.1 Formulation

Essentially, the linear regression predicts an intersection's $\Delta SSTT$ score using predictor variables which characterize the currently signalized intersection in the real network. Recall that the higher the $\Delta SSTT$, the more likely an intersection is to benefit local congestion under reservation control beyond signal control. Results presented in Section 5 show that the first mentioned $\Delta SSTT$ ranking method's $\vec{z}$ solutions perform quite well making this score favorable to predict as it also encapsulates localized congestion effects.

Predictor variables were chosen based on descriptive power of the real signalized intersection as well as ease of collection. All predictor variables are described in Table 1 and are moderately easy to obtain from city and transportation authorities, or through simulation. Cumulative through, left, and right turn demand are the only variables obtained through DTA simulation and are later shown to be the most significant. Link length variables are considered because previous studies indicate that FCFS reservation inefficiencies are exacerbated by queue spillback onto close-proximity roads and intersections. Several signal and traffic control-specific variables such as cycle length and number of unrestricted turning movements are considered in an effort to surface inefficiencies in current controls.

The general regression formula is as follows in Equation 7, where $\Delta SSTT^*$ is the predicted "effective sub-system travel time" score, $\vec{\beta}$ is the vector of estimated variable coefficients, $\vec{X}$ is the vector of predictor variables, and $FFTT$ (free-flow travel time) is the regression constant.

$$\Delta SSTT^* = FFTT + \vec{\beta} * \vec{X} \tag{7}$$

4.2.2 Regression training

The dataset used to estimate variable coefficients consists of $|E|$ entries, each of which contains an intersection i's mentioned predictor variables and $\Delta SSTT$ score. The $\Delta SSTT$ for each i is obtained using the $\Delta SSTT$ ranking method from Section 4.1. We train two separate regressions using our testbed network data and a different downtown network's data, and apply both to the same testbed network. The non-testbed-trained regression is estimated to evaluate the extensibility of the regression to intersections in other networks and the testbed-trained regression assesses data-fitting. Section 5.2 details a regression model trained on the Dallas network.

TABLE 1 Multilinear Regression Predictor Variables Considered

Predictor Variable	**Variable Description**	**Units**
Number of phases	The total number of signal phases across a cycle	Number of phases
Cycle length	The time of one complete signal phasing cycle	seconds
Number of moves	The total number of non-restrictive turning movements for the intersection. Turning movements are defined by an approach link and an exit link.	Number of turning movements
Number of through turns	The total cumulative through demand of the intersection across all approaches	Number of vehicles
Number of left turns	The total cumulative left turn demand of the intersection across all approaches	Number of vehicles

Number of right turns	The total cumulative right turn demand of the intersection across all approaches	Number of vehicles
Minimum length	The minimum length of a link entering or exiting the intersection	Length in feet
Maximum length	The maximum length of a link entering or exiting the intersection	Length in feet
Average length	The average length of a link entering or exiting the intersection	Length in meters
Minimum link capacity	The minimum capacity of a link entering or exiting the intersection	Number of vehicles/hour
Total link capacity	The total cumulative capacity of all links entering or exiting the intersection	Number of vehicles/hour

4.3 Genetic Algorithm

This section presents a genetic algorithm (GA) which, unlike previous methods, attempts to directly solve the bi-level optimization problem heuristically. We implement a GA that evaluates and alters configurations of the same network using DTA and a "survival of the fittest" policy to determine improvement search directions to iteratively approach an optimal $\vec{z}$ solution. We begin with a general introduction to GAs followed by the formulation of our own GA model. In addition to finding $\vec{z}$ solutions with fixed reservation proportions (α), we formulate an "unconstrained" GA which allows for change in α.

Although this method by far requires the most computation time of any other method presented due to solving DTA many times, it generally provides the best $\vec{z}$ solutions at each reservation proportion.

4.3.1 A background on genetic algorithms

A genetic algorithm is a class of metaheuristic computational methods inspired by genetic evolution used to solve constrained and unconstrained optimization problems. The algorithm starts with an initial population of individuals, measures each's performance, and then iteratively creates new and better performing generations by combining the best traits of older generations. The algorithm then theoretically ends with the best performing individual.

4.3.2 Formulation

This section details our implementation of a GA to directly and heuristically solve the bi-level optimization problem stated in Equations 1 – 4. We first detail our constrained GA which is used for the bulk of experimentation, and then present the modified unconstrained version.

At the root of our GA implementation, each individual n in the population N is a different feasible configuration, $\vec{z}_n$ of the same network. Each individual possesses $|E|$ total genes which are defined by $z_i^n \in \{0, 1\}\ \forall i \in E$ and are what the GA modifies during initial population generation, crossover, and mutation. $TSTT_n$ is used as an individual's fitness value (or effectiveness), and is found by solving DUE ($\mathrm{F}(\vec{z}_n)$) using DTA to obtain $\vec{x}_n$ and then

$TSTT(\vec{x}_n, \vec{z}_n)$. Because this GA is constrained, we enforce $\sum_i z_i^n = r \;\; \forall n \in N$, where r is the number of required reservations found by $r = |E| * \alpha$. The following is an overview of the algorithm's steps.

1. *Initial population*: Generate an initial population of h randomly generated individuals, each satisfying $\sum_i z_i^n = r$;
2. *Population evaluation:* Calculate the $TSTT$ $\forall n \in N$ and then rank N in order of $TSTT$. Store the individual with the lowest $TSTT$ as the best.
3. *Parent selection*: Select the best performing proportion k of N as eligible parents and eliminate the bottom $1 - k$ proportion. Randomly choose $|N| * (1 - k)$ pairs of parents from the eligible list, removing parents as they are chosen;
4. *Crossover*: To create a new child, iterate through each z_i of both selected parents. For the constrained GA, randomly choose an i to look at. If $z_i^{Parent\ 1} = z_i^{Parent\ 2}$, then give z_i^{Child} the same control. Else, use the crossover probability p, shown by Equation 8 below, to determine the child's control.

   p is a linear pdf that creates a $p \in [0.5, 1.0]$ and gives the child a higher probability of inheriting the higher performing parent's control as the difference in $TSTT$ between the two parents increases.

$$p = 0.5 + 0.5 * \frac{|TSTT_{Parent\ 1} - TSTT_{Parent\ 2}|}{TSTT_{all-signals} - TSTT_{all-reservations}} \tag{8}$$

   Do this until the r limit has been reached for the child and assign everything else as signals;
5. *Mutation*: Each new child is chosen to be mutated with probability m. If chosen, each z_i^n of the individual has a probability b of being switched to the opposite control. This is to introduce randomness into the population and avoid falling into a local minimum;
6. *Children fitness evaluation*: Find the $TSTT$ of each new child and add them to the set of parents to create the new N and re-rank N. The individual with the lowest $TSTT$ is stored and the algorithm loops back to Step 3 for u iterations.

Next, the unconstrained GA essentially follows the same steps as the constrained, however, the initial population has no $\sum_i z_i^n = r \;\; \forall n$ constraint and each z_i^n has an equally likely chance of being $\{0\}$ or $\{1\}$. Then, at each crossover and mutation step, every z_i is considered. The unconstrained GA theoretically approaches the highest-performing network configuration which yields the minimum possible $TSTT$, however we later show during experimentation that it falls into a possible local minimum and is eventually outperformed by the "constrained" GA and $\Delta SSTT$ ranking at even lower reservation proportions.

Because the GA uses no intersection-specific characteristics or performance measures, it essentially just provides a highly effective $\vec{z}$ and doesn't offer many "pro-reservation" generalizations. For this reason and long computation times, GA solutions are primarily used as a benchmark for high $\vec{z}$ performance and is the main method used to identify visual control-placement trends seen in the testbed network in Section 5.5.

5 EXPERIMENTAL RESULTS

This section presents experimental results of testing the $\Delta SSTT$ ranking, linear regression ($\Delta SSTT^*$) ranking, and GA methods on the large-scale city network of downtown Austin, TX. All methods obtain feasible mixed-configurations ($\vec{z}$) of reservations and signals in the network in an effort to reduce congestion and minimize $TSTT$, evaluated using DTA.

We first show network-specific implementations of each method and their $TSTT$ results. We then compare the methods in terms of effectiveness and efficiency. We finally link visual and quantitative network-wide intersection trends, finding better reservation placement in consecutive chains on highly trafficked streets.

The downtown Austin network used for all experimentation, shown in Figure 3, contains 1,247 links, 546 nodes (174 signalized intersections), 171 zones, and 62,783 vehicle trips over a 4-hour observation period. This network includes several large arterials and a large downtown grid. This is a useful testbed as flow on the grid is primarily restricted by intersections, and the large network allows for alternative route choices. In addition, to train a regression, we use the downtown Dallas network containing 152 signalized intersections and 167,592 vehicle trips over a 4-hour period. The DTA models used in this paper are described in Section 2.2 and solved using the method of successive averages to a 2% gap, defined in Equation 9. The $shortest\ path\ time$ refers to a total travel time experienced if all demand were to be loaded onto the simulation's current shortest paths.

$$gap = \frac{TSTT - shortest\ path\ time}{TSTT} \tag{9}$$

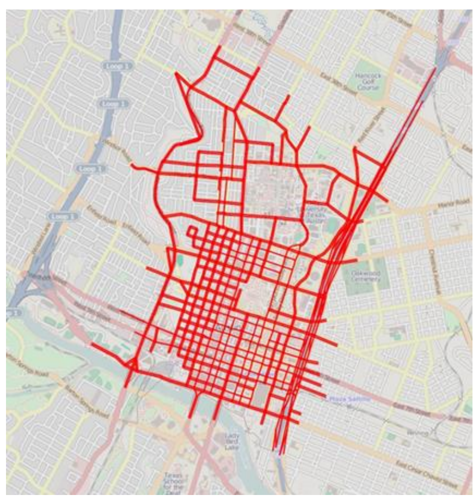

FIGURE 3 Downtown Austin network

Experiments were run for every method at each $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$ (defined in Section 3). For comparison, we simulate the Austin network with both all-signals and all-reservations yielding $TSTT$s of 6443.22 hrs and 4560.14 hrs respectively, labeled in Figure 4. We also test a random configuration method in which, for each α, we evaluate 10 randomly generated $\vec{z}$ solutions and average their $TSTT$s, also shown in Figure 4 as Random. This method is used as a benchmark given that effective methods should at least beat complete randomization.

5.1 $\Delta SSTT$ Ranking Results

This ranking method uses subnetwork simulation results to assign a $\Delta SSTT$ score to each eligible intersection in a network to capture a localized benefit to travel times under reservation vs. signal control. The 174 eligible intersections were then ranked in order of descending scores and the top $\alpha \times 174$ are assigned as reservations and the rest signals.

The $\Delta SSTT$ ranking method performed well on the Austin network, improving $TSTT$ beyond the all-reservation case with just over a 40% reservation proportion and clearly outperforming the Random method. This trend continued as the $\Delta SSTT$ configurations decreased in $TSTT$ at a decreasing rate as α increased.

All previous experimental studies with this network showed improved travel times, with no signs of paradoxical inefficiencies associated with reservation-based control (*13*, *4*). This experiment's results show that such paradoxes can exist in a large-scale DTA network with more congestion benefits than the all-reservation case at less than half the reservation control. This also supports the validity of using $\Delta SSTT$ scores to train a regression, detailed in the next section.

5.2 Multilinear Regression Scoring ($\Delta SSTT^*$ Ranking) Results

The $\Delta SSTT^*$ ranking method uses linear regression to predict an intersection's $\Delta SSTT$ score ($\Delta SSTT^*$) given a set of significant but easily obtainable predictor variables, $\vec{X}$. Two regressions are estimated, one regressing Dallas intersection data and the other regressing Austin data. The purpose of using a Dallas regression to predict Austin scores is to test transferability of one city's reservation intersection behavior to another's making deployment easier in practice. This may also surface common trends seen in variables. The purpose of an Austin regression predicting its own scores is to validate the linear trend assumption. $\Delta SSTT$ training data is obtained as described in Section 4.1 and predictor variables are obtained from the City of Austin and simulation, described in Section 4.2.

TABLE 2 Dallas-Trained Multilinear Regression Summary

Variable	β (coefficient)	Std. error	t-score
(Constant)	-717.3	-717.3	-717.3
Cycle length	3.286	3.286	3.286
Number of moves	9.495	9.495	9.495
Number of Through turns	0.261	0.261	0.261
Number of left turns	0.43	0.43	0.43
Number of right turns	0.414	0.414	0.414
Minimum length	0.409	0.409	0.409

Table 2 details the Dallas-trained regression model which includes only the significant predictors of $\Delta SSTT$ from the pool in Table 1. Relative significance of variables was evidenced from t-values at a 95% confidence level ($|t_{var}| \geq 1.645$). The model had an R = 0.868, R^2 = 0.754, adjusted R^2 = 0.752, and standard error of the estimate = 360.818. It is evident that cycle length and all three turning demand variables proved to be significant predictors in the model.

The cycle length coefficient is positive implying an increase in $\Delta SSTT$. Demand is expectedly significant and positive as major arterials tend to have higher through demands and large queue spillback at peak times. A positive coefficient suggests an increase in $\Delta SSTT$ with increased turning demand, implying more benefit under reservation control. The regression indicates that heavily demanded intersections perform better as reservations.

Two experiments were run on Austin's network, each with intersections ranked according to either the Austin-trained or Dallas-trained regression. Both $\Delta SSTT^*$ rankings performed similarly, as shown in Figure 4. Although $TSTT$s steadily decreased as α increased, this was to be expected and the $\Delta SSTT^*$ rankings performed worse than even the Random method. The result shows a linear trend cannot be fit to the $\Delta SSTT$ scores. Complex interdependencies between proximal intersections most likely attribute to this non-linear trend.

5.3 Genetic Algorithm Results

The GA takes in a set of model parameters and iteratively tends towards optimal $\vec{z}$ solutions with minimal travel times. In this paper, we use a custom Java GA code to create our model. Model parameters used in the Austin network experiments include an initial population $h = 100$, eligible parent proportion of the population $k = 0.75$, individual mutation probability $m = 0.1$, and gene mutation probability $b = 0.07$. Given parameters were found based on trial-and-error methods and computation time assumptions.

Results in Figure 4 show that the GA overall obtained the best results. The GA mostly outperformed the $\Delta SSTT$ ranking method with larger improvements over the method at lower reservation proportions, however the two came close in $TSTT$ performance and the $\Delta SSTT$ method marginally beat the GA at $\alpha = 0.8$. At just under $\alpha = 0.4$, the GA also outperforms the all-reservation case.

An additional unconstrained GA case was run which attempted to solve the bi-level optimization problem defined in Section 3 using any proportion of reservations. This unconstrained GA therefore approaches the optimal α as well. The resulting configuration had $\alpha = 0.86$ and a $TSTT$ of 4229.2 hours which was marginally outperformed by both the constrained GA and the $\Delta SSTT$ ranking method at a lower $\alpha = 0.8$.

Figure 5 shows the GA's performance for the unconstrained, $\alpha = 0.2$, and $\alpha = 0.4$ cases over 100 iterations, with the latter two showing slightly more of a flattening in $TSTT$. Though the steeper convergence graph may imply opportunity for more improvement, Figure 6 shows a relatively steady increase in reservation proportion over the iterations, possibly leading to a local minimum.

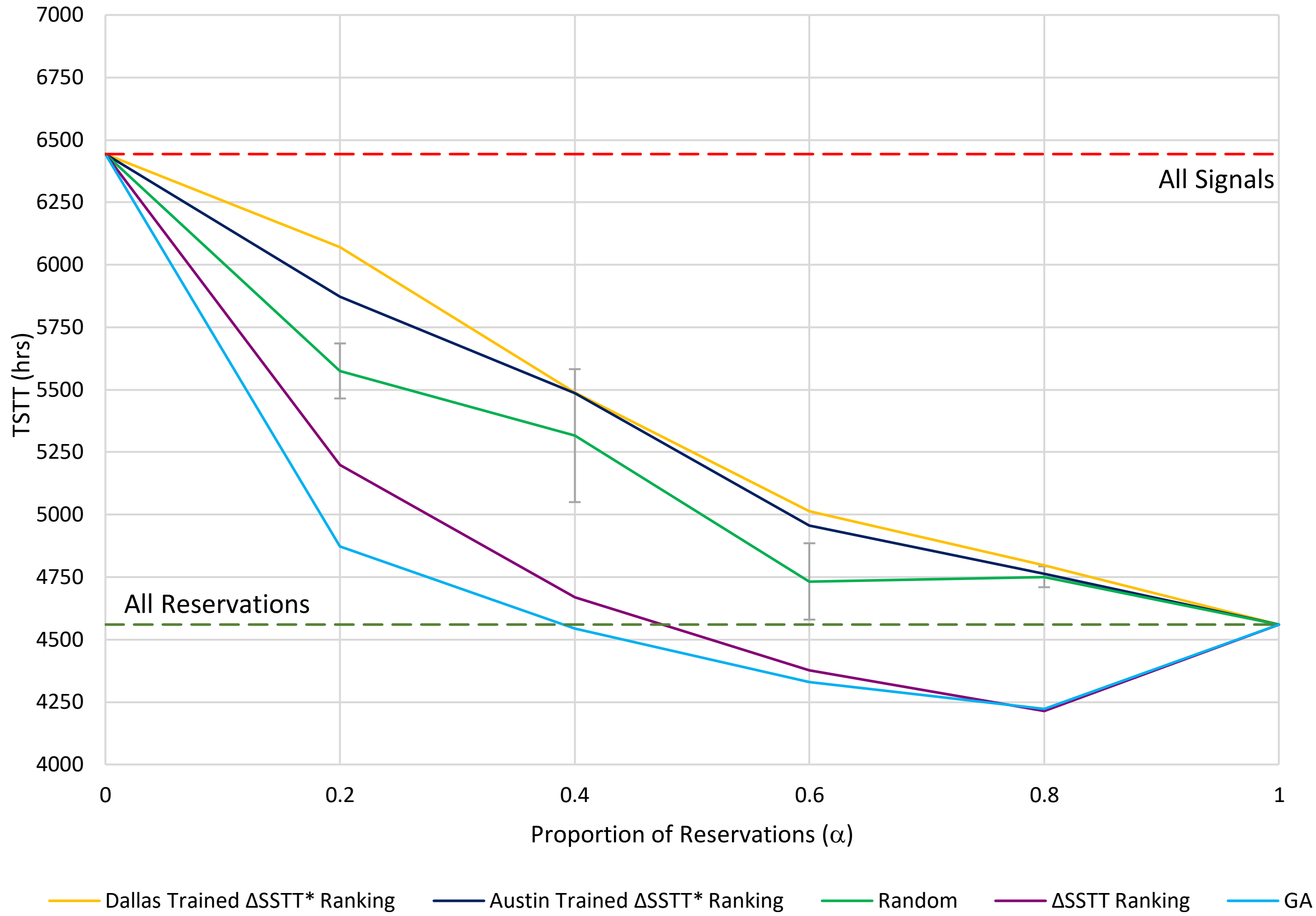

FIGURE 4 Downtown Austin results summary

5.4 Comparative Performance of Methods

As shown by results, the GA obtained the highest performing and most effective $\vec{z}$ which had the lowest $TSTT$s, the $\Delta SSTT$ ranking method obtained very close travel times, and finally the $\Delta SSTT^*$ ranking method had the worst travel times. However, the most effective methods were not necessarily the most efficient methods.

Though the GA provided results with the most system-wide benefit, it was by far the most computationally expensive method. 100 iterations of the GA meant 5100 runs of DTA (100 initial population + 50 new children/iteration) to solve DUE on the same large-scale network. At an average run's convergence time of 15 min/run, a single GA result requires about 22 hours of computation. On the other hand, the $\Delta SSTT$ ranking method achieved results similar to the GA and is much more time efficient. Although we are running DTA on 174 subnetworks under both controls, the single-intersection subnetworks each converge in 2-3 seconds, making for a conservative computation time of 17.5 minutes to obtain all scores. Finally, the $\Delta SSTT^*$ method is the most time efficient of the three as it only entails applying a regression equation to a data set, but the discovered solutions perform worse than even randomly generated ones. However, the regressions revealed important "pro-reservation" intersection characteristics which may allow for development of better methods.

Note that Section 5 does not include the d_{score} ranking method as it did not yield significant results. Turning demand coefficients were found based on GA result data and intersections were ranked based on the calculated d_{score}'s. This method was outperformed by the $\Delta SSTT$ ranking method, however performed better than randomized configurations. The d_{score} method's minimal computation time does not outweigh the predictive power of the $\Delta SSTT$ ranking. Because of this and because the method didn't reveal any additional "pro-reservation" intersection characteristics, it was not tested further. For reference, at $\alpha = 0.2$ and 0.4 the d_{score} method gave a $TSTT$ of 5458.2 hrs and 4950.0 hrs respectively.

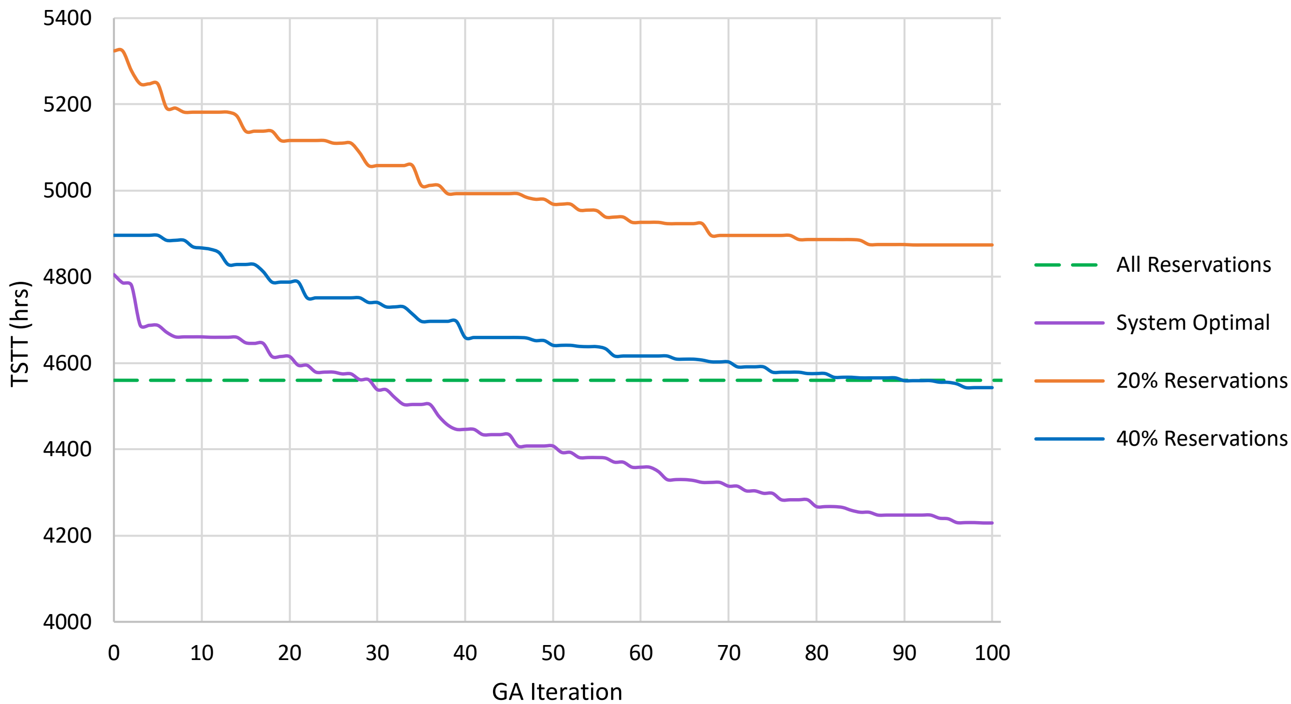

FIGURE 5 GA performance over 100 iterations

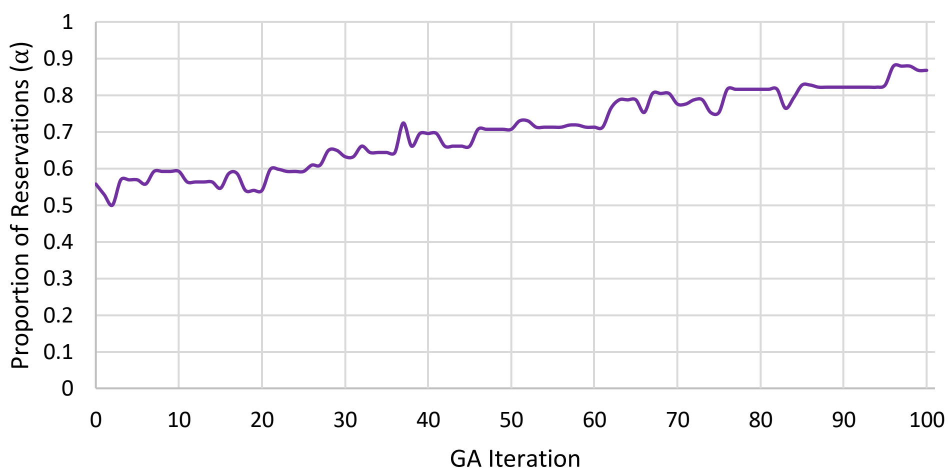

FIGURE 6 Unconstrained GA variation of α over 100 iterations

5.5 Trends in Reservation-based Intersection Placement

Because finding many quantitative trends and metrics in reservation placement is difficult, it is hard to develop deployment strategies solely on these metrics to be used by transportation planners and policymakers. Using visual observations and observed quantitative data, we are able to generalize trends in reservation placement. We take the highest performing $\vec{z}$ solutions from the constrained GA and place them on an Austin city street map.

Figure 7 shows mappings of the GA's resulting configurations with reservation (TBR) intersections in green and signalized intersections in red. Even though all GA experiments are independent of each other, we see that every reservation from the $\alpha = 0.2$ (35 reservations) case remained a reservation (except for 1) in the $\alpha = 0.4$ (70 reservations) case. This overlap supports similar configuration patters seen in $\Delta SSTT$ ranking results. We notice that reservation intersections were typically kept together, typically in consecutive chains or corridors. We also notice that these chains are along highly congested corridors in the peak periods such as 15^{th} St, Cesar Chavez St, Lamar Blvd, Congress Ave and MLK Blvd. GA results show almost 5.2 times the number of through turns on average at reservation intersections compared to signalized intersections and 2 to 4 times the number of left and right turns.

As we move to higher reservation proportions, the reservation chains began to intersect. On the right map of Figure 7, reservation chains going from 15^{th} St, MLK Blvd, Cesar Chavez St and others go directly to large orthogonal arterial and freeway roads (Lamar and I35 frontage road).

These reservation chains are seemingly placed at these locations to promote progression of major arterial streets and avoid potential FCFS inefficiencies previously seen (*12*). With multiple reservations in a row, a progression similar to that of pre-timed signals is possible and could prevent queue spillback onto smaller streets as many attempt to enter arterials during peak periods.

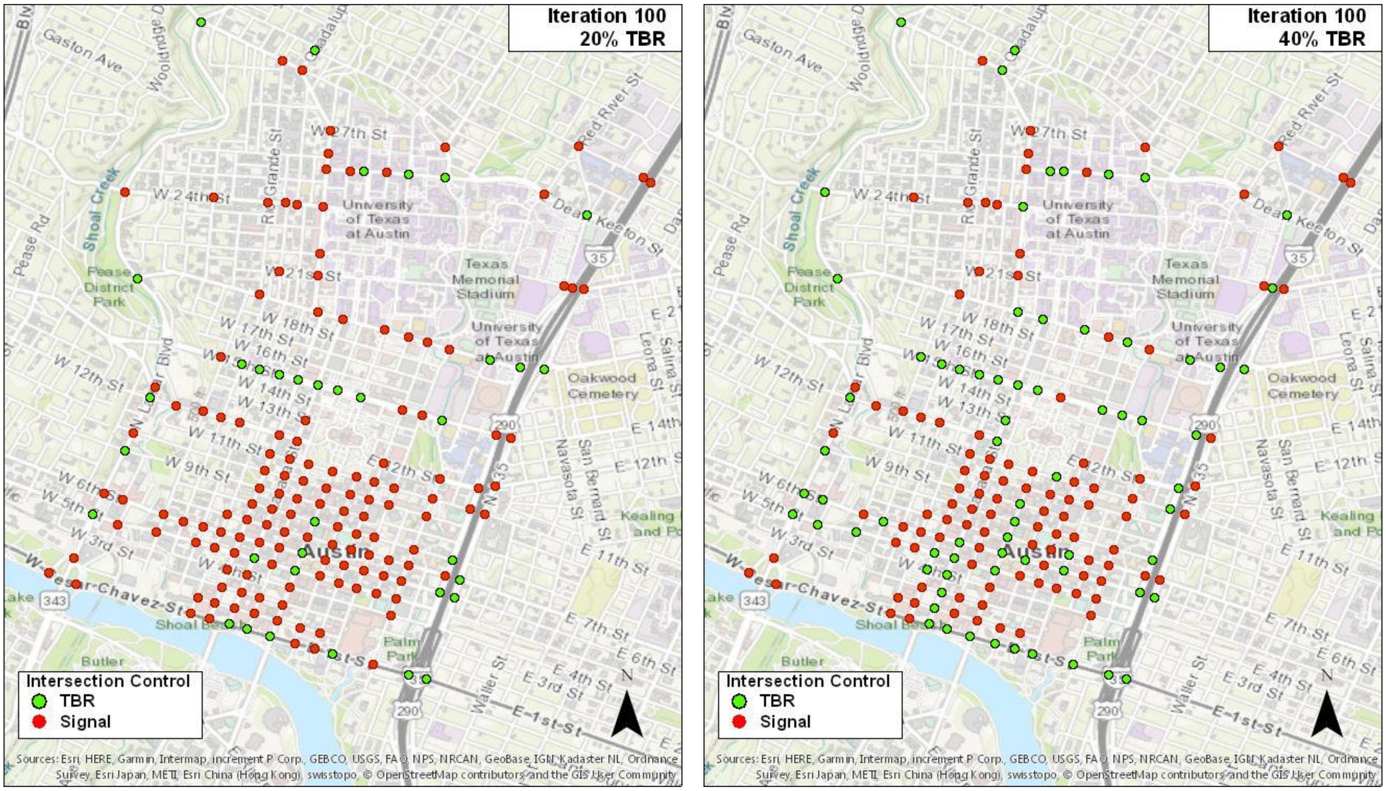

FIGURE 7 Control configurations found by the constrained GA with $\alpha = 0.2$ (left) and $\alpha = 0.4$ (right)

6 CONCLUSIONS

This paper presented and tested methods for finding favorable mixed-control configurations of FCFS reservation-based and signalized intersections in the large-scale city network of downtown Austin. The optimization problem of minimizing $TSTT$ is challenging as it requires solving for DTA, so we proposed several heuristic methods. We present three different methods for obtaining favorable network configurations including an effective sub-system travel time ranking, a multilinear regression intersection ranking, and a genetic algorithm.

First, a ranking method assigns scores to intersections ($\Delta SSTT$) which represent a differential performance measure of the individual intersection under reservation vs. signal control in terms of travel time. Austin test intersections were then ranked accordingly and results show the method worked well to improve travel times, outperforming the all-reservation case but with just over 40% reservations. This method proved more tedious than applying a readily available regression equation, however is still relatively quick.

Next, a multilinear regression was trained with data from a separate downtown Dallas network, to test extensibility of the regression to other networks, with its predictor variables being easily attainable intersection characteristics. Significant variables were primarily turning demand-related except for signal cycle length. Austin intersection $\Delta SSTT^*$ scores were predicted using the regression and were ranked accordingly. However, when tested in simulation, regression ranking results did not perform well and were outperformed by a random intersection selection method. The dependent variable was concluded to not fit a linear trend, however if correct, a regression equation could prove useful as it would allow any set of signalized intersections to be ranked easily.

Finally, a genetic algorithm to successively create better performing configurations is proposed. The GA solved DTA to evaluate fitness over 100 iterations, proving to be very computationally expensive requiring nearly 22 hours to complete. This method also only gives a solution and no insight into significant “pro-reservation” characteristics. However, the GA provided the lowest $TSTT$ results, just marginally lower than the $\Delta SSTT$ ranking method, also beating the base all-reservation case with 60% less smart intersections.

Mapping the GA results revealed a placement of reservation intersections in chains of consecutive reservations along very highly congested roads at peak hours. This most likely was to provide progression along large arterials and mitigate paradoxical effects seen with FCFS reservations. These trends and congestion benefits can be very useful in terms of planning and policy, especially with the deployment of reservations into our infrastructure.

These results and methods motivate the need for further analysis of reservation performance trends and intersection characteristics for proper deployment techniques. Although FCFS performs well in some situations (*4, 8, 10*), it does worse than signals in others and these results can be taken further to develop system optimal control policies. Although improvements were seen in mixed-configurations, further mesoscopic modeling studies on other reservation-based control policies would be likely more efficient. Additionally, the assumption of a 100% CAV penetration rate may not be achieved until well into the future. For this reason, further experimentation needs to be done using hybrid-reservation control as some work has shown its inefficiency compared to fully autonomous reservation control (*16*).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Data-Supported Transportation Operations and Planning Center and the National Science Foundation under Grant No. 1254921

CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: R. Patel, P.Venkatraman; data collection: R. Patel, P. Venkatraman; analysis and interpretation of results: R. Patel, S.Boyles; draft manuscript preparation: R.Patel. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Van Arem, B., Van Driel, C.J., Visser, R., 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. IEEE Trans. Intell. Transport. Syst. 7 (4), 429–436.

2. Kesting, A., Treiber, M., & Helbing, D. (2010). Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *368*(1928), 4585-4605.

3. Shladover, S., Su, D., & Lu, X. Y. (2012). Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (2324), 63-70.

4. Patel, R., Levin, M. W., & Boyles, S. D. (2016). Effects of autonomous vehicle behavior on arterial and freeway networks. *Transportation Research Record: Journal of the Transportation Research Board*, (2561).

5. Levin, M.W., Boyles, S.D., 2016. Intersection auctions and reservation-based control in dynamic traffic assignment. Transport. Res. Rec.: J. Transport. Res. Board (2497), 35–44

6. Braess, P. D. D. D. (1968). Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, *12*(1), 258-268.

7. Daganzo, C. F. (1998). Queue spillovers in transportation networks with a route choice. *Transportation Science*, *32*(1), 3-11.

8. Dresner, K., & Stone, P. (2004, July). Multiagent traffic management: A reservation-based intersection control mechanism. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2* (pp. 530-537). IEEE Computer Society.

9. Dresner, K., & Stone, P. (2006, July). Traffic intersections of the future. In *Proceedings of the National Conference on Artificial Intelligence* (Vol. 21, No. 2, p. 1593). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

10. Fajardo, D., Au, T. C., Waller, S., Stone, P., & Yang, D. (2011). Automated intersection control: Performance of future innovation versus current traffic signal control. *Transportation Research Record: Journal of the Transportation Research Board*, (2259), 223-232.

11. Li, Z., Chitturi, M., Zheng, D., Bill, A., & Noyce, D. (2013). Modeling Reservation-Based Autonomous Intersection Control in VISSIM. *Transportation Research Record: Journal of the Transportation Research Board*, (2381), 81-90.

12. Levin, M. W., Boyles, S. D., & Patel, R. (2016). Paradoxes of reservation-based intersection controls in traffic networks. *Transportation Research Part A: Policy and Practice*, *90*, 14-25.

13. Levin, M.W., Boyles, S.D., 2016. Intersection auctions and reservation-based control in dynamic traffic assignment. Transport. Res. Rec.: J. Transport. Res. Board (2497), 35–44.

14. Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, *28*(4), 269-287.

15. Daganzo, C. F. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79-93.

16. Sharon, G., Boyles, S. D., & Stone, P. (2017). Intersection Management Protocol for Mixed Autonomous and Human-Operated Vehicles. *Transportation Research Part C: Emerging Technologies (Under submission TRC-D-17-00857)*.